AI & SOCIETY
https://doi.org/10.1007/s00146-026-02967-1

REVIEW

Conditions of benefits and risks when algorithmic technology is implemented for public sector policing and fraud detection: a systematic literature review

Spencer Poodiack Parsons[1] · **René Torenvlied**[1]

Received: 26 May 2025 / Accepted: 28 February 2026
© The Author(s) 2026

Abstract
Across the globe, cities, states, and nations are moving forward with implementing artificial intelligence and machine learning technologies for policing and fraud detection, motivated by an optimistic view that technology can improve the functioning of government. Yet, their implementations have sometimes led to disastrous outcomes for citizens, leading many to view these technologies with deep skepticism. We bridge this divide by conducting a multi-disciplinary systematic literature review ($n = 157$) focusing on specific conditions for benefits and risks associated with the implementation of artificial intelligence and machine learning technologies in predictive policing and fraud detection and prevention. We find that the social science literature generally and, more specifically, the field of Public Administration focuses on the risks of algorithmic technology, and we identify conditions of risks, such as when algorithmic systems are designed with a model of threat. The engineering literature instead focuses on potential benefits, and we identify conditions of benefits, including alignment with legal and policy objectives. We integrate these conditions into a socio-technical governance framework that conceptualizes technical system quality, human–technology interaction, and institutional context as interacting mechanisms shaping both decision outcomes and institutional legitimacy. This framework highlights the importance of closer collaboration between data scientists and social scientists to design and implement policing and fraud detection systems that better align with public and societal values.

Keywords Artificial intelligence · Machine learning · Predictive policing · Fraud detection

1 Introduction

In the United States, open government job positions related to artificial intelligence doubled in 2024 compared to the previous two years (Alder 2024). This increase in hiring signaled a broad government interest in adopting algorithmic technologies, such as artificial intelligence and machine learning. Although this adoption is not uniform across countries, there are clear reasons why a government would want to employ algorithms (Wang et al. 2022). One reason would be a government's interest in protecting its citizenry from undesirable, malevolent, or even criminal activities.

The goal of increased effectiveness in protecting citizens led to the development of algorithmic-based predictive policing technologies, beginning with the implementation of CompStat in New York City in 1994 (Tulumello and Iapaolo 2022). Predictive policing systems also align with government efficiency, as they can potentially optimize the distribution of police resources in complex environments (Benbouzid 2019). Additionally, predictive policing systems are built to reduce and prevent crime, a prominent demand of citizens (Domínguez and Scartascini 2022; Picasso and Cohen 2018). There is empirical evidence demonstrating potential *benefits* from implementing predictive policing. A randomized controlled field trial found a 7.4% decrease in crime in neighborhoods using algorithmic-based predictive policing versus no decrease in similar neighborhoods that did not use the technology (Mohler et al. 2015).

A related technology is algorithmic-based fraud detection and prevention systems, which originated in the private sector and were pioneered by companies like AT&T in the late

✉ Spencer Poodiack Parsons
s.j.poodiack-parsons@utwente.nl

René Torenvlied
r.torenvlied@utwente.nl

[1] University of Twente, Enschede, Netherlands

Published online: 10 March 2026

1990s (Becker et al. 2010). Today, algorithmic-based welfare fraud detection systems have been implemented by governments in the United States, the United Kingdom, and the Netherlands (Giest and Klievink 2024; Marr 2018; Wieringa 2023). Fraud detection algorithms have helped the United States (U.S.). Treasury Department recovered $1 billion from check fraud (Egan 2024). Furthermore, fraud detection algorithms may offer opportunities to improve equity, as biases in these technologies can, in principle, be audited and modified more readily than human biases (Kroll et al. 2017).

However, the implementation of these technologies is not without *risks*. For example, the "MiDAS" algorithm in Michigan, which was used to predict welfare fraud, made erroneous predictions 93% of the time (Wykstra 2020). In 2015, the Dutch government began using an algorithmic system called "SyRI" to detect and prevent welfare fraud (Wieringa 2023). The system was disastrous and led to numerous unintended consequences, including tens of thousands of families pushed into poverty and more than a thousand children forced into foster care (Heikkilä, 2022). Scholars who re-built in-use predictive policing algorithms and applied them to neighborhoods in the United States found that Black residents would be targeted by police at twice the rate of white residents (Lum and Isaac 2016).

Although the use of fraud detection algorithms could save governments billions, these systems could also lead to severe unintended consequences, leaving policymakers and professionals in an arduous quagmire. The prospect of severe unintended consequences may paralyze policymakers and prevent them from reaping the potential benefits of predictive policing algorithms. Therefore, it is crucial to understand the specific conditions under which the benefits and risks of predictive policing algorithms emerge. In the present literature review, we aim to identify those conditions and, thus, offer policymakers a roadmap to better navigate these benefits and risks. The goal of moving past general benefits and general risks towards identifying specific conditions from the research literature informs our main research question:

How do extant studies assess the conditions under which benefits and risks emerge from implementing algorithmic technology for policing and fraud detection and prevention?

To answer this question, we conduct a Systematic Literature Review (SLR). We employ the 'Preferred Reporting Items for Systematic Reviews and Meta-Analyses' (PRISMA) approach as a modern approach to systematic literature reviews. The PRISMA approach facilitates transparent and complete reporting of systematic reviews, which can be reproduced by other scholars (Page et al. 2021). We connect diverse academic fields, including Computer Science/Data Science, Public Administration, as well as the broader social sciences. We organize our review by dividing this multi-disciplinary corpus into the (broadly formulated) categories of *conceptual, engineering/technology,* and *empirical* studies.

Public Administration is an interdisciplinary field that deserves specific attention in the paper because policing is mainly the prerogative of governments. Much of the Public Administration literature focuses on the risks of algorithmic technology (Kakabadse et al. 2003; Wirtz et al. 2021). These risks are often framed in terms of potential harm to key public sector values, such as accountability or equity (Busuioc 2021; Young et al. 2019; Jørgensen and Bozeman 2007). Yet, scholars have also proposed that modern algorithmic technology, such as artificial intelligence, may reconcile this omnipresent debate, allowing administrators to be increasingly efficient while also elevating public values (Vogl et al. 2019).

In the present study, we find that contemporary social science and, more specifically, Public Administration literature have focused on the risks of algorithmic technology, while the Computer Science/Data Science literature has focused on potential benefits. We make several contributions. Predictive policing and fraud detection and prevention are critical domains of government policy that require great attention from both a social and technical perspective. Our study is one of the first systematic literature reviews that integrates critiques from the social sciences with the study of models from the Computer Science/Data Science perspective. Moreover, our study moves the discussion past identifying risks and benefits to uncovering specific conditions that can lead to benefits or risks. This results in actionable information for politicians, policymakers, practitioners and administrators, charged with designing and implementing algorithmic technologies for predictive policing and fraud detection and prevention. In doing so, we integrate the identified conditions into a socio-technical governance framework that conceptualizes technical system quality, human–technology interaction, and institutional context as interacting mechanisms shaping both decision outcomes and institutional legitimacy.

2 Methods

The conceptualization of modern algorithmic technology, specifically artificial intelligence, has garnered significant debate (Straub et al. 2023). Currently, artificial intelligence is conceptualized in a wide variety of ways (van Noordt and Misuraca 2022), most of which are grounded in machine learning algorithms (Anastasopoulos and Whitford 2019), particularly variants of artificial neural networks (Sousa et al. 2019). In the present literature review, we aim to include a range of conceptualizations relevant to our research questions and take a pragmatic approach: we generally refer to 'algorithmic technology' as artificial intelligence and machine learning technologies.

Algorithmic technology has many use cases, and in the present paper, we focus specifically on predictive policing and fraud detection and prevention. Some scholars consider fraud detection and prevention under the predictive policing conceptual umbrella (Macnish et al. 2020). Other scholars note that fraud detection algorithms have their origin in the private sector (Becker et al. 2010). We follow Macnish et al. (2020) in analyzing these uses in tandem because they have the unique ability to cause harm to civilian populations if used improperly (Wykstra 2020; Heaven 2020).

The types of research surveyed in the multi-disciplinary corpus can be divided into three broad categories: conceptual, engineering/technology, and empirical research. Although broad, this categorization assists in analyzing a diverse array of scholarship. We deemed a study *empirical* if it systematically drew conclusions based on empirical observations as evidence, and *conceptual* if it did not. If the study aimed to build a prediction model, it was classified as *engineering/technology*. This same process was repeated for the benefits and risks sections of this study. We then followed our conceptualization of benefits and risks and analyzed the critical connections and divergences of our corpus, intending to synthesize new knowledge (Torraco 2005).

An important additional characteristic of the reviewed literature concerns the evidentiary basis of the identified conditions. Because our corpus spans multiple disciplines, the nature of the supporting evidence varies in methodological design and inferential scope. We distinguish four evidentiary bases. First, conceptual or normative analyses, including legal and ethical scholarship which advance theoretical arguments without systematic empirical testing. Second, engineering performance studies that assess model accuracy or system performance under simulated or technical conditions. Third, empirical studies that examine associations in perceptions, behaviors, or outcome data, including surveys, interviews, and quantitative analyses of administrative or enforcement data, identifying patterns without isolating causal effects. Fourth, empirical case studies that analyze specific real-world implementations in depth and examine how algorithmic systems operate within particular institutional contexts, but do not employ formal causal identification strategies.

These evidentiary bases rely on distinct methodological logics and should not be interpreted as comparative demonstrations of epistemological "truth". Accordingly, the "conditions" identified in this review reflect recurring patterns across heterogeneous forms of evidence rather than uniformly established causal determinants of realized benefits or risks.

2.1 Eligibility criteria

For this systematic literature review, we decided not to define temporal limits to which articles are included. It is important to follow historical developments and not eliminate classic works due to their early publication date. Furthermore, we are interested in assessing the development of scholarship over time. We therefore developed a list of inclusion criteria based on the research aims of this systematic literature review:

1. Algorithmic technology should be central to the research aims of the included study. Because scholarly attention sometimes is framed towards fashionable concepts, such as artificial intelligence, some articles briefly mention algorithmic technology in the abstract, introduction, or conclusion, while focusing on other topics. Such articles were excluded.
2. We focus on the public sector and not the private industry. Algorithmic technology receives a staggering amount of research attention in the private sector (a search for research on algorithmic technology in the private sector results in over 150,000 results on SCOPUS). These articles were removed.
3. The articles must include a discussion of the benefits or risks of algorithmic technology within the public sector. Generic overviews without substantive conclusions were not included.
4. As we specifically focus on predictive policing, fraud detection and prevention, articles not related to these topics were excluded.

2.2 Search and database strategy

We endeavored to be as thorough and transparent as possible by employing the PRISMA approach to systematic literature reviews (Uttley et al. 2023). We designed a three-step search strategy for a general literature search in the SCOPUS platform. SCOPUS is more accurate than Google Scholar and has more coverage than Web of Science (Falagas et al. 2008). In a first step, we searched for studies assessing the implementation of algorithmic technology in the public sector in the context of policing and fraud detection and prevention—including keywords for algorithmic technology such as "algorithms", "artificial intelligence", and "machine learning". In a second step, we captured articles that focus on the public sector, including search terms that point towards the public sector and government. Step three of the search strategy narrows the body of literature down using search terms related to predictive policing and fraud detection, and prevention.

In addition to the general literature, we more specifically evaluated the Public Administration literature on the benefits and risks of the implementation of algorithmic technology in the public sector. We used the Web of Science platform for this search, which can limit searches to journals specifically within the field of Public Administration. We searched for

all articles within this field related to algorithmic technology (using the same search terms as in the general search).

Table 1 shows that the concepts and terms chosen resulted in 500 articles on the SCOPUS platform and an additional 23 articles on Web of Science for Public Administration literature. As Fig. 1 shows, relevant research has been published since 1987, but is still very skewed towards the contemporary. In the field of Public Administration, every article discovered has been published since 2019.

More broadly, we selected Scopus and Web of Science as our primary databases due to their broad interdisciplinary coverage and inclusion of peer-reviewed research across domains relevant to this study. Both platforms offer standardized metadata and structured indexing, which enhance the transparency and reproducibility of the search process. The exact search strings we employed, shown in Table 1, allow for straightforward replication of this systematic literature review.

2.3 Selection process

The search strategy resulted in 523 records: 500 in the general search and 23 for the field of Public Administration specifically. There was no overlap between these two corpuses. As shown in Fig. 2, 163 of the reports were then retrieved and assessed after screening. In the general literature, 358 reports were excluded, and in the Public Administration literature, only 2 reports were excluded. Roughly 48 percent of the 358 articles excluded were due to not being about policing or fraud detection and prevention. Approximately 34 percent were excluded for their focus on the private sector, and about 17 percent were excluded due to algorithmic technology not being central. One article was excluded for quality concerns, and one was excluded because it contained no discussion of either benefits or risks. The two articles in the Public Administration corpus were excluded during the screening phase for being, remarkably, about the private sector. Unlike what we observed for the multi-disciplinary corpus, a significantly higher share of research outputs in the Public Administration literature utilized empirical methodologies.

For the final review, six studies that passed the initial screening did not end up in the corpus. The most common reason was that the article focused on private sector outcomes, such as firm performance, rather than the public sector. The selected 157 articles present a highly relevant, diverse, and interesting collection of research. The entire process flow of the selection procedure is presented in Fig. 2. To support the reliability and consistency of the analysis, the initial coding was carried out by the lead author using a structured template developed in consultation with the research team. Coded materials were then reviewed and cross-checked by the co-author, and any discrepancies or ambiguities were resolved through collaborative discussion. This process helped ensure consistent interpretation and reduced the potential for individual bias in the synthesis.

3 Document characteristics

As conceptualized, we considered a wide range of scholarship, including numerous disciplines and types of research outputs. This resulted in a diverse array of document types being reviewed. The most common output of the general SCOPUS search was journal articles ($n = 84$), but also reviewed was a varied slate, which included book chapters ($n = 8$), conference papers ($n = 46$), and reviews ($n = 2$). All selected articles were then coded and analyzed in Microsoft Excel. For the field of Public Administration, we exclusively retrieved articles from Web of Science.

In the general SCOPUS search, we surveyed a heterogeneous corpus of literature to assess the conditions for benefits and risks when implementing algorithmic technology in policing and fraud detection and prevention. This included many engineering/technology articles, as well as scholarships from fields such as Law, Philosophy,

Table 1 Search terms used in literature review

Platform	Concepts and terms used for search	Full search code	Results
	Multi-Disciplinary		
SCOPUS	**Algorithmic Technology** Algorithm, Artificial Intelligence, Machine Learning **Public Sector** Public Administration, Government **Fraud/Policing** Fraud, Police/ing	(**TITLE-ABS-KEY** (algorithm* **OR** "artificial intelligence" **OR** "machine learning" **OR** "ai") AND (**TITLE-ABS-KEY** (fraud* **OR** police* **OR** policing) **AND TITLE-ABS-KEY** ("public admin*" **OR** govern*)))	500
	Public Administration		
Web of Science	**Actionable Intelligence** Algorithm, Artificial Intelligence, Machine Learning **Public Administration** Limited to Public Administration Journals	**TI =** (algorithm* **OR** artificial intelligence **OR** machine learning **OR** AI) AND **AB =** (algorithm* **OR** artificial intelligence **OR** machine learning **OR** AI) **AND KP =** (algorithm* **OR** artificial intelligence **OR** machine learning **OR** AI)	23

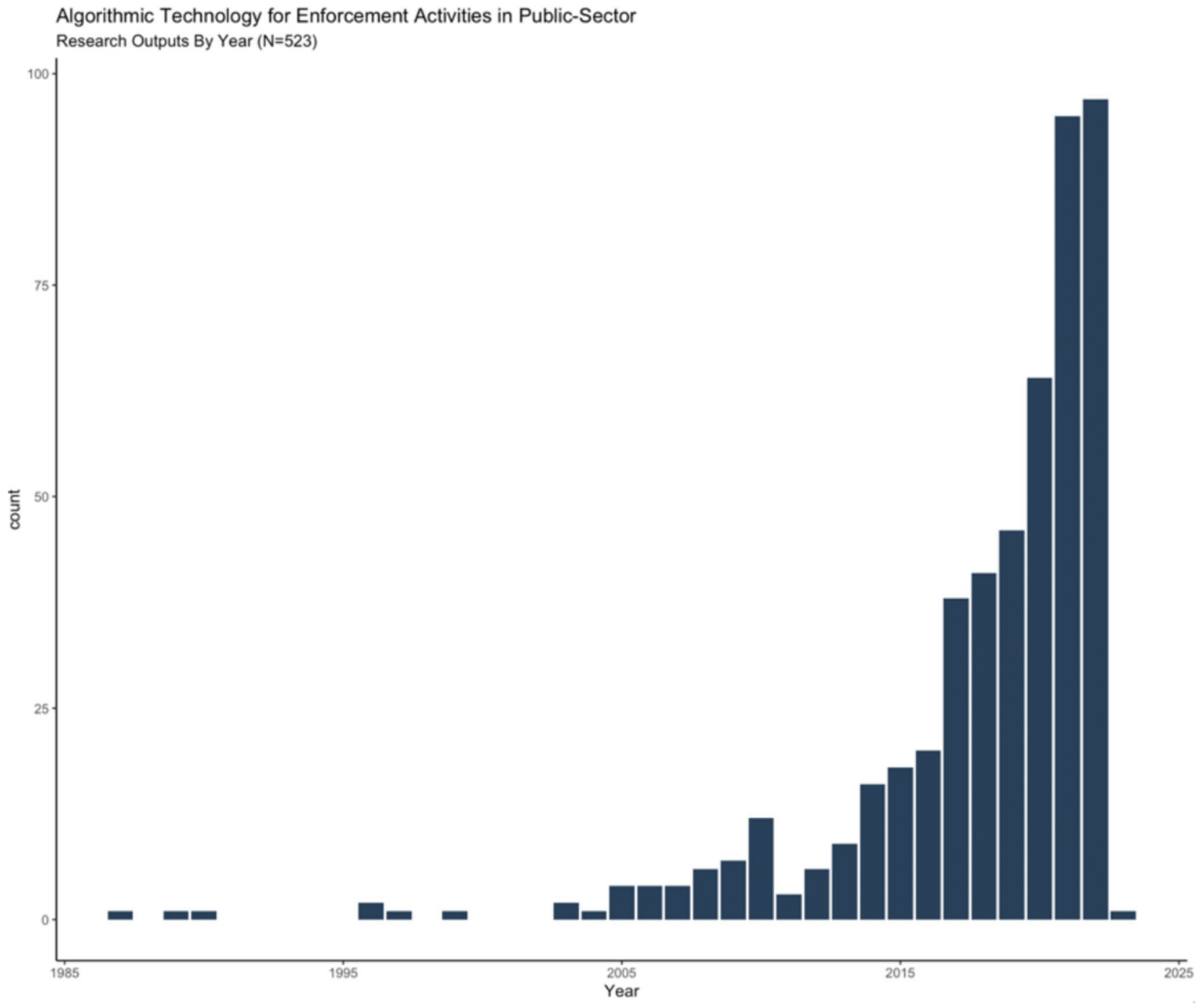


Fig. 1 Research outputs by year

and Economics. The most common source was the journal *Advances in Intelligent Systems and Computing*. This engineering/technology journal was the home for ($n = 6$) of the included articles. In all, the general search included 113 different research sources from across the academic world. This indicates that interest in predictive policing and algorithmic fraud detection and prevention is scattered across a wide range of academic journals and (sub-) disciplines. In the specific search within the field of Public Administration, only two out of 15 journals had more than one article included in the Review.

Our systematic literature review focuses on English-language academic literature, and thus excludes gray literature such as government reports, NGO publications, community-based documentation, and non-English sources. This exclusion may introduce publication and language bias, favoring well-documented, higher-profile systems and underrepresenting cases from low-resource or marginalized contexts. Future work could incorporate a broader range of materials to support more diverse and globally representative findings.

4 Benefits of algorithmic technology

4.1 Evidence from conceptual research

Our general SCOPUS search was designed to review the benefits and risks of the implementation of algorithmic technology in the public sector in the context of policing and fraud detection and prevention. Conceptual research is

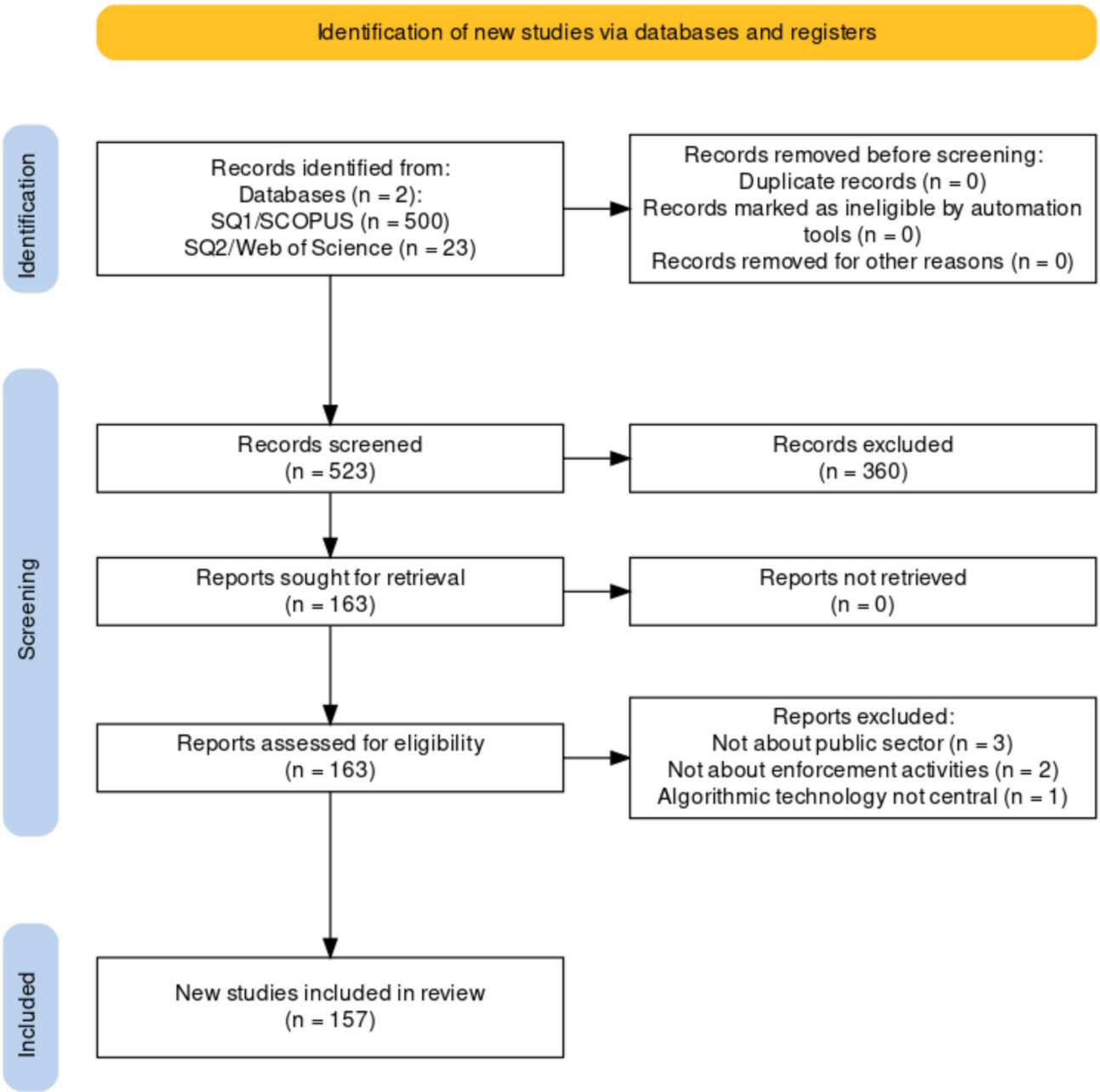


Fig. 2 Flow diagram of the included studies

common and imperative in new realms of scholarship (Vargo and Lusch 2004), such as artificial intelligence and machine learning, as the base of scholarship is likely to be underdeveloped from a theoretical perspective. Importantly, the base of conceptual scholarship retrieved in this systematic literature review leaned heavily towards the risks. Just ($n = 8$) of the relevant conceptual studies focused on benefits compared to ($n = 37$) on risks.

Resonating with the perspective of Herbert Simon, multiple studies argued that algorithmic technology could improve police performance and efficiency (Wu and Sun 2022; Li and Jiang 2020; Margetts and Dorobantu 2019). Other research outputs offered similar arguments regarding improving efficiency in the context of detecting and preventing fraud (Vaquero García and Cadaval Sampedro 2022). One prescient article from 1990 discussed the bounty of promising use cases for algorithmic technology in the context of law enforcement, including predicting that technology would "recognize patterns of criminality" by 2000 (Hernandez 1990). Finally, an article that combined expertise from

both legal and computer science fields argued that algorithmic technology using the latest methods could potentially be superior to human police agents since the biases and flaws of the technology can be modified and improved in ways difficult for humans (Kroll et al. 2017).

4.2 Evidence from engineering/technology research

An important body of relevant scholarship in the engineering/technology literature pointed to significant algorithmic innovations that could radically alter the performance and efficiency of public sector enforcement agencies. Fraud detection and prevention is a very active field of study in the engineering/technology literature. A diverse slate of algorithmic methods has been developed and tested in the service of public sector fraud detection and prevention. These techniques include a K-nearest neighbor (KNN), which is a clustering approach, as well as an artificial neural network (ANN) (Chhabra Roy and Prabhakaran 2023; Al-witwit and Ibrahim 2020). Using an ANN can allow the analyst to increase model accuracy by capturing non-linear relationships in the data. The effectiveness of the approach of utilizing an ANN was borne out by the model reporting a 95 percent accuracy rate in catching insurance fraud (Al-witwit and Ibrahim 2020). Other prediction models designed for fraud detection and prevention tested a host of machine learning algorithms to assess which performed best. One study testing different models predicting social security fraud found that a Random Forest Classifier (a collection of decision trees) achieved an impressive 99.3 percent accuracy rate (Anil Kumar et al. 2022). For an administrator concerned with public sector organizational efficiency, these articles point to significant opportunities for the implementation of algorithmic technology to detect and prevent fraud.

Scholarship in the predictive policing field also points to the potential for increased efficiency by the implementation of algorithmic technology. In one example, the authors used an innovative two-stage approach, in which they first clustered the data using KNN clustering and subsequently predicted criminal activity using linear regression (Guevara and Santos 2021). Guevara and Santos (2021) found they could predict the location of more than 80 percent of crimes committed in the city of Quito, Ecuador. A different research team used an XGBoost algorithm, which is also in the decision tree family, to predict the macro causes of crime, which resulted in a 93 percent accuracy rate (Santos Junior et al. 2021). The results from both predictive policing as well as fraud detection and prevention demonstrate that there is much potential to improve government enforcement efficiency through the implementation of algorithmic technology.

It is important to note that these high-accuracy results were achieved in controlled test environments, where the data is generally carefully curated and pre-processed for the specific use case. In contrast, real-world public sector implementations often face data quality issues, such as incompleteness, bias, or lack of timeliness, which can lead to a significant drop in performance. As a result, algorithmic accuracy in practice may fall short of expectations and limit operational relevance (Ludwig and Mullainathan 2021; Udoh 2020).

4.3 Evidence from empirical research

To our surprise, there has been only very limited quantitative empirical research done on such an important and topical subject. Only $n = 27$, roughly 17 percent of the articles that passed the initial screening, were empirical. Only $n = 10$, or 6 percent of the articles tested a hypothesis using quantitative data.

A few articles provided empirical support for *perceptions* of benefits. Multiple scholars found that enforcement officers involved in both policing and fraud detection and prevention contended that algorithmic technology could allow them to be more efficient and effective at their jobs (Mohd Noor and Mansor 2019; Jen-Hwa Hu et al. 2005). This finding shows that end-users of these technologies in public organizations foresee potential benefits from their implementation. Such perception is important because it would be hard to imagine the benefits of algorithmic technology being realized if enforcement officers themselves would not support the implementation.

One article showed that public sector employees, such as police officers, were aware of and attempted to respond to relevant concerns found in academic research (Veale et al. 2018). Veale et al. (2018) noted that government practitioners in their study did not use race-related variables because they tried to avoid potentially biased results. Swiss civil servants were enthusiastic about algorithmic technology but also concerned that many important legal and societal questions remain unanswered (Simmler et al. 2023). Finally, one article went outside of public sector employees themselves and found that citizens would find decisions based on algorithmic technology more legitimate if humans were involved, particularly in highly pivotal contexts such as approval for Medicare or Medicaid (Waldman and Martin 2022).

In general, empirical research demonstrated that professionals working towards enforcing the laws could potentially support algorithmic technology under the conditions that: (a) it made them more efficient at their work while (b) guaranteeing some level of equity and (c) potential legal and ethical risks are minimized. The empirical evidence was very sparse

when it came to understanding citizens' views on potential benefits. However, Waldman and Martin (2022) pointed to the potential condition that ensuring human governance over algorithmic decision-making can lead to increased decision legitimacy.

4.4 Conditions for potential benefits

Our discussion in the previous sections already hinted at some conditions under which the implementation of algorithmic technology in policing and fraud detection and prevention is expected to have benefits. In this section, we discuss those conditions more specifically. The engineering/technology literature is clear that, at least from a theoretical perspective, there are many potential benefits to the implementation of algorithmic technology in the public sector in the context of policing and fraud detection and prevention. From the literature, we identified four specific conditions: technical efficacy, internal support, alignment with legal and policy objectives, and public support.

The first condition is *technical efficacy*. We define technical efficacy as the degree to which a machine learning or algorithmic system accurately and reliably performs its intended task within ideal or controlled conditions. In both predictive policing as well as fraud detection and prevention use cases, researchers were able to devise innovative machine-learning methods and demonstrate accuracy rates as high as 99.3 percent (Anil Kumar et al. 2022). If enforcing the laws were a purely technical problem, there is little doubt that governments and societies would strongly benefit from the implementation of algorithmic technology.

After ensuring the technical efficacy of the modeling approach, the second condition to potentially realizing benefits is *internal support* for the technology. The police and those who investigate fraud do not operate inside a computer simulation where all the complexities of society can be "datafied". Therefore, when thinking about the conditions for benefits, it is important to move beyond assessing the technology alone. In the conceptual corpus, multiple scholars have argued that algorithmic technology could improve policing or fraud detection and prevention. Empirical research buttresses this rationale, showing that many enforcement officers had a positive perception of algorithmic technology (Wu and Sun 2022; Li and Jiang 2020; Margetts and Dorobantu 2019; Mohd Noor and Mansor 2019; Jen-Hwa Hu et al. 2005).

The third condition we identified is *alignment with legal and policy objectives*. One article argues that technological advances allow algorithmic systems to be designed so that they "better align with legal and policy objectives" (Kroll et al. 2017). As the "SyRI" algorithm in the Netherlands demonstrated, many unintended consequences arise due to potential illegalities of the system not being considered prior to implementation (Hussain 2020). This alignment will be especially relevant as methods such as Explainable Artificial Intelligence (XAI) continue to develop (Gohel et al. 2021).

The fourth condition we identified is *public support*. Predictive policing and fraud detection and prevention algorithms often operate under a cloud of secrecy, which further erodes citizen trust when their consequences come to light (AlgorithmWatch 2020). A potential mechanism for increasing public support for algorithmic decision-making in contexts such as healthcare is to retain human governance over these systems (Waldman and Martin 2022). Also, it will be important to consider issues like public accountability, as well as decision legitimacy, and in general to ensure that society is a stakeholder in the design process of these technologies.

To conclude, while there are no panaceas for ensuring that algorithmic technology is beneficial to public organizations, the extant literature demonstrates that technical performance, internal and external support, and the use of methods to align policy, legal, and technology goals are expected to constitute conditions for benefits derived from the implementation of algorithmic technology in policing, fraud detection, and prevention. We also observed, however, that extant research has not (yet) based those expectations on solid and objective empirical evidence. Hence, it remains largely unknown how and to what extent those conditions affect the benefits of implementing algorithmic technology in policing, fraud detection, and prevention. As we elaborate in Sect. 7, these conditions do not operate in isolation but interact with each other in promoting (or harming) public sector decision-outcomes and institutional legitimacy. Table 2 summarizes the conditions identified as potentially leading to benefits and specifies the evidentiary basis supporting each claim, distinguishing among conceptual or normative analyses, engineering performance studies, empirical studies examining associations in perceptions, behaviors, or outcome data, and empirical case studies.

5 Risks of algorithmic technology

5.1 Evidence from conceptual research

Our general SCOPUS search was also designed to also review the risks associated with the implementation of algorithmic technology for fraud detection and prevention and predictive policing. While benefits were sparsely mentioned, there was deep and thorough coverage of risks in the literature we retrieved. A lens of normative pessimism is most often, but not exclusively, taken by scholars from the social sciences.

Table 2 Conditions that can lead to benefits

Condition	Potential benefit	Source	Evidence base
Technical efficacy	The utilization of accurate and innovative systems can increase enforcement efficiency	Anil Kumar et al. (2022) and Guevara and Santos (2021)	Engineering performance study
Internal support	End users of the technology are motivated to use and improve the algorithmic systems	Mohd Noor and Mansor (2019) and Hu et al. (2005)	Empirical study examining associations in perceptions, behaviors, or outcome data
Alignment with legal and policy objectives	Reduce unintended consequences	Kroll et al. (2017)	Conceptual or normative analysis
Public support	Increase citizen trust in these technologies which can decrease the impact of negative outcomes	Akkaya and Krcmar (2019) and Waldman and Martin (2022)	Empirical study examining associations in perceptions, behaviors, or outcome data

The first and most notable argument was the risk of algorithmic technology trampling on the legal rights of citizens. In the context of Europe, it has been reasoned that the current regulatory framework does not give clear guidance on how legal rights are balanced when the police use algorithmic technology (Roksandic et al. 2022). Studies also maintained that existing regulations in the United States were not up to the task of regulating these technologies (Rich 2016). Other scholars went further, concluding not just that legal risks were unclear based on existing regulations but also that legal rights are already being threatened today. One U.S.-based scholar argued that even though predictive policing algorithms are not necessarily designed to be racist, they lead to racist outputs and should be challenged under the 14th Amendment (O'Donnell 2019). One study argued that fundamental legal rights would be endangered even with a predictive policing algorithm with a 99% accuracy rate (Berman 2018).

Infringements on general legal rights are certainly not the only potential risk. More specifically, loss of privacy is discussed by scholars as a significant potential risk, noting that algorithmic technology-based systems review data from a huge swath of citizens, including many who will never commit any sort of crime or violation (Guo and Kennedy 2023; Niculescu-Dincă 2021). Some researchers claim that algorithmic technology threatens to classify citizens as measurable "data points" rather than individuals with agency and rights (Wood 2019). A more banal but well-noted risk arises from a lack of technical efficacy: malicious outcomes may arise simply because the algorithms do not work as well as intended (technically) and are often over-hyped (Edler Duarte 2021; Degeling and Berendt 2018).

Some academics have moved beyond pointing out risks and offer guidance on reducing these risks by suggesting innovative ethical frameworks. One scholar mentioned that designers of algorithmic systems should move from a "model of threat" to an "ethics of care" (Asaro 2019). Another research team compiled a five-point ethical framework and demonstrated how this approach could have reduced harm in situations such as the disastrous use of the "SyRI" algorithm for welfare fraud detection in the Netherlands (van Bruxvoort and van Keulen 2021).

5.2 Evidence from engineering/technology research

Remarkably, we were unable to identify any articles from the engineering/technology fields that explicitly considered the risks of the implementation of algorithmic technology in the context of policing and fraud detection and prevention. Although this finding may have resulted from potential limitations in our search strategy (see discussion), we interpret this lack of risk-focused research as potential evidence of a techno-optimist approach within parts of the engineering/technology literature.

5.3 Evidence from empirical research

Several notable concerns surfaced in the empirical literature on the risks of the implementation of algorithmic technology. Some scholars conclude that the effectiveness of algorithmic technology is unclear. For example, empirical evidence following the implementation of predictive policing algorithms did not find evidence of reductions in crime (Tulumello and Iapaolo 2022). Other scholars empirically reveal risks. An empirical study based on publicly available policing data provided evidence that bias could occur due to highly flawed and imbalanced data (Udoh 2020). Multiple articles focused on risks that could potentially occur due to algorithmic technologies operating in an interdependent socio-technical system. One article found that algorithmic technology may give practitioners a false sense of certainty that previously did not exist in an enforcement context (Hartmann and Wenzelburger 2021). A qualitative case study of the Berlin police department revealed that officers became

devoted followers of an omnipresent algorithm, thereby effectively depriving the organization of their expertise and discretion (Meijer et al. 2021).

Only a few researchers moved outside the realms of agencies to understand perceived risks in society. One article found that only 29 percent of German citizens believed an algorithm that would inspect all welfare applications for fraud would be a positive development (Akkaya and Krcmar 2019). Additionally, researchers argue that fraud detection algorithms increasingly harm individuals who are already marginalized in society (Nikidehaghani et al. 2022). This is also true of policing algorithms; for example, scholars who analyzed facial-recognition systems across over 1000 cities within the United States found increased racial arrest disparities in policing among cities using these systems (Johnson et al. 2022).

5.4 Conditions for potential risks

It is important to better understand the conditions that may lead to legal and human rights challenges rather than exclusively focus on the risks. Below, we summarize three conditions of potential risks we identified from the literature: (a) designing systems with a model of threat, (b) discriminatory or inaccurate model output, and (c) novel human–machine interactions. The risks these conditions create can take the form of potential constitutional legal challenges. For example, racist outputs of predictive policing algorithms (even if not using race as a predictor in the model) could lead to 14th Amendment challenges under the United States Constitution (O'Donnell 2019). Risks can also be related to deep normative human rights concerns, such as treating individuals as "data points" rather than as individuals with agency and rights (Wood 2019).

The first condition for potential risks is when the algorithmic technology is designed based on a *'model of threat*.' One argument presented in the literature is that many systems today are designed with a "model of threat" in mind, which leads to increasingly aggressive socio-technical interactions between enforcement officers, algorithmic technology, and suspected criminals (Asaro 2019). A notable example is the Dutch welfare fraud detection algorithm "SyRI", which has been widely regarded as a policy failure. One of the key reasons for this outcome was its targeted deployment in so-called "problem areas," where it was used to flag perceived threats of social welfare fraud (Rachovitsa and Johann 2022).

A second condition for potential risks arises from *inaccurate model outputs*. The predominant example for risks to arise is that the data the model is trained on can be biased due to the reflection of previous discriminatory enforcement practices, as is common in situations using historical policing data in the United States (Johnson et al. 2022). However, research shows that even if the data captures no previously existing biases, the use of an algorithm can be problematic simply due to poor-quality data (Udoh 2020). Poor-quality data can result in significant harm to citizens. For example, the Michigan "MiDAS" social welfare fraud detection algorithm, which was wrong in 93% of cases, caused over a thousand people to file for bankruptcy (Wykstra 2020).

The third condition of potential risks from the implementation of algorithmic technology arises due to the challenges of *interacting human agents and algorithmic technology* in a novel and interdependent manner. Even if the technology was designed and implemented perfectly, risks could arise due to the interaction of human enforcement officers and algorithmic technology. Enforcement officers, especially those without a background in statistics or data science, may feel a false sense of certainty when utilizing algorithmic technology, even though these technologies generally operate in the realm of probabilities (Hartmann and Wenzelburger 2021). This false sense of precision could lead to a voluntary loss of discretion and expertise of enforcement officers at the hands of an omniscient algorithmic system. Additionally, the introduction of algorithmic technology may alter institutional incentives and thereby reshape human behavior. For example, after the NYPD began using CompStat, Bronstein (2015) observed that "police department activity quotas reduce police officer discretion and promote the use of enforcement activity for reasons outside of law enforcement's legitimate goals". This suggests that algorithmically driven performance metrics can pressure supervisors to impose de facto quotas, encouraging officers to prioritize easily quantifiable, low-level offenses over more substantive and potentially difficult policing objectives.

Table 3 summarizes the conditions identified as potentially leading to risks and specifies the evidentiary basis supporting each claim, distinguishing among conceptual or normative analyses, engineering performance studies, empirical studies examining associations in perceptions, behaviors, or outcome data, and empirical case studies.

6 Public administration research on algorithmic technology

Rather than focus on policing and fraud detection and prevention solely, as we did in the previous section, we included all articles from Public Administration journals, which focused on the benefits and risks of the implementation of algorithmic technology within the public sector.

6.1 Benefits of algorithmic technology

Public Administration scholars were far more focused on the risks related to algorithmic technology in the public

Table 3 Conditions that can lead to risks

Condition	Potential risk	Source	Evidence base
Designing systems with a model of threat	Overly aggressive enforcement behavior	Asaro (2019)	Conceptual or normative analysis
Discriminatory or inaccurate model output	Legal challenges to using the system or the model simply does not work as intended	O'Donnell (2019) and Udoh (2020)	Multiple evidentiary bases: conceptual or normative analysis; empirical study examining associations in perceptions, behaviors, or outcome data
Novel human–machine interactions	Modify the behavior of enforcement agents in unforeseen ways, which can lead to unintended consequences	Hartmann and Wenzelburger (2021) and Meijer et al. (2021)	Empirical case study

sector than on the potential benefits. However, some scholars have drawn attention to the potential benefits of algorithmic technology in the public sector. These studies are quite heterogeneous, with use cases ranging from machine learning-supported peace negotiations to technology-aided bureaucratic discretion (Arana-Catania et al. 2022; Young et al. 2019). The study on technology-aided bureaucratic discretion theorized that algorithmic technology could optimize discretion by increasing scalability, decreasing cost, and improving quality (Young et al. 2019). The authors position their theoretical work in contrast to empirical work that focuses on algorithmic technology curtailing bureaucratic discretion.

Interestingly, one study empirically corroborates the existence of a "privacy paradox:" citizens individually do not place much of a premium on their own personal privacy, despite being concerned more generally (Willems et al. 2023). The results of this study contrast major privacy concerns associated with the implementation of algorithmic technology in the public sector (European Commission 2020). Related is a study that offers a "Federated Machine Learning" framework. Such a framework attempts to navigate the many data-sharing restrictions prevalent within the public sector and produces actionable insights (Balta et al. 2021).

At a higher level of abstraction, one study provided a holistic review of algorithmic technology in the public sector (specifically artificial intelligence) and identified ten potential areas of value creation, which include everything from virtual agents to cognitive robotics and autonomous systems (Wirtz et al. 2021). Related are the results of a different study, which reports that algorithmic technology is not so much adopted within an instrumental neo-liberal (new public management, new public governance) administrative framework but rather within a digital government-era framework (Ojo et al. 2019).

6.2 Risks of algorithmic technology

Similar to our general SCOPUS review, we observed more normative pessimism than techno-optimism in publications in the Public Administration outlets retrieved through our Web of Science review. Many concerns are expressed about risks ranging from macro-level governance concerns to questions regarding micro-level employee engagement and support of the technologies. At the macro-level, scholars report many algorithmic governance challenges. Some articles emphasize that existing public sector institutions have not been designed with technologies of the future in mind, and new institutions need to be created (Taeihagh 2021; Filgueiras 2022).

Other articles ask how governments can be accountable to citizens with algorithmic technology involved or even making fully automated decisions (Bolingford et al. 2020). Other researchers attempted to propose solutions to many of the evident risks of the implementation of algorithmic technology in the public sector. One research output recommended that policymakers pursue AI projects that contribute to the "open government goals" that include an increasingly transparent public sector (Gao and Janssen 2020). Scholars, when reviewing public sector projects, have noted a lack of user engagement in thinking about ethics in algorithmic technology projects and offered a social mutuality framework to resolve this challenge (Koskimies and Kinder 2024). Finally, one researcher believes algorithmic decision-making is incompatible with principles of individual justice, which exist as stated aims in many criminal justice systems (Binns 2022).

The organizational-, or meso-level, is the most developed area of research in assessing the risks of algorithmic technology in the public sector. One article surveys public sector employees in Taiwan and reports that managers are more supportive of artificial intelligence than staff (Huang

et al. 2022). Scholars also find that introducing algorithmic-based agents into work teams presents many of the same challenges as introducing new human employees. For example, the challenge of how to equitably divide up work tasks (Maragno et al. 2023). In the Swiss criminal justice system, it was found that technological and organizational factors were more important than environmental factors for the success of algorithmic initiatives in the public sector (Neumann et al. 2024). Finally, researchers argue that social trust and network structures are critical to public sector organizations making use of emerging algorithmic technologies (Wilson and Broomfield 2022).

At the micro-level, scholars have assessed citizen and public sector employee preferences regarding the use of algorithmic technology. One experiment found that citizens trust algorithmic-based decisions more when their rationale can be explained, but if the decisions are merely accessible, then citizens will not be as trusting (Grimmelikhuijsen 2023). However, a cross-national survey reported that citizens are often reluctant to the government's use of algorithmic technology in the public sector in general and may be open to banning these technologies if they are perceived to lead to negative externalities, such as increased unemployment (Ehret 2022). As for internal perspectives, public sector employees in China, when surveyed, felt that algorithmic technology reduced their discretion (Wang et al. 2024).

7 Socio-technical governance framework of interacting conditions

Our review identified three interrelated sets of conditions shaping algorithmic outcomes in public sector applications: (1) technical system quality, (2) human–technology interaction, and (3) institutional context. Each contributes differently to the balance of risks and benefits, and they often interact in compounding ways. As detailed in Tables 2 and 3, the identified conditions draw on heterogeneous evidentiary bases and therefore should not be interpreted as equivalently established causal determinants. We conclude this section by unifying these conditions as three interdependent elements of a socio-technical system that jointly shape decision outcomes and institutional legitimacy.

The first set of conditions relates to the *quality of the technical system*. Three conditions identified in the literature are particularly relevant: technical efficacy, inaccurate model outputs, and designing systems with a model of threat. The quality of the algorithmic model is a key factor in achieving efficiency and reducing risks such as discrimination, bias, or overly aggressive enforcement. Technical efficacy, which is defined as how well a system performs its intended task under ideal or controlled conditions, appears to be an important enabling condition for realizing benefits. As the literature shows, highly accurate systems can achieve impressive results in test environments (Anil Kumar et al. 2022; Guevara and Santos 2021). However, high accuracy alone is not sufficient. When a system is designed with a built-in assumption of threat, it can generate aggressive or punitive outcomes even if it is technically sound. Improving model quality and avoiding threat-based design are therefore crucial first steps in bridging the gap between optimism about potential benefits and real-world concerns about risks.

The second set of conditions concerns the *interaction between human agents and algorithmic systems*. Two conditions identified in the literature are particularly relevant: novel human–machine interactions and internal support. The introduction of algorithms into decision-making processes can reshape enforcement behavior in unexpected ways, sometimes leading to unintended and potentially harmful consequences. At the same time, it is plausible that when enforcement agents gain experience with algorithmic advice and retain sufficient autonomy and discretionary authority, internal support may develop and foster more effective use and iterative improvement of these systems. Despite its importance, our review found that research on human–technology interaction in this context remains limited. Only a small number of Public Administration studies have examined this dimension, focusing either on perceptions of legitimacy or frontline employee experiences with algorithmic tools (Grimmelikhuijsen 2023; Wang et al. 2024; Huang et al. 2022).

The third set of conditions relates to the *institutional context*. Two conditions from the literature are particularly salient: alignment with legal and policy objectives and public support. These conditions are critical for reducing unintended consequences and sustaining the legitimacy of algorithmic systems within public organizations. When algorithms are introduced without legal clarity or meaningful public engagement, trust can erode and harm may result. There is evidence, for example, that public support for algorithmic fraud detection is weaker than that of political leaders (Akkaya and Krcmar 2019). Alignment with legal and policy objectives is equally important. As Public Administration scholar Dwight Waldo noted, "Things are not simply efficient or inefficient. They are efficient or inefficient for given purposes, and efficiency for one purpose may mean inefficiency for another" (Waldo 1948). Although written long before modern algorithmic systems, Waldo's warning remains highly relevant. Administrators must carefully consider the specific institutional contexts, values, and objectives guiding implementation and weigh potential efficiency gains against possible unintended consequences.

The perspective of socio-technical systems provides a useful framework for integrating the three sets of conditions identified in our review. First introduced by Trist and Bamforth (1951), socio-technical systems theory emphasizes the

interdependence between technological and social components. Since then, this approach has been developed across fields including science and technology studies (STS), engineering, and philosophy, while maintaining its core insight: new technologies reshape social relations, and social structures in turn shape technological effects (Kudina and van de Poel 2024). Scholars have increasingly embraced this perspective in their study of algorithms and AI (Johnson and Verdicchio 2025; Kudina and van de Poel 2024; Riesen 2025). Our specific approach is motivated by Kroes et al. (2006), who note that socio-technical systems are constituted by social institutions, human agents, and technical artefacts, all of which are integral to the effectiveness of the system as a whole.

Within this socio-technical perspective, the conditions identified in our review translate directly into three interrelated dimensions. Technical system quality incorporates the condition of technical efficacy, defined above as the degree to which an algorithmic system accurately performs its intended task under controlled conditions (Anil Kumar et al. 2022; Guevara and Santos 2021), as well as risk conditions related to discriminatory or inaccurate model output (O'Donnell 2019; Udoh 2020) and designing systems with a model of threat (Asaro 2019). Human–technology interaction captures the condition of internal support, reflected in enforcement officers' willingness to adopt and improve algorithmic systems (Mohd Noor and Mansor 2019; Jen-Hwa Hu et al. 2005), alongside the risk condition of novel human–machine interactions, which may reshape discretion, generate false certainty, or alter professional behavior in unforeseen ways (Hartmann and Wenzelburger 2021; Meijer et al. 2021). Institutional context encompasses alignment with legal and policy objectives (Kroll et al. 2017) and the condition of public support (Akkaya and Krcmar 2019; Waldman and Martin 2022), both of which influence whether algorithmic systems are perceived as legitimate and normatively appropriate within a given governance environment.

Viewed through this lens, technical system quality, human–technology interaction, and institutional context are not independent predictors of success or failure. Rather, they operate as interacting mechanisms within a broader socio-technical governance system. Technical design shapes the information available to decision-makers; human actors interpret and operationalize that information; and institutional norms determine how resulting decisions are evaluated, contested, and legitimized. These mechanisms are mutually conditioning: variation in one dimension shapes the functioning of the others and, in combination, structures the overall performance of the system. Together, these dynamics determine both substantive decision outcomes and the institutional legitimacy of algorithmic governance in policing and fraud detection. A summary of our framework is presented in Fig. 3.

8 Discussion

The research analyzed in this systematic literature review produced a wide range of results. On one end of the spectrum, we observe a techno-optimist engineering/technology literature, bursting with innovative artificial intelligence and machine-learning methods that could potentially transform the public sector. On the other side of the spectrum, we observe a social science literature dominated by scholars who adopt either a (normative) pessimistic perspective or build on theory and, albeit scarce, empirical evidence that points to a range of risks that should make administrators in public organizations trepidatious about the implementation of algorithmic technology. Our socio-technical governance framework helps reconcile this divide by demonstrating how technical, human, and institutional factors jointly shape outcomes.

Building on this framework, several implications follow for the study and governance of algorithmic systems in policing and fraud detection. System design choices should not be treated as neutral technical decisions but as interventions that embed normative assumptions and shape downstream institutional dynamics. The case of the Michigan "MiDAS" system illustrates how flawed data, combined with a fraud detection framework operating under a presumption of guilt, produced widespread false accusations and severe social consequences (Wykstra 2020; Charette 2018). This example underscores how model inaccuracy and threat-based design assumptions can interact, amplifying risks in practice.

At the same time, the framework highlights that decision outcomes emerge not only from technical performance but from how algorithmic recommendations are interpreted, operationalized, or deferred to by human actors. Even technically effective systems may generate unintended consequences if they reshape professional discretion or induce forms of overreliance that alter decision-making behavior (Hartmann and Wenzelburger 2021; Meijer et al. 2021). Moreover, institutional alignment and public legitimacy remain central to sustainable implementation. Systems that conflict with legal norms or lack public support may fail despite internal or technical strengths, as legitimacy depends not only on performance but also on perceived fairness, accountability, and alignment with public values. Future research should therefore examine how specific implementation choices, including transparency, oversight, and stakeholder engagement, influence long-term institutional outcomes. Answering these questions will also require a

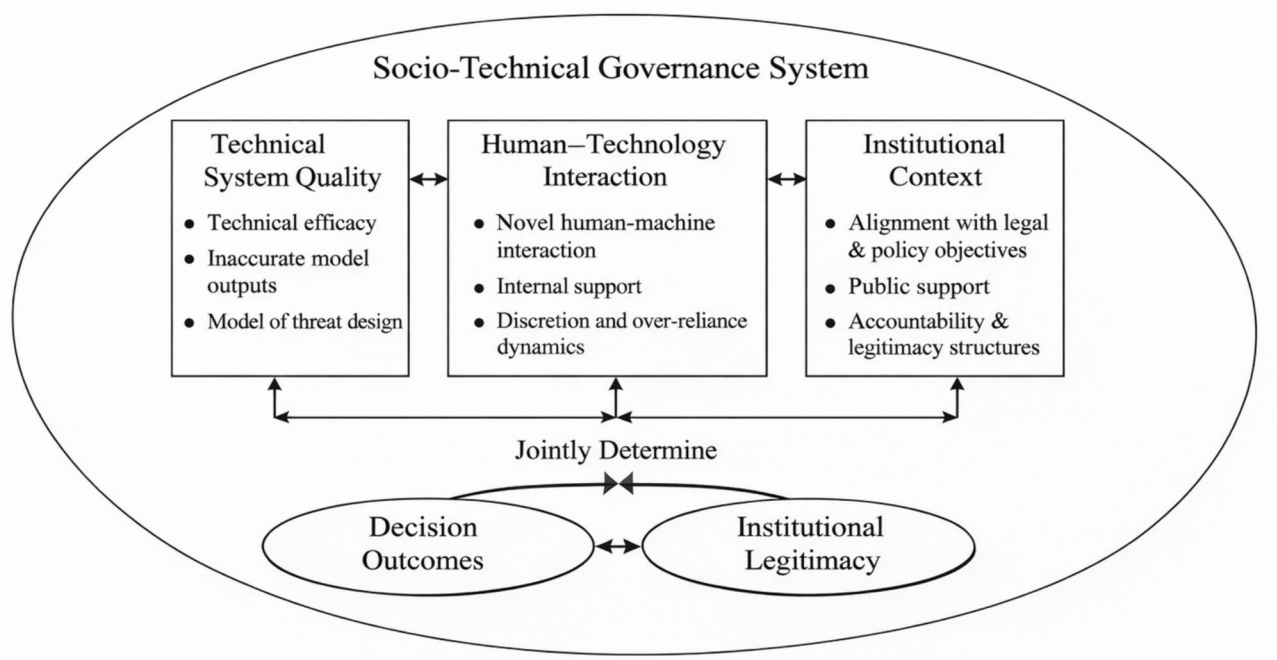


Fig. 3 Socio-technical governance framework integrating technical, human–technology, and institutional conditions in algorithmic policing and fraud detection

stronger empirical foundation than the current literature provides.

Only a small proportion of studies within the overall corpus employed empirical methodologies (~ 17%). This finding is mirrored in literature reviews undertaken by other scholars in the context of both predictive policing and AI generally (Meijer and Wessels 2019; Wirtz et al. 2021). Within this limited corpus of empirical research, most studies focus on *perceptions*, be they employee or citizen perceptions of algorithmic technology (Akkaya and Krcmar 2019; Grimmelikhuijsen 2023; Wang et al. 2024; Huang et al. 2022). While important, studying perceptions does not fully address the conditional real-world impact of algorithmic technology. For example, the seemingly fundamental question of whether implemented predictive policing systems have reduced crime cannot be definitively answered by the present study.

Furthermore, the existing corpus of theoretical and conceptual studies lacks methodological diversity, particularly when it comes to rigorous analytical and simulation-based approaches. We recommend that scholars draw on methods such as game theory and agent-based modeling to develop sharper, more empirically testable hypotheses about when algorithmic technologies are likely to produce risks or benefits. This is especially important given the context-dependent nature of our findings: the same technology may generate benefits under certain conditions and risks under others. Rigorous modeling can help clarify these dynamics by systematically exploring how key variables, such as discretion, bias, or system accuracy, interact across different implementation scenarios. For example, agent-based modeling could be used to theorize how human decision-makers respond differently to algorithmic versus human advice, particularly when the algorithmic model is inaccurate or biased, and to examine how these micro-level interactions scale up to shape macro-level outcomes.

Studies have identified both benefits and risks of using algorithmic technologies for predictive policing and fraud detection. Gaining deeper insights into these context-specific benefits and risks would enable policymakers to make well-informed decisions, supported by a rigorous cost–benefit analysis. For such an evaluation—both politically and practically—it is essential to identify the specific conditions under which these risks and benefits emerge, assuming that governments will continue to adopt these technologies.

As with any academic study, our systematic literature review has certain limitations. We selected only two academic databases out of the many possible sources of scholarship, which may have resulted in some relevant articles not being assessed as part of this SLR. It is also possible that our inclusion criteria restricted potentially relevant articles (Uttley et al. 2023). Even though these possibilities can never be prevented, our careful selection of keywords in combination with the broad databases used makes it rather unlikely. In addition, our article classification (conceptual, empirical, engineering/technology) was done practically,

and reasonable experts can likely find some areas of disagreement. This SLR could also have been influenced by publication bias, meaning articles that did not include a "statistically significant" finding were not published and therefore not included in the review (Dalton et al. 2016).

Although this review identifies recurring "conditions" associated with benefits and risks, these should not be interpreted as universally established causal determinants. As detailed in the Methods section and reflected in Tables 2 and 3, the supporting evidence varies substantially in methodological design and inferential scope. Much of the empirical literature consists of studies examining associations in perceptions, behaviors, or outcome data, as well as in-depth case studies of specific implementation contexts, none of which employ formal causal identification strategies. Engineering studies typically evaluate model performance under controlled conditions rather than real-world institutional implementation. Conceptual and normative analyses advance important theoretical arguments but do not provide systematic empirical validation. Accordingly, the conditions identified in this review represent synthesized patterns across heterogeneous evidentiary bases rather than confirmed causal mechanisms. Future research employing more rigorous causal designs is needed to test these relationships directly in real-world settings.

Similarly, the body of evidence reviewed reflects a notable degree of contextual diversity. The studies span different methodological approaches, disciplinary traditions, and national settings. For instance, the review includes perception-based research from both the Netherlands and Taiwan, suggesting that similar concerns and institutional dynamics emerge across varied cultural and administrative contexts (Grimmelikhuijsen 2023; Huang et al. 2022). Although the English-language scope of the review may introduce some geographic bias, the recurring patterns identified here appear to extend beyond any single country or institutional environment.

While our focus has been on the public sector, many of the conditions identified in this review, such as technical efficacy, novel human–technology interaction, and alignment with relevant objectives, are likely to be relevant across other domains where algorithmic systems are used. Sectors such as healthcare and finance also rely on complex human–technology configurations and face similar challenges in balancing efficiency, fairness, and accountability. Future research could explore how these conditions manifest in different institutional environments and whether similar patterns of risk and benefit emerge when algorithmic systems are implemented outside the public sector.

Author contributions S. P.: conceptualization, data curation, writing—original draft, writing—review and editing, visualization R.T.: writing—original draft, writing—review and editing, supervision, project administration.

Funding No specific funding was received for the research, authorship, or publication of this manuscript.

Data availability No datasets were generated or analysed during the current study.

Declarations

Conflict of interest The authors declare that they have no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

Akkaya C, Krcmar H (2019) Potential use of digital assistants by governments for citizen services: the case of Germany. In: Proceedings of the 20th annual international conference on digital government research, pp 81–90. https://doi.org/10.1145/3325112.3325241

Al-witwit SSI, Ibrahim AA (2020) Improving operational efficiency of government using artificial intelligence. IOP Conf Ser Mater Sci Eng 928(2):022014. https://doi.org/10.1088/1757-899X/928/2/022014

Alder M (2024) How the Biden administration is tackling diversity in federal AI hiring. FedScoop. https://fedscoop.com/how-the-biden-administration-is-tackling-diversity-in-federal-ai-hiring/

AlgorithmWatch (2020) How Dutch activists got an invasive fraud detection algorithm banned. https://algorithmwatch.org/en/syri-netherlands-algorithm/

Amaro S (2021) Dutch government resigns after childcare benefits scandal. CNBC. https://www.cnbc.com/2021/01/15/dutch-government-resigns-after-childcare-benefits-scandal-.html

Anastasopoulos LJ, Whitford AB (2019) Machine learning for public administration research, with application to organizational reputation. J Public Admin Res Theory 29(3):491–510. https://doi.org/10.1093/jopart/muy060

Arana-Catania M, van Lier F-A, Procter R (2022) Supporting peace negotiations in the Yemen war through machine learning. Data Policy 4:e28. https://doi.org/10.1017/dap.2022.19

Asaro PM (2019) AI ethics in predictive policing: from models of threat to an ethics of care. IEEE Technol Soc Mag 38(2):40–53. https://doi.org/10.1109/MTS.2019.2915154

Balta D, Sellami M, Kuhn P, Schöpp U, Buchinger M, Baracaldo N, Anwar A, Ludwig H, Sinn M, Purcell M, Altakrouri B (2021) Accountable federated machine learning in government: engineering and management insights. Electronic participation, vol 12849, pp 125–138. Springer International Publishing, Berlin. https://doi.org/10.1007/978-3-030-82824-0_10

Becker RA, Volinsky C, Wilks AR (2010) Fraud detection in telecommunications: history and lessons learned. Technometrics 52(1):20–33. https://doi.org/10.1198/TECH.2009.08136

Benbouzid B (2019) To predict and to manage. predictive policing in the United States. Big Data Soc 6(1):2053951719861703. https://doi.org/10.1177/2053951719861703

Berman E (2018) A government of laws and not of machines. SSRN Electron J. https://doi.org/10.2139/ssrn.3098995

Binns R (2022) Human judgment in algorithmic loops: individual justice and automated decision-making. Regul Gov 16(1):197–211. https://doi.org/10.1111/rego.12358

Bolingford I, Bagaric M, Bull M, Hunter D, Stobbs N (2020) Is Australia ready for AI on the bench? J Judic Admin 30(1):3–18. https://doi.org/10.3316/agispt.20201020038408

Bronstein N (2015) Police management and quotas: governance in the CompStat era. Columbia J Law Soc Probl 48(4):543–581

Busuioc M (2021) Accountable artificial intelligence: holding algorithms to account. Public Admin Rev 81(5):825–836. https://doi.org/10.1111/puar.13293

Charette R (2018) Michigan's MiDAS unemployment system: algorithm alchemy created lead, not gold. IEEE Spectrum. https://spectrum.ieee.org/michigans-midas-unemployment-system-algorithm-alchemy-created-lead-not-gold

Chhabra Roy N, Prabhakaran S (2023) Internal-led cyber frauds in Indian banks: an effective machine learning–based defense system to fraud detection, prioritization and prevention. Aslib J Inf Manag 75(2):246–296. https://doi.org/10.1108/AJIM-11-2021-0339

Dalton JE, Bolen SD, Mascha EJ (2016) Publication bias: the elephant in the review. Anesth Analg 123(4):812–813. https://doi.org/10.1213/ANE.0000000000001596

Darroch G (2022) Van Rij admits institutional racism is a problem at the Dutch tax office. DutchNews.Nl. https://www.dutchnews.nl/news/2022/05/van-rij-admits-institutional-racism-is-a-problem-at-dutch-tax-office/

de Graaf G, Paanakker HL (2015) Good governance: performance values and procedural values in conflict. Am Rev Public Admin 45(6):635–652. https://doi.org/10.1177/0275074014529361

Degeling M, Berendt B (2018) What is wrong about Robocops as consultants? A technology-centric critique of predictive policing. AI Soc 33(3):347–356. https://doi.org/10.1007/s00146-017-0730-7

Domínguez P, Scartascini C (2022) Willingness to pay for crime reduction: evidence from six countries in the Americas (IDB Publications Working Paper No. 12488). Inter-American Development Bank. https://doi.org/10.18235/0004531

Edler Duarte D (2021) The the making of crime predictions: sociotechnical assemblages and the controversies of governing future crime. Surveill Soc 19(2):199–215. https://doi.org/10.24908/ss.v19i2.14261

Egan M (2024) AI helped Uncle Sam catch $1 billion of fraud in one year. And it's just getting started. CNN Business. CNN. https://www.cnn.com/2024/10/17/business/ai-fraud-treasury/index.html

Ehret S (2022) Public preferences for governing AI technology: comparative evidence. J Eur Public Policy 29(11):1779–1798. https://doi.org/10.1080/13501763.2022.2094988

European Commission. Joint Research Centre (2020) AI Watch, artificial intelligence in public services: overview of the use and impact of AI in public services in the EU. Publications Office. https://doi.org/10.2760/039619

Falagas ME, Pitsouni EI, Malietzis GA, Pappas G (2008) Comparison of PubMed, Scopus, Web of Science, and Google Scholar: strengths and weaknesses. FASEB J 22(2):338–342. https://doi.org/10.1096/fj.07-9492LSF

Filgueiras F (2022) New Pythias of public administration: ambiguity and choice in AI systems as challenges for governance. AI Soc 37(4):1473–1486. https://doi.org/10.1007/s00146-021-01201-4

Fountain JE (2022) The moon, the ghetto and artificial intelligence: reducing systemic racism in computational algorithms. Gov Inf Q 39(2):101645. https://doi.org/10.1016/j.giq.2021.101645

Gao Y, Janssen M (2020) Generating value from government data using AI: an exploratory study. In: Viale Pereira G, Janssen M, Lee H, Lindgren I, Rodríguez Bolívar MP, Scholl HJ, Zuiderwijk A (eds) Electronic government, vol 12219. Springer International Publishing, Berlin, pp. 319–331. https://doi.org/10.1007/978-3-030-57599-1_24

Giest SN, Klievink B (2024) More than a digital system: how AI is changing the role of bureaucrats in different organizational contexts. Public Manag Rev 26(2):379–398. https://doi.org/10.1080/14719037.2022.2095001

Gohel P, Singh P, Mohanty M (2021) Explainable AI: current status and future directions (arXiv:2107.07045). https://doi.org/10.48550/arXiv.2107.07045

Grimmelikhuijsen S (2023) Explaining why the computer says no: algorithmic transparency affects the perceived trustworthiness of automated decision-making. Public Admin Rev 83(2):241–262. https://doi.org/10.1111/puar.13483

Guevara C, Santos M (2021) Crime prediction for patrol routes generation using machine learning. In: Herrero Á, Cambra C, Urda D, Sedano J, Quintián H, Corchado E (eds) 13th international conference on computational intelligence in security for information systems (CISIS 2020), vol 1267. Springer International Publishing, Berlin, pp 97–107. https://doi.org/10.1007/978-3-030-57805-3_10

Guo Z, Kennedy L (2023) Policing based on automatic facial recognition. Artif Intell Law 31(2):397–443. https://doi.org/10.1007/s10506-022-09330-x

Hartmann K, Wenzelburger G (2021) Uncertainty, risk and the use of algorithms in policy decisions: a case study on criminal justice in the USA. Policy Sci 54(2):269–287. https://doi.org/10.1007/s11077-020-09414-y

Heaven W (2020) Predictive policing algorithms are racist. They need to be dismantled. MIT Technol Rev. https://www.technologyreview.com/2020/07/17/1005396/predictive-policing-algorithms-racist-dismantled-machine-learning-bias-criminal-justice/

Heikkilä M (2022) Dutch scandal serves as a warning for Europe over risks of using algorithms. POLITICO. https://www.politico.eu/article/dutch-scandal-serves-as-a-warning-for-europe-over-risks-of-using-algorithms/

Hernandez AP (1990) Artificial intelligence and expert systems in law enforcement: current and potential uses. Comput Environ Urban Syst 14(4):299–306. https://doi.org/10.1016/0198-9715(90)90004-D

Huang H, Kim KC (Casey), Young MM, Bullock JB (2022) A matter of perspective: differential evaluations of artificial intelligence between managers and staff in an experimental simulation. Asia Pac J Public Admin 44(1):47–65. https://doi.org/10.1080/23276665.2021.1945468

Hussain H (2020) Dutch court finds SyRI algorithm violates human rights norms in landmark case. Human Rights Pulse. https://www.humanrightspulse.com/mastercontentblog/dutch-court-finds-syri-algorithm-violates-human-rights-norms-in-landmark-case

Jen-Hwa Hu P, Lin C, Chen H (2005) User acceptance of intelligence and security informatics technology: a study of COPLINK. J Am Soc Inf Sci Technol 56(3):235–244. https://doi.org/10.1002/asi.20124

Jada I, Mayayise TO (2024) The impact of artificial intelligence on organisational cyber security: an outcome of a systematic literature review. Data Inf Manag 8(2):100063. https://doi.org/10.1016/j.dim.2023.100063

Johnson TL, Johnson NN, McCurdy D, Olajide MS (2022) Facial recognition systems in policing and racial disparities in arrests. Gov Inf Q 39(4):101753. https://doi.org/10.1016/j.giq.2022.101753

Johnson DG, Verdicchio M (2025) The sociotechnical entanglement of AI and values. AI Soc 40(1):67–76. https://doi.org/10.1007/s00146-023-01852-5

Jørgensen TB, Bozeman B (2007) Public values: an inventory. Admin Soc 39(3):354–381. https://doi.org/10.1177/0095399707300703

Kakabadse A, Kakabadse NK, Kouzmin A (2003) Reinventing the Democratic Governance Project through information technology? A growing agenda for debate. Public Admin Rev 63(1):44–60. https://doi.org/10.1111/1540-6210.00263

Anil Kumar KM, Bhargava AR, Abawajy J (2022) Detection of false income level claims using machine learning. Int J Mod Educ Comput Sci 14(1):65–77. https://doi.org/10.5815/ijmecs.2022.01.06

Koskimies E, Kinder T (2024) Mutuality in AI-enabled new public service solutions. Public Manag Rev 26(1):219–244. https://doi.org/10.1080/14719037.2022.2078501

Kroes P, Franssen M, van de Poel I, Ottens M (2006) Treating sociotechnical systems as engineering systems: some conceptual problems. Syst Res Behav Sci 23(6):803–814. https://doi.org/10.1002/sres.703

Kroll J, Huey J, Barocas S, Felten E, Reidenberg J, Robinson D, Yu H (2017) Accountable algorithms. Univ Pa Law Rev 165(3):633–705

Kudina O, van de Poel I (2024) A sociotechnical system perspective on AI. Minds Mach 34(3):21. https://doi.org/10.1007/s11023-024-09680-2

Kumar S, Ahmed R, Bharany S, Shuaib M, Ahmad T, Tag Eldin E, Rehman AU, Shafiq M (2022) Exploitation of machine learning algorithms for detecting financial crimes based on customers' behavior. Sustainability 14(21):13875. https://doi.org/10.3390/su142113875

Li L, Jiang L (2020) Research on city public security administration based on intelligent policing. In: 2020 international conference on urban engineering and management science (ICUEMS), pp 692–695. https://doi.org/10.1109/ICUEMS50872.2020.00151

Ludwig J, Mullainathan S (2021) Fragile algorithms and fallible decision-makers: lessons from the Justice System. J Econ Perspect 35(4):71–96. https://doi.org/10.1257/jep.35.4.71

Lum K, Isaac W (2016) To predict and serve? Significance 13(5):14–19. https://doi.org/10.1111/j.1740-9713.2016.00960.x

Macnish K, Wright D, Jiya T (2020) Predictive policing in 2025: a scenario. In: Jahankhani H, Akhgar B, Cochrane P, Dastbaz M (eds) Policing in the era of AI and smart societies. Springer International Publishing, Berlin, pp 199–215. https://doi.org/10.1007/978-3-030-50613-1_9

Maragno G, Tangi L, Gastaldi L, Benedetti M (2023) AI as an organizational agent to nurture: effectively introducing chatbots in public entities. Public Manag Rev 25(11):2135–2165. https://doi.org/10.1080/14719037.2022.2063935

Margetts H, Dorobantu C (2019) Rethink government with AI. Nature 568(7751):163–165. https://doi.org/10.1038/d41586-019-01099-5

Marr B (2018) How the UK government uses artificial intelligence to identify welfare and state benefits fraud. Forbes. https://www.forbes.com/sites/bernardmarr/2018/10/29/how-the-uk-government-uses-artificial-intelligence-to-identify-welfare-and-state-benefits-fraud/

Meijer A, Lorenz L, Wessels M (2021) Algorithmization of bureaucratic organizations: using a practice lens to study how context shapes predictive policing systems. Public Admin Rev 81(5):837–846. https://doi.org/10.1111/puar.13391

Meijer A, Wessels M (2019) Predictive policing: review of benefits and drawbacks. Int J Public Admin 42(12):1031–1039. https://doi.org/10.1080/01900692.2019.1575664

Mohd Noor NRA, Mansor N (2019) Exploring the adaptation of artificial intelligence in whistleblowing practice of the internal auditors in Malaysia. Procedia Comput Sci 163:434–439. https://doi.org/10.1016/j.procs.2019.12.126

Mohler GO, Short MB, Malinowski S, Johnson M, Tita GE, Bertozzi AL, Brantingham PJ (2015) Randomized controlled field trials of predictive policing. J Am Stat Assoc 110(512):1399–1411. https://doi.org/10.1080/01621459.2015.1077710

Molina AD (2009) Values in public administration: the role of organizational culture. Int J Organ Theory Behav 12(2):266–279. https://doi.org/10.1108/IJOTB-12-02-2009-B007

Moore MH (1995) Creating public value: strategic management in government. Harvard University Press, Cambridge

Neumann O, Guirguis K, Steiner R (2024) Exploring artificial intelligence adoption in public organizations: a comparative case study. Public Manag Rev 26(1):114–141. https://doi.org/10.1080/14719037.2022.2048685

Ngai EWT, Hu Y, Wong YH, Chen Y, Sun X (2011) The application of data mining techniques in financial fraud detection: a classification framework and an academic review of literature. Decis Support Syst 50(3):559–569. https://doi.org/10.1016/j.dss.2010.08.006

Niculescu-Dincă V (2021) Theorizing technologically mediated policing in smart cities: an ethnographic approach to sensing infrastructures in security practices. In: Nagenborg M, Stone T, González Woge M, Vermaas PE (eds) Technology and the city, vol 36. Springer International Publishing, Berlin, pp 75–100. https://doi.org/10.1007/978-3-030-52313-8_5

Nikidehaghani M, Andrew J, Cortese C (2022) Algorithmic accountability: Robodebt and the making of welfare cheats. Acc Audit Account J 36(2):677–711. https://doi.org/10.1108/AAAJ-02-2022-5666

Norga A (2021) 4 benefits and 4 drawbacks of predictive policing. Liberties.eu. https://www.liberties.eu/en/stories/predictive-policing/43679

O'Donnell RM (2019) Challenging racist predictive policing algorithms under the Equal Protection Clause. N Y Univ Law Rev 94(3):544–581

Ojo A, Mellouli S, Ahmadi Zeleti F (2019) A realist perspective on AI-era public management*. In: Proceedings of the 20th annual international conference on digital government research, pp 159–170. https://doi.org/10.1145/3325112.3325261

Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, Shamseer L, Tetzlaff JM, Akl EA, Brennan SE, Chou R, Glanville J, Grimshaw JM, Hróbjartsson A, Lalu MM, Li T, Loder EW, Mayo-Wilson E, McDonald S, Moher D (2021) The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. https://doi.org/10.1136/bmj.n71

Pearsall B (2010) Predictive policing: the future of law enforcement? NIJ J 266:16–19

Picasso E, Cohen MA (2018) Valuing the public's demand for crime prevention programs: a discrete choice experiment (SSRN Scholarly Paper No. 3015649). Social Science Research Network. https://doi.org/10.2139/ssrn.3015649

Profiling and SyRI (2015) PILP. https://pilpnjcm.nl/en/dossiers/profiling-and-syri/. Retrieved 9 Nov 2022

Rachovitsa A, Johann N (2022) The human rights implications of the use of AI in the digital welfare state: lessons learned from the Dutch *SyRI* case. Hum Rights Law Rev 22(2):ngac010. https://doi.org/10.1093/hrlr/ngac010

Riesen E (2025) A sociotechnological-system approach to AI ethics. AI Soc 40(8):6231–6243. https://doi.org/10.1007/s00146-025-02335-5

Rich ML (2016) Machine learning, automated suspicion algorithms, and the Fourth Amendment. Univ Pa Law Rev 164(4):871–929

Roksandic S, Protrka N, Engelhart M (2022) Trustworthy artificial intelligence and its use by law enforcement authorities: where do we stand? In: 2022 45th jubilee international convention on information, communication and electronic technology (MIPRO), pp 1225–1232. https://doi.org/10.23919/MIPRO55190.2022.9803606

Santos Junior RV, Coelho JVV, Cacho NAA (2021) A Macrocause classification model for violent crime analysis in the field of public safety based on machine learning techniques. In: 2021 IEEE international smart cities conference (ISC2), pp 1–7. https://doi.org/10.1109/ISC253183.2021.9562842

Wykstra S (2020) Government's use of algorithm serves up false fraud charges. Undark Magazine. https://undark.org/2020/06/01/michigan-unemployment-fraud-algorithm/

Simmler M, Brunner S, Canova G, Schedler K (2023) Smart criminal justice: exploring the use of algorithms in the Swiss criminal justice system. Artif Intell Law 31(2):213–237. https://doi.org/10.1007/s10506-022-09310-1

Simon HA (1947) Administrative behavior: a study of decision-making processes in administrative organization. Macmillan, London

de Sousa WG, de Melo ERP, Bermejo PHDS, Farias RAS, Gomes AO (2019) How and where is artificial intelligence in the public sector going? A literature review and research agenda. Gov Inf Q 36(4):101392. https://doi.org/10.1016/j.giq.2019.07.004

Steinglass M (2013) Dutch uproar over Bulgarian benefit fraud. Financial Times. https://www.ft.com/content/7afd3bd6-bcac-11e2-b344-00144feab7de

Straub VJ, Morgan D, Bright J, Margetts H (2023) Artificial intelligence in government: concepts, standards, and a unified framework. Gov Inf Q 40(4):101881. https://doi.org/10.1016/j.giq.2023.101881

Taeihagh A (2021) Governance of artificial intelligence. Policy Soc 40(2):137–157. https://doi.org/10.1080/14494035.2021.1928377

Torraco RJ (2005) Writing integrative literature reviews: guidelines and examples. Hum Resour Dev Rev 4(3):356–367. https://doi.org/10.1177/1534484305278283

Trist EL, Bamforth KW (1951) Some social and psychological consequences of the longwall method of coal-getting: an examination of the psychological situation and defences of a work group in relation to the social structure and technological content of the work system. Hum Relat 4(1):3–38. https://doi.org/10.1177/001872675100400101

Tulumello S, Iapaolo F (2022) Policing the future, disrupting urban policy today. Predictive policing, smart city, and urban policy in Memphis (TN). Urban Geogr 43(3):448–469. https://doi.org/10.1080/02723638.2021.1887634

Udoh ES (2020) Is the data fair?: An assessment of the data quality of algorithmic policing systems. In: Proceedings of the 13th international conference on theory and practice of electronic governance, pp 1–7. https://doi.org/10.1145/3428502.3428503

Uttley L, Quintana DS, Montgomery P, Carroll C, Page MJ, Falzon L, Sutton A, Moher D (2023) The problems with systematic reviews: a living systematic review. J Clin Epidemiol 156:30–41. https://doi.org/10.1016/j.jclinepi.2023.01.011

van Bruxvoort X, van Keulen M (2021) Framework for assessing ethical aspects of algorithms and their encompassing socio-technical system. Appl Sci Basel 11(23):11187. https://doi.org/10.3390/app112311187

van Noordt C, Misuraca G (2022) Exploratory insights on artificial intelligence for government in Europe. Soc Sci Comput Rev 40(2):426–444. https://doi.org/10.1177/0894439320980449

Vaquero García A, Cadaval Sampedro M (2022) Reflexiones y propuestas para combatir la corrupción pública en la administración local. Rev Esp Transparencia 15:181–205. https://doi.org/10.51915/ret.219

Vargo SL, Lusch RF (2004) Evolving to a new dominant logic for marketing. J Mark 68(1):1–17. https://doi.org/10.1509/jmkg.68.1.1.24036

Veale M, Van Kleek M, Binns R (2018) Fairness and accountability design needs for algorithmic support in high-stakes public sector decision-making. In: Proceedings of the 2018 CHI conference on human factors in computing systems, pp 1–14. https://doi.org/10.1145/3173574.3174014

Vogl TM, Seidelin C, Ganesh B, Bright J (2019) Algorithmic bureaucracy: managing competence, complexity, and problem solving in the age of artificial intelligence (SSRN Scholarly Paper No. 3327804). https://doi.org/10.2139/ssrn.3327804

Waldman A, Martin K (2022) Governing algorithmic decisions: the role of decision importance and governance on perceived legitimacy of algorithmic decisions. Big Data Soc 9(1):205395172211004. https://doi.org/10.1177/20539517221100449

Waldo D (1948) The administrative state: a study of the political theory of American public administration. Ronald Press, Faridabad

Waldo D (1952) Development of theory of democratic administration. Am Polit Sci Rev 46(1):81–103. https://doi.org/10.2307/1950764

Wang G, Xie S, Li X (2024) Artificial intelligence, types of decisions, and street-level bureaucrats: evidence from a survey experiment. Public Manag Rev 26(1):162–184. https://doi.org/10.1080/14719037.2022.2070243

Wang Y, Zhang N, Zhao X (2022) Understanding the determinants in the different government AI adoption stages: evidence of local government chatbots in China. Soc Sci Comput Rev 40(2):534–554. https://doi.org/10.1177/0894439320980132

Wieringa M (2023) "Hey SyRI, tell me about algorithmic accountability": lessons from a landmark case. Data Policy 5:e2. https://doi.org/10.1017/dap.2022.39

Willems J, Schmid MJ, Vanderelst D, Vogel D, Ebinger F (2023) AI-driven public services and the privacy paradox: do citizens really care about their privacy? Public Manag Rev 25(11):2116–2134. https://doi.org/10.1080/14719037.2022.2063934

Wilson C, Broomfield H (2022) Learning how to do AI: managing organizational boundaries in an intergovernmental learning forum. Public Manag Rev. https://doi.org/10.1080/14719037.2022.2055119

Wirtz BW, Langer PF, Fenner C (2021) Artificial intelligence in the public sector—a research agenda. Int J Public Admin 44(13):1103–1128. https://doi.org/10.1080/01900692.2021.1947319

Wood MA (2019) Algorithmic tyranny: *Psycho-Pass* , science fiction and the criminological imagination. Crime Media Cult 15(2):323–339. https://doi.org/10.1177/1741659018774609

Wright BE (2015) The science of public administration: problems, presumptions, progress, and possibilities. Public Admin Rev 75(6):795–805

Wu Y, Sun L (2022) Deep learning algorithm for grassroots police professional and technical talents construction and incentive. In: Pei Y, Chang J-W, Hung JC (eds) Innovative computing, vol 935. Springer Nature, Singapore, pp 699–704. https://doi.org/10.1007/978-981-19-4132-0_89

Young M, Himmelreich J, Honcharov D, Soundarajan S (2021) The right tool for the job? Assessing the use of artificial intelligence for identifying administrative errors. In: DG.O2021: the 22nd annual international conference on digital government research, pp 15–26. https://doi.org/10.1145/3463677.3463714

Young MM, Bullock JB, Lecy JD (2019) Artificial discretion as a tool of governance: a framework for understanding the impact of artificial intelligence on public administration. Perspect Public Manag Govern. https://doi.org/10.1093/ppmgov/gvz01

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.