

Artificial Intelligence and Decision-Making: the question of Accountability

Francesco Gualdi

London School of Economics and Political Science
F.Gualdi@lse.ac.uk

Antonio Cordella

London School of Economics and Political Science
A.Cordella@lse.ac.uk

Abstract

Public sector organizations literature has addressed the influence of AI on decision-making process, looking mainly at rationalization and efficiency. However, recent adoptions of AI have been challenged because of their discriminatory nature. As a result, questions emerged on the accountability of AI supported decision-making processes in the public sector. This research sheds light on how AI transforms decision-making processes in the public sector and hence on their accountability. The paper illustrates that AI adoptions lead to the emergency of technological entanglements – assemblages – which might impact upon AI accountability. Building on the findings of some of the most controversial and discussed cases of AI adoption in the public sector – COMPAS in the US and UKVI in the UK – the paper makes the case for a new approach to AI supported public sector decision-making accountability.

1. Introduction

Public sector organizations adopt Artificial Intelligence (AI) systems to improve the efficiency of the provision of public services [1]. The Public Administration (PA) has deployed AI tools in order to reach better informed and more rapid decision-making in multiple service domains: from policing to home office, from criminal justice to healthcare. AI in fact formalises decision-making processes rationalising the decision-making tree in the algorithmic code. Hence, AI reduces administrative discretionary [2] explicating formal and informal rules into the script of the technology [3]. Accordingly, AI encodes into the scripts of the algorithm formal and informal organisation processes which create a new technological accountability which supersedes formal and informal mechanisms of accountability. Decisions taken by AI algorithms follow the logic designed into the script of the technology, which imposes a technological order [4] that transcends rational (the

logic of consequence) and institutional (logic of appropriateness) decision-making logics [5]. In the public sector, these AI adoptions are mostly studied looking at the legal and normative transformations they trigger, however the impacts these transformations have on the accountability of the action of the PA is mostly undermined. To shed light on this crucial aspect related to the adoption of AI in the public sector the paper addresses the following research question:

RQ: what is the accountability of AI supported decision-making process in the public sector?

To answer this question the paper builds on the theory of assemblages [6]. The theory helps to unfold the intertwined nature of AI systems and organization and legal actions which ultimately determine the accountability of the actions of public sector organizations.

2. Background

Artificial Intelligence adoptions in the public sector are not equally researched as those in the private sector [7]. However, scholars have investigated the topic and found that AI systems can benefit public service provision: Wirtz et al. [8] offer a comprehensive review of the most recent researches in this domain.

A robust stream of literature accounts for how AI reduces the red tape and improves the efficiency of the administrative procedures [9]. AI favours cost savings and better resource allocation [10]. AI systems have been successfully adopted in different domains of the public sector to support service provisions and to rationalise administrative tasks [1, 11]. Despite these positive and encouraging outcomes, obstacles and challenges are also identified as direct consequence of AI adoptions in public sector. Overall, scholars accounted for potential problems related to AI governance [12]; from AI ethical standards [13]; to AI's socially acceptance and trust [14]. Despite the attention given to the relationships between ICT and the formalisation of public organisation practices and processes in literature [15, 16], the how AI impacts

these formalisations and the associated decision-making processes has not been extensively researched yet. In this paper we first discuss how AI algorithms operate and their impact on decisions taken in the PA. Hence, we question their impacts on PA decision-making accountability. Recent publications have addressed the problem of AI accountability looking at auditing [17] and algorithmic explainability [18]. While insightful, these contributions do not help shedding light on why AI might challenge the accountability of public sector decision-making processes.

Given the space limits imposed by this submission, we cannot enter in an in-depth discussion of the complex concept of accountability. For the research purpose, we rely on Bovens' [19] definition of accountability as a "mechanism that involves an obligation to explain and justify an actor conduct" to those who can scrutinize and judge the actor's conduct and order sanctions. Formal and informal mechanisms of accountability can be identified [19]. Given the focus (and the space limitations) of this research, and the fact that informal mechanisms of accountability do not fully capture the formalisations imposed by ICT systems [20], we will focus on formal mechanisms of accountability. Hence, we investigate and question the accountability of the techno-legal mechanisms that result from the negotiation between formal, legal mechanisms of accountability with AI systems.

The increased concerns with regards to the discriminating effects of PA decisions mediated by AI systems [1] raises the call to question AI accountability i.e. to explain and justify how AI impacts PA decisions. The lack of accountability of AI decision-making systems have raised concerns on the use of AI to support public sector service delivery [21]. These concerns mainly deal with ethical and legal aspects [22].

The search for AI accountability in this context has been framed as a human/technology control dilemma [8]: humans are not accountable since they are not in full control of AI [23]; or humans are accountable for the technology despite their control over AI [24]. Alternatively, many scholars have focused on transparency as crucial determinant of AI accountability. Hence, scholars focus on technological transparency, i.e. the need to open the "black box" involving decision-making based on AI systems [25], others call for the introduction of designers' "explanation" as reliable practice to improve the algorithmic transparency [18, 26].

Building on this background, this paper aims to make a step further to explain and justify, hence question the accountability of decision-making processes supported by AI in the public sector. We

argue that to fully unpack the accountability of AI-based decision-making process, it is important to shed light on the complex effects of AI systems on the decision-making logics followed by public sector organisations. This research shows that technology is not a neutral tool of transformation: rather, technology reflects but also transcend decision-making drivers of the context where it is deployed [27]. Accordingly, AI systems adopted to support decision-making processes in the PA generate organizational, legal, and institutional, transformations that deeply impact on the logics driving public decision-making and hence on services production and delivery [6, 28].

3. How technology shapes and transforms practices: functional simplification and closure

The literature that analyses how technology engrains and transforms social and institutional practices into the code of technology [29, 30] is very valuable to understand and depict how the socio-institutional transformations engendered by AI adoptions in the public sector impact on the accountability of the decision-making processes. This literature does not specifically study AI but offers an in-depth account of how information and communication technology (ICT) works and impacts on organization practices, which very well applies to AI. Rooted in Luhmann's [31] work on technology, this literature challenges the widespread assumption that understanding ICT is a matter of means and ends [29], and sheds light on the specific trajectories by which ICT transforms existing relations in the contexts of its application.

ICT has properties that structure the causal connections between the organizational practices, processes and activities it mediates [29, 31]. ICT re-maps the organizational procedures and practices it mediates [31] in the predefined logical sequences of actions [32] unique of the code that structures the functioning of each specific software application. By so doing, ICT constructs into the software functionalities sets of structured sequences and interdependences that standardise the execution of organizational procedures and processes. The effect is that ICT imposes regulatory frameworks that frame organizational activities, providing stable and standardized means of interaction and coordination [29]. These regulatory frameworks are shaped into the technical functionalities of the systems. Work sequences and flows are embedded into the technological functions, reducing the procedural complexity of the work in causal or instrumental

relations that are standardized and stabilized in the technical scripts. This happens as the result of the processes of functional simplification and closure [33, 29]. These are the processes by which ICT engrains organization's activities, procedures, or operations into its codes. The code that governs the functions of an ICT system has specific functional requirements that only allow to design activities, processes, or operations to be executed in specific ways [29]. Whenever a public administration designs and deploys an ICT solution in an organizational context, the normative, institutional, and legal logics governing the action of the organization have to be re-written into the code of the ICT. In other words, as explained by Lanzara [6], institutional codes which regulate the functioning of an existing organization are reframed to accommodate the unique code of the chosen ICT system. Moreover, the complexity of all the normative, legal, behavioural and social values that shape the causal connections that determine how an organization works cannot be fully accommodated in the design of an ICT solution. It is not possible to rewrite and translate all these causal connections into the scripts of ICT systems. These casual connections have to be reduced – or, functionally simplified: the language of automation needs to enclose the composite of procedures, actions and relations in standardized and replicable logical sequences [29] which govern the function of the ICT system. In order to successfully replicate the practices reduced into the code of the ICT, the causal connections already simplified need to be functionally closed. That is, the reduction of complexity needs to be isolated from the external context, in order to avoid possible interference from social actors and execution. While functional simplification is obtained through the reconstruction of existing relations in instrumental set of actions [29], functional closure refers to the application of a “protective cocoon” which safeguards the causal connections reduced and allows their replicability over time [29, p.192].

Functional simplification and closure guarantee that ICT reproduces its standards and executes its functions consistently in different contexts of application. Though, the normative, legal, behavioural and social values that are functionally simplified and closed in the ICT change their features for ever. They lose their original characteristics because functional simplification and closure rewrite them in new, stabilised sequences of causal interdependences that make them very difficult to be changed [29]. These sequences of causal interdependences, structured in the code of the ICT, have the ability and the power to redesign organizational processes and procedures. Hence, ICT creates new paths of interdependences that

produce new sets of regulations in the contexts of their application [30]. Therefore, every time ICT systems are deployed to support decision-making processes, they replicate and reproduce the functionally simplified and closed patterns over time [30]. The more the decision-making processes are digitised, the more they get standardised into stable causal and instrumental relations [33], which make them very difficult to modify. By so doing, ICT becomes a regulative regime [34]: it incorporates and reproduces practices and actions in its codes, which carry the power and the authority to regulate how organization's decision-making processes are structured.

4. The creation of assemblages

Functional simplification and closure describe the way by which institutional bodies and norms are transformed (and reduced) in ICT standard. Therefore, the deployment of ICT in public sector organizations produces relevant institutional transformations beyond [6, 30] the impact on rationalisation and effectiveness.

In order to better account for the different impacts of ICT on public sector decision-making processes, it is valuable to consider that to deploy ICTs in complex organizations means to juxtapose a single regulative regime – the one carried by ICT – with other regulative regimes such as normative, legal, institutional and cultural, that constrain how the organization take decisions [34]. ICT adoptions occur either in coordination or collision with other regulative regimes, such as those imposed by normative prescriptions, legal codes, bureaucratic logics, cultural and social systems of values. However, once ICT is deployed it has to negotiate with the other regulative regimes – normative, legal, institutional and cultural – that already regulate how the organization takes decisions. The impact of ICT on organization decision-making is therefore the outcome of the negotiation of the regulative regime carried by ICT with others which already shape organization decision-making processes. This negotiation is complex and time-consuming, but it is important to fully appreciate how ICT impacts and transforms public sector organizations and the way in which these organizations take decisions.

In order to adequately assess how the negotiations between the regulative regime of ICT and the other regulative regimes occur, we have to clearly identify the mechanisms which shape these negotiations. Contini and Cordella [35] observe that the different regulative regimes do not mix or simply overlap: they maintain a degree of autonomy when they encounter in a specific context.

To better understand how different regimes negotiate the way in which the organizations take decisions we have to investigate how the different configurations carried by the regimes overlap, interconnect, clash or intertwine. To explain these negotiations, Lanzara [6] provides the useful concept of “assemblages”. Assemblages are the product of multiple determinations that are not reducible to a single logic: functional linkages and other public sector practices and actions equally count in working out new institutional ecology. According to Lanzara [6], an assemblage is a particular institutional configuration, generated by the negotiation of ICT with existing normative, legal, institutional and cultural arrangements. Every time an ICT is deployed in a public sector organization with the aim to achieve a specific objective, it engages with existing normative, legal, institutional and cultural arrangements which already govern how the organization works. The arrangements are regulated by norms, laws and behaviours which structure the administrative activity and define boundaries of operations. Once ICT is deployed within a specific context, it doesn't impact on an unregulated context. ICT layers on existing regimes which already regulate the organization decision-making processes [6].

The negotiation among the different regimes produces techno-legal-institutional assemblages that are by nature loosely structures [6]. The compatibility of the different regimes is limited and complicated, because the different regimes carry different logics, patterns and systems of values [6, 32]. The negotiations among these regimes involve multiple dimensions and trajectories that are difficult to intergrade and that are not necessarily compatible [6]. This further explains why the adoptions of ICTs in the public sector are challenging, difficult to be managed, and why the outcomes of the transformations they facilitate are also difficult to predict and control [36].

To unfold the impacts of AI on public sector decision-making process, and on the production and delivery of public services, is therefore important to analyse (a) how AI functionally simplifies and closes organization practices and (b) how the negotiation between the regulative regime carried by AI and those already present in the public organization unfolds.

5. AI: functional simplification and closure

Artificial intelligence provides new means to analyse and process information [37] which transform the way in which public sector organizations produce and provide services [8, 1]. Artificial intelligence relies on algorithmic data processing. Algorithms

structure problem solving in formalised and pre-defined rules to be implemented in step by step computational operations [38]. The use of AI in public sector organizations is therefore associated with transformations in data processing that follow the predefined rules formalised in step by step computational operations which constitute the kernel of the algorithmic functions at the core of the AI system. The formalisation of organization processes offers univocal means to execute them, which facilitates and leads the automations of organization decision-making processes. The structured rules that predefine how data processing is executed allow to delegate the execution of organization decisions to the AI system – i.e. this result in the automation of organization processes. The automation of organization processes redefines fundamental organization practices, delegating the execution of tasks that were previously under the responsibility of organization actors, to the AI algorithms [39]. The automation of organization tasks brought by the adoption of AI systems also reshapes the distribution of control and authority within and among organization functions. The responsibility and accountability of organization tasks are not any more exclusively under the traditional organizations' functions and people in charge of it, but rather in the structure of the algorithm that now executes these tasks [7]. The rules that govern the fractioning of the algorithms at the core of the AI system captivate organization decision-making processes. These rules delegate organizations' decision responsibilities and accountability [40] to technological artefacts that enforce pattern of action that are unique, and not necessarily compatible with the contextual normative, administrative and legal governing logics.

6. AI: the assemblage

AI redesigns organizations' decision-making rules and regulations into the functional sequences and interdependences proper of the algorithm. By doing so, it follows the design principle specific of the chosen technology, which can diverge from the trajectory proper of the specific organization's decision-making rules and regulations [7, 6]. The tension that emerges between the different decision-making logics governing the functioning of the AI algorithms and of the organization rules and regulations has implications that are unique given the normative and regulatory nature of public sector organizations. The formalisation of the computational mechanisms that govern the decision-making logic of the AI transforms how the public administration takes decisions, which has impacts on how public services

are produced and provided. Public sector organizations have in fact a unique and very specific role: to enable policies that have been selected by elected leaders [41]. Public sector organizations offices should deploy these policies and thus deliver public services taking decisions that follow the specific prescriptions of the policies. To effectively follow these prescriptions, public sector decision-making processes are organized and regulated on the base of well-defined normative and legal principles. Normative and legal principles provide a frame of action that guarantee effective and coherent decision-making processes. Inscripting normative and legal values into the algorithmic formalizations of AI embeds intrinsic transformations in the decision-making process of public sector organizations [3]. AI supported public sector decisions are defined by the structured predefined computational sequences proper of the algorithm, rather than by the normative and legal framework that govern the driving policy. Changing the process by which decisions are taken directly impacts the values the services generate [42].

Thus, the contribution of AI systems to public sector decision-making must be assessed considering the profound transformation they have on the organization's procedures which govern the public sector decisions [43]. Whenever AI is adopted to sustain the decision-making process in the public sector, such transformations cannot be neglected to evaluate the outcomes of the decision-making and the consequences it has on public services. Accordingly, responsibility and accountability need to be adequately assessed and evaluated against this background.

7. Methodology

In order to illustrate our argument, we discuss two different examples of AI adoption in the public sector. We rely on the exploratory case study approach [44, 45], using secondary data to understand how the algorithm at the core of the functioning of the AI negotiated its functioning with existing norms and laws governing the decision-making processes of the public administration. Exploratory case study approach is relevant to investigate contexts where there aren't clear and single set of outcomes [44]. Through the selected cases, we aim to demonstrate that assemblages emerge from AI adoptions call for a more nuanced understanding of the accountability of AI supported decision-making processes in the public sector. The examples under investigation rely on different AI adoptions to support decision-making processes in the public sector. Data collection includes

secondary sources, mainly governmental documents, Parliamentary inquiries and media releases.

For each case, data collection was undertaken in order to identify the main aspects of each single case study prior to the definition of the research question and hypotheses. Following the exploratory case study approach, the research question was formulated after the data were collected. On the basis of qualitative data analysis and collection, we were able to formulate our research interest, which has turned out to be an explanation of the impact the algorithm at the core of an AI system has on public sector decision-making accountability. Case study research is the most suitable approach for examining a phenomenon in its natural settings [45] and therefore the ideal vehicle for gaining a deeper understanding of the political, social, and technical factors that influence the accountability of AI supported decision-making processes in the public sector.

8. Cases

8.1 UKVI: Speedy boarding for white people?

Since 2015, Home Office UK Visas and Immigration (UKVI) adopted an AI algorithm to sustain the decision-making process in the VISA applications service. The AI is used to streamline applications: it assigned a different colour to applications in relation to the applicant's perceived risk rating. Green (low risk), Amber (medium risk) and Red (high risk). The AI predicts applicants' risk on the basis of criteria such as nationality, ethnicity and age. Once assessed, applications are sent to caseworkers who further scrutinize the files and take the final decision. The AI system has a relevant impact on application processing time, and most important on the application's chances to be approved: figures show that 96.36% of Green applications are accepted. In contrast, 81.08% of Amber applications and to 48.59% of Red applications are granted a VISA. In other words, less than one application out of two, if labelled Red, is granted a VISA [46]. The streamlining activity which assigned a Green, Amber or Red tag to the VISA applications has an impact on the organization of the whole decision-making process. The AI algorithm increased the productivity of the offices which led managers to set a daily target: officials are supposed to produce 75 decisions on Green, 35 on Amber and 25 on Red applications a day. The official report by the Independent Chief Inspector of Borders and Immigrations defined these targets "challenging", and the report emphasized that due to the high expected productivity, officials often relied on the AI

algorithm classification without a proper review on the decisions adopted [46].

Civil society organizations, political forces and media shed a light on the algorithmic assessment process, claiming that it was perpetrating discriminations [47, 48]. Moreover, the report [46] revealed that “(a) tool that streams applications into Green, Amber, Red clusters carries the risk of ‘confirmation bias’” (p.19). In other words, the system perpetrated existing biases related to nationality and hence ethnic. NGOs critics to the policy claimed that the outcome of this procedure was to create a “Fastlane” for white people, while BAME groups had less chances to see their applications approved [49].

The case became subject of a parliamentary debate, where the UK government rejected accusations of racism, stating that system worked in line with prescriptions from the Equalities Act of 2010 [50]. In particular, the Home Office said that the algorithm is used exclusively to streamline applications, and that the final decision is taken by human officials. The purpose of the algorithm is to increase efficiency, saving time and cutting costs in the whole procedure. NGOs and opposition parties asked the government to disclose additional information related to the algorithm logic and functioning [48]. As of January 2020, the government opposed to reveal the algorithm functioning details, such as the factors and the weights that determine the potential risk, or the updating of the algorithm itself [47]. Most recent data show that as of June 2019, UKVI rejected circa 400.000 applications out of 3.3 million received.

8.2 State v Loomis

In 2013, US citizen Eric Loomis was arrested in Wisconsin for driving a car involved in a former shooting case. Loomis pleaded guilty and was accordingly sentenced six years of detention. However, the sentence was not exclusively based on the judge’s evaluation of Loomis’ criminal record: a specific AI system produced a score which expressed Loomis’ potentiality to re-offend. The judge did not decide the sentence exclusively on the score, but they took into account the input produced by the AI system. The software, called COMPAS (Correctional offender management profiling for alternative sanctions), was based on an algorithm developed by business company Northpoint (now Equivant). The algorithm was able to predict the possibility to re-offend of a single citizen, on the basis of historical data and the answers to a questionnaire [51]. Loomis was assessed as an individual with high-risk to re-offend. What is relevant in this case is how the algorithm worked: it did not assess Loomis’ individual recidivist, rather, the

algorithm elaborated the prediction on the basis of a comparison between Loomis’ data and a group of individuals with similar characteristics [52].

Loomis challenged the sentence, claiming that the algorithm-based software was a clear violation of the right to due process [53]. In particular, Loomis’ case highlighted that (a) the algorithm was biased because it utilized information such as gender and race, (b) the algorithm used to predict Loomis’ recidivism by aggregating data on “recidivism risk for groups similar to the offender” [52], (c) that the use of COMPAS did not allow the defendant to access algorithm details. In other words, Loomis challenged the validity and accuracy of a system which sentenced citizens on the basis of an algorithm which could not be accessed by the defendant because it was protected by property rights and patents.

The case was submitted to the Wisconsin Supreme Court, which rejected the defendant claims, pointing out that (a) the risk assessment procedure was mostly correct, (b) the score was only one of the factors taken into consideration by the judge to formulate the sentence, (c) that Loomis could have access to the criteria used by COMPAS, but not on the algorithm mechanisms because they were covered by property rights [54]. However, the Court recommended a careful use of AI-based tools in sentencing: as Justice Roggensack’s concurring opinion demonstrated, judges are entitled to “consider” AI-based information, not to “rely” on them to formulate sentences [54].

However, this collides with the US Supreme Court (i.e. *Townsend v Burke*) ruling which recognized that the due process right to a fair sentencing procedure included “the right to be sentenced on the basis of accurate information”. To this end, the Court in *State v Skaff* went on to underscore that a defendant must be given “means” to investigate and verify the information. This rises a fundamental question: how can the information produced by an AI system be investigated and verified? Hence, how can the AI be used to support courts’ rulings in the US?

9. Discussion

The selected examples offer the opportunity to analyze how AI shapes decision-making in the public sector. Despite their differences, the examples offer relevant insights on the way AI systems work and impact on decision-making processes underpinning public service provision.

The cases show that the choice to implement an AI system to profile individuals on the basis of preset criteria designed in the algorithm produces inputs for the decision-making process that distort the outcome

of the decisions. These inputs are in fact not neutral: rather, they carry values that reflect the way in which AI systems functionally simplify and close the data. To match the needs of the algorithm, data about individuals (immigrants, asylum seekers, convicted felon) are fetched from different databases. In the case of UK from databases which contain information about known patterns in immigration abuses per age or nationality, etc. [46]; in the case of USA from police records, residential records, employment records, etc. [51]. The data sets fetched from the different databases are weighted to determine different risks associated with a profile. The weighting of the variable in the algorithm is defined by its design and the logical constraints of the chosen scripts at the core of the AI system [55]. Moreover, the AI profiling algorithm combines data owned by different agencies. These data are structured following the specific needs of each agency, regardless on whether the data belong to a convicted felon or and immigrant applying for a VISA. The logic which functionally simplified and closed the original data sets (which reflect the needs of the single agency for which the data were initially produced) is neglected when the data are weighted in the algorithm used to predict the level of risk the individual might create for the UK or US authorities. Moreover, the AI algorithm has functions which define unique data patterns which univocally determine how the data are combined and used [55]. The weighting of the different data and the definition of the patterns among data are also functionally simplified and closed data sets and data relations [56]. The weighting produces sets of causal connections that structure the outcomes of data interactions and processing and hence the profiling of the individual.

For example, in COMPAS the Violent Recidivism Risk Scale is calculated as follows [57]:

$$s = a(-w) + a_{\text{first}}(-w) + h_{\text{violence}}w + v_{\text{edu}}w + h_{\text{nc}}w$$

where s is the violent recidivism risk score, w is a weight multiplier, a is current age, a_{first} is the age at first arrest, h_{violence} is the history of violence, v_{edu} is vocation education level, and h_{nc} is the history of noncompliance. The weight w is “determined by the strength of the item’s relationship to person offense recidivism” [57].

Since the UK government refuses to give any details of the factors and relationships used to determine the risk it is not possible to provide a snapshot of how risk is calculated in UKVI case. However, we know that nationality, residence, academic qualification, language proficiency, financial situation, and age are among key criteria used by the algorithm to process applications.

In both cases, data are de-contextualized from the relevant context, and the attributes which define an

individual’s profile are abstracted from the context where they were produced and fragmented into independent data units – such as data cells in the database. As the examples show, the AI systems design in their code a set of new data linkages and relations which abstract individuals’ background and history to accommodate the logic of the algorithm. The AI algorithm constructs a unique profile of the individuals designed in the pattern of relationships that are constructed in the code of the algorithm which does not necessary reflect the real profile of the single person. In the UK or USA cases the risk identified by the algorithm is not the individual risk but the output of the data relationships constructed in the algorithm. The algorithm provides univocal ways to infer how different individual characteristics such as gender, age, nationality, combine to determine the social risk. The action of reducing causal connections is noteworthy necessary: without this action, the algorithm would not be able to work and produce the profiling. Also, to be effective, the set of causal connections must be isolated from the external context and made stable over time; hence it has to be functionally closed. Therefore, the process by which the AI profiling tool reconstructs the individual’s history is “cocooned” into the script of the algorithm. The procedure by which the AI processes the data stored in different databases is engrained in the script of the technology and cannot be modified if not following a formal complex technical procedure [58].

Profiling the individuals following the script of the algorithm changes the underline rules of the decision-making process. As the selected cases show, the profiling follows the logic of the script and not necessarily those prescribed by the law or by the human-based procedure. Indeed, all projects were criticized because the profiling failed to adhere to the existing legal and administrative framework. When the algorithm functionally simplifies data relationships and data interdependences, it creates a new assemblage where the regulatory regime of the law is intertwined with regulatory regime of the technology. In other words, the decisions on VISA streamlining and on justice sentencing are negotiated between the regulatory regimes of the law and of the technology.

The cases reveal that the outcome of the decision-making process impacted by AI systems is different from the decision-making made by a human agent. In fact, the algorithm does not assess the data that belong to a single person in the light of individual characteristics and experiences; rather, it elaborates the outcome of a preset collective experience reconstructed according to technological standards. The results are indeed difficult to connect to the backgrounds of the individuals that are profiled. As a

consequence, the AI prediction of an individual to be harmful for the British or American society depends on the outputs of a profiling tool which works on the basis of a world which is reconstructed by functional simplification and closure made by the AI algorithm. Those who have to take the final decision will only have access to data that are decontextualized. Hence, they will not be able to assess the individual case. The human-mediated final decision-making process emerges as a new assemblage where decision makers take decisions using abstract data and data combined by the AI, instead of relying on individual data and circumstances as prescribed by legal norms and rules.

Building on these findings, we aim to critically evaluate the concept of accountability when it relates to AI-based decision-making. The two selected cases show that the outcome of the decision-making process informed by AI has been formally challenged. The UK government rejected allegations of discriminating practices related to AI, and so did the Wisconsin Supreme Court. Interestingly, public institutions built their counter-argument on the basis of two considerations: (a) the algorithm had an ancillary role to the decision-making, and the public sector final decision relied on the human evaluation and (b) it was denied the right to access algorithms used to support decision-making.

With respect to the first point, public institutions claimed that the decision-making process is made of two functions, automation and human evaluation. These two functions are not only separated, but reflect a hierarchical difference, because the human evaluation might overrule the information produced by AI. However, since the mechanisms that govern the functioning of the algorithm are not made public and transparent, the decision makers cannot know how and why a specific profiling is made. They have to take decisions using inputs different than those they had to rely upon had the AI system not been in place. The AI systems make recommendations using patterns that are not known to the decisions makers. In both cases, the decision makers have to judge and justify if they do not follow the AI recommendations, but not if they agree with it.

For instance, in the UKVI case, public managers in charge of the final decisions on VISA are not expected to review a decision which was in line with the streaming rating [46]. In other words, Green applications approved, and Red applications rejected remained undisputed. When asked about possible allocating errors by the AI algorithm, public managers showed confidence on the accuracy of the algorithm streaming rating [46] even if they, as well as the public, are not given any access or explanation on how the algorithm works. Hence, the human decision

makers' actions are constrained by the functions executed by AI algorithm. Decision makers receive black boxed inputs to structure their decision-making. Even if the final decision is not taken by the AI, the AI produces key inputs for the decision makers that are by nature undisputable. The functional simplification and closure and the algorithmic interdependent create assemblages that are very difficult if not impossible to untangle [56].

Therefore, when it comes to accountability, we posit that to fully understand who is accountable for decisions that rely on inputs generated by AI systems, we have to consider the whole assemblage, where responsibilities are shared by the technology and by humans. Although valuables, contributions that underline that human agents should take full responsibility for the machines [59: pp. 19-20] do not capture the profound transformative power that AI holds and deploys whenever adopted in a decision-making process, even if only to assist the decision makers. Responsibility is not exclusive: it is shared, precisely because of the nature of assemblages created by the adoption of technology. The structuration of assemblages blurs the boundaries of accountability, because the negotiation between existing regimes reshapes the decision-making process: technology and humans activities intertwine in a new entanglement that has to be adequately accounted. Accordingly, it is necessary to evaluate the entire process with all the inputs that might come from technology and humans, and make the new process accountable as a whole.

With respect to the second point – the fact that public sources often deny the access to algorithms details – we do not neglect the relevance of property rights and patents. However, if the public sector decides to rely on private firms' consultancy to elaborate algorithms which might serve in public sector decision-making, we do believe that the public sector has additional duties of accountability to the citizens. On this point, some scholars have advanced possible solutions in order to provide the best possible transparency without releasing algorithms' details [51, 26]. Although valuable, these contributions do not completely capture the profound transformations that AI produces in the decision-making process. Through this paper, we have investigated cases where the outcome of the AI supported decision-making had a severe impact on citizens' life. Given the increasing relevance of the AI impact on public service provision, there is a need for a better understanding of the consequences of AI adoption to inform public sector decision-making processes. To this purpose, and following many public bodies and institutions already claiming for it [46, 59], we believe that opening the algorithm, and exposing to the public scrutiny the

steps of the decision-making AI supports will be a substantial action towards a better accountability. The light shall be directed on the accountability of the assemblage constituted by technological, institutional and legal dimension and not independently on each dimension.

10. Conclusions

AI in public sector organizations is often deployed to achieve the objectives of rationalization and efficiency. However, given the increasing diffusion of AI tools, public managers should better consider the multidimensional impact of AI on the normative and legal frameworks which underpin public administration decision-making processes. Through this research, we shed light on the impact of AI on the decision-making processes in the public sector. The cases of UKVI and State v Loomis offer interesting insights on the impacts of the adoption of AI on the accountability of public sector organizations decision-making. AI functionally simplifies and closes administrative processes. As a consequence, AI has impacts that transform the logic underlining the decision-making processes. The cases show that AI profiles in a way which is quite different from what is prescribed by legal and normative frameworks. Hence, AI can produce outcomes which do not reflect relevant factual circumstances. When these outcomes are used as inputs to support public sector decision-making, issues of accountability arise. As the examples show, institutions denied allegations of discriminations regarding the algorithms. At the same time, a request for further disclosure of algorithms' details has been turned down. Algorithms can be seen as powerful carrier of modernization in public organizations: however, their capacity to create value has to be closely scrutinized. To this purpose, the research reveals that unpacking how algorithms work is needed to adequately understand the impact AI has on the accountability of public sector decision-making.

11. References

[1] W. G. de Sousa, E. de Melo, P. Bermejo, R. Farias, and A. Gomes, "How and where is artificial intelligence in the public sector going? A literature review and research agenda", *Government Information Quarterly*, 36(4), 2019.

[2] P. A. Busch and H. Z. Henriksen, "Digital discretion: A systematic literature review of ICT and street-level discretion", *Information Polity*, 23(1), 2018, pp. 3-28.

[3] M. Bovens and S. Zouridis, "From street-level to system-level bureaucracies: How information and communication technology is transforming administrative discretion and constitutional control", *Public Administration Review*, 62(2), 2002, pp. 174-184.

[4] S. Timmermans and M. Berg, "The practice of medical technology", *Sociology of health & illness*, 25(3), 2003, pp. 97-114.

[5] March, J. G., and J. P. Olsen, *Rediscovering Institutions*, The Free Press, New York, US, 1989.

[6] Lanzara, G. F., "Building digital institutions: ICT and the rise of assemblages in government". In Contini, F., and Lanzara, G.F. (Eds), *ICT and Innovation in the Public Sector: European Studies in the making of eGovernment*, London, UK: Palgrave MacMillan, 2009, pp. 9-48.

[7] K. Martin, "Ethical implications and accountability of algorithms", *Journal of Business Ethics*, 160(4), 2019, pp. 835-850.

[8] B. Wirtz, J. Weyerer, and C. Geyer, "Artificial Intelligence and the Public Sector - Applications and Challenges", *International Journal of Public Administration*, 42:7, pp. 596-615, 2019.

[9] W. D. Eggers, T. Fishman, and P. Kishnani, "AI-augmented human services: Using cognitive technologies to transform program delivery". Edited by Deloitte, C. Wernick and C. Wolff, 2017.

[10] H. Mehr, H. Ash, and D. Fellow, "Artificial intelligence for citizen services and government". *Ash Cent. Democr. Gov. Innov. Harvard Kennedy Sch.*, no. 1-12, 2017.

[11] G. Kouziokas, "The application of artificial intelligence in public administration for forecasting high crime risk transportation areas in urban environment". *Transportation Research Procedia*, 24, 2017, pp. 467-473.

[12] U. Gasser and V. Almeida, "A layered model for AI governance". *IEEE Internet Comp.*, 21(6), 2017, pp. 58-62.

[13] B. Mittelstadt, P. Allo, M. Taddeo, S. Wachter and L. Floridi, "The ethics of algorithms: Mapping the debate", *Big Data & Society*, 3(2), 2016, pp. 1-21.

[14] J. Zerilli, A. Knott, J. Maclaurin, and C. Gavaghan, "Transparency in algorithmic and human decision-making: is there a double standard?", *Philosophy & Technology*, 32(4), 2019, pp. 661-683.

[15] T. Schillemans and M. Busuioc, "Predicting public sector accountability: From agency drift to forum drift", *Journal of Public Administration Research and Theory*, 25(1), 2015, pp. 191-215.

[16] J. Carlo Bertot, P.T. Jaeger and J.M. Grimes, "Promoting transparency and accountability through ICTs, social media, and collaborative e-government", *Transforming Government: People, Process and Policy*, Vol. 6 No. 1, 2012, pp. 78-91.

[17] C. Sandvig, K. Hamilton, K. Karahalios, and C. Langbort, "Auditing algorithms: Research methods for detecting discrimination on internet platforms", 64th Annual Meeting of the ICA, May 22, 2014, Seattle, WA, USA.

[18] M. Oswald, "Algorithm-assisted decision-making in the public sector: framing the issues using administrative law rules governing discretionary power", *Phil. Trans. R. Soc. A* 376: 20170359, 2018.

[19] M. Bovens, "Analysing and Assessing Accountability: A Conceptual Framework", *European Law Journal*, 13(4), 2007, pp. 447-468.

[20] B. S. Romzek, K. LeRoux, and J. M. Blackmar, "A Preliminary Theory of Informal Accountability among Network Organizational Actors", *Public Administration Review*, 72(3), 2012, pp. 442-453.

- [21] K. de Fine Licht and J. de Fine Licht, “Artificial intelligence, transparency, and public decision-making”, *AI & Society*, 2020.
- [22] D. Helbing et al., “Will Democracy Survive Big Data and Artificial Intelligence?”, in: Helbing D. (Ed.), *Towards Digital Enlightenment*, 2019, Springer, Cham.
- [23] A. Matthias, “The responsibility gap: Ascribing responsibility for the actions of learning automata”, *Ethics and Information Technology*, 6(3), 2004, pp. 175–183.
- [24] De George, R. T., *The ethics of information technology and business*. Malden, MA: Blackwell Publishing, 2003.
- [25] L. Floridi, J. Cowls, M. Beltrametti, R. Chatila, P. Chazerand, V. Dignum, and B. Schafer, “AI4People - an ethical framework for a good AI society: opportunities, risks, principles, and recommendations”, *Minds and Machines*, 28(4), 2018, pp. 689-707.
- [26] F. Doshi-Velez, M. Kortz, R. Budish, C. Bavitz, S. Gershman, D. O'Brien, and A. Wood, “Accountability of AI under the law: The role of explanation”. *arXiv preprint arXiv:1711.01134*, 2018.
- [27] A. Cordella and F. Iannacci, “Information systems in the public sector: The e-Government enactment framework”, *The Journal of Strategic Information Systems*, 19(1), 2010, pp. 52-66.
- [28] A. Cordella, and C. M. Bonina, “A public value perspective for ICT enabled public sector reforms: A theoretical reflection”, *Government Information quarterly*, 29(4), 2012, pp. 512-520.
- [29] J. Kallinikos, “The order of technology: complexity and control in a connected world”, *Accounting, management, and information technologies*, 15 (3), 2005, pp. 185-202.
- [30] Kallinikos, J., *The Consequences of Information: Institutional Implications of Technological Change*, Edward Elgar, Cheltenham, UK, 2006.
- [31] Luhmann, N., *Risk: a sociological theory*, Transaction Publishers, New Brunswick, New Jersey, 2005.
- [32] C. Ciborra and O. Hanseth, “From Tool to Gestell: Agendas for Managing the Information Infrastructure”, *Information Tech. and People*, 11 (4), 1998, pp. 305-327.
- [33] A. Cordella and N. Tempini, “E-government and organizational change: Reappraising the role of ICT and bureaucracy in public service delivery”, *Government Information Quarterly*, 32(3), 2015, pp. 279-286.
- [34] J. Kallinikos, “Institutional complexity and functional simplification: the case of money claim online service in England and Wales”. In Contini, F. and G. F. Lanzara, (Eds), *Ict and Innovation in the Public Sector: European Studies in the making of eGovernment*, Palgrave MacMillan, London, UK, 2009, pp. 174-210.
- [35] F. Contini and A. Cordella, “Law and technology in civil judicial procedures”. In Brownsword, R., E. Scotford, and K. Yeung, (Eds), *The Oxford handbook of the law and regulation of technology*, Oxford University Press, Oxford, UK, 2016, pp. 246–268.
- [36] M. Hildebrandt, “Saved by Design? The Case of Legal Protection by Design”, *Nanoethics*, 11(3), 2017, pp. 307-311.
- [37] Brynjolfsson E. and McAfee, A., *The second machine age: Work, progress, and prosperity in a time of brilliant technologies*, WW Norton&Company, New York, 2014.
- [38] Conte, S. D. and C. De Boor, *Elementary numerical analysis: an algorithmic approach*. Society for Industrial and Applied Mathematics, Philadelphia, US, 2017.
- [39] G. von Krogh, “Artificial intelligence in organizations: New opportunities for phenomenon-based theorizing”, *Academy of Manag. Discoveries*, 4 (4), 2018, pp. 404-409.
- [40] A. Shrestha and A. Mahmood, “Review of deep learning algorithms and architectures”, *IEEE Access*, 7, 2019, pp. 53040-53065.
- [41] J. P. Olsen, “Maybe it is time to rediscover bureaucracy”, *Journal of Public Administration Research and Theory*, 16(1), 2006, pp. 1-24.
- [42] A. Cordella, “E-government: towards the e-bureaucratic form?”, *Journal of Information Technology* 22, 2007, pp. 265–274.
- [43] Moore, M. H., *Creating public value: Strategic management in government*, Harvard University Press, Harvard, US, 1995.
- [44] Yin, R. K., *Case study research: Design and methods*, Sage Publications, Thousand Oaks, US, 2003.
- [45] P. Baxter and S. Jack, “Qualitative case study methodology: Study design and implementation for novice researchers”, *The qualitative report*, 13(4), 2008, pp. 544-559.
- [46] Independent Chief Inspector of Borders and Immigration (ICIBI), “Report on entry clearance processing operations in Croydon and Istanbul, July 2017”.
- [47] H. Warrell, “Home Office under fire for using secretive visa algorithm”, *Financial Times*, 09/06/2019.
- [48] H. McDonald, “AI system for granting UK visas is biased, Rights groups claim”, *The Guardian*, 29/10/2019.
- [49] C. Threipland and O. Rickett, “Price and prejudice: automated decision-making and the UK government”, *The Justice Gap*, 17/03/2020.
- [50] “Visa Processing Algorithms”, transcript, The House of Commons, 19th June 2019.
- [51] H. Liu, C. Lin, and Y. Chen, “Beyond State v Loomis: artificial intelligence, government algorithmization and accountability”, *International Journal of Law and Information Technology*, 27(2), 2019, pp. 122-141.
- [52] “State v. Loomis - Wisconsin Supreme Court Requires Warning Before Use of Algorithmic Risk Assessments in Sentencing”, 130 *Harv. L. Rev.* 1530.
- [53] K. Freeman, “Algorithmic injustice: How the Wisconsin Supreme Court failed to protect due process rights in State v. Loomis”, *North Carolina Journal of Law & Technology*, 18(5), 75, 2016.
- [54] *State v. Loomis*, 881 N.W.2d 749 (Wis. 2016).
- [55] S. Faraj, S. Pachidi and K. Sayegh, “Working and Organizing in the Age of the Learning Algorithm”, *Information and Organization* (28), 2018, pp. 62-70.
- [56] W. Samek, T. Wiegand and K.-R. Müller, “Explainable Artificial Intelligence: Understanding, Visualizing and Interpreting Deep Learning Models”, *arXiv preprint arXiv:1708.08296*, 2017.
- [57] Northpointe, “Practitioner’s Guide to COMPAS core”, 22 March 2015.
- [58] Norman, D., *The Design of Everyday Things*, Basic Books, 2013.
- [59] The Committee for Standards in Public Life, “Artificial Intelligence and Public Standards”, February 2020.