Algorithmic Profiling in Public Employment Services: a Systematic Review on the Effects on Caseworkers and Jobseekers

Mareike Sirman-Winkler, Daria Szafran and Sonja Mei Wang

June 17, 2024

Manuscript prepared for the 5th Street-level Bureaucracy Conference,
18-20 June 2024, Aalborg University

Algorithmic Profiling in Public Employment Services

A Systematic Review on the Implications for Caseworkers and Jobseekers

Mareike Sirman-Winkler (WZB Berlin Social Science Center)[1]
Daria Szafran (University of Mannheim)[2]
Sonja Mei Wang (University of Wuppertal)[3]

Draft Version[4]. 1st June 2024. Please do not cite without permission.[5]

Abstract:
For decades, Public Employment Services have used algorithmic systems to profile jobseekers. Algorithmic profiling systems aim to improve the efficiency of PES. However, some systems have faced criticism from academia and civil rights organizations. This criticism is centered around the limited discretion caseworkers have due to system design and the difficulty jobseekers face in rejecting an assigned profiling score. For this reason, the chances and risks of algorithmic profiling are currently the subject of intensive research, particularly in light Artificial Intelligence (AI). To date, however, relatively little is known about the actual impact of algorithmic profiling on caseworkers and jobseekers. Country-specific case studies often deal with this topic, but they are scattered across disciplines. That is why we conduct a pre-registered systematic literature review to identify those articles. Our study investigates the impact of algorithmic profiling on caseworkers and jobseekers. Specifically, we examine the implications of algorithmic profiling on (1) efficiency, (2) accountability, (3) transparency, and (4) contestability. The synthesis of nine empirical studies, featuring algorithmic tools in European countries, show that the intention (e.g., increased efficiency) and the actual impact of profiling tools diverge. Algorithmic profiling increases the administrative burden and its use strongly depends on the acceptance of caseworkers. It also redefines caseworkers' discretionary power, knowledge management and skill formation. Jobseekers face inclusion challenges if they do not have sufficient resources (such as digital skills) to access digital public services. But profiling can also contribute to transparency which is a prerequisite for being able to contest algorithmic decisions. This study enhances the understanding of the role of digital tools for the work of street-level bureaucrats and for citizens. The findings can contribute to the public debate about digitalization in public administration by systematically identifying and compiling a differentiated picture of the impact and risks of algorithmic profiling. The results can contribute to evidence-based policy making in the age of digital government.

Word count: 8,802

---

[1] Mareike Sirman-Winkler, WZB Berlin Social Science Center, Reichpietschufer 50, 10785 Berlin, Germany; E-Mail: mareike.sirman-winkler@wzb.eu. ORCID: 0000-0002-8608-9936. (Corresponding author)

[2] School of Social Sciences, University of Mannheim, A5, 6, 68159 Mannheim, Germany; E-Mail: daria.szafran@uni-mannheim.de. ORCID: 0000-0001-6387-7190

[3] University of Wuppertal, Institute for Sociotechnical Systems, Gaußstraße 20 42119 Wuppertal, Germany; E-mail: swang@uni-wuppertal.de, ORCID: 0000-0003-1219-8117.

[4] This manuscript is part of an ongoing research project.

[5] Acknowledgement: The views expressed in this study are those of the authors. Any remaining errors are the sole responsibility of the authors.

1 INTRODUCTION

In many countries, Public Employment Services (PES) have introduced algorithmic profiling systems. They are used, e.g., to assess jobseekers' chances on the labor market and select suitable labor market policy instruments, like further training measures (Körtner & Bonoli, 2022). Algorithmic profiling systems have been introduced for reasons such as improving efficiency of counseling (Allhutter et al., 2020) or personalizing services (Sztandar-Sztanderska & Zielenska, 2018). However, the intention (e.g., increased efficiency) and the actual impact of profiling tools can diverge. In some cases, algorithmic profiling led to discrimination. That is why some of these systems have been subject to criticism by academia, civil rights organizations and other groups. Criticism includes difficulties of rejecting an assigned score and limited caseworker discretion due to system design (Allhutter et al., 2020). Moreover, concerns have been raised that systems which are intended to be used as decision support by PES staff can run on an almost automated basis, as argued by Niklas et al. (2015).

For this reason, the chances and risks of algorithmic profiling are currently the subject of intensive research, particularly in light of Artificial Intelligence (AI). Algorithmic profiling systems differ vastly from each other in the method used for calculating chances, e.g., statistical methods or AI methods. This can have an impact on transparency of the respective system and explainability of the calculated outcome. If explainability is low, then interpretability is also impacted. Understanding how an algorithmic system arrives at a certain outcome is necessary to understand possible discrimination. In the case of unequal treatment, the transparency of algorithmic systems could provide jobseekers with an opportunity to argue against an assigned profiling score. With the increased use of algorithmic profiling systems in recent years, caseworkers' discretion to decide on measures and funding available to jobseekers has changed. Assigned outcomes can have a real impact on a jobseeker's access to resources (for example, trainings), making this an important field of research for an informed public discourse.

To date, however, relatively little is known about the actual impact of algorithmic profiling on caseworkers and jobseekers. We consider that this is important for evidence-based policy-making, in particular in the light of current rapid developments in the field of AI. Against this background, our research question is: What are the implications of algorithmic profiling for caseworkers and jobseekers?

Specifically, we want to examine the implications of algorithmic profiling for (1) efficiency as perceived by the caseworkers, (2) accountability, and (3) transparency as well as (4) jobseekers' possibilities to contest the decision of an algorithmic profiling tool (contestability).

Algorithmic profiling tools have been the focus of country-specific case studies. These studies, however, are oftentimes scattered across disciplines. Therefore, we conduct a qualitative systematic literature review to identify and synthesize relevant articles. The focus of our systematic review are qualitative studies. Qualitative methods make it possible to gain an in-depth understanding of how algorithmic profiling is used in practice and how tools are perceived by stakeholders who are directly impacted by their introduction and use (Braun and Clarke 2021). Thematic Analysis: A Practical Guide. London: Sage. The goal is to enrich the discussion about the heterogeneity of algorithmic profiling systems, their impact on caseworkers and jobseekers, and how this information can inform the development of future systems, taking into account the views of these stakeholders (Sweet & Moynihan, 2007). We also focus on various algorithmic profiling systems which have been introduced for reasons such as improving efficiency of counseling (Allhutter et al., 2020) or personalizing services (Sztandar-Sztanderska & Zielenska, 2018).

The remainder of this article is structured as follows: Section 2 presents the theoretical background. Section 3 presents the research design. The empirical results are presented and discussed in Section 4. Section 5 summarizes the results and discusses their implications.

2 THEORETICAL BACKGROUND

2.1 Algorithmic profiling in Public Employment Services

A wide range of profiling approaches exist: There is rule-based profiling, caseworker-based profiling, statistical profiling, and artificial intelligence-based profiling (AI-based profiling) (Desiere et al., 2019; Desiere & Struyven, 2021)[6]. According to

[6] There is a large body of literature on different profiling approaches used in OECD countries (Barnes et al. 2015; Desiere et al. 2019; Loxha and Morgandi 2014). Barnes et al. (2015) describe the same approaches as Desiere et al. (2019), with the addition of soft-profiling which is defined as "a combination of eligibility rules, caseworker discretion, administrative data and more subjective, qualitative assessments and psychological screening tools". Loxha and Morgandi (2014) propose an analytical framework of profiling approaches, based on the level of caseworker discretion and the complexity of the information flow, which consists of caseworker-based profiling, rules-based

Desiere et al. (2019), rule-based profiling is based on administrative eligibility criteria such as age and educational level, and was used in the previous profiling system of Flanders (Belgium) until 2018. In this system, the final decision was taken by the caseworkers, as they could accept or correct the automated classification. Caseworker-based profiling is based on the judgment of caseworkers, and is used in countries such as Germany and Switzerland, although there is some element of statistical profiling in the German PES through a software tool used to calculate labor market chances, which has been criticized for its opacity (AlgorithmWatch, 2019). Statistical profiling is based on statistical models to predict the chances of finding new employment (Desiere et al., 2019).

Algorithmic profiling is "any form of automated processing of personal data in order to analyze or predict personal aspects of individuals" (Scott et al., 2022, p. 2). In the context of PES, it is used "to assess jobseekers, allocate resources, and evaluate further steps" (Scott et al., 2022, p. 2). These algorithmic profiling approaches differ in terms of the data and methods used, caseworker discretion, and the outcomes of the profiling systems. For example, in the Dutch system, jobseekers fill out an online questionnaire which calculates a probability of finding new employment within 12 months (Desiere & Struyven, 2021), whereas in the currently suspended Austrian system profiling takes place during an in-person interview in which the counselor is informed of how the system categorizes the job seeker. In the Austrian system, profiling outcomes are used for resource allocation and data is entered into the system during an in-person counseling conversation, whereas the percentage of the Dutch system is used to decide which job seekers should be invited to in-person counseling. Profiling systems can also differ in terms of caseworker discretion, ranging from full caseworker discretion, to decision support, where the final decision rests with the caseworker, to automated decision systems, where no caseworker discretion is possible.

2.2 Implications of algorithmic profiling for Public Employment Services

A critical promise of supporting street-level bureaucrats' decision-making with

profiling, and data-assisted profiling. Caseworkers play a central role in data-assisted profiling, as they are responsible for job seeker segmentation and resource allocation, but they use data more intensively than in caseworker-based profiling (Loxha and Morgandi 2014). Regarding the types of profiling proposed by Loxha and Morgandi (2014) and Barnes et al. (2015) it has to be taken into account that they were published in 2014 and 2015, respectively, and in recent years AI tools have become more pervasive in society.

algorithmic tools is to enhance the efficiency of public service delivery for users and clients (Bullock, 2019; Zuiderwijk et al., 2021). Predicting jobseekers' labor market chances, identifying suitable jobs and providing tailored further training recommendations are complex tasks. However, algorithmic profiling tools (if they are AI-based) can infer patterns from thousands of employment histories of similar jobseekers and thus make accurate recommendations, freeing up caseworkers' time and resources. At the same time, jobseekers can benefit from such profiling systems, as more appropriate training can help them get back into a (possibly better-paying) job in the long run. The convenience and time saved by relying on algorithmic recommendations is an obvious metric for evaluating the benefits of such a system. In this sense, both caseworkers and jobseekers would benefit from more efficient services. At this point, the question arises as to why PES do not take full advantage of the promised efficiency gains from algorithmic profiling by integrating such systems into their work processes. The answer may be related to how such automated tools affect administrative accountability.

A particular implication that algorithmic profiling brings to public administration is the challenge of maintaining accountability (Busuioc, 2021; König & Wenzelburger, 2021b). Accountability in this context means that an agency or public servant is able to explain or justify his or her decisions (Bannister et al., 2020; Busuioc, 2021). However, this can be challenging when AI is used in administrative decisions, such as in profiling. Caseworkers may no longer understand the decision-making of the AI system and as a result may lose control over the outcomes. This 'responsibility gap' (Wirtz et al., 2018) can call the accountability of administrative decisions into question (König & Wenzelburger, 2021a; Krafft et al., 2020). The tension arises from the pressure to use AI (e.g., by governments or citizens) while maintaining accountability which, Busuioc (2021) emphasizes, "is a hallmark of bureaucratic legitimacy and one that administrators cannot outsource or renege on". Ultimately, accountability is also the definition of a legal status as to who is responsible for the decisions of ADM systems (Wirtz et al., 2018).

Criticism has also been raised in relation to transparency. Berman (2023), e.g., discusses the explainability of the algorithmic profiling system of the Swedish PES which is based on a machine learning (ML) model. He notes that a more opaque ML-based model was chosen over a statistical model. The two versions have almost the same prediction accuracy, but the statistical model has a much higher transparency, raising the question of why this was the case. In the Dutch profiling system, jobseekers fill out an online questionnaire (Wijnhoven & Havinga, 2014). The results are only visible to caseworkers, jobseekers are not informed of the outcome and thus cannot interpret any possible outcomes.

Another aspect discussed in the context of algorithmic profiling is the affected subjects' possibility to contradict or revoke the algorithmic decision. This right to contest is ensured in the European context by Article 22 of the EU's General Data Protection Regulation (GDPR). It constitutes a legal safeguard for individuals who are subject to algorithmic decisions and fosters their active engagement in the decision-making process (Klutzz et al., 2022). However, such legal measures aimed at protecting citizens' fundamental rights against potentially discriminating algorithmic outcomes are inevitably post-hoc (Almada, 2019). Additionally, contestability should be incorporated already throughout the process of software design, as an integral part of algorithmic decision-making tools. (Alfrink et al., 2022) reviewed recent literature on AI contestability by design to provide a framework consisting of various build-in system features and development practices that aim at increasing the contestability of algorithmic systems. The features include, for example, opportunities for active human oversight and correction, the provision of explanations that involve the social, organizational, and legal context of an automated decision, or the possibility to request an intervention. The practices enhancing contestability identified by Alfrink et al. (2022) span the complete lifecycle of an algorithmic system ranging from early stages of AI system development (e.g., developers should anticipate the consequences of their tools), through the testing phase (e.g., developers should ensure that their AI systems are actually responsive to contestation) to the monitoring after AI systems have been deployed (e.g., constant control of system performance for unfair outcomes should be ensured).

The theoretical implications of algorithmic profiling lead us to examine how algorithmic profiling actually affects people today. To this end, we look at the perspectives of caseworkers and jobseekers. The perspective on the social implications is also recommended in existing research on AI in public employment services (e.g., (Bloch Haug, 2022).

3 METHOD

We conducted a systematic literature review (Petticrew & Roberts, 2006) of qualitative studies on the implications of algorithmic profiling tools on both caseworkers and jobseekers. The review is based on the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) (Page et al., 2021). The corpus of studies was analyzed qualitatively using the thematic synthesis approach according to Thomas and Harden (2008). Before starting our systematic search, we preregistered our exact approach on Open Science Framework (OSF)[7]. We used the PRISMA checklist (Table A2) to report the review process (Moher et al., 2009).

3.1 Developing the research question

A systematic literature review requires a profound examination of the research question (Booth et al., 2016). Booth et al. (2016) emphasize that the nature of the research question is essentially influenced by the objective and focus of the review. Based on this, they distinguish three types of research questions: effectiveness questions, methodology questions, and conceptual questions. Since our research interest is on the effects of algorithmic profiling on caseworkers and jobseekers, our research question is an effectiveness question.

Petticrew and Roberts (2006) recommend to specify the review question by breaking it down into sub-questions. When asking a question about effectiveness, they recommend using the PICOC method (population, intervention, comparison, outcomes, context) in order to consider the components of the question. In our paper, the ‘population’ consists of stakeholders, who are directly affected by the implementation of algorithmic profiling systems in PES, i.e., caseworkers as well as jobseekers. The ‘intervention’ is the use of algorithmic profiling as a set of tools to make PES more efficient. Like Starke et al. (2022), we did not specify ’comparison’, e.g., between statistical profiling and caseworker-based profiling (because we do not consider it beneficial to our topic). The ‘outcomes’ are the effects on the relationship between jobseekers and caseworkers, as described by these groups. Because these effects, especially in social science research, may not be dichotomous but multi-layered, Petticrew and Roberts (2006) recommend considering not only the success of

[7] We used the 66 item Generalized Systematic Review Registration Form by Van den Akker et al. (2023). The (anonymized) link to preregistered study is: https://osf.io/2xtdj/?view_only=841da9e24e9843b1aab5926b94a1a8ea

an intervention but its broader 'context'. The context arises from the nature of the subject under investigation: the public employment services. Following Starke et al. (2022), we did not specify the context any further, e.g., by limiting it to a specific geographical region, in order to include as much empirical data as possible. Based on these criteria, we derived the following research question to guide our systematic review: *What are the implications of algorithmic profiling for caseworkers and jobseekers in the context of Public Employment Services?* We answer this question with respect to the criteria we discussed above, namely efficiency, accountability, transparency and contestability.

3.2 Literature search and study selection

To optimize our literature search strategy, we adopted an automated approach to identify relevant keywords using R (R Core Team 2023) and the package *litsearchr* (Grames et al., 2019). This approach encompasses several steps, which we describe in the following. First, we conducted a naive search on December 6, 2023 in two databases (EBSCO and Web of Science) using our initial set of keywords (see Table A1) that we identified applying the "pearl-growing" method (Booth et al., 2016, p. 314). For this purpose, we determined three papers as pearls, chosen based on the venue they were published, high citation count for the field of study, and recency of publishing (Ammitzbøll Flügge et al., 2021; Petersen et al., 2021; Sztandar-Sztanderska & Zielenska, 2020). Starting from those pearls, we used backward- and forward-searching through citations and added additional papers as sources for our keywords (Table A1). The naive search yielded 1,774 unique articles. The R package *litsearchr* allowed us then to extract potential search terms from the articles' titles, abstracts, and tagged keywords using a function that approximates the *Rapid Automatic Keyword Extraction* (RAKE) algorithm.

Next, we created a keyword co-occurrence network that we quantitatively assessed to detect a cutoff point of keyword importance, which allowed us to remove terms that were not central to our field of study. This resulted in 3,031 potential search terms that we further reduced by screening out the terms containing words unrelated to the topic of our study. Finally, we manually reviewed the remaining 1,642 potential search terms. The manual revision yielded 15 relevant keywords that we additionally included in our search string. The final string that we used for our database search is:

(street?level bureaucra OR street?level work* OR case?work* OR case?manage* OR front?line work* OR front?office OR unemploy* OR job?seeker* OR assistance recipient* OR client*) AND (profiling OR profiling system OR algorithmic profiling OR classification of jobseeker* OR algorithmic tool* OR algorithmic decision?making OR decision support OR automated decision?making OR street?level algorithm* OR data work OR artificial intelligence OR ai OR machine?learning OR ml) AND (government service* OR employment service* OR public employment service* OR employment agenc* OR public employment agenc* OR public service* OR job placement)*

The final database search was conducted on February 9, 2024 in the following databases[8]: Web of Science, ProQuest, APA PsycNet, ACM Digital Library, and Google Scholar. The databases were chosen based on their relevance for algorithmic systems and qualitative research.

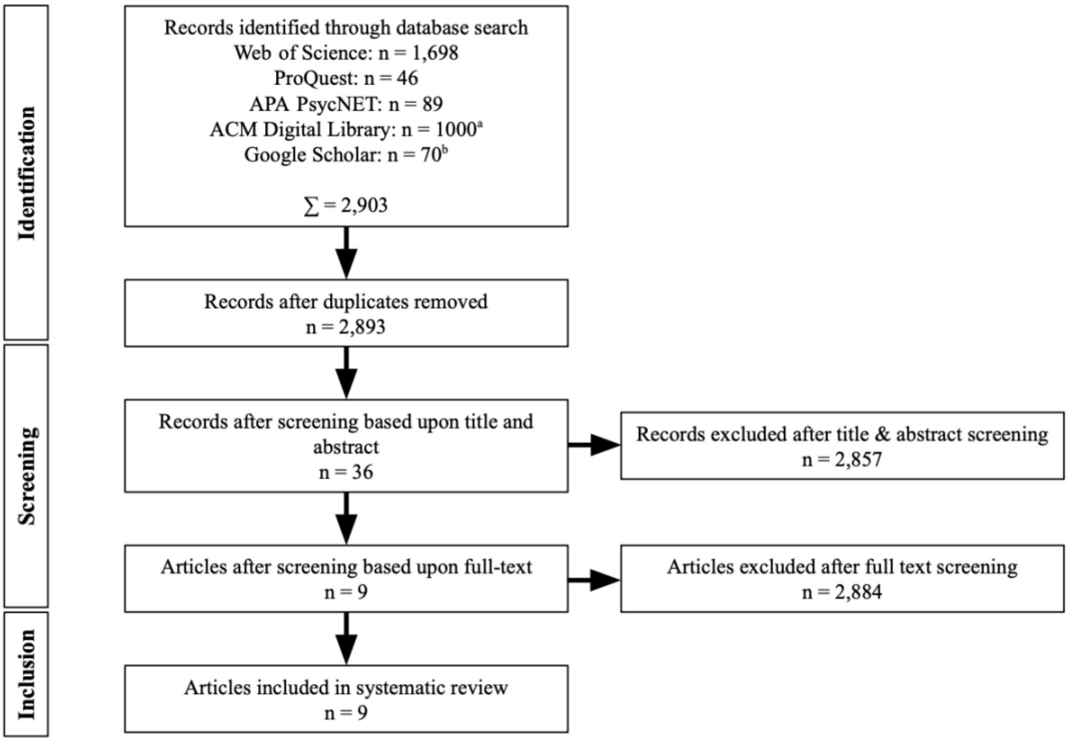
Before we conducted the database search, we established eight inclusion criteria for our literature review to which we added one more during the screening process. First, we include only publications written in English. Since most papers are written in English, this inclusion criterion allows for comparison and inter-rater reliability testing in the process of coding, which ultimately maximizes the transparency of our results. Second, we only include peer-reviewed journals or peer-reviewed conference proceedings. Third, given the context specified in the previous subsection, we only include papers about PES. Fourth, following the previously identified population, we only include papers that focus on either caseworkers or jobseekers, or both stakeholder groups simultaneously. Fifth, we analyze only papers that apply a qualitative method, either purely qualitative or in a mixed-methods approach. In contrast to other approaches, qualitative methods allow an in-depth exploration of the opinions, emotions, and reactions of research subjects. Accordingly, results from qualitative studies provide a sound basis that enables us to systematically investigate the effects of algorithmic profiling tools on the affected stakeholder groups. Sixth, we decided to only include studies which generated original data, e.g., through interviews. Finally, to set a time frame for our review, we include papers published up until the date of the final database search, however, not prior to 1997[9]. During the screening process, we found papers that met all our inclusion criteria but were not focusing on algorithmic profiling, which is why we added the criterion "is about algorithmic profiling".

[8] Three databases that we initially identified as relevant had to be excluded during the search. We could not use our search string in the ScienceDirect and IEEE Xplore databases due to its length. The search in the Scopus database using our search string did not yield any results.

[9] With the introduction of the US initiative "Worker Profiling and Reemployment Services" in 1993, the US became one of the pioneers of statistical profiling (Bloch Haug, 2022: 455). Stephen A. Wandner described the profiling system in 1997 in his article "Early re-employment for dislocated workers in the United States" (International Social Security Review).

The PRISMA flow diagram (see Figure 1) presents the process of study selection.

Figure 1: PRISMA flow diagram for systematic literature review



During the database search, we identified a total of 2,903 articles in five data bases, of which 2,893 remained after deduplication. These articles were divided among three screeners for initial screening by their titles, abstracts, and keywords. After the initial screening, we excluded 2,857 articles because they did not fulfill the inclusion criteria. This left us with a total of 36 articles for full-text screening, which again were divided between three screeners. After full-text screening, we excluded an additional 27 articles (leaving a total of 2,884 excluded articles). In the end, 9 studies were included in our systematic literature review.[10]

3.2 Thematic synthesis

The studies were analyzed by using the thematic synthesis approach by Thomas and Harden (2008), which is a method for synthesizing the research findings of multiple qualitative studies (Thomas & Harden, 2008, p. 2). In a first step, descriptive themes are generated by grouping similar codes together to identify patterns and

[10] This manuscript is part of an ongoing research project. Further studies can be added manually.

commonalities across the data. In the second step, based on the interpretation (synthesis) of descriptive themes, analytical themes are developed (Flemming & Noyes, 2021, p. 6). The advantage of a thematic synthesis compared to other approaches is its ability to develop explicit and implicit themes within the data while maintaining the context of the analyzed studies (Noyes et al., 2017, p. 6). The goal of a thematic synthesis is to gather findings from multiple studies to generate new findings.

We combined the thematic synthesis approach by Thomas and Harden (2008) with the inductive coding approach by (Mayring, 2014; Mayring, 2022). Inductive coding is a qualitative data analysis technique which allows researchers to analyze data without a predefined theoretical framework (Mayring, 2022, p. 84ff.). It was chosen because our research question is explorative in its nature (Mayring, 2022, p. 104) and in order to capture the different perspectives towards algorithmic profiling. We only coded the relevant parts of the studies (Mayring, 2014, p. 79).

4 Findings

In this section, we present the findings based on a thematic synthesis of nine[11] empirical studies. First, we outline the characteristics of the studies. Then, we present the analytical themes, based on the descriptive themes we inductively found in the data (for an overview of the descriptive and analytical themes see Table A3).

4.1 General characteristics of the reviewed articles

Table 1 provides an overview of the main characteristics of the analyzed studies. All studies were published between 2016 and 2023 in peer-reviewed journals or peer-reviewed conference proceedings. The studies were conducted in Western countries, namely Sweden, Denmark, France, Germany, Netherlands and the United Kingdom. The studies focus on the perceptions and experiences of caseworkers in municipal employment offices only or of caseworkers and (former) jobseekers. These perceptions and/or experiences were captured through case studies, combined with qualitative interviews, observations, and other ethnographic methods such as co-design workshops.

[11] This paper is part of an ongoing research project. It reflects the status as of March 5, 2024. Other relevant papers that did not appear in the search will be included in the final analysis.

Table 1: Characteristics of the studies

	Author(s)	Year	Country	Qualitative Method	Population	Context and Scope	Algorithmic Tool
A	Bernhard and Wihlborg	2022	Sweden	Case study including interviews and observations	Caseworkers	PES and social insurance	RPA tool for algorithmic decision-making for job placement
B	Boulus-Rødje	2018	Denmark	Ethnography	Caseworkers and jobseekers	Municipal jobcenter	Computational artifacts including algorithmic profiling tools
C	Delpierre et al.	2023	France	Case study including interviews and observation	Caseworkers	Municipal jobcenter	Statistical profiling tool
D	Dolata et al.	2020	Germany	Case study including interviews and observation	Caseworkers and jobseekers	PES	Case and knowledge management systems
E	Holten Møller et al.	2020	Denmark	Workshop	Caseworkers	Municipal jobcenter	Algorithmic decision-making systems for job placement
F	Kersing et al.	2022	Netherlands (Rotterdam)	Case study including interviews and observation	Caseworkers	Municipal jobcenter	Data dashboard
G	Nagy	2016	United Kingdom	Observation (ethnography)	Caseworkers and jobseekers	Jobcenter (United Kingdom)	Information and communication technology (ICT)
H	Petersen et al.	2021	Denmark	Case study including interviews	Caseworkers	Municipal jobcenter	Artificial Intelligence
I	Scott & Wang et al.	2022	Germany	Co-design workshop	Jobseekers	PES	Algorithmic decision-making systems

4.2 Implications of algorithmic profiling for caseworkers

4.2.1 Administrative burden

In the studies examined, we were able to see that algorithmic profiling was introduced with the intention of increasing efficiency, for example by reducing the administrative burden on caseworkers. Contrary to the widespread assumption that automation technologies lead to a reduction of administrative burden, we find that in the case of algorithmic profiling the administrative burden is actually increased. First, we find evidence that caseworkers make additional efforts to maintain the flow of information. Caseworkers have to make decisions based on the information available about the jobseekers. However, this information is sometimes scattered in individual documents, as the following quote from case F (Netherlands) shows: “Work coaches must manually go through files in the main registration database to find this information. They do not always have time to do this due to the high caseloads and work pressure.” (case F: 10). Algorithmic profiling systems also lead to dividing jobseekers into groups according to their labor market chances. In the case of Denmark, these groups are insufficient in their informative value, which is why caseworkers – in order to do their job and successfully profile jobseekers – add additional information the system does not capture (like “mentally ill”) (case H: 16).

In the case of Denmark, case B goes on to show that the case handling system has some limitations, such as "[...] lack of support for reading several notes simultaneously, making the task tedious and time-consuming [...]" (case B: 866). The lack of user-friendliness leads to additional efforts being made to maintain the flow of information about the jobseekers: “Denise developed a workaround whereby she has a 'master summary document' in Word for every citizen she has, containing copies of her running notes from [the case handling system].” (case B: 853).

4.2.2 Acceptance

The introduction of profiling systems does not mean that all caseworkers use the systems to the same extent. This is exemplified in the case of France: “Some [advisors] simply ignore [the tools] and only enter the data into the system afterwards; some use the tools to check their own diagnosis, and others delegate most of the diagnosis to them.” (case C: 8). Barriers to the adoption rate of algorithmic profiling tools in daily working routines include the perceived usefulness of such tools: “Among counselors, profiling was seen as useless.” (case C: 7). This attitude seems to depend on socio-demographic characteristics such as the age and employment status of the caseworkers, as in the case of France, older caseworkers who had a public employment status were more reluctant to use profiling tools than younger caseworkers: “They unanimously emphasized that they did not need statistical information to diagnose the unemployed. They did not rate the diagnostic tools highly and some even rejected them.” (case C: 10).

The improvement in the usage rate was attributed to “[...] significant changes in employment status and recruitment policies at Pôle Emploi [French PES] [which] favored the acceptance of statistical profiling in working practice.” (case C: 9). In addition to the recruitment policy, the framing was also changed in France, as a member of the Pôle Emploi management team describes “[...] the term 'profiling' was abandoned because it created so much tension [...]. [The] way things are presented to the consultants is key: they see it as a decision-making tool or as a substitute for their expertise.” (case C: 8).

4.2.3 Algorithmic profiling redefines caseworkers' discretionary power

In the papers examined, we found evidence that algorithmic profiling redefines the caseworkers' discretion. This means that algorithmic profiling can both curtail and expand the discretionary power of caseworkers.

The factors that tend to limit the scope for decision-making (barriers) include implementation processes that outsource certain tasks previously handled within the PES. For example, in Sweden, jobseekers have to verify themselves digitally via an eID to access certain services. Case A (12) notes:

> *„Nevertheless, the improved security at the log-in stage prevents the front-line case workers at AF [PES] from supporting the clients since e-IDs are provided by other organisations. Here, too, some clients' refusal to even attempt to use digital devices and a personal e-ID further constrained the discretion of the front-line case workers. [...] Even when the front-line case workers have a high degree of discretion and high service ambition, they are not able to fully support the clients to use RPA to make the work at the agency more efficient."*

(The low adoption rate of digital devices by jobseekers can also be seen in case I.) If the interaction between caseworkers and jobseekers decreases, caseworkers also have less opportunity to align their actions with the needs of the jobseekers. Case I emphasizes the importance of face-to-face contact: "[Participants] also mentioned that their request for funding was accepted upon convincing the counsellor in a face-to-face meeting." (case I: 2145). Case C highlights another barrier to caseworker discretion: „[New recruits] get 9 days of training, which focuses on mastering the IT systems and no longer includes any human or social sciences. […] '[T]hey've started to downgrade the training of employment advisers. They are turning them into production and recording agents […]." (case C: 9). Delpierre et al. (2023) conclude in the context of AI: "Even if AI 'does not replace the employment adviser' as Pôle Emploi project actors say, the adviser's role remains considerably reduced." (case C: 9). This new role is also discussed in the case of Sweden: "I have a role called 'customer resource'. […] We […] are instructed to help as many clients as possible to access the online services. […] We are a resource for the clients, we are not employment administrators. This is new for everyone […]." (case A: 384).

However, algorithmic systems do not exclusively curtail caseworker's discretionary scope. Dolata et al. (2020) note that "[T]echnology contributes to the enablement of consultants as well" (Dolata et al., 2020, p. 686). In the papers we reviewed, we found that redefining the decision-making scope can be to the advantage of the caseworker in so far as they can live out their pro-social attitude more strongly: "It can be seen as an enhanced socially- and service-oriented discretion, compensating for the limitations given of digital discretion" (Bernhard & Wihlborg, 2022, p. 387). In the case of Germany (case D: 686) it was found that

> *"[w]ith reviewing partial information provided by the system, [caseworkers] can establish a picture of the person they are talking to. This empowers the consultant to interact with the clients in a more personal way, along with their desired job objectives. At the same time, the results make clear that the technology can impact the frontline consultation only as far as the consultant allows for this influence."*.

Digital systems can also be prone to errors, for example if they can only be operated by those jobseekers who have digital skills, as with the self-service terminal in Sweden. In this case, caseworkers can use their "[…] digital discretion in new ways in order to be able to work alongside the RPA to bring all clients into the system." (case A: 387).

If decisions are merely suggested, as in the case of Denmark, this can also stimulate what one caseworker calls "'artistic freedom' – the possibility to flexibly interpret these categories and their application in practice" (case B: 867). This is also the case of Netherlands where the "[…] job counsellors […] make [] the final decision on jobseekers' situations […]" (case I: 2146).

4.2.4 Caseworkers' resources: Case knowledge management and skill formation

The flow of case knowledge is very important in order to successfully support jobseekers. Caseworkers often exchange case-specific knowledge informally. In the case of Denmark (case B: 862), it was found that increased automation has led to the previous informal flow of knowledge being impaired:

> *"Half of the caseworkers' time is spent on collecting and assembling large amounts of information from different systems and sources, and producing information that records interactions and decisions. Yet, this information is not used by front-line workers for reflections upon existing practices. Typically, knowledge about local experiences is exchanged across professional groups either in an informal manner (e.g., during breaks) or during the weekly/biweekly cross−/departmental meetings."*

This is also confirmed by Petersen et al. (2021), who studied Denmark as well:

> *"Hence, of great importance is their concern with the epistemology of their knowledge when classifying citizens. Making their descriptions representable and traceable to AI would, as reported in this study, take the classifications out of the*

> *human field of accountability and the actual situations in which the decisions they represent are undertaken." (case H: 21)*

An obstacle to the use of profiling systems is that caseworkers have sufficient skills to deal with the systems. In the Netherlands (case F: 10), the following was noted:

> *"Some work coaches do not know the exact definition of the terms used in the dashboard and therefore register incorrectly. […] But they also feel that they lack the capacities (mainly analytical insight) and digital skills to work with the dashboard yet. This is also acknowledged by some of the work coaches: 'But I would have given myself, when it comes to capacities, I would have given myself a 6 out of 10.'".*

In France (case C: 8), it was found that explanations help in the case of missing skills:

> *"Albert, a former ANPE adviser, summarized the change in this way: 'Before, we had the raw score but without any explanation, without knowing which variables explained the score. Now, we have information to help with interpretation'. And a Pôle Emploi executive stated: 'One of the difficulties is getting the adviser to use the algorithm correctly […] That's why we enrich what it says, adding the reasons why it thinks this or that. This is a real plus for the advisers, and we hope they will come on board more easily, too'."*

At the same time, however, skills are also strengthened, for example by not having all information on jobseekers in one place, as is the case of Denmark (case B: 854):

> *"'We haven't had any system that captures all information in one place, and it is a problem for the target group that we work with… So it has been an obstacle… a complication for us […] because one must be a detective in order to find all the information, and to know where to look and how to operate with these systems'.".*

4.3 Implications of algorithmic profiling for jobseekers

4.3.1 Implications of algorithmic profiling for the inclusion of jobseekers

When analyzing the papers, we identified certain barriers to the access of jobseekers to certain applications and resources. One is the lack of digital tools and digital skills among jobseekers, as in the case of Sweden (case A: 382f.): "However, there were still obvious challenges when meeting clients at the front office who do not have the ability, or the personal technology (tablet or mobile phone) needed to use the services." Secondly, it is the unwillingness to use digital tools (case A: 384): "There are some [customers] who stated: 'No, business with the computer is not for me' [...]". Language can also be an obstacle, as a caseworker notes that almost all services are exclusively available in Swedish, although English as an international language understood by most people would facilitate the use, and personal translation and guidance were described as "not always correct and clear" (case A: 385).

In addition to the technical or qualification-related requirements of jobseekers in the context of algorithmic profiling, there are further implications for jobseekers in terms of equal treatment and impartiality. In the case of Denmark (case E: 6), Holten Møller et al. (2020) conclude that "Our process with caseworkers demonstrated that the concept of value metrics, so important in the design of algorithmic systems, is not monolithic and tends to be oversimplified." They also point out that "oversimplification in classification can manifest itself in algorithmic decision support systems as discrimination against certain individuals". In the case of Sweden, Bernhard and Wihlborg (2022, p. 384) conclude:

> *"Since personal and case sensitive information is less transparent for the staff, there is also the potential to make case management more impartial with a focus on legislative duty-oriented values. The ambition to design more advanced and efficient systems is in line with the complex legislative framework and with the intentions of the welfare policies behind the specific social insurance scheme to be inclusive and impartial."*

Petersen et al. (2021) also point out the ethical limits of the use of algorithmic profiling in the case of Denmark (case H:). They emphasize that only humans are able to understand human nature:

> *"To the caseworkers, it is the kind of data that only professional workers can act on. They are to do with judgements that only people make about each other: about character, intention, reliability, good faith and the rest. If we believe the caseworkers, judgement of character cannot - and should not - be summarised in a bullet list, for example. To our knowledge, these insights have not previously been reported in the literature. [...] In the context of this research, risk predictions of long-term unemployment were defined by the municipality as a problem that AI could solve. However, the caseworkers are sceptical of the idea that anyone or anything ought to predict people's futures." (Case H: 21)*

They conclude that "[a]s we have shown empirically in this paper, it is not only a matter of technology that plays a role in the implementation of AI for decision support, but also caseworkers' moral judgments about what data is considered problematic to collect." (case H: 22).

4.3.2 Implication of algorithmic profiling on the transparency of the process and the results

Contrary to assumptions, we find evidence that the use of automation techniques can also lead to increased transparency for jobseekers through direct contact with the tools (Sweden, case C: 9):

> *"Some experiments now involve jobseekers in collecting information through the Pôle Emploi web portal, via a multitude of applications. This makes it possible to establish direct machine-human contact; analysis of the data entered leads to a wide range of suggestions regarding jobseeking approaches to be favored, relevant training courses, recommended employment channels, comments on CVs, etc."*

Scott et al. (2022) (case I: 2145) come to a similar conclusion when they say that "[t]he implications of algorithmic system outputs must be understood and communicated by the creators and full documentation of data used and design decisions should be available."

5 Discussion and conclusion

The goal of this systematic literature review was to examine how algorithmic profiling impacts caseworkers and jobseekers in Public Employment Services. We did that by systematically analyzing and synthesizing empirical studies that empirically show the actual implications of algorithmic profiling tools. Our literature review summarizes the scattered literature on these implications. In doing so, it provides an empirically grounded contribution to the discussion on the opportunities and risks of technological innovations in PES.

Based on our literature review, we draw the following conclusions: Although algorithmic profiling continues to be developed and deployed, there are relatively few empirical studies that focus on the experiences of caseworkers and jobseekers (at least not in the literature published in English). We identified nine empirical studies that analyze the actual implications of algorithmic profiling. The summary of results is presented in this section. This section also discusses the implications of these findings for efficiency, accountability, transparency, and contestability as core public values in the context of e-government.

From the perspective of caseworkers, algorithmic profiling tends to increase the administrative burden. This is because such systems can hinder the flow of case-specific information, requiring caseworkers to make additional efforts to maintain it. How intensively caseworkers use these systems varies greatly, both within and between PES. The acceptance of these systems is determined by perceived benefits, caseworker age, and how the systems are framed by government and PES management. Despite the assumption that algorithmic profiling leads to greater efficiency in resource allocation, we found that there are other variables that could have a more lasting effect on efficiency. Bannister and Connolly (2014, p. 13) state that efficiency is one of the core public sector values in the context of e-government. Algorithmic profiling can contribute to the efficiency of administration, “as it offers a much less expensive mode of access than face to face or telephone” (Bannister et al., 2020, p. 124). However, its efficiency is questionable as it tends to increase the administrative burden and is not universally accepted among PES caseworkers.

From the caseworker's perspective, we also found evidence of a redefinition of discretion (this finding is in line with Marienfeldt (2024)). The reduction in interpersonal interaction between caseworker and jobseeker may initially lead to a reduction in the discretionary power. The studies reviewed also show that the introduction of profiling systems is accompanied by more statistical recording, documentation and evaluation tasks, which can change the role and nature of caseworkers’ job. On the other hand, the limitations of profiling systems can also lead caseworkers to reinterpret their discretionary power. Our analysis shows that they use their discretion (even) more for pro-social behavior, so that, for example, disadvantaged jobseekers can be helped (even more). Algorithmic profiling must also be discussed in the context of case knowledge management and skill development. The previously discussed increased administrative burden associated with profiling tools (see previous paragraph) can mean that caseworkers have less time to share their experiences. This affects the essential internal flow of case-specific information between caseworkers. In addition to experience, skills are an important resource for caseworkers. In some cases, they lack the skills, e.g., digital skills, to integrate profiling systems into their work. Explanations of the variables, for example, can help here. However, dealing with (user-unfriendly) profiling systems also strengthen skills such as the ability to collect and combine data from different sources (“detective skills”). Redefining caseworkers’ discretionary power, knowledge management, and skills has implications for the accountability of administrative action.

Accountability is another core public value in the context of e-government (Bannister & Connolly, 2014, p. 13). In the case of algorithmic profiling the concept ‘algorithmic accountability’ refers to “accepting the responsibility for actions and decisions” (Lepri et al., 2017, p. 605). The implementation of algorithmic profiling implies that administrative decision-making is delegated to algorithms. That can involve a trade-off between the benefits of automatization and a potential loss in human agency and oversight.

Like caseworkers, jobseekers are not a homogeneous group, but differ in their (digital) skills. If jobseekers do not have sufficient skills, they may not be able to use certain applications in the PES (terminals, online applications, etc.). A lack of knowledge of the language used in the context of the PES and can be another barrier. technical equipment (smartphone, computer) can also be a barrier to access. In the case of (imminent) unemployment, jobseekers may not always have the financial resources to buy the necessary technical equipment.

In addition to the technical or skill requirements for jobseekers in the context of PES, there are further implications for jobseekers in terms of equal treatment and impartiality. Oversimplification of classifications can lead to discrimination against jobseekers. Through barriers to access resources, algorithmic profiling has implications for the inclusion and equal treatment of jobseekers and thus for the value of fairness, which is also part of the core set of values in the context of e-government (Bannister & Connolly, 2014, p. 13). At the same time, caseworkers in this case (as described above) try to use their discretion to help jobseekers, e.g., by translating and coaching. Letting machines evaluate aspects of human nature (such as jobseekers' intention and reliability) can also be perceived as unfair in the eyes of jobseekers. Here, caseworkers are certain that these predictions cannot and should not be performed by a machine.

The introduction of algorithmic profiling and related systems has implications for the transparency of processes and decisions in PES. On the one hand, co-production processes, in which jobseekers are in direct contact with digital tools, can lead to greater transparency regarding the personal data processed. On the other hand, jobseekers also need to be able to understand the decision-making logic and outcomes, which would require more transparency. Transparency can be described as another core public value in the context of e-government (Bannister & Connolly, 2014, p. 13). It is "[…] cited as a value that can be delivered in radically new ways by ICT" (Bannister & Connolly, 2014, p. 14). As we have seen in our analysis, "the impact of ICT on transparency remains ambivalent" (Bannister & Connolly, 2014, p. 121)

In an extreme case of algorithmic decision-making, for example, bureaucratic agents can no longer be held responsible for administrative decisions, a phenomenon referred to as the 'responsibility gap' (Wirtz et al., 2018) or 'accountability deficit' (Bannister et al., 2020). Thus, developing measures to maintain algorithmic accountability is essential to the design of profiling systems in public administration. After all, policy makers should be aware that technology is not value free (Miller, 2021), "but rather, its implementation is driven by perceived values" (Bannister & Connolly, 2014).

Our study has a number of limitations. First, this work gives an overview of English language publications on qualitative work conducted in PES worldwide. We did not cover non-English language publications for transparency and reliability reasons, although in such a field of research some publications might only be published in the respective countries' language. Second, we did not include publications using quantitative methods, future systematic reviews could focus on such publications as well. Third, the selection of our studies for analysis is not entirely neutral, as the starting point for empirical case studies in the field of profiling in PES is often a scandal (as in the case of the Austrian PES). There may also be successful profiling systems that include the perspectives of our stakeholder groups. However, we have not (yet) found these studies. Four, certain implications may be better investigated using quantitative methods, such as the implications for efficiency. Five, we have also discussed how profiling systems impact caseworkers and jobseekers. However, there are other stakeholder views in the context of profiling systems that we have not covered in this systematic review, which are policy makers and developers of systems. These stakeholder groups are not directly involved after deployment, which means that there are no immediate effects on them. Nonetheless, their perspective plays an important role in this issue.

References

Alfrink, K., Keller, I., Kortuem, G., & Doorn, N. (2022). Contestable AI by Design: Towards a Framework. *Minds and Machines*. https://doi.org/10.1007/s11023-022-09611-z

AlgorithmWatch. (2019). Automating Society 2019. In.

Allhutter, D., Cech, F., Fischer, F., Grill, G., & Mager, A. (2020). Algorithmic Profiling of Job Seekers in Austria: How Austerity Politics Are Made Effective. *Frontiers in Big Data*, *3*.

Almada, M. (2019, Juni 17, 2019). Human intervention in automated decision-making: Toward the construction of contestable systems.*ICAIL '19*

Ammitzbøll Flügge, A., Hildebrandt, T., & Møller, N. H. (2021). Street-Level Algorithms and AI in Bureaucratic Decision-Making: A Caseworker Perspective. *Proceedings of the ACM on Human-Computer Interaction*, *5*(CSCW1), 40:41–40:23. https://doi.org/10.1145/3449114

Bannister, F., & Connolly, R. (2014). ICT, public values and transformative government: A framework and programme for research. *Government Information Quarterly*, *31*(1), 119-128. https://doi.org/10.1016/j.giq.2013.06.002

Bannister, F., Connolly, R., Giest, S., & Grimmelikhuijsen, S. (2020). Administration by algorithm: A risk management framework. *Information Polity*, *25*(4), 471-490. https://doi.org/10.3233/ip-200249

Berman, A. (2023, 2023). Why Does the AI Say That I Am Too Far Away From the Job Market?

Bernhard, I., & Wihlborg, E. (2022). Bringing all clients into the system – Professional digital discretion to enhance inclusion when services are automated. *Information Polity*, *27*(3), 373-389. https://doi.org/10.3233/ip-200268

Bloch Haug, K. (2022). Structuring the scattered literature on algorithmic profiling in the case of unemployment through a systematic literature review. *International Journal of Sociology and Social Policy*, *43*(5/6), 454-472. https://doi.org/10.1108/ijssp-03-2022-0085

Booth, A., Sutton, A., & Papaioannou, D. (2016). *Systematic Approaches to a Successful Literature Review*. SAGE.

Braun, Virginia, and Victoria Clarke. 2021. *Thematic Analysis: A Practical Guide*. London: Sage.

Bullock, J. B. (2019). Artificial Intelligence, Discretion, and Bureaucracy. *American Review of Public Administration*, *49*(7).

Busuioc, M. (2021). Accountable Artificial Intelligence: Holding Algorithms to Account. *Public Adm Rev*, *81*(5), 825-836. https://doi.org/10.1111/puar.13293

Delpierre, A., Demazière, D., & El Fatihi, H. (2023). The stealth legitimization of a controversial policy tool: Statistical profiling in French Public Employment Service. *Regulation & Governance*. https://doi.org/10.1111/rego.12541

Desiere, S., Langenbucher, K., & Struyven, L. (2019). *Statistical profiling in public employment services: An international comparison*.

Desiere, S., & Struyven, L. (2021). Using Artificial Intelligence to classify Jobseekers: The Accuracy-Equity Trade-off. *Journal of Social Policy*, *50*(2), 367-385. https://doi.org/10.1017/S0047279420000203

Dolata, M., Schenk, B., Fuhrer, J., Marti, A., & Schwabe, G. (2020). When the System Does Not Fit: Coping Strategies of Employment Consultants.

Computer Supported Cooperative Work (CSCW), *29*(6), 657-696. https://doi.org/10.1007/s10606-020-09377-x

Flemming, K., & Noyes, J. (2021). Qualitative Evidence Synthesis: Where Are We at? *International Journal of Qualitative Methods*, *20*. https://doi.org/10.1177/1609406921993276

Grames, E. M., Stillman, A. N., Tingley, M. W., & Elphick, C. S. (2019). An automated approach to identifying search terms for systematic reviews using keyword co-occurrence networks. *Methods in Ecology and Evolution*, *10*(10), 1645-1654. https://doi.org/10.1111/2041-210X.13268

Holten Møller, N., Shklovski, I., & Hildebrandt, T. T. (2020). *Shifting Concepts of Value* Proceedings of the 11th Nordic Conference on Human-Computer Interaction: Shaping Experiences, Shaping Society,

Klutzz, D. N., Kohli, N., & Mulligan, D. K. (2022). Shaping Out Tools: Contestability as a Means to Promote Responsible Algorithmic Decision Making in the Professions. In K. Martin (Ed.), *Ethics of Data and Analytics*. CRC Press.

König, P. D., & Wenzelburger, G. (2021a). The legitimacy gap of algorithmic decision-making in the public sector: Why it arises and how to address it. *Technology in Society*, *67*. https://doi.org/10.1016/j.techsoc.2021.101688

König, P. D., & Wenzelburger, G. (2021b). The legitimacy gap of algorithmic decision-making in the public sector: Why it arises and how to address it. *Technology in Society*, *67*, 101688. https://doi.org/https://doi.org/10.1016/j.techsoc.2021.101688

Körtner, J., & Bonoli, G. (2022). Predictive algorithms in the delivery of public employment services. In D. Clegg & N. Durazzi (Eds.), *Research handbook of labour market policy in rich democracies*. Edward Elgar Publishing.

Krafft, T. D., Zweig, K. A., & König, P. D. (2020). How to regulate algorithmic decision-making: A framework of regulatory requirements for different applications. *Regulation & Governance*, *16*(1), 119-136. https://doi.org/10.1111/rego.12369

Lepri, B., Oliver, N., Letouzé, E., Pentland, A., & Vinck, P. (2017). Fair, Transparent, and Accountable Algorithmic Decision-making Processes. *Philosophy & Technology*, *31*(4), 611-627. https://doi.org/10.1007/s13347-017-0279-x

Marienfeldt, J. (2024). Does digital government hollow out the essence of street-level bureaucracy? A systematic literature review of how digital tools' foster curtailment, enablement and continuation of street-level decision-making. *Social Policy & Administration*. https://doi.org/10.1111/spol.12991

Mayring, P. (2014). *Qualitative content analysis: theoretical foundation, basic procedures and software solution*. Klagenfurt. Qualitative content analysis: theoretical foundation, basic procedures and software solution

Mayring, P. (2022). *Qualitative Inhaltsanalyse*. Beltz Verlagsgruppe. http://www.content-select.com/index.php?id=bib_view&ean=9783407258991
https://content-select.com/de/portal/media/view/6230dc9f-d158-4bda-a1f2-1d0db0dd2d03

Miller, B. (2021). Is Technology Value-Neutral? *Science, Technology, & Human Values*, *46*(1), 53-80. https://doi.org/10.1177/0162243919900965

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & Group, P. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ*, *339*, b2535. https://doi.org/10.1136/bmj.b2535

Niklas, J., Sztandar-Sztanderska, K., & Szymielewicz, K. (2015). *Profiling the Unemployed in Poland: Social and Political Implications of Algorithmic Decision Making*.

Noyes, J., Booth, A., Flemming, K., Garside, R., Harden, A., Lewin, S., Pantoja, T., Hannes, K., Cargo, M., & Thomas, J. (2017). Cochrane Qualitative and Implementation Methods Group guidance series-paper 3: methods for assessing methodological limitations, data extraction and synthesis, and confidence in synthesized qualitative findings. *J Clin Epidemiol*, *97*, 49-58. https://doi.org/10.1016/j.jclinepi.2017.06.020

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., . . . Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *International Journal of Surgery*, *88*, 105906. https://doi.org/10.1016/j.ijsu.2021.105906

Petersen, A. C. M., Christensen, L. R., Harper, R., & Hildebrandt, T. (2021). "We Would Never Write That Down": Classifications of Unemployed and Data Challenges for AI. *Proceedings of the ACM on Human-Computer Interaction*, *5*(CSCW1), 102:101–102:126. https://doi.org/10.1145/3449176

Petticrew, & Roberts. (2006). *Guide-of-systematic-reviews-in-social-sciences*.

Scott, K. M., Wang, S. M., Miceli, M., Delobelle, P., Sztandar-Sztanderska, K., & Berendt, B. (2022). *Algorithmic Tools in Public Employment Services: Towards a Jobseeker-Centric Perspective* 2022 ACM Conference on Fairness, Accountability, and Transparency,

Starke, C., Baleis, J., Keller, B., & Marcinkowski, F. (2022). Fairness perceptions of algorithmic decision-making: A systematic review of the empirical literature. *Big Data & Society*, *9*(2). https://doi.org/10.1177/20539517221115189

Sweet, M., & Moynihan, R. (2007). *Improving population health: the uses of systematic reviews*.

Sztandar-Sztanderska, K., & Zielenska, M. (2018). Changing social citizenship through information technology. *Social Work & Society*, *16*(2).

Sztandar-Sztanderska, K., & Zielenska, M. (2020). What Makes an Ideal Unemployed Person? Values and Norms Encapsulated in a Computerized Profiling Tool. *Renegotiating Social Citizenship –Democracy in Welfare Service States /*, *18*(1).

Thomas, J., & Harden, A. (2008). Methods for the thematic synthesis of qualitative research in systematic reviews. *BMC Med Res Methodol*, *8*, 45. https://doi.org/10.1186/1471-2288-8-45

Wijnhoven, M. A., & Havinga, H. (2014). The Work Profiler: A digital instrument for selection and diagnosis of the unemployed. *Local Economy*, *29*(6-7), 740-749. https://doi.org/10.1177/0269094214545045

Wirtz, B. W., Weyerer, J. C., & Geyer, C. (2018). Artificial Intelligence and the Public Sector—Applications and Challenges. *International Journal of Public Administration*, *42*(7), 596-615. https://doi.org/10.1080/01900692.2018.1498103

Zuiderwijk, A., Chen, Y.-C., & Salem, F. (2021). Implications of the use of artificial intelligence in public governance: A systematic literature review and a research agenda. *Government Information Quarterly*, *38*(3). https://doi.org/10.1016/j.giq.2021.101577

Annex

Table A1: Keywords used for the naive search

Cluster 1: population	Cluster 2: intervention	Cluster 3: context
street-level bureaucra* caseworker* unemploy* jobseeker*	profiling algorithmic profiling classification of jobseekers algorithmic tool* algorithmic decision-making automated decision-making street-level algorithm* data work artificial intelligence AI human oversight	public employment services public services

Note: Keywords sorted into three clusters which are combined with the AND operator between clusters and the OR operator within clusters. The article from which these keywords originated from are cited. Not all of these articles fit our inclusion criteria, so some of the articles were only used to identify a relevant keyword.

Table A2: PRISMA checklist

Topic	**#**	**Checklist Item**	**Reported in section #**
TITLE			
Title	1	Identify the report as a systematic review.	Title
ABSTRACT			
Abstract	2	See the PRISMA 2020 for Abstracts checklist	Abstract
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of existing knowledge.	1
Objectives	4	Provide an explicit statement of the objective(s) or question(s) the review addresses.	1
METHODS			
Eligibility criteria	5	Specify the inclusion and exclusion criteria for the review and how studies were grouped for the syntheses.	3
Information sources	6	Specify all databases, registers, websites, organisations, reference lists and other sources searched or consulted to identify studies. Specify the date when each source was last searched or consulted.	3
Search strategy	7	Present the full search strategies for all databases, registers and websites, including any filters and limits used.	3
Selection process	8	Specify the methods used to decide whether a study met the inclusion criteria of the review, including how many reviewers screened each record and each report retrieved, whether they worked independently, and if applicable, details of automation tools used in the process.	3
Data collection process	9	Specify the methods used to collect data from reports, including how many reviewers collected data from each report, whether they worked independently, any processes for obtaining or confirming data from study investigators, and if applicable, details of automation tools used in the process.	3

Topic	#	Checklist Item	Reported in section #
Data items	10a	List and define all outcomes for which data were sought. Specify whether all results that were compatible with each outcome domain in each study were sought (e.g. for all measures, time points, analyses), and if not, the methods used to decide which results to collect.	3
	10b	List and define all other variables for which data were sought (e.g. participant and intervention characteristics, funding sources). Describe any assumptions made about any missing or unclear information.	N.A.
Study risk of bias assessment	11	Specify the methods used to assess risk of bias in the included studies, including details of the tool(s) used, how many reviewers assessed each study and whether they worked independently, and if applicable, details of automation tools used in the process.	N.A.
Effect measures	12	Specify for each outcome the effect measure(s) (e.g. risk ratio, mean difference) used in the synthesis or presentation of results.	N.A.
Synthesis methods	13a	Describe the processes used to decide which studies were eligible for each synthesis (e.g. tabulating the study intervention characteristics and comparing against the planned groups for each synthesis (item 5)).	3
	13b	Describe any methods required to prepare the data for presentation or synthesis, such as handling of missing summary statistics, or data conversions.	3
	13c	Describe any methods used to tabulate or visually display results of individual studies and syntheses.	3
	13d	Describe any methods used to synthesize results and provide a rationale for the choice(s). If meta-analysis was performed, describe the model(s), method(s) to identify the presence and extent of statistical heterogeneity, and software package(s) used.	3
	13e	Describe any methods used to explore possible causes of heterogeneity among study results (e.g. subgroup analysis, meta-regression).	3
	13f	Describe any sensitivity analyses conducted to assess robustness of the synthesized results.	N.A.
Reporting bias assessment	14	Describe any methods used to assess risk of bias due to missing results in a synthesis (arising from reporting biases).	3
Certainty assessment	15	Describe any methods used to assess certainty (or confidence) in the body of evidence for an outcome.	3
RESULTS			
Study selection	16a	Describe the results of the search and selection process, from the number of records identified in the search to the number of studies included in the review, ideally using a flow diagram.	3, Figure 1
	16b	Cite studies that might appear to meet the inclusion criteria, but which were excluded, and explain why they were excluded.	N.A.
Study characteristics	17	Cite each included study and present its characteristics.	4, Table 1
Risk of bias in studies	18	Present assessments of risk of bias for each included study.	N.A.
Results of individual studies	19	For all outcomes, present, for each study: (a) summary statistics for each group (where appropriate) and (b) an effect estimate and its precision (e.g. confidence/credible interval), ideally using structured tables or plots.	N.A.
Results of syntheses	20a	For each synthesis, briefly summarise the characteristics and risk of bias among contributing studies.	5

Topic	#	Checklist Item	Reported in section #
	20b	Present results of all statistical syntheses conducted. If meta-analysis was done, present for each the summary estimate and its precision (e.g. confidence/credible interval) and measures of statistical heterogeneity. If comparing groups, describe the direction of the effect.	N.A.
	20c	Present results of all investigations of possible causes of heterogeneity among study results.	N.A.
	20d	Present results of all sensitivity analyses conducted to assess the robustness of the synthesized results.	N.A.
Reporting biases	21	Present assessments of risk of bias due to missing results (arising from reporting biases) for each synthesis assessed.	N.A.
Certainty of evidence	22	Present assessments of certainty (or confidence) in the body of evidence for each outcome assessed.	N.A.
DISCUSSION			
Discussion	23a	Provide a general interpretation of the results in the context of other evidence.	5
	23b	Discuss any limitations of the evidence included in the review.	5
	23c	Discuss any limitations of the review processes used.	5
	23d	Discuss implications of the results for practice, policy, and future research.	5
OTHER INFORMATION			
Registration and protocol	24a	Provide registration information for the review, including register name and registration number, or state that the review was not registered.	Registration on February 9 2024 on Open Science Framework (OSF) Link: osf.io/2xtdj
	24b	Indicate where the review protocol can be accessed, or state that a protocol was not prepared.	N.A.
	24c	Describe and explain any amendments to information provided at registration or in the protocol.	N.A.
Support	25	Describe sources of financial or non-financial support for the review, and the role of the funders or sponsors in the review.	N.A.
Competing interests	26	Declare any competing interests of review authors.	One of the authors wrote one of the papers included in the analysis. This author will not code her own paper.
Availability of data, code and other materials	27	Report which of the following are publicly available and where they can be found: template data collection forms; data extracted from included studies; data used for all analyses; analytic code; any other materials used in the review.	N.A.

Table A3: Themes based on inductive coding

Theme (Code)	Anchor Example	Descriptive Themes	Analytical Themes
"(limited) information"	"Nevertheless, some work coaches say that the information in the dashboard is still too limited to make a good selection for most vacancies because they miss information about work experience, the kind of education, or preferences for certain jobs or branches, skills, capacities, and so forth. Some of this information is available in the online questionnaire but must be registered manually in the main registration database. This information in the main registration database is not transferred to the dashboard automatically. Work coaches must manually go through files in the main registration database to find this information. They do not always have time to do this due to the high caseloads and work pressure." (Case F: 10) "Although recognizing the limits and pitfalls of match groups, this explains why caseworkers still use them in their internal communication and compensate for their limited information by adding layers of informal nuances which they feel are necessary to 'do their job'. As we have seen, this knowledge takes the form of classifications, such as 'heavy 2.2', combined with other membership categories, such as 'mentally ill'." (Case H: 16)	Additional efforts to maintain the flow of information	Administrative Burden
"workaround"	"Denise developed a workaround, whereby she has a 'master summary document' in Word for every citizen she has, containing copies of her running notes from CHS2. This document provides her with quick and general overview of all her interactions with the citizen, compensating for the lack of time she has and the limited usability friendliness of CHS2." (Case B: 853) "More importantly, I have pointed out how there are more than 20 computational artifacts used by the different professional groups in this jobcentre. I have examined CHS2, the main case-handling system, and identified various limitations. This includes, lack of support for reading several notes simultaneously, making the task tedious and time-consuming, and leading some caseworkers to develop workarounds (i.e., the development of a 'master document')." (Case B: 865f.)		
"acceptance"	"Yet there is no generalized acceptance: advisers use the tools in divergent ways. Some simply ignore them, inputting data to the system afterwards; some use the tools to check their own diagnosis, and others delegate the bulk of diagnosis to them. […] A second condition is staff turnover at Pôle Emploi. In our fieldwork, we detected a link (to be tested via an extensive survey) between the use of profiling tools and adviser seniority. Those who began working pre-merger, have more than 20 years' experience and public employment status unanimously stressed that they had no need for statistical information in order to diagnose the unemployed. They did not rate the diagnostic tools highly, and some even dismissed them. […] This suggests that the significant changes in employment status and recruitment policy12 at Pôle Emploi in the decade following its creation13 have favored the acceptability of statistical profiling within working practice." (Case C: 8f.)	Dissemination of algorithmic profiling	Acceptance
"term"	"As a result, the term 'profiling' was abandoned because it created so much tension: it created more problems than solutions. It's quite easy to understand that the way things are presented to advisers is central: they see it as an aid to decision-making or a replacement for their expertise. (Member of the Pôle Emploi management team). Nowadays, the emphasis is on "personalized diagnosis" (or even "personalization") of jobseekers. This vocabulary seems to be widely shared: 'Today, we talk about 'diagnoses' of the jobseeker, which means making sure the jobseeker is directed toward the right service, the one that's a good match for his or her situation.' (Adviser and trade union representative)." (Case C: 7f.)	Reframing	
"useless"	"We have had some developments. Not among politicians and cabinet members, it's in the services, on the technical aspects of profiling. We've had quite a few debates, with foreign experiences too. And doubts have been raised, more and more. […] I think we tried to bring this up at sub-directorate level, saying: be careful, statistical modeling poses a lot of problems. (Former member of the Directorate of Studies and Statistics at the Ministry of Labor). Among advisers, profiling was considered useless: "…internally, we gradually saw that advisers were no longer relying on it much at all, less and less, compared to the beginning. In a few years it was over, done with" (Senior Manager, working at PES at the time)." (Case C: 7)	Perceived usefulness	
"resentment"	"In short, the "advisers" category has thus become more heterogeneous: the proportion of "old" advisers most reluctant to use profiling is decreasing, while newcomers do not share their resentment. Today, those most prone to resistance against profiling are in the minority, and this in turn favors the development of profiling in frontline work." (Case C: 10)	Age	
"interpretability"	"Interpretability: The implications of algorithmic system outputs must be understood and communicated by the creators and full documentation of data used and design decisions should be available." (Case I: 8)		

Table A3: Themes based on inductive coding (continued)

Theme (Code)	Anchor Example	Descriptive Themes	Analytical Themes
"discretion"	"These Swedish public agency case studies show that the discretion of public front-line case workers is reframed when RPA is implemented. The case workers changed their strategies towards clients who cannot use the systems themselves." (Case A: 14) "The introduction of RPA has been shown to result in a shift towards a digital mode of discretion, which can have a positive effect on civil servants' discretionary practices." (Case A: 11)	Redefinition of decision-making conditions	Discretionary Power
"expertise"	"The consultants actively overcome the curtailing elements by employing external tools, adjusting the amount of computer use during the consultation or by employing the client's own resources when encouraging them to search for information. Those combinations require creativity and experience beyond what the systems offer." (Case D: 30)	Expertise & Creativity	
"intervention"	"During the fact-finding workshop, the reliance on the intervention of job counsellors to make the final decision on jobseekers' situations was considered as a positive aspect of the Dutch Work Profiler, as was the conceptualization of its scores as a conversation starter between the counsellor and the jobseeker." (Case I: 2146)	Right of final decision	
„downgrade"	"[New recruits] get 9 days of training, which focuses on mastering the IT systems and no longer includes any human or social sciences. For a unionized adviser, these changes downgrade both the work itself, and the profession: 'Adapting mathematical and statistical tools to humanity is problematic […] And that's why, in my opinion, they've started to downgrade the training of employment advisers. They are turning them into production and recording agents, which makes it possible to develop computer and technological developments and to direct diagnosis toward simplification and quantification'." (Adviser at ANPE, then Pôle Emploi)." (Case C: 9)	Redefinition of caseworker's role	
"rationalization"	"Finally, we theorize a specific form of rationalization, called professional rationalization, in which the shaping of advisers de facto neutralizes the autonomy left to them in their work." (Case C: 10)		
"role"	"I have a role called 'customer resource'. I meet the clients at the Customer Square which was set up in 2019 and I guide them to solve their questions using e-services if possible. We (my colleagues and I) are instructed to help as many clients as possible to access the online services. This is because AF's management wants to emphasize a new way for us to work. We are a resource for the clients, we are not employment administrators. This is new for everyone; we have to learn . . . (Interview at AF)" (Case A: 384) "For jobseekers assessed by these applications as "autonomous enough," contact with advisers may be remote-only, or delayed. Even if AI "does not replace the employment adviser" as Intelligence Emploi project actors say, the adviser's role remains considerably reduced." (Case C: 9)		
"Personal interaction"	"With reviewing partial information provided by the system, they can establish a picture of the person they are talking to. This empowers the consultant to interact with the clients in a more personal way, along with their desired job objectives. At the same time, the results make clear that the technology can impact the frontline consultation only as far as the consultant allows for this influence." (Case D: 686)	Personal interaction	
"face-to-face"	"This participant is talking about specific coaching that is available for refugees and migrants in Germany. They also mentioned that their request for funding was accepted upon convincing the counsellor in a face-to-face meeting." (Case I: 2145)		
"interpretability"	"Interpretability: The implications of algorithmic system outputs must be understood and communicated by the creators and full documentation of data used and design decisions should be available." (Case I: 8)		

Table A3: Themes based on inductive coding (continued)

Theme (Code)	Anchor Example	Descriptive Themes	Analytical Themes
"(case) knowledge"	"Half of the caseworkers' time is spent on collecting and assembling large amounts of information from different systems and sources, and producing information that records interactions and decisions. Yet, this information is not used by front-line workers for reflections upon existing practices. Typically, knowledge about local experiences is exchanged across professional groups either in an informal manner (e.g., during breaks) or during the weekly/biweekly cross–/departmental meetings." (Case B: 22) "Hence, of great importance is their concern with the epistemology of their knowledge when classifying citizens. Making their descriptions representable and traceable to AI would, as reported in this study, take the classifications out of the human field of accountability and the actual situations in which the decisions they represent are undertaken." (Case H: 21) "Furthermore, their unwillingness to formalise classifications is also reinforced in the light of AI. As long as AI does not have the information needed, it cannot purposefully make predictions about citizens, which suits them well. As we have seen, caseworkers believe that data is even more recalcitrant to 'proper understanding' when viewed 'from afar', without reference to the real-world character of actual decision-making. During our observations and interviews, the caseworkers firmly expressed their belief that the imminent introduction of AI techniques is representative of precisely this move. They also worry that it might remove the boundaries in sharing and changing information as it will be accelerated and perpetuated by machines." (Case H: 19)	Knowledge loss	Caseworkers' resources: Case knowledge management and skill formation
"enabling"	"The dashboard has an enabling role because it is a tool that helps work coaches with a caseload manager identity to deal with their new tasks. The dashboard can analyze data and give an overview of the caseload. This was not possible in the main registration base were finding specific information must be done manually." (Case F: 11)	Generation of new knowledge	
"interpretation"	"Albert, a former ANPE adviser, summarized the change in this way: 'Before, we had the raw score but without any explanation, without knowing which variables explained the score. Now, we have information to help with interpretation'. And a Pôle Emploi executive stated: 'One of the difficulties is getting the adviser to use the algorithm correctly […] That's why we enrich what it says, adding the reasons why it thinks this or that. This is a real plus for the advisers, and we hope they will come on board more easily, too'." (Case C: 8)	Provision of explanation	
"digital skills"	"Some work coaches do not know the exact definition of the terms used in the dashboard and therefore register incorrectly. […] But they also feel that they lack the capacities (mainly analytical insight) and digital skills to work with the dashboard yet. This is also acknowledged by some of the work coaches: 'But I would have given myself, when it comes to capacities, I would have given myself a 6 out of 10.' " (Case F: 10)	Skills & Competences	

Table A3: Themes based on inductive coding (continued)

Theme (Code)	Anchor Example	Descriptive Themes	Analytical Themes
"language barriers"	However, some of the front-line case workers also have competence in languages other than Swedish and thus have other resources for coaching clients. The informants discussed this in terms such as: 'The language … almost all of our e-services are in Swedish, and this is problematic. One might wish that there was more in English at least because it is an international language that most people understand. But the information you find when you are logged in to 'My Pages' is always only in Swedish. Our personal translation and guidance are not always correct and clear.' (Interview at AF) This shows how the front-line case workers coach clients to use RPA and related digital systems. They need competence not only as front-line case workers in their specific area, but also to support the clients' digital competences and to coordinate and coach clients in more general terms about welfare policies and related issues within Swedish society. " (Case A: 385)	Coaching	Inclusion
"(access) thresholds"	"Additionally, and in particular at AF, workers showed clear socially-oriented values to work around the digital systems to encourage and include those clients who did not use the digital services. It was obvious that the more automated the systems became, the higher the thresholds were for the most vulnerable clients and the more the front-line workers developed socially-oriented strategies." (Case A: 11) "When clients encounter new thresholds to accessing public welfare services due to RPA, the front-line case workers develop new competences using their discretion to help the clients navigate and use RPA. […] The standardisation of RPA forces the front-line case workers to increase their personal service flexibility, referring to service-oriented values. […] Instead, the front-line case workers develop their competences in line with socially- and service-oriented values to support clients." (Case A: 14f.)	Overcoming access barriers	
"challenges"	"However, there were still obvious challenges when meeting clients at the front office who do not have the ability, or the personalI. Bernhard and E. Wihlborg / Bringing all clients into the system 383 technology (tablet or mobile phone) needed to use the services." (Case A: 382f.)		
"(lack of) competences"	"The welfare schemes managed by both agencies target clients in vulnerable situations, who commonly have fewer resources and lower levels of competence. Therefore they require even more personal guidance and support. One front-line case worker said: There is a need for a lot of human resources when helping clients to use this stuff. Now we are better at doing this. It looks a little different in different offices even though we use the same system, the same rules, the same everything, but there are differences in terms of how you organise things and which staff you put in which places" (Case A: 10) "'There are some [clients] who say, "No, that business with computers isn't anything for me", or "I really have tried all the possibilities for getting an e-ID but it's not possible because I can't", or "I'm not able to become a customer of a bank [providers of e-ID]'. (Interview at AF)" (Case A: 12)		
"red flag"	"Caseworker1: 'Not necessarily.. In job placement caseworkers are not necessarily agreeing.. and a red flag could also be interpreted into.. something that will not necessarily help this particular individual.. but just be interpreted as.. let me get this particular individual out the door.. [..] Sometimes other caseworkers were simply too busy.. to take the required action.' […] She also raised caution about the implications of a "red flag," which could prompt a caseworker to avoid a concrete individual due to the perceived extra work involved." Case E: 7)	Stigmatization & impartiality	
"discrimination"	"Our process with caseworkers demonstrated that the concept of value metrics, so important in the design of algorithmic systems, is not monolithic and tends to be oversimplified. We show how and why we cannot and should not reduce the concept of value metrics to performance measurements alone. Prior research shows that such oversimplification in		

Table A3: Themes based on inductive coding (continued)

	classification can materialize within algorithmic decision-support systems as discrimination against specific people [11].” (Case E: 6)		
“impartial”	As illustrated by the description of the situation at AF above, RPA demands more secure log-ins and this constrains what can be seen as the clients’ discretion. They also showed the clients at the contact centre how information security has increased and that personal information is no longer shared among the staff, instead remaining ‘hidden’ in the system. Since personal and case sensitive information is less transparent for the staff, there is also the potential to make case management more impartial with a focus on legislative duty-oriented values. The ambition to design more advanced and efficient systems is in line with the complex legislative framework and with the intentions of the welfare policies behind the specific social insurance scheme to be inclusive and impartial.” (Case A: 384)		
“judgements”	“To the caseworkers, it is the kind of data that only professional workers can act on. They are to do with judgements that only people make about each other: about character, intention, reliability, good faith and the rest. If we believe the caseworkers, judgement of character cannot - and should not - be summarised in a bullet list, for example. To our knowledge, these insights have not previously been reported in the literature.” (Case H: 21) “In the context of this research, risk predictions of long-term unemployment were defined by the municipality as a problem that AI could solve. However, the caseworkers are sceptical of the idea that anyone or anything ought to predict people’s futures.” (Case H: 21)	Ability to understand human nature	
“moral(e)	“As we have empirically shown in this paper, it is not only a question of the technicality that matters when implementing AI for decision-support but also caseworkers’ moral judgements about what data is considered problematic to record.” (Case H: 22)	Data protection and ethical considerations	
“direct machine (…) contact”	“Some experiments now involve jobseekers in collecting information through the Pôle Emploi web portal, via a multitude of applications. This makes it possible to establish direct machine-human contact; analysis of the data entered leads to a wide range of suggestions regarding jobseeking approaches to be favored, relevant training courses, recommended employment channels, comments on CVs, etc.” (Case C: 9)	Transparency of process	Transparency