Perceived challenges and opportunities of machine learning applications in governmental organisations
an interview-based exploration in the Netherlands

Delfos, Jeroen; Zuiderwijk, Anneke; Van Cranenburgh, Sander; Chorus, Caspar

DOI
10.1145/3560107.3560122

Publication date
2022

Document Version
Final published version

Published in
Proceedings of the 15th International Conference on Theory and Practice of Electronic Governance, ICEGOV 2022

Perceived challenges and opportunities of machine learning applications in governmental organisations: an interview-based exploration in the Netherlands

Jeroen Delfos
J.Delfos@tudelft.nl
Delft University of Technology
Delft, Netherlands

Anneke Zuiderwijk
A.M.G.Zuiderwijk-vanEijk@tudelft.nl
Delft University of Technology
Delft, Netherlands

Sander van Cranenburgh
S.vanCranenburgh@tudelft.nl
Delft University of Technology
Delft, Netherlands

Caspar Chorus
C.G.Chorus@tudelft.nl
Delft University of Technology
Delft, Netherlands

ABSTRACT

As the application of machine learning (ML) algorithms becomes more widespread, governmental organisations try to benefit from this technology. While ML has the potential to support public services, its application also introduces challenges. Several scholars have described the possible opportunities and challenges of ML applications in the public sector conceptually. However, it is not yet investigated if and how these concepts materialise and are perceived by end-users in the public sector when ML is applied. Therefore, it is neither clear whether these concepts are valid, nor what regulation could be introduced to address them effectively. This empirical study's objective is to shed light on how challenges and opportunities of governmental use of ML algorithms are perceived by Dutch professionals in the public sector. We attain our objective by conducting interviews with twelve professionals from Dutch executive and supervisory organisations in the public sector that respectively use ML and supervise the use of ML. Results show that ML is used primarily for improvements in the accuracy and speed of public task execution. Furthermore, interviewed professionals experience several barriers for ML implementation as well as risks following from the use of ML. The implications of these findings for practice are discussed, as well as opportunities for further research.

CCS CONCEPTS

• Computing methodologies → Machine learning; **• General and reference → General conference proceedings**; **Empirical studies**.

KEYWORDS

Machine Learning, Artificial Intelligence, Public Sector, Government, Interviews, Challenges, Opportunities

ICEGOV 2022, October 4–7, 2022, Guimarães, Portugal

ACM ISBN 978-1-4503-9635-6/22/10.
https://doi.org/10.1145/3560107.3560122

ACM Reference Format:
Jeroen Delfos, Anneke Zuiderwijk, Sander van Cranenburgh, and Caspar Chorus. 2022. Perceived challenges and opportunities of machine learning applications in governmental organisations: an interview-based exploration in the Netherlands. In *15th International Conference on Theory and Practice of Electronic Governance (ICEGOV 2022), October 4–7, 2022, Guimarães, Portugal.* ACM, New York, NY, USA, 8 pages. https://doi.org/10.1145/3560107.3560122

1 INTRODUCTION

Machine learning (ML) algorithms[1] are at the core of several highly successful companies, such as Alphabet and Meta. Following their example, the public sector and governments have been increasingly using ML applications to support their tasks over the past years. With the help of ML, governments try to improve contact with and services to citizens. For example, virtual assistants help with tax applications [1] and national statistics are available through the 'Google Assistant'[35]. Furthermore, ML algorithms have been introduced in fraud detection practices[26] and in the judiciary system[3]. In the Netherlands, a survey pointed out that 50% of all responding government agencies use algorithms, 84% of which stated they were using ML algorithms [10]. [12] shows that 45% of 142 researched governmental agencies in the United States use Artificial Intelligence (AI)[2] or ML and "the pace of AI/ML development in government seems to be accelerating" [12, p.91].

The adoption of ML applications is understandable, as public services may become cheaper and more effective [21]. Decisions can be better informed through the use of data analysis [43]. However, in recent papers, it has been argued that ML applications may introduce new challenges such as the unfair treatment of citizens due to biased input data [22], accountability gaps [41] and breaches of personal privacy [4]. Other studies have pointed out that the fear of these challenges can in turn lead to the underuse of ML applications [15], leading to missed chances of improvement of governmental services.

[1] Machine learning algorithms are those algorithms that can automatically detect patterns in data. They differ from less complex algorithms by the characteristic that ML algorithms need little specification on how the pattern detection task must be executed [34]

[2] We use the following definition of AI in this paper: "the capability of a computer system to show human-like intelligent behaviour characterised by certain core competencies, including perception, understanding, action, and learning" [44, p. 599]

Over the past few years, a political debate has risen about the danger of algorithms in the public sector. Just like in many other countries, politicians in the Netherlands called for "transparent and controllable" algorithms [29], public availability of the "functioning and source code of algorithms or methods of analysis with a considerable impact on citizens" [38] and guidelines for public use of algorithms [24]. In reaction to this debate, the European Commission proposed AI regulation [13] and the coalition agreement of the in 2021 formed Dutch cabinet contains the goal of installing an "algorithm watchdog" [40]. However, for any AI regulation to be effective, it should be aimed at mitigating potential risks that are introduced with the use of ML algorithms while not severely hampering its benefits [30]. Furthermore, it should be clear to lawmakers what opportunities and risks they need to include in their trade-offs for legislation development [2].

Little is known about the challenges and opportunities of governmental use of ML in practice. Multiple scholars have been studying the challenges and opportunities of ML, some of them focusing on governmental use in particular. The larger body of research has a conceptual character [15, 18, 32], as opposed to presenting empirical data. Furthermore, some literature synthesises opinions and observations of academic experts [11, 41]. However, very little research can be found on the materialisation of challenges and opportunities of ML as perceived by professionals in public practice, the end-users of these tools. This empirical insight is important since it can validate or contest conceptual studies and generate new concepts and theories about the challenges and opportunities of ML. Such insight may help to design effective regulations for ML and assist supervisors in enforcing these regulations.

To arrive at these insights, interviews with twelve professionals working at the national government of the Netherlands were conducted. Six professionals work at organisations implementing ML to support their executive duties, six professionals work at organisations that have a supervisory role regarding these organisations using ML algorithms. The Dutch context is interesting since the government is highly digitised [37] and is at the frontier of openly sharing what algorithms are being used in the public sector [14].

2 BACKGROUND

Since the beginning of this century, the application of ML algorithms has been surging [19, 28, 46]. Firstly, this surge goes hand in hand with the increasing amount of available data, due to the digitisation of data which was earlier stored analogously such as visual content (on film) and writing (on paper) [8]. Secondly, administration, formerly absent or documented on paper, is increasingly done digitally. Thirdly, new types of data are emerging, such as social media data, personal location data obtained from mobile phones and log files from clicks on websites, primarily flowing from devices and sensors [39]. Furthermore, the internet increases the possibilities to collect and share data [8], while cheaper storage capacity enables saving all this data [19]. ML algorithms are able to analyse this data and reveal information hidden in this data [28]. Over the past years, ML applications have been empowered by increased computational power and the development of new algorithms and theories [19].

Governments can benefit from the developments of ML, as they have large amounts of data available. Public organisations gather data of citizens in order to execute public tasks, but the capacity is lacking to process all data by hand. When ML algorithms can be applied, opportunities lie in an increase in decision-making accuracy, an increase in the speed with which information tasks can be executed and a reduction in costs by replacing human capacity [21]. Several public tasks have the potential to benefit from ML. Some tasks involve the prediction of a variable, such as crime risk [20], recidivism [7] or cyber-attacks [36]. Other tasks involve the identification of people or objects, such as (per)ocular recognition for border security [25] and face recognition for forensic applications [27]. ML can help with prioritisation, for example when detecting fraud in tax declarations [26] or risk scoring of financial transactions to detect possible money laundering [6]. Contact with citizens can be automated with the help of ML, for example in chatbots [1] or voice assistants [35].

Wirtz et al. presents four dimensions of challenges for AI[3], in an attempt to capture the broad range of challenges that may occur when implementing AI in the public sector [44]. First, there is the technological dimension, involving the safety, quality, feasibility of AI systems and the expertise needed to implement AI systems. Second, there are societal challenges, involving (the lack of) trust in AI-supported systems [23] and the replacement of the human workforce with machines [5]. Third, there are ethical challenges such as discrimination against groups of citizens due to biased input data [22] and the possible constraining effect of AI on human decision making [17]. Fourthly, there are regulatory challenges due to its potential autonomy, making it difficult to scrutinise an algorithm [18] or hold people accountable for its working [42]. Furthermore, AI can be used to gather or analyse data without the consent of citizens, intruding on their privacy [45].

It is important to note that challenges are not the same as risks. Risks are comprised of a potential occurrence of some event, the consequences of this event with some probability and the valuation of these consequences [31]. For example, Brynjolfsson and Mitchell notes that ML will alter our economies and labour, but further understanding is needed to know what are the consequences of specific uses of ML [5]. Furthermore, if mitigation strategies are in place, potentially harmful events can be prevented. For example, Engstrom et al. proposes solutions that can be helpful for the challenge of scrutinising algorithms and accountability gaps [12]. Lastly, the valuation of any probable harmful consequences can be compensated by the potential advantages of the use of ML [2].

The four dimensions of Wirtz et al. provide a very complete overview of AI challenges [44]. We will use these dimensions in our research as a framework to analyse the perceived challenges of ML. The three main benefits of ML Maciejewski for the public sector will be used to analyse the opportunities [21].

3 RESEARCH APPROACH

To achieve our research goal, we use interviewing as our main research method. Interviewing is chosen because it aligns well with the exploratory character of this study. Interviewing allows for 'mutual exploration' of the interview subject and for 'investigation

[3]Note that these challenges are not specific for ML, but for the broader concept of AI

Table 1: Overview of interviewees

ID	Interviewee position	Organisation type
I1	Manager	Executive organisation
I2	Data scientist	Executive organisation
I3	Data scientist	Executive organisation
I4	Manager	Executive organisation
I5	Data scientist	Executive organisation
I6	Data scientist	Executive organisation
I7	Manager	Supervising agency
I8	Director	Supervising agency
I9	Researcher	Supervising agency
I10	Manager	Supervising agency
I11	Manager & advisor	Supervising agency

of causation' [16, p. 125]. By posing open-end questions, the interviewer leaves room for unexpected answers and is able to acquire a deeper understanding of the subject. A *semi-structured* interview will be performed, using a set of predefined questions, which are presented in appendix A. The interview questions were sent to the interviewee beforehand.

Eleven interviews were conducted with twelve professionals[4] from Dutch executive organisations that use ML and supervisory agencies supervising the application of ML by these executive organisations. These groups offer two interesting perspectives, as they must respectively comply with and enforce on proposed AI regulation. Such regulation can only be effective when aimed at risks and the trade-off between risks and opportunities of ML.

Both professionals with deep technological knowledge about ML, as well as professionals with managerial tasks related to ML were interviewed. In this way, we explore both the challenges of the technology and of the technology in its socio-technical context. Table 1 presents a list of the characteristics of the professionals that were interviewed.

With the permission of the interviewees, voice recordings were made of each interview. These voice recordings were subsequently transcribed and the interviewees were given the possibility to check the transcriptions. Qualitative data analysis was performed using the Software ATLAS.ti (version 9). First, a round of coding was done by reading each interview line by line and marking important statements. The first round of coding was done in a 'data-driven' manner, as the codes follow from the raw data, i.e. the transcripts, as opposed to (pre)defining codes from theory [9]. Secondly, axial coding was performed, by grouping codes into categories to capture relationships between related codes [33]. These categories include the dimensions used by [44] to classify perceived challenges of ML and the main benefits of ML as described by [21] to analyse ML opportunities. Thirdly, a second round of coding was done, consisting of merging overlapping codes, and checking for consistent coding over the different interviews. Two of the authors were involved in the data analysis.

Starting the analysis process in a data-driven manner allowed for exploring all opportunities and challenges that could be found in the statements of the interviewees. Only after listing all these opportunities and challenges, they were placed in the proposed frameworks. By doing so, it is possible to critically reflect on the frameworks.

[4]Ten interviews were conducted with one interviewee and one interview was conducted with two interviewees

4 RESULTS

This section presents the results of the analysis of the twelve interviews. The opportunities, respectively the risks, of ML that the interviewees perceive in their organisation and the public organisations that they supervise. Tables 2 and 3 show all perceived opportunities and risks with the ID of the interviewees corresponding to Table 1.

4.1 Perceived opportunities of ML

In this section, we analyse what professionals in the Dutch government perceive as the opportunities of the use of ML algorithms. Table 2 shows that eleven distinct opportunities were mentioned. The opportunities were categorised using the three main benefits distinguished by Maciejewski [21]. One opportunity did not fit any of the categories, and was assigned to a fourth category, named 'transparency'.

Five opportunities were mentioned that are related to accuracy improvements. The most observed opportunity is to process previously unused data. ML has no problem with large sets of data, whereas processing these data sets was impossible with the scarce human capacity available in the organisations of the interviewees. Three interviewees mentioned that inspection tasks were carried out mainly by taking a random sample of the to be inspected population: *"In our domain it is always about finding the needle in a haystack" [I1]*. ML algorithms can be used to make better-informed choices on who or what to inspect. New patterns can be found in data with the help of ML, that were missed by humans: *"we see a lat of opportunities that these algorithms can help find patterns that our individual professionals do not see" [I1]*. ML also helps to inform policymakers, which can get more information out of available data with ML. One interviewee mentioned that ML helps in making decisions more objectively, whereas previously 'gut feeling' decisions were not uncommon.

The interviewed professionals perceived four opportunities that relate to the improvement of the speed with which public information tasks can be executed. Three professionals see that ML leads to better allocation of human capacity, for example when capacity can be assigned to tasks with a relatively high reward: *"What is the most promising case? Which case should we handle first?" [I4]*. Furthermore, ML can support decision-makers with a better information position, which makes the processing of work easier and therefore faster: *"We can do things with camera's and other sensors, like microphones" [I6]*. Two interviewees mentioned that there are repetitive tasks that are experienced as boring and employees faced with these tasks tend to lose focus and work slower. ML can take over these tasks and therefore increase speed. One interviewee mentioned that ML is employed to work on the same work as is being done by humans, in order to extend capacity.

Three interviewees mentioned an opportunity that had a direct relation to reducing costs of operations. This opportunity lies in the replacement of human personnel. It must be noted however that almost all perceived opportunities can have an indirect effect on costs.

Besides benefits related to the accuracy, speed and cost, two interviewees mentioned that ML leads to more transparent and traceable documentation of decisions, in comparison to human decision making: *"With human decision-making it is often far less clear what variables are considered" [I9].*

4.2 Perceived challenges of ML

In this section, we analyse what professionals in the Dutch government perceive as the challenges of the use of ML algorithms. The four dimensions of AI challenges of Wirtz et al. are used as a framework for this analysis [44]. We categorised the challenges mentioned by interviewees using these four dimensions. For four challenges we were not able to assign them to any of the four dimensions and were categorised as 'other' challenges. We observe different types of challenges. Firstly, there are barriers that make it difficult for organisations to use systems using ML on a daily basis in core processes. Secondly, there are risks that follow from using ML in public organisations. These risks include potential hazards to the quality or the continuity of the execution of public tasks. An overview of all challenges mentioned by interviewees and whether they were viewed as a risk or a barrier can be found in Table 3.

While all interviewees, working at executive organisations, were experimenting with using ML algorithms, only one interviewee indicated that these algorithms were integrated into core organisational processes. The others used ML for incidental data analyses but were faced with barriers that challenged a more structural application of ML. The quality of the data was mentioned by most interviewees as a barrier for ML implementation. A reason for low data quality is that a lot of data comes from registrations which are made not with the aim to analyse them, but because of legal obligations. Furthermore, registrations on crime and violations of regulations are incomplete. Four interviewees indicated that existing IT infrastructure poses barriers for the implementation of ML. Integration in existing systems is often difficult, but necessary because ML is used to help in existing work processes. Furthermore, ML algorithms are developed mostly in open-source languages, that are not easily integrated into the IT infrastructure, and IT staff is not used to working with such applications: *"Machine learning needs constant development. Our organisation is not prepared for this" [I6].* The lack of investments in the development of ML is mentioned both as a barrier for ML implementation, as well as a risk. Four interviewees reported that investments should be increased in order to be able to go from experiments to implementation. These investments concern both IT infrastructure and personnel costs. Risks follow from the observation that there is capacity needed for maintaining the quality of algorithms, once implemented. The GDPR[5] was mentioned twice as a barrier. Data that is available cannot be used due to legal constraints, although societal benefits are expected from using this data. The last barrier is seen in getting colleagues on board for working with ML. Professionals in management and working with the outcomes lack trust in the proposed ML algorithms: using ML *"means that they have to change their way of working. We can not force them to do so" [I3].*

[5]General Data Protection Regulation

Besides barriers, the interviewees see risks of the use of ML algorithms in public organisations. Six risks are related to the implementation of ML technology. Knowledge about ML is mentioned six times as a risk. More specifically, interviewees mention a lack of knowledge from the side of managers, the people responsible for the organisations using ML, as a risk: *"The risk lies in the gap between the knowledge of data analysts and decision-makers" [I1].* Another risk that involves knowledge is seen in external hire. Professionals with ML knowledge are scarce, and external hire can help with this. However, knowledge is not maintained and leaves when the hiring period is done. A monodisciplinary team is also a risk. Two interviewees mentioned that this can lead to ML of lower quality since the focus on important aspects can be missing. Lastly, one interviewee mentioned that ML applications run on infrastructure that is delivered by a small number of companies. A calamity at one of these companies can cause the failure of a lot of ML applications at once: *"It's a winner-takes-all market, regarding the supply of infrastructure. That leads to new vulnerabilities" [I10].*

Three mentioned risks relate to changes in society and human interaction due to ML. Firstly, it is seen as a risk that professionals do not accept ML outcomes. These professionals do not agree with a situation in which part of their work is being outsourced to a machine and do not take ML outcomes into consideration or only do this when the outcomes confirm their own logic: *"it is a risk that they only follow the model when they agree and strengthen their own judgement" [I3].* Secondly, one interviewee mentioned that it is a risk that society will not agree on the use of ML. Thirdly, it is a risk that professionals lose skill and knowledge by leaning on ML outcomes.

Three risks are seen by interviewees regarding the ethics of ML. Five of the interviewees see risks related to bias in input data used for ML algorithms. This bias can lead to discriminatory practices, for example when certain parts of the population are overrepresented in registrations. One interviewee mentions the risk that citizens might not be aware that the government use ML. This might make them feel fooled or not being taken seriously, as human interaction can be preferred over interaction with a machine. Lastly, it is mentioned that automated decision making following from ML leaves no room for personal circumstances of citizens, that would be relevant and considered when a human decision-maker would be involved.

The category of law and regulations concerns the governance of ML. Eight interviewees see risks in this theme regarding the opaqueness of some ML algorithms. ML algorithms like neural networks might produce results that cannot be logically explained by humans. Interviewees working at executive organisations are well aware that they have a responsibility to be able to explain how decisions came to be. However, they do state that they have difficulties determining what is a good explanation and what is the lower limit of explaining a decision. Six interviewees see a risk of difficulties overturning decisions that were made by ML algorithms. Citizens should be in the position to contest a decision. To enable this, civil servants should be able to understand the process that led to a decision and feel confident to overturn this decision as they see fit. In order to contest a decision it also needs to be clear who is responsible for the decision, which can become opaque when ML systems are used: *"You, as a citizen, can not just call someone for an explanation for an automated decision" [I3].* This risk of opacity of

Table 2: Opportunities of ML as perceived by Dutch public professionals

Category	Opportunities mentioned by interviewees
Accuracy of decision making	Process data that was not being used due to scarce human capacity [I1, I3, I5, I6, I10]
	Informed inspections instead of random samples [I1, I3, I7]
	Find new patterns in data [I1, I5, I6]
	Make policies based on better information [I1, I2]
	Make decisions more objectively [I7]
Acceleration of information tasks	Better allocate human personnel [I1, I2, I4]
	Work from a better information position [I2, I6]
	Let computers work on things that are boring for humans [I6, I10]
	Divide tasks between computers and humans to extend capacity [I6]
Reduction of decision-making costs	Replacement of human personnel [I4, I6, I9]
Transparency	Transparent and traceable documentation of decisions [I5, I9]

Table 3: Challenges for ML as perceived by Dutch public professionals

Category	Challenges mentioned by interviewees	Risk or barrier
Technology implementation	Knowledge about ML is missing on a managerial level [I1, I2, I3, I7, I8, I10]	Risk
	Data quality is not sufficient for ML applications [I2, I3, I4, I5, I10]	Barrier
	IT Infrastructure is not ready for ML applications [I2, I3, I5, I6]	Barrier
	Investments in the development of ML algorithms are missing [I1, I2, I9, I10]	Risk + barrier
	There is not enough capacity and knowledge in IT staff to take ML applications into production [I5, I6]	Barrier
	Data cannot be used for ML due to the GDPR [I3, I11]	Barrier
	Professionals do not trust ML to take over work [I2, I3]	Barrier
	Important aspects of ML development can be overlooked in monodisciplinary teams [I2, I9]	Risk
	Knowledge about ML comes from external hire and disappears when hired personnel leaves [I2]	Risk
	ML algorithms can contain design flaws [I7]	Risk
	Infrastructure used for ML is provided by a small number of companies, making it vulnerable in case of calamities [I10]	Risk
Society	ML algorithms may not be accepted by professionals working with ML outcomes [I3, I5, I6]	Risk
	ML algorithms may not be accepted by society [I2]	Risk
	Professionals lose skill due to ML taking over work [I6]	Risk
Ethics	Biased data leads to biased ML outcomes with discriminatory effects [I3, I4, I5, I8, I11]	Risk
	Citizens are not aware that ML is used when they interact with public organisations [I8]	Risk
	Automated decision making following from a ML algorithm leaves no room for personal circumstances of citizens [I11]	Risk
Law & Regulations	Some ML algorithms are opaque in their working [I1, I4, I5, I6, I7, I8, I10, I11]	Risk
	Outcomes of ML cannot be overturned or questioned [I3, I7, I8, I9, I10, I11]	Risk
	Organisations lose control of processes if ML is delivered by a third party [I6, I7, I9, I10, I11]	Risk
	It is unclear who is responsible for the quality and the outcomes of ML algorithms [I1, I2, I3, I10]	Risk
	ML can breach the privacy of citizens [I2]	Risk
Other	The speed with which ML can process tasks makes for a larger impact of a mistake, compared to manual processing of tasks [I8, I11]	Risk
	The hype of ML leads to the use of ML while other tools are better fitted for the job [I3]	Risk
	Connected systems using ML can cause cascading effects that are difficult to oversee [I10]	Risk
	While trying to be transparent about the use of ML, citizens can game the system [I3]	Risk

responsibilities for the quality and the outcomes of ML algorithms is mentioned by four interviewees. Five interviewees see it as a risk that third parties deliver ML systems to public organisations. This can lead to more opaqueness in how the algorithms work due to disclosure restrictions, but also to more opaqueness in responsibilities. One interviewee mentioned privacy as a risk of ML, while we saw earlier that privacy regulations impose a barrier for the application of ML.

Four risks were mentioned by interviewees that did not match any of the dimensions as described [44], but do relate to the implementation of ML. Firstly, the speed with which ML can process tasks is an opportunity but also leads to larger consequences when a mistake is made. Secondly, the hype about ML can lead to the implementation of ML while other tools might be more appropriate: *"People want something with AI, it doesn't matter what" [I3].* While this is not a risk of ML itself, it is indicated that this can lead to misuse of ML. Projects that involve ML have a higher chance to make a claim on innovation budgets, compared to less complex solutions that might perform better for the job at hand. Thirdly, when multiple systems work with ML, these systems can influence each other. These cascading events can have unforeseen consequences. Lastly, it is mentioned that citizens can 'game the system' when they know how a ML algorithm operates.

5 DISCUSSION, LIMITATIONS AND FURTHER RESEARCH

The interviewees were questioned in an exploratory manner, posing open questions and leaving room for side-tracks that were deemed interesting to the interviewees. The results of these interviews show that professionals in Dutch public organisations recognise both the opportunities and challenges of ML algorithms as described in conceptual studies. The used categories for both the opportunities and challenges are neither mutually exclusive nor commonly exhaustive. Perceived opportunities and challenges were categorised as we saw best fitted. However, categorisation is inevitably arbitrary and subjective for some parts. Furthermore, interviewees mentioned an opportunity and several risks that could not be attributed to the categories of the conceptual frameworks used during the analysis. Although the lists of opportunities and challenges give a broad overview of what is perceived, some perceived opportunities and challenges may have been missed, due to the limited amount of interviews conducted.

Several perceived risks relate to each other and are in themselves not necessarily a risk. For example, citizens may not be aware that a public organisation is using ML to interact with them. This can result in a lower trust in this public organisation when they do find out about this. Furthermore, when asked about the risks of ML interviewees tend to mention factors that could have a mitigating effect on possible risks, such as investing in knowledge about ML and its risks.

It should be noted that this study included professionals from the Netherlands. The Dutch context might differ from that of other countries. For example, sectors that are privatised in the Netherlands can be public in others, shifting the range of public use of ML. Furthermore, the implementation phase of ML varies amongst countries. Countries with a high level of digitisation in public services, like the Netherlands, will have more opportunities for using ML, as opposed to countries with a lesser digitised government. Lastly, values may differ between countries. The European Union has been at the forefront regarding for example data privacy regulation, signalling a high value for privacy. Differences in values, like privacy, might give a different perception of the opportunities and challenges of ML.

The list of challenges is considerably longer than the list of opportunities. However, it must be noted that this is not quantitative research, so this does not lead to the conclusion that the risks outweigh the opportunities. The opportunities are the reason for the interviewees that work at an executive organisation to develop ML applications while being conscious of the mentioned challenges.

We have several recommendations for further research that can build on the results presented in this study. Firstly, we were able to structure the mentioned opportunities and challenges, using the concepts presented in Maciejewski, Wirtz et al. [21, 44]. However, a comparison between conceptual studies and the empirical results of this research is an avenue for further research. Secondly, a more rigid approach to analysing risks can be adopted. The risks mentioned in this research are overlapping and influence each other. It is still relatively unclear what are the hazardous events we want to prevent, what are the consequences of these events and how do we value these events. Thirdly, further research can explore the opportunities and challenges in countries that have different characteristics regarding public services, digitisation rate and values. Fourthly, a more quantitative approach can be used to be able to prioritise challenges.

6 CONCLUSION AND RECOMMENDATIONS

This study's objective was to shed light on how challenges and opportunities of governmental use of ML algorithms are perceived by Dutch professionals in the public sector. We have done this by interviewing twelve professionals from Dutch public organisations, both with supervisory and executive functions. This research gives an empirical contribution to existing literature with a predominantly conceptual character.

We identified eleven distinct opportunities, mentioned by the interviewees. Using data that is now not used due to limited human capacity is the most mentioned opportunity. While interviewees do mention several other opportunities related to improvements in accuracy and speed of public services, opportunities in reducing costs are sparsely mentioned. Additionally, there is an opportunity in making transparent and traceable documentation of decisions due to the use of ML. We identified 26 challenges, mentioned by the interviewees. The interviewees see two types of challenges, barriers for and risks of ML algorithms. Firstly, there are risks that the use of this technology brings to the quality of public services, citizens, or society as a whole. Secondly, there are challenges that impose a barrier to the implementation of ML. One risk that was recognised broadly by the interviewees was the possibility that ML algorithms may be too opaque for a proper explanation of outcomes. The interviewees also perceive little guidance on what level of explanation is sufficient.

The majority of the interviewees had experience with ML for incidental analyses but were faced with these barriers when trying to implement systems using ML for day-to-day use. It is recommended for legislators and policy makers to take the distinction between barriers and risks into account. Barriers will prevent or slow the implementation of ML, making it harder to benefit from this technology. If policymakers want to get the benefits of ML, they should remove the barriers that prevent the further application of ML. Risks are the challenges that have the potential to inflict harm. As effective regulation of ML should aim at the risks of this technology [30], it is key to take these risks into consideration when drafting regulations.

A INTERVIEW QUESTIONS

A.1 Procedural Questions

These questions are asked after the consent form has been filled in and signed. By asking these questions we make sure that the interviewee has been well informed before giving consent. Besides this, the main points of the consent form will be briefly explained.

- Did you read the consent form "Interviews on the perceived risks of machine learning in government"?
- Do you have any remaining questions about the consent form?

A.2 Contextual Questions

These questions serve as context for the rest of the interview and will not address the goal of the interview itself.

- For what organisation do you work?
- What is your role in this organisation?
- In what way does your work relate to the topic of machine learning?

A.3 In-depth Questions

This part will answer the research question of this research. This is also the part where we will accommodate for for elaboration on unforeseen talking points, which may be of interest for the research of this thesis.

Parts of some of the questions have text in brackets. For this parts, the formulation of the question will be different for each of the interviewee categories as listed in 1.

- How would you describe the phase of machine learning application within [your organisation/ the organisations you supervise]?
- What kind of machine learning applications does [your organisation/ the organisations you supervise] use?
- In what tasks that [your organisation/ the organisations you supervise] performs can machine learning applications be of use?
- What opportunities can machine learning applications yield for [your organisation/ the organisations you supervise]?
- What are the risks of implementing machine learning applications for [your organisation/ the organisations you supervise]?
- What can be done to mitigate the risks regarding machine learning applications
- Do you see a role for supervisory agencies to prevent or mitigate risks regarding machine learning applications
- Do you see a role for supervisory agencies to promote the opportunities of machine learning applications?

REFERENCES

[1] Australian Taxation Office. 2015. Introducing Alex, our new web assistant. \url{https://beta.ato.gov.au/Tests/Introducing-Alex–our-new-web-assistant}.

[2] Julia Black and Andrew Murray. 2019. Regulating AI and machine learning: setting the regulatory agenda (complementar). *European Journal of Law and Technology* 10, 3 (2019), 1–17. http://eprints.lse.ac.uk/102953/4/722_3282_1_PB.pdf

[3] Tim Brennan, William Dieterich, and Beate Ehret. 2009. Evaluating the Predictive Validity of the Compas Risk and Needs Assessment System. *Criminal Justice and Behavior* 36, 1 (1 2009), 21–40. https://doi.org/10.1177/0093854808326545

[4] Dennis Broeders, Erik Schrijvers, Bart van der Sloot, Rosamunde van Brakel, Josta de Hoog, and Ernst Hirsch Ballin. 2017. Big Data and security policies: Towards a framework for regulating the phases of analytics and use of Big Data. *Computer Law & Security Review* 33, 3 (6 2017), 309–323. https://doi.org/10.1016/j.clsr.2017.03.002

[5] Erik Brynjolfsson and Tom Mitchell. 2017. What can machine learning do? Workforce implications. *Science* 358, 6370 (12 2017), 1530–1534. https://doi.org/10.1126/science.aap8062

[6] Zhiyuan Chen, Le Dinh Van Khoa, Ee Na Teoh, Amril Nazir, Ettikan Kandasamy Karuppiah, and Kim Sim Lam. 2018. Machine learning techniques for anti-money laundering (AML) solutions in suspicious transaction detection: a review. *Knowledge and Information Systems* 57, 2 (11 2018), 245–285. https://doi.org/10.1007/s10115-017-1144-z

[7] Alexandra Chouldechova. 2017. Fair Prediction with Disparate Impact: A Study of Bias in Recidivism Prediction Instruments. *Big Data* 5, 2 (2017), 153–163. https://doi.org/10.1089/big.2016.0047

[8] Kenneth Cukier and Viktor Mayer-Schoenberger. 2013. The rise of big data: How it's changing the way we think about the world. *Foreign Affairs* 92, 3 (2013), 28.

[9] Jessica T. DeCuir-Gunby, Patricia L. Marshall, and Allison W. McCulloch. 2011. Developing and Using a Codebook for the Analysis of Interview Data: An Example from a Professional Development Research Project. *Field Methods* 23, 2 (5 2011), 136–155. https://doi.org/10.1177/1525822X10388468

[10] Sophie Doove and Daniëlle Otten. 2018. *Verkennend onderzoek naar het gebruik van algoritmen binnen overheidsorganisaties.* Technical Report. Statistics Netherlands, Den Haag. https://www.cbs.nl/nl-nl/maatwerk/2018/48/gebruik-van-algoritmen-door-overheidsorganisaties

[11] Yogesh K. Dwivedi, Laurie Hughes, Elvira Ismagilova, Gert Aarts, Crispin Coombs, Tom Crick, Yanqing Duan, Rohita Dwivedi, John Edwards, Aled Eirug, Vassilis Galanos, P. Vigneswara Ilavarasan, Marijn Janssen, Paul Jones, Arpan Kumar Kar, Hatice Kizgin, Bianca Kronemann, Banita Lal, Biagio Lucini, Rony Medaglia, Kenneth Le Meunier-FitzHugh, Leslie Caroline Le Meunier-FitzHugh, Santosh Misra, Emmanuel Mogaji, Sujeet Kumar Sharma, Jang Bahadur Singh, Vishnupriya Raghavan, Ramakrishnan Raman, Nripendra P. Rana, Spyridon Samothrakis, Jak Spencer, Kuttimani Tamilmani, Annie Tubadji, Paul Walton, and Michael D. Williams. 2021. Artificial Intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. *International Journal of Information Management* 57, July (4 2021), 101994. https://doi.org/10.1016/j.ijinfomgt.2019.08.002

[12] David Freeman Engstrom, Daniel E Ho, Catherine M Sharkey, and Mariano-Florentino Cuéllar. 2020. Government by algorithm: Artificial intelligence in federal administrative agencies. *NYU School of Law, Public Law Research Paper* 20-54 (2020). https://doi.org/10.2139/ssrn.3551505

[13] European Commission. 2021. Proposal for a Regulation of the European Parlement and of the Council Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts.

[14] Luciano Floridi. 2020. Artificial Intelligence as a Public Service: Learning from Amsterdam and Helsinki. *Philosophy & Technology* 33, 4 (12 2020), 541–546. https://doi.org/10.1007/s13347-020-00434-3

[15] Luciano Floridi, Josh Cowls, Monica Beltrametti, Raja Chatila, Patrice Chazerand, Virginia Dignum, Christoph Luetge, Robert Madelin, Ugo Pagallo, Francesca Rossi, Burkhard Schafer, Peggy Valcke, and Effy Vayena. 2018. AI4People—An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations. *Minds and Machines* 28, 4 (12 2018), 689–707. https://doi.org/10.1007/s11023-018-9482-5

[16] Gary Eugene Gorman, Peter Robert Clayton, Sydney J Shep, and Adela Clayton. 2005. *Qualitative research for the information professional: A practical handbook* (2nd ed.). Facet Publishing, London.

[17] Marijn Janssen, Martijn Hartog, Ricardo Matheus, Aaron Yi Ding, and George Kuk. 2022. Will Algorithms Blind People? The Effect of Explainable AI and Decision-Makers' Experience on AI-supported Decision-Making in Government. *Social Science Computer Review* 40, 2 (4 2022), 478–493. https://doi.org/10.1177/0894439320980118

[18] Marijn Janssen and George Kuk. 2016. The challenges and limits of big data algorithms in technocratic governance. *Government Information Quarterly* 33, 3 (7 2016), 371–377. https://doi.org/10.1016/j.giq.2016.08.011

[19] Michael I. Jordan and Tom M. Mitchell. 2015. Machine learning: Trends, perspectives, and prospects. , 255–260 pages. https://doi.org/10.1126/science.aaa8415

[20] Georgios N. Kouziokas. 2017. The application of artificial intelligence in public administration for forecasting high crime risk transportation areas in urban environment. In *Transportation Research Procedia*, Vol. 24. Elsevier B.V., Volos, 467–473. https://doi.org/10.1016/j.trpro.2017.05.083

[21] Mariusz Maciejewski. 2017. To do more, better, faster and more cheaply: using big data in public administration. *International Review of Administrative Sciences* 83, 1_suppl (3 2017), 120–135. https://doi.org/10.1177/0020852316640058

[22] Ninareh Mehrabi, Fred Morstatter, Nripsuta Saxena, Kristina Lerman, and Aram Galstyan. 2019. A Survey on Bias and Fairness in Machine Learning. https://doi.org/10.48550/ARXIV.1908.09635

[23] Albert Meijer and Stephan Grimmelikhuijsen. 2020. Responsible and Accountable Algorithmization: How to Generate Citizen Trust in Governmental Usage of Algorithms. In *The Algorithmic Society* (1st ed.). Routledge, London, 53–66. https://doi.org/10.4324/9780429261404-5

[24] Jan Middendorp and Nico Drost. 2019. Motie van de leden Middendorp en Drost over voorwaarden voor het ontwikkelen van een richtlijn voor het gebruik van algoritmes door overheden [Motion]. https://www.tweedekamer.nl/kamerstukken/moties/detail?id=2019Z12879&did=2019D26547

[25] Ishan Nigam, Mayank Vatsa, and Richa Singh. 2015. Ocular biometrics: A survey of modalities and fusion approaches. *Information Fusion* 26 (11 2015), 1–35. https://doi.org/10.1016/j.inffus.2015.03.005

[26] César Pérez López, María Delgado Rodríguez, and Sonia de Lucas Santos. 2019. Tax Fraud Detection through Neural Networks: An Application Using a Sample of Personal Income Taxpayers. *Future Internet* 11, 4 (3 2019), 86. https://doi.org/10.3390/fi11040086

[27] P. Jonathon Phillips, Amy N. Yates, Ying Hu, Carina A. Hahn, Eilidh Noyes, Kelsey Jackson, Jacqueline G. Cavazos, Géraldine Jeckeln, Rajeev Ranjan, Swami Sankaranarayanan, Jun-Cheng Chen, Carlos D. Castillo, Rama Chellappa, David White, and Alice J. O'Toole. 2018. Face recognition accuracy of forensic examiners, superrecognizers, and face recognition algorithms. *Proceedings of the National Academy of Sciences* 115, 24 (6 2018), 6171–6176. https://doi.org/10.1073/pnas.1721355115

[28] Junfei Qiu, Qihui Wu, Guoru Ding, Yuhua Xu, and Shuo Feng. 2016. A survey of machine learning for big data processing. *EURASIP Journal on Advances in Signal Processing* 2016, 1 (12 2016), 67. https://doi.org/10.1186/s13634-016-0355-x

[29] Rens Raemakers and Kees Verhoeven. 2018. Het bericht 'Algoritme voorspelt wie fraude pleegt bij bijstandsuitkering' [Parliamentary questions]. https://www.nrc.nl/nieuws/2018/04/08/algoritme-voorspelt-wie-fraude-pleegt-bij-bijstandsuitkering-a1598669

[30] Chris Reed. 2018. How should we regulate artificial intelligence? *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 376, 2128 (9 2018), 20170360. https://doi.org/10.1098/rsta.2017.0360

[31] W. D. Rowe. 1975. *An Anatomy of Risk.* Washington D.C.

[32] Laurie A. Schintler and Rajendra Kulkarni. 2014. Big Data for Policy Analysis: The Good, The Bad, and The Ugly. *Review of Policy Research* 31, 4 (7 2014), 343–348. https://doi.org/10.1111/ropr.12079

[33] Cliff Scott and Melissa Medaugh. 2017. Axial Coding. In *The International Encyclopedia of Communication Research Methods*. Wiley, 2455 Teller Road, Thousand Oaks California 91320 United States, 1–2. https://doi.org/10.1002/9781118901731.iecrm0012

[34] Shai Shalev-Shwartz and Shai Ben-David. 2014. *Understanding machine learning: From theory to algorithms* (1st ed.). Cambridge University Press, New York.

[35] Statistics Netherlands. 2022. "Praat met het CBS", de voice assistent van het CBS. https://www.cbs.nl/nl-nl/over-ons/innovatie/project/praat-met-het-cbs-de-voice-assistent-van-het-cbs

[36] Nan Sun, Jun Zhang, Paul Rimba, Shang Gao, Leo Yu Zhang, and Yang Xiang. 2019. Data-Driven Cybersecurity Incident Prediction: A Survey. *IEEE Communications Surveys & Tutorials* 21, 2 (2019), 1744–1772. https://doi.org/10.1109/COMST.2018.2885561

[37] United Nations. 2020. *E-Government Survey 2020 - Digital Government in the Decade of Action for Sustainable Development: With addendum on COVID-19 Response.* Technical Report. United Nations, New York. 364 pages.

[38] Kees Verhoeven, Kathalijne Buitenweg, and Michiel van Nispen. 2018. Motie van het lid Verhoeven c.s. over openbaarheid van de werking en de broncode van algoritmen en analysemethoden [Motion]. https://www.tweedekamer.nl/downloads/document?id=75ae1e3c-3c10-4d8a-b8af-5673c9cfcbd8

[39] Richard L Villars, Carl W Olofson, and Matthew Eastwood. 2011. Big data: What it is and why you should care. *White Paper, IDC* 14 (2011), 1–14.

[40] VVD, D66, CDA, and ChristenUnie. 2021. Omzien naar elkaar, vooruitkijken naar de toekomst - Coalitieakkoord 2021-2025. https://www.rijksoverheid.nl/documenten/publicaties/2022/01/10/coalitieakkoord-omzien-naar-elkaar-vooruitkijken-naar-de-toekomst

[41] Meredith Whittaker, Kate Crawford, Roel Dobbe, Genevieve Fried, Elizabeth Kaziunas, Varoon Mathur, Sarah Mysers West, Rashida Richardson, Jason Schultz, and Oscar Schwartz. 2018. *AI now report 2018.* AI Now Institute at New York University.

[42] Maranke Wieringa. 2020. What to account for when accounting for algorithms: A systematic literature review on algorithmic accountability. In *Conference on Fairness, Accountability, and Transparency (FAT* '20)*. ACM New York, NY, USA, Barcelona, 1–18. https://doi.org/10.1145/3351095.3372833

[43] Andy Williamson. 2014. Big Data and the Implications for Government. *Legal Information Management* 14, 4 (12 2014), 253–257. https://doi.org/10.1017/S1472669614000553

[44] Bernd W. Wirtz, Jan C. Weyerer, and Carolin Geyer. 2019. Artificial Intelligence and the Public Sector—Applications and Challenges. *International Journal of Public Administration* 42, 7 (5 2019), 596–615. https://doi.org/10.1080/01900692.2018.1498103

[45] Samuel Yeom, Irene Giacomelli, Matt Fredrikson, and Somesh Jha. 2018. Privacy Risk in Machine Learning: Analyzing the Connection to Overfitting. In *2018 IEEE 31st Computer Security Foundations Symposium (CSF)*, Vol. 31. IEEE, 268–282. https://doi.org/10.1109/CSF.2018.00027

[46] Lina Zhou, Shimei Pan, Jianwu Wang, and Athanasios V Vasilakos. 2017. Machine learning on big data: Opportunities and challenges. *Neurocomputing* 237 (5 2017), 350–361. https://doi.org/10.1016/j.neucom.2017.01.026