“We Would Never Write That Down”: Classifications of Unemployed and Data Challenges for AI

ANETTE C. M. PETERSEN, IT University of Copenhagen, Denmark
LARS RUNE CHRISTENSEN, IT University of Copenhagen, Denmark
RICHARD HARPER, Lancaster University, England
THOMAS HILDEBRANDT, University of Copenhagen, Denmark

This is a pre-print of an article published in Proceedings of the ACM on Human-Computer Interaction. The final authenticated version is available (from April 2021) at: https://doi.org/10.1145/3449176

Please cite as: Anette C. M. Petersen, Lars Rune Christensen, Richard Harper, and Thomas Hildebrandt. 2021. “We Would Never Write That Down”: Classifications of Unemployed and Data Challenges for AI. In *Proceedings of the ACM on Human-Computer Interaction,* Vol. 5, CSCW1, Article 102 (April 2021), 26 pages, https://doi.org/10.1145/3449176

ABSTRACT: This paper draws attention to new complexities of deploying artificial intelligence (AI) to sensitive contexts, such as welfare allocation. AI is increasingly used in public administration with the promise of improving decision-making through predictive modelling. To accurately predict, it needs all the agreed criteria used as part of decisions, formal and informal. This paper empirically explores the informal classifications used by caseworkers to make unemployed welfare seekers ‘fit’ into the formal categories applied in a Danish job centre. Our findings show that these classifications are documentable, and hence traceable to AI. However, to the caseworkers, they are at odds with the stable explanations assumed by any bureaucratic recording system as they involve negotiated and situated judgments of people’s character. Thus, for moral reasons, caseworkers find them ill-suited for formal representation and predictive purposes and choose not to write them down. As a result, although classification work is crucial to the job centre’s activities, AI is denuded of the real-world (and real work) character of decision-making in this context. This is an important finding for CSCW as it is not only about whether AI can ‘do’ decision-making in particular contexts, as previous research has argued. This paper shows that problems may also be caused by people’s unwillingness to provide data to systems. It is the purpose of this paper to present the empirical results of this research, followed by a discussion of implications for AI-supported practice and research.

KEYWORDS: Unemployment; Invisible work; Moral reasoning; Data work; Predictions; AI.

1 INTRODUCTION

> “Internally, we divide 2.3’s into *‘heavy’* and *‘light’*. ‘*Light’* are those you expect can handle a part-time job [...] or possibly be ‘ready to work’ if they get the right support. Then there’s the other group. If you call it *‘heavy’,* it really doesn’t sound very nice, but...
> They are the ones where medical conditions, the citizen’s behaviour, and other things point in the direction that it’s not very realistic [to get a job]. This is not something I can communicate. They’re all categorised as 2.3’s.”

In job centres in Denmark, caseworkers must categorise unemployed welfare seekers based on their ‘readiness’ to work. Caseworkers have two categories to choose from: those who are *‘ready to work’* and those who are *‘ready for activation measures’* needed to become employable. The categories are also legally known as match group ‘2.2’ and ‘2.3’[1] and are designed to get people off benefits and back into employment [10]. The problem is that, in practice, welfare seekers often do not ‘fit’ into prescriptive distinctions. Many get stuck in a ‘grey area’ of the unemployment system

[1] Match groups has since moved from §2 to §6 in the law text and been renamed to match group ‘6.2’ and ‘6.3’.

where they are not well enough to manage a 'normal' job, but on the other hand, they cannot document a permanent inability to work either. To operate in the grey areas of their work, caseworkers make their own classifications of the individuals they seek to assist. However, for reasons we will demonstrate in this paper, they choose not to record this work - despite it being intrinsic to their activities, enabling them to process cases in practical ways. The caseworkers' classificatory work could be useful as training data for AI systems. However, as the introductory quote clarifies, it remains invisible to the bureaucratic record, and hence AI too.

AI systems are increasingly entering the field of public administration with the promise of improving decision-making [60]. In Denmark's capital region, they are already working to 'assist' caseworkers with predictions about unemployed citizens [24, 76], such as their risk of entering long-term unemployment [35, 62]. To succeed at this task, the systems need to be based on a solid understanding of how decisions about individuals are, in fact, made. Research in CSCW has previously raised concerns about the information available to systems. Informal practices are often excluded from formal representation, and typically only the 'formal' criteria of decision-making are made visible to systems [13, 18, 56, 61]. Thus, technologies are routinely seen as incomplete 'ordering devices' [18]. Through ordering, they enable (and shape) what is made visible and what is not [e.g. 13, 47, 61, 81, 84]. But as they have incomplete information, they order poorly.

At first glance, this makes the appeal of AI systems all the greater, since they can handle complexity and combine different criteria in ways that distinguish them from earlier technologies. This could include factoring in informal classifications. The 'smartness' of AI systems lies not only in their ability to process big amounts of data, beyond the scale of humans [32], as they can process heterogeneous data and identify many patterns within it, scale often being useful but not always a prerequisite. The new possibilities that afford AI are demanding considerable reflection on what this might lead to [36]. In this paper, we argue that it is not just a question of the technicality that matters for success to arise with AI. As Barocas and Selbst [5] point out, *'What a model learns depends on the examples to which it has been exposed'.* Bigger data is not always better data, and larger volumes of data do not always enable diverse patterns to emerge if the data are suspect [20]. However large a dataset, if the data are misrepresentative, or in some other way poor, the outputs of their analysis will be similarly poor [9]. It might even matter that some data are unusable, unavailable, or unrecorded since this will limit what AI systems can learn from. Resulting absences might raise more questions than answers. Data that is processed might also reflect biases in society and can affect classes of people in consistently unfavourable ways [5, 7, 33, 50, 72]. Data is not merely a matter of what machines learn from; it is also a matter of humans who feed machines the data [22].

A long time ago, Bowker and Star [18] called attention to the moral decisions involved when creating and maintaining classification systems, many of which end up in computer systems. Up to this day, these concerns have mainly been addressed from the perspective of policymakers [e.g. 36, 51] or system developers [e.g. 1, 15]. A common belief is that decisions about categorisation are no longer made at 'street-level', but undertaken by 'system-level bureaucrats' who fit street-level categories to computer level ones [15]. However, this perspective ignores an intricate interdependence on the street level and how computer-level categories are often dependent on human judgements made by practitioners [79]. The category work used to make data starts at the street-level of, for example, caseworkers. However, concerns about data are rarely addressed from caseworkers' perspective. We have limited knowledge about the work going into producing data and its consequences for the resulting datasets, used by AI systems, remains unknown.

In this paper, we build on previous work on these aspects of classification systems to address new complexities of deploying AI to sensitive contexts, such as welfare allocation. We report on caseworkers' decisions when classifying unemployed citizens on welfare benefits in a Danish job centre. We consider both the information generated by caseworkers and the data made available to AI, as a method for understanding the possibilities and limitations of these systems. Unlike previous studies, we find that the difficulties of implementing AI may not be as reliant on the features of the technology itself; such as the adequacies of training, the labelling of data sets or the depth of machine learning. Instead, as we will demonstrate, problems may also be caused by people's unwillingness to provide data. Studying how caseworkers make these decisions is important and timely as AI may be introduced in the hope of automating and augmenting this work. In either case, how well it might do so deeply depends on the adequacy of the data provided. The questions we seek to answer are: *Why are caseworkers reluctant to record all data on unemployed citizens that they themselves use in their classification work, and what are the implications for AI as a decision-support tool in this context?*

We begin this task by describing the related work reviewed for this study, followed by its setting and methods. We then present our findings and conclude with a reflection of what they mean for the role of AI in welfare services, and elsewhere. For this study, we apply a broad understanding of AI and simply characterise it as processing data for the purposes of pattern analysis. How AI does this or whether the patterns in questions are for welfare provision or some other task is immaterial. Although we fully understand that AI systems can be varied and 'multiple', we want to take their potential seriously by understanding how they may constitute and become constituted in and through caseworkers' practice [22]. Our concern is first and foremost with that of practice, with AI as a consequent of that practice.

2 RELATED WORK

As socio-technical systems of classification, the technicalities of AI are in many ways related to the social history of ordering of various types [22]. In the following, we take on this position as we revisit previous research on classification and how classification systems, with their increasingly predictive power, may contribute to, or intensify, social processes of order and control [23]. We are also interested in how categories matter for individuals, not just things that constrain them. We link studies across CSCW, HCI and STS to address more current concerns about AI in the area of inquiry. Finally, we bring forth new challenges of implementing AI in these and other sensitive decision-making contexts.

2.1 Sorting out and 'making up' people

To classify is part of being human and as previous research shows, we have always sorted people into 'kinds' as a way of navigating spaces and making sense of the world [29, 30, 31, 40, 81]. Classification also works to serve institutional needs. Public organisations, such as job centres, need to classify those they serve to determine their economic support and, pivotal to this, is to make certain individuals 'legible', by which is meant appropriately classified. As noted by Garsten and Jacobsson [40], legibility is crucial, as it allows staff in job centres, namely caseworkers, to verify, control, sanction, reward, follow-up, evaluate, or compare intervention programmes about welfare seekers. This is by no means a neutral process but one informed by organisational priorities and political aspirations and much more besides [40]. Leaving aside their source for the moment; the use of these categories, the process of making individual cases legible, ensures the categories have a continuing life [88]. Administrative categories that make individuals 'legible', reinforce and revitalise the very standards they articulate. What is effectually considered 'normal' is what is 'legible'.

There are, however, different ways of approaching categories and their uses analytically. One way is to look at categories themselves, assuming that categories impose themselves in people. The philosopher of science, Hacking [42], for example, connects the emergence of a statistical society to the idea of 'making up people' in that *"Human beings and human acts come into being hand in hand with our invention of the categories labelling them".* Classifications, he argues, furthermore 'loop back' as they shape those being classified. They become whom we have defined them to be, which in turn confirms our classification, and leads to further classification. Classifications are sociologically *performative* as they contribute to constituting further actions and expectations of those classified [40, 53]. In this view, to classify is highly consequential for those who are being classified – and especially if they do not fit into universal standards and 'match' the explicit assumptions made about them. As noted by Bourdieu [14] *'assembling in one place a population homogeneous in its dispossession also has the effect of accentuating dispossession'.* Classifications do not only label people. They may also alter future outcomes and determine people's fate [46] or lead them to be thrown around the system [13], discriminated against or left behind.

Another way to analyse category work is to look at how people appropriate categories in their own reasoning. As Douglas [28, 29] reminds us (albeit in a very different context), classifications are to be understood in their context of use by those whose business it is. It is individuals who sort out how local circumstances are seen to match prescribed categories and whether those categories need reformulating. Often, street-level bureaucrats, such as caseworkers [54], function in situations too complex to reduce to prescribed responses. Instead, they use their 'informal' powers of discretion to interpret and modify formal rules before making decisions about intervention of one kind or another [11, 89]. Their decisions are informed not only by bureaucratic rules and standards but also by collective judgements [66] about, for instance, moral standards [49, 59] that underscore the application of categories. These categories are expressed in everyday language and organise how the world is understood in terms of categories.

Sacks [74] also noted how people use categories as 'devices' that link types to doings. For example, the type, 'normal worker', is used to imply things about behaviour. It is not just a person's label (this person is type X) but links the person to kinds of behaviour (i.e., because this person is type X they are likely to do type Y behaviours). Building on these insights about categories in use, Maynard [58] finds that such language practices are intrinsic to organisational life. In law settings, public defenders rely on how certain 'types' of people behave in certain ways. For example, 'someone very poor' is likely to steal 'an item of necessity' because 'she needed it', and not because 'she was in too much of a hurry to pay for it'. Similarly, Keddell [49] found that in the case of child protection services, social workers often describe parents in terms of mental illness to convey reasons for their 'lack of culpability'. In regards to membership categorisation, this keeps them out of the category 'blameworthy' and maintains them in the category of a 'good' (but struggling) parent. According to Sacks [74], the characteristics used to classify people are not 'natural kinds', simply awaiting 'discovery' [31]. They are reflexively constitutive of everyday life. People use membership categorisation devices to make the world meaningful, classified, and ordered in ways that make sense to them. Hence, Sacks and others in this tradition, look at how people use categories in their particular life situation.

Both the social view of categories and their contextual use, speak to the case of job centres in Denmark. The categories used here can be thought of as social constructs and hence imposing meaning, but how they are used in particular instances also opens up reasoning as a human act in particular places. To categorise someone as 'employable' implies a 'universal' distinction, such as employability and non-employability, but is deployed in reference to a particular person [83]. The analysis and development of technologies that rely on categorical representations, like AI, can be confounded by this distinction, with the categories sometimes giving an 'allure of objectivity and inevitability'. This can make iterating the systems difficult. It can also beg the question about which categories they should articulate [7]. For the same reasons, as we will turn to next, 'classification systems' are often found limited in their capability to capture any 'matter out of place' [29].

2.2 Technologies and incomplete 'ordering devices'

In the context of public administration, divergence from bureaucratic order is typically treated as non-compliance, but, as Garfinkel shows, rules and practice generally have a complex relationship with each other. In Good Organisational Reasons for Bad Clinic Records', Garfinkel [38] famously points to a gap between organisational ordering patterns and what it actually takes to describe practice:

> "The documents' meanings are altered as a function of trying to assemble them into the record of a case [...] Thus an effort to impose a formal rationale on the collection and composition of information has the character of a vacuous exercise because the expressions which the so ordered documents will contain will have to be 'decoded' to discover their real meaning in the light of the interest and interpretation which prevails at the time of their use." [38]

A bureaucratic rule saying something *should* be done may change the produced account of that work, while the work *itself* might remain the same [71]. That is, there may be good reasons to work around formal systems, which is also a central focus of CSCW research. At the core of CSCW is an emphasis on understanding the details of social settings to inform technology design. Studies of 'invisible' work have already proved seminal in amplifying our understanding of professional work and proven useful in giving voice to the performance of tasks that is often left unacknowledged or unnoticed by others [e.g. 16, 18, 56, 57, 61, 84]. For example, Suchman [85] sought to apply Garfinkel to the question of machine manuals (in particular, Xerox photocopiers) and noted how the manuals only made practical sense when used in the right situation. In themselves, such things do not have adequate meaning. They are too abstract. However, as Garfinkel [39] noted, this 'looseness' allows their generalised meaning to be appropriated for particular contexts. In other words, manuals, like rules, are designed to be interpreted. In Suchman's words, *'the efficiency of plans as representations comes precisely from the fact that they do not represent those practices and circumstances in all of their concrete detail'* [85].

Building on these insights, several studies have examined the problem space between the information made accessible to systems through formal documentation and the broader set of information constituting of the 'know-how' actually used in some situations. Manuals and guides are resources and not descriptions of work, but how they are resources is worthy of investigation. From the work of Suchman [86] and Schmidt and Wagner [75], prescriptive technologies arguably achieve their efficacy not despite, but because of what they leave unspecified. A formal plan's inherent underspecification affords the space of action needed for its realisation [87]. A slightly different argument is put forward by Bowker and Star [18] who refer to classification systems as 'ordering devices' [78] and find that an emphasis on 'order' can ignore the informal and 'invisible' work that has gone into creating and maintaining order. In their view; in practice, every standard is overdetermined and incomplete. Tinkering, repairing, subverting, or circumventing prescriptions of standards are necessary to make them work [52]. From this perspective, categories are not resources to individuals, as Sacks would have it, but things to be 'resisted'.

These two distinct but related views have driven a great deal of research. For example, prior studies have empirically investigated the invisible work involved in making classification schemes ‘fit’ into local arrangements. Martin et al. [56] find that work requires continuous in-situ decisions and workarounds by operators which, among other things, involve the creation of new categories [56]. These may be valuable and crucial to the actual conduct of the work process – yet they are not visible outside their context of use. In the CSCW community, these categories have become known as ‘residual categories’ (i.e. ‘other’, ‘nowhere else classified’ categories). Residual categories are not represented within any given classification system, yet, classification systems often have to rely on residual categories to render themselves complete [57]. For example, in their analysis of the International Classification of Diseases, Bowker and Star [18] point to the contingencies and contests that went into the classification of viruses and the surprising non-existence of ‘old age’ as a formal cause of death. Research in hospitals also shows that ‘subtle categories’ are used to sort out patients with potential cancer [61]. These include phrases and concepts like *‘patient lost 20 kg’* or *‘weight loss’*. Both create a definition *not* supported by the ‘formal’ categories, which typically presume that the existence of cancer is ‘clear’.

Here, as elsewhere, caseworkers know that multiple factors determine whether a person is processed one way or another, and sometimes only show themselves after extensive inquiry. Uncritical reliance on technology might therefore hide the complexity of real-world (and real work) decision-making [57]. According to Bowker, technologies affect what will, and what will not be made visible [18]. Ultimately, they “*operate through being invisibly exclusionary*” [17]. This, in turn, led many CSCW researchers to attribute issues of visibility to the technologies and conclude that we need to build systems that are better at taking into consideration the informal (and ‘invisible’ practices), based on the implicit view that systems essentially *shape* what is made visible and what is not [e.g. 13, 47, 61, 81, 84].

While a focus on technological incompleteness has been a useful basis for a considerable body of research to date, they are simultaneously limited by the technologies of the time they were written. In light of recent advances in AI and its growing use in public administration [60], we need new perspectives to understand the challenges involved when implementing AI in sensitive decision-making contexts.

2.3 From expert systems to machine predictions

The interest in CSCW with plans and situated action draws on early discussions with the AI community and a critique of the ‘office automation’ movement from the 1970s [85]. At this time, computer systems were fundamentally ‘dumb’ [44]. Information was organised in strict and inflexible hierarchies, and they were incapable of adjusting criteria to real-world (and real work) realities. What mattered practically was not the issue of what words really ‘meant’, but how they could be defined in formal terms. With recent advances in AI, information can now be organised in more flexible ways. Bits of data can have multiple associations with other data, and categories can change over time [41]. Thus, it is no longer sufficient to say that systems are simply ‘incomplete’ or unable to support people in doing their tasks, as previous research reports. For instance, residual and subtle categories such as *‘old age’* [18] and *‘patient lost 20 kg’* [61] may not be formally specified by practitioners, but this does not in itself exclude them from being included in the treatment of a patient. With current AI technology, their meaning may be derived from other patterns in the data and still be considered a cause of death or a sign of cancer. Recent machine learning (ML) trends include unsupervised learning, a type of ML that looks for *previously undetected* patterns in datasets with no pre-existing labels. However, even with more flexible forms of databases, classification and categorisation remain vitally important since there is a premediated order necessary for algorithms to work; information must be formalised so that algorithms can act on it without any regular human intervention or oversight [41].

Attributes like someone's 'real' age may be appropriately formalised and objectified in machinery, but challenges quickly emerge when the existence of stable explanations are taken as given [77]. Additional value may be added in the interaction between people, which could cause problems to computers as they cannot *experience* the world as human beings and the dynamics involved in these contexts [27]. Examples from medical research include the development of ML tools to predict sepsis in patients. Here, the authors found that in practice, there was no standard way of diagnosing the disease. In return, they ask *"What should constitute an explainable algorithm in clinical practice when the definition and underlying pathophysiology of sepsis are incompletely understood in the first place?" [77].* It is significant for technology design that *"if a description is not there, then intentional actions under that description cannot be there either"* [42] and once descriptions are created, they may become difficult to 'unthink' [82]. Transparency of information is not merely a matter of reporting or disclosing information about an already existing description of a person. It also creates the person it seeks to make transparent [36]. Technology, then, becomes a crucial, shaping element in that it helps bring people into being [8]. It may also intensify social processes of classification and control [23]. The design itself is a locus for political action, and through making classifications visible outside their context of use, they can be used politically [68].

Friedman, Bannon and others remind us that there is no such thing as value-free design [3, 4, 37, 63]. Technologies used to support public sector services particularly *"reflect values from the very political context in which they are borne"* [90]. Looking at the public sector and beyond, current AI systems are primarily driven by economic incentives and efficiency ideals, and they are routinely designed to profile people and predict their futures [23, 62, 64, 65, 70]. In these contexts, poor people on public benefits often become the test subjects [33]. Notably, recent findings from a participatory design set up in a job centre show that risk prediction of long-term unemployed did not fit the practices of caseworkers [62]. Instead, caseworkers wanted to shift attention from the individual towards the organisation and use AI to predict waiting time in cases, such as the time it takes to receive medical records on citizens. This finding adds to previous concerns of risk prediction that focuses on negative outcomes based on negative inputs, which may, in turn, drive negative actions. Instead, Brown and colleagues [21] ask for predictive models to invert risk factors into positive variables, such as the likelihood of not-failure.

Data that are used predictively may also reflect biases in society and negatively affect groups of people [5, 33, 72]. Related work on classification and predictive AI finds this to be deeply problematic. For example, in 'Race After Technology', Benjamin [7] challenges the position that predictive methods are generally beneficial to society. Exemplified through a machine-learning algorithm used to predict crime zones, she addresses racial profiling problems and asks if people in these zones will automatically be perceived as suspicious? [7]. The closer the particular prediction is to broad categories, such as race and gender, the more troubling it seems. It may be even more problematic if the social categories that underlie the training data are externally assigned, and the role of self-knowledge is ignored in the process. This has been the focus of recent critical studies, such as automatic gender recognition, where gender is generally seen as an essentialist binary in which there are two categories [50].

Data is often used by AI tools to draw normative distinctions of people [33] and from this and other examples above, it can be argued that technologies may in some situations become solutions to what are, in reality, social and political problems [6]. When dealing with data-dependent AI that learns from real-word attributes, derived from human activities, about human matters - we must take these and other issues into account. However, as public administrations increasingly embrace AI, these perspectives often get lost in the process, and the work going into producing training data goes unaddressed. Before presenting our empirical findings on this matter, we will first consider the research setting and our study methods in more detail.

3 RESEARCH SETTING

This paper's research setting is a major municipal job centre in the capital region of Denmark looking to experiment with AI to predict, and thus prevent, long-term unemployment of its citizens. The municipality is a known frontrunner in adopting new technology to support decision-making activities and has previously experimented with AI in areas unrelated to unemployment. At the time of conducting our research, the job centre constituted a Head of Employment, administrative staff and four departments handling cases concerning cash benefits: 1) allowance and availability, 2) job and company, 3) job and competencies, and 4) job and resources. Caseworkers within these departments consisted of a small number of trained social workers and a larger group of caseworkers with different backgrounds. These included previous experience as a sales manager, unemployed, graduate etc. All caseworkers attend mandatory courses, but no prior experience in the field is required.

The role of job centres in Denmark is to provide a unitary employment system offering one-stop access to all citizens. Their main task is to establish a quick and efficient match between job seekers and companies [26]. It is a legal responsibility of the caseworkers in job centres to decide whether a person seeking cash benefits can take a full-time job within three months of unemployment. If the answer is yes, the person is placed in a 'ready to work' match group, also legally known as '2.2'. If the person faces challenges beyond unemployment and is found (currently) unable to work, they are considered 'ready for activation measures' and placed in match group '2.3' [10]. Activation measures are designed to get people 'job ready' and may include training courses to assess work capability and skills. Illustrative examples of who may be an 'obvious' fit for match group 2.2 and 2.3 are provided below. Anyone who does not fit into these match groups is considered a 'grey area'.

> Match group 2.2: Anne is a 32-year-old woman who recently graduated with a bachelor's in marketing. Anne is passionate about finding a full-time job where she can use her new skills. However, a lack of professional experience proves it difficult. Anne's caseworker signs her up for an internship as an administrative assistant in the metal industry. At first glance, both Anne and the caseworker are sceptical if it is the 'right fit', but it turns out that the chemistry is good. Anne increasingly develops her skills and as she settles into her new role, the company offered her a full-time job.

> Match group 2.3: Carsten is a 50-year-old male recently diagnosed with schizophrenia. Carsten attended ten years of primary school but received no further education and never held a permanent job. He lived for a few years in a high support care home and has struggled with alcohol and hard drugs. Today he is using hash to self-medicate his mental condition. The caseworker finds no reasonable doubt that he is not (yet) ready to take on a full-time job, so there is no need to do a more thorough assessment.

Everyone in match group 2.2 is required to be actively job seeking. The caseworkers told us that in practice, it means that they must submit at least two job applications per week. If failing to do so, the caseworker must sanction the citizen (i.e. stop their payments for some time) to keep control of welfare benefits. According to the caseworkers, and as illustrated in the example above, it is far more challenging to force actions upon or sanction 2.3's. Per definition, they face challenges beyond employment that must be considered. If an applicant is below the age of 30 or fails to meet the requirements for cash benefits, they may be put into one of 12 other match groups as per the

Danish law of active job-creation effort. These will not be given further attention in this paper as our research focused exclusively on cases involving match group 2.2 and 2.3, and those who fall somewhere between the two categories (i.e. the *grey area*). We chose these cases in collaboration with the municipality as they are high in number and often complex, thus calling attention to the need for careful examination before considering the introduction of AI.

3.1 AI in the job centre

This study is part of an interdisciplinary research project where caseworkers' decision-making activities are ethnographically examined across Danish municipalities to inform and evaluate the development and use of AI tools for decision-support. When this study occurred, the caseworkers' primary system was a case management system called 'Momentum'. Momentum is developed and maintained as a cloud service by one of the research project's industrial partners. Furthermore, internal project members from computer science concerned with AI development were given access to extract de-identified data for use as part of the research project. Collaboration between internal project members and the job centre was initiated at the project's beginning in 2018, and the goal was to implement the learnings from the research project on an ongoing basis until its completion in 2021, meaning that the caseworkers would be using AI for decision-support on real cases at some point within this timeframe.

When conducting the research for this paper, our fellow project members from computer science were simultaneously doing experimental studies on whether it was possible to predict long-term unemployment on anonymised real cases from the job centre's match group 2.2. and 2.3 citizens. The goal of predicting long-term unemployment was based on a wish expressed by the municipality. A 'good' outcome was defined as employment within a year (as per the law on active job creation efforts [10]) and the idea was to investigate if it was possible to train classifications to predict the risk of a negative outcome (i.e. long-term unemployment). The data sources used for the initial studies concerned the events that may happen during a case: 1) the laws regulating the process and the domain knowledge of lawyers, 2) the IT systems used for case handling and the domain knowledge of system developers and 3) the workflows carried out by caseworkers and their domain knowledge. The information on the latter was to be obtained directly from our empirical, ethnographically informed practice studies. However, the experiment showed multiple challenges involved in performing such a prediction.

The challenges addressed in this paper was the finding that data was missing to train the models – possibly hindering successful implementation of the new technologies in the future. Key to the early models developed by our project members was the match group history of citizens (i.e., 2.2 and 2.3). Initially, they used law texts and database registrations to gain more information about match groups. Still, it was soon discovered that crucial information was missing on caseworkers' reasoning behind placing or moving citizens in or between match groups. For instance, as the developers told us; the most common registered cause of match group closing was 'other' or 'change to other match group' – which is not indicating much and thus, not very useful for predictive purposes. This became a practical challenge in the labelling of traces and the prediction of long-term employment since the developers, among other things, needed to know if a citizen's case ended with a job.

Based on these insights, the goal of this study was to investigate the criteria used by caseworkers when making decisions, formal and informal, to gain a better understanding of how citizens are moved in and between match groups, how these activities are communicated and thus, how they may (and may not) be supported by AI. Our interests were thereby broader and more far-reaching than those associated with the above AI experiments. It was not our goal to determine neither long-term employment nor whether that was a 'good' or 'useful' prediction to make. We will elaborate further in the methods section below.

4 METHODS

Our study was performed by conducting ethnographically informed field studies [69] with caseworkers handling cases concerning recipients of cash benefits in match group 2.2 or 2.3. The first and second authors visited and revisited the site over seven months between fall 2018 and spring 2019 for a total of four weeks. Together with industrial partners and internal team members from the research project, we participated in training courses on current case management systems and workshops with both caseworkers and developers of AI systems to establish common grounds for intervention. Additionally, the first author observed and interviewed multiple caseworkers, participated in courses, meetings between caseworkers and citizens, and analysed documents outlining organisational structures, policies, and laws.

Observations and interviews were typically structured in two phases. On the first day, we would observe a caseworker in administrative tasks, meetings with citizens, and conversations with colleagues. We would be following them in all of their activities that day, going for walks with them, and having lunch with them. The second day was followed up with further observations and semi-structured interviews with the caseworker lasting approximately 1 to 1.5 hours each. Field notes were written during observations. Interviews were used to interpret the findings from other data sources better and dig deeper into caseworkers' practices. In total, we observed and interviewed seven caseworkers in different age groups, of which five were women, and two were men. Their background included two trained social workers, one teacher, one graduate in political science/economics, one sales manager, one substance abuse counsellor, and one caseworker with previous experience from other job centres. Two of the caseworkers shared their own experience of having been a part of the welfare system before. Everyone expressed a feeling of being part of an 'unfair' system and wanting to make a difference. We also engaged in conversations with both citizens, managers, and security guards in the job centre. Interviews were audio-recorded and transcribed for structured analysis in NVivo 12 [67].

Data gathering took a relatively unstructured form to begin with and developed into a more structured and strategic form towards the end of the study [43]. This process enabled us to continuously gather, categorise, compare and contrast common themes and significant issues found in the data. Our ongoing interpretation was used to inform the direction of the fieldwork [48]. During the early phases of data collection, we discovered a vital misalignment between the factors going into classifying citizens and the data available to AI developers to generate predictive recommendations. As previously stated, this insight led us to analytically identify the information used as part of this work and what becomes visible or remains invisible to technological tools [81].

Our ethnographic study is inspired by ethnomethodological perspectives [39]. Therefore, the primary attention is on caseworkers' *reasoning practices* - the reasons they offer to explain and describe their classification work [34]. In choosing this as our research focus, we also follow the recommendation by Harper et al. [45] suggesting that *'reasons provide the bedrock of how choices are seen, accounted for and ignored'*. While we attend to AI issues, our main concern is with the pragmatics of classifications that emerged from the observations and interviews with caseworkers. This involves taking a theoretical unmotivated approach to activities and *'looking just to see what people are doing'* [73]. Therefore, it is not our role to decide what things are, what matters, what is important, trivial, or even right or wrong, but to ascertain *how* things are made sense of by those who are doing them [25]. Additionally, in taking an ethnomethodological approach, we also consider AI in how it is 'already embedded' in particular circumstances [22]. At the time of conducting our field studies, all caseworkers in the job centre were familiar with the ongoing AI development and its implementation in practice in the near future and that the goal of our empirical work, ultimately, was to look for implications for design. This also means that AI can be seen as somewhat already introduced to the job centre in how it already occupies caseworkers' minds, making them act and react in specific ways to it [22]. Recognising this helps us understand AI's current role in the job centre and understand the negotiations involved in developing AI tools in these contexts and the possible resistance to these processes.

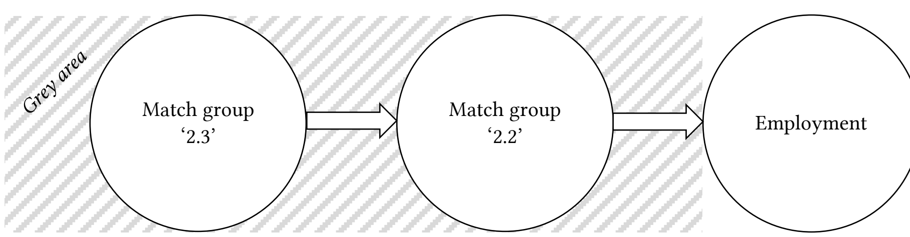
5 FINDINGS

As we move on to present this study's findings, we will first account for the formal aspects of categories used by caseworkers and describe their practical limitations. From this, we will draw attention to the informal classifications created and used by caseworkers to address these practical concerns. Lastly, we will present the reasoning offered by the caseworkers as they decide whether to disclose these data and hence, to make them available to AI for predictive purposes.

5.1 Classifying the unemployed

While the formal task of caseworkers is to help unemployed find jobs and companies finding new employees, in reality, they deal with many people who have been in and out of the system for years, some of which never had a job. Legally, cash benefits are considered temporary assistance and the legal system is designed to get as many off benefits as possible and (back) into work. The evolving nature of these regulations further supports this. Whereas once there were five 'ways' to be unemployed and receive cash benefits, there are now two [55]. More recent governmental strategies seek to move an additional 20,700 Danish recipients of cash benefits from match group 2.3 towards 2.2 [2]. Caseworkers undertake decisions about placing unemployed citizens into these ever-reducing number of match groups. Their decision-making is based on what one might say are normative models. The models presuppose that any 'normal' person would be willing to work for 'normal' reasons and have an economic incentive to do so. This normative rationale is reflected in the Danish law on active job creation efforts as illustrated in figure 1 below, where not only the goal is to bring everyone on cash benefits closer to employment, but everyone is characterised as being employable, people's motivations are understood as given and the general capacity to being functional is routine. We might say that the law 'encourages' some ways of seeing and being (and 'working') and outlaw others. As previously noted, citizens are legally characterised as either 'ready to work' (match group 2.2) or 'ready for activity measures' (match group 2.3), leaving a grey area between binary designations. A grey area is included in the below figure for illustrative purposes, but with no intent to block out its many shades.

Figure 1: The legal rationale behind match groups for recipients of cash benefits.



Whether it is normative or otherwise, caseworkers need to classify citizens to determine their employability. However, they know a lot is at stake for those who are being classified. Therefore, it is also a big responsibility for the caseworkers to ensure that each citizen is in the 'right' match group. To live up to this responsibility, caseworkers need the appropriate knowledge required to get citizens clarified and processed purposefully - and at the right time. When meeting someone for the first time, and when the need to evaluate past decisions arise, caseworkers have to answer critical questions such as if that person is ready to actively search for a job or participate in activation measures to get them closer to the job market. Technically, this decision can be made by using the legal definitions described earlier. Still, there are only a few options to choose from, and the logic of the binary system often has little to do with reality. As one caseworkers explained:

> "Your only challenge in life is that you don't have a job. That's basically what it means to be 'ready to work'."

Thus, a central problem caseworkers encounter is that the options available in the system are too narrow, and they find themselves struggling to make them work. The caseworkers told us that in practice, most cases are multi-faceted and do not 'fit' into intangible values and one-sided characteristics:

> "We constantly experience [grey areas] as we have to do with people and people cannot be placed into categories [...], but you have to do it."

Making decisions about people's rights and duties without having all the information needed can be problematic and fraught with consequences. When people enter the job centre, they often face severe circumstances that have not yet been clarified. As we were told, *"you may have woken up with a depression*". Regardless, every argument for match group 2.3 and against employment needs to be backed up with documentation. This includes medical records describing addictions, mental illness, or other challenges citizens' face, making them unable to work. Retrieving these records can take more time than caseworkers have, and the records may not always be comprehensive enough for their purpose. Other times, citizens themselves may not agree to get an examination in the first place, or they simply fail to show up at the doctor's office. For whatever reason, the caseworkers, as a result, repeatably and often unwillingly find themselves classifying citizens as '2.2's' even though they find citizens unfit and unprepared to meet the demands placed upon them with the category. One caseworker described the challenges of such a case:

> "How can he be [a 2.2]? He hasn't been at work for eight years. A 2.2 is someone we evaluate can return to the job market within three months, and if he hasn't been on the job market for eight years, it has to tell us something's wrong. Then it's us who don't have the correct information. Documentation is missing. I don't know why he's a 2.2. I can't just make him a 2.3."

The caseworker needs only to recognise that the citizen has been unemployed for eight years, to determine the unsuitability of relying on the match group for adequate information. The caseworkers explained that due to cases like these, they think of match groups as something that mainly defines the rights and duties of a citizen. Still, it does not necessarily tell them anything about the person behind or whether they believe he or she belongs to the match group. It says more about the work that needs to be done to accommodate or 'fit' the rules and universal concepts. Similarly, moving citizens from one match group to another does not necessarily mean that a change in the citizen's situation has occurred. Recall that the different 'ways' to be unemployed and receive cash benefits in Denmark were recently reduced and constrained. Indeed, historical and political circumstances play a part in both the creation and maintenance of categories. Furthermore, it is (and has to be) in the management's interest to meet the increasing political demands of getting a higher number of unemployed into labour. Therefore, the desire to move a citizen to match group 2.3 is also made difficult by managerial restrictions. In setting restrictions, at least it 'looks' as if the job centre is working towards that goal.

Being categorised as a '2.3' is arguably preferable to those unable or unwilling to take a job since the match group comes with fewer obligations and less punishment. Besides, where everyone in match group 2.2 must be actively job seeking, 2.3's are, per definition, facing challenges beyond unemployment which must be considered. Then again, one of the caseworkers told us that in practice: *'You have to be almost dead before you can become a 2.3'.* There are several problems associated with this since citizens with challenges beyond unemployment do not become 'ready to work' just because they get a different label attached to them. Instead, and as previously noted, it often means that they get thrown around the system or left behind; not that they end up getting a job. We were informed that those who are 'unfit' for work often complain that they become more ill after signing up for welfare, because of the increased level of pressure that comes with the demands and expectations induced by categories such as '2.2' and '2.3'.

Arguably, the formal categories treat citizens as mere abstractions, with little or no attention to their social and historical context. As we have seen, caseworkers find it too *'difficult to put them into the boxes [the job centre] wants them to fit into.'* This is reflected in their work practices, where knowledge and insights about citizens are supported with additional information to be meaningful. As we will show next, the match groups cannot (and do not) stand alone. Instead, their meaning is negotiated and contested as part of situated practices and for practical purposes.

5.2 Judging people's character

We found that the formal and institutional match groups serve a useful function in processing cases during our observations and interviews with caseworkers. While the caseworkers recognise the limits and pitfalls of match groups and the predictive purposes they may be used for, they do not abandon them in their internal communication altogether. Instead, we observed that they navigate the formal categories and compensate for their limited information by adding layers of informal knowledge. This takes the form of adjectives which give the nuance to the categories implied in otherwise blunt nouns. To the caseworkers, this is a necessary requisite to distinguish citizens from each other:

> "We have to find our own words or ways to describe the citizens [...] because you cannot just say 'a 2.3' and understand what a 2.3 is. A 2.3 can be many different types of people."

By adding to the vocabulary, the caseworkers use their 'informal' powers of discretion to interpret and modify formal rules in ways they believe are necessary to make them actionable. Through language and categorical work, caseworkers make their 'clients' somehow 'fit' the framework. From our observations of internal meetings and caseworkers' daily interactions with colleagues, as well as follow-up interviews, we discovered a common practice of referring to citizens in terms of: *'light'*, *'heavy'*, *'good'*, *'bad'*, *'permanent'*, *'the better'* and *'the best'* 2.2's and 2.3's. Examples include *'heavy 2.2's'* who, according to the caseworkers, are those who do not act like 2.2's - alluding to the legal definition of groups and their long ('heavy') or close ('good' or 'light') distance to match groups and, ultimately, employment. Concerning the *'heavy 2.2's'*, a caseworker described the necessity of classifying this 'type' of people:

> "It's a special match group because it's those who are not suitable for other places. They're a bit in east, west, north and south with their problems."

Caseworkers' classifications mark a clear mismatch between the formal categories and the caseworkers own discretionary judgments about people. These judgements help the caseworkers see the difference between what we might call a view 'from afar' (see: figure 1) and a view 'from within'. The view from within is often based on sensory impressions obtained from frequent and mandatory meetings with citizens. The caseworkers told us that observations of citizens are crucial for obtaining a full understanding of the person behind the 'labels'. This includes phrases such as *"he smells of alcohol"* or *"she gets upset when the conversation turns to her health"*. In making these distinctions, the caseworkers reveal more of people's characteristics and provide a better foundation based on which they can choose the most appropriate actions. As one caseworker explained: *'there are just so many factors to solve, beyond helping them find a job'*.

For instance, those who have been in the system for a long time are also someone who might not get a job, and a 'heavy 2.2' will not get the same treatment as a 'good 2.2'. For a practical example, a caseworker responsible for running a course on 'job skill development' explained that she only invited 'the best 2.3's' as they are more likely to become 2.2's and thus, in a situation where job skills are considered more prevailing. We also learned that in addition to separating those who, for instance, are 'light' from those who are 'heavy', the caseworkers further recognised the groups of citizens and the teams working with them as different, by distinguishing between different types of 'light' and 'heavy'. Besides helping the caseworkers better define people's distance to the job market and the factors affecting this, this also served as a valuable guideline in the distribution of cases based on caseworkers' preferences and skills. As one caseworker in team *'heavy 2.3's'* described:

> "Each of us has our key competencies. Some are better at handling *'the difficult, mentally ill'*, others are better at handling *'those with an addiction'*, some handle those we call *'the psychopaths'*. It can be *'those who are violently aggressive'*, have *'a violent background'*, those who are *'diagnosed psychopaths.'*"

Membership categories such as 'heavy' and 'addict' create adjectives which, according to the caseworkers, inject necessary nuances into the categorisation work. By combining formal categories with classifications that they feel 'go together', they have a way of keeping certain citizens from being treated as 'ready to work'. If someone is classified as a 'heavy' and 'mentally ill', they may be seen as lacking culpability and therefore maintain a construction as 'good' but 'struggling' person. This, in turn, affects the way caseworkers approach the situation and choose appropriate actions. Recall that everyone in match group 2.2 must apply for a minimum of two jobs per week; otherwise, they must face sanctions. Regardless, and as we have shown, the law is not the only factor influencing how choices are arrived at. Caseworkers are just as concerned with the individual circumstances of the case. A caseworker dealing with 'heavy 2.2's provided us with an example to illustrate this. The case involved a citizen who formally belonged to match group 2.2, but who failed to comply with the legal requirements of job applications:

> "I've had quite a few 'heavy' citizens, in quotation marks. Someone like Lene (name changed for confidentiality) [...] Just getting up in the morning, I think, is a big challenge for her [...] Lene has to apply for two jobs per week, but I told her; *'instead, apply for two during the next two weeks'*. Because she's not going to apply for four, I'd rather say; *'we'll give you this goodwill, and I understand you, and I hear you and so on'* [...] It's a grey area because the law says I can't do that. The law states that you must apply for two jobs a week as a minimum [...], but I achieve nothing from contacting [the team responsible for

sanctioning citizens] to let them know that Lene hasn't applied for the jobs she was supposed to. If I did that, Lene would have finally snapped. Let's just slowly try a different approach instead, before we start sanctioning this too."

There may be different reasons why someone is considered 'heavy', as the quote above suggests. Every situation requires interpretation through listening and understanding and if someone is already facing a difficult time and at a breaking point, there is no reason to take away their benefits, regardless of what the law says. The law has to be interpreted as it is being applied to concrete contexts. This also implies that caseworkers do not see classifications such as 'difficult' or 'heavy' as inherently negative or positive. Instead, the labels signal that people in groups such as 'heavy' or 'difficult' are there for different reasons. When using the classifications 'heavy 2.3' and 'violently aggressive' in combination, the caseworkers explained how these might also serve to protect themselves when meeting citizens they believe display these personality traits and to make sure these 'types' of cases are distributed to someone who is 'better at handling them' to use the caseworkers' own words. Classifications, then, are used by caseworkers in such a way to form a co-membership with other categories in a situational relevant 'device'. That is to say, the use of particular classifications such as 'heavy' and 'mentally ill' are heard to go together within the device 'good but struggling' (one suggests the other) whereas violent or aggressive is different. A caseworker explained that if a citizen seems aggressive, it might lead them to think that *'they are just not very likely to get a job'*, but for reasons different than if the citizen were classified in other ways.

What the caseworkers do, essentially, is making it *practical* to categorise citizens. Indeed, there is an economy to this, in that judgements are not endlessly nuanced but sufficiently so to enable caseworkers to 'do their job'. The classifications elaborated with and through the adjectives, are there for practical purposes - getting the individuals in question into the 'formal' system. We may call them 'negotiated terms'. The classifications made by the caseworkers are used throughout casework to help the caseworkers create a better image of the citizens when determining their needs and support, as well as any other measures needed to be taken into consideration to process the case.

Below, we have included a table summarising key characteristics used and reused by caseworkers in combination and collaboration when making sense of citizens and to make the formal categories 'work'. The categories presented in this table are grouped based on the combination of words used by caseworkers to describe citizens. For example, 'good' 2.2's are considered able to work, but within this category, there is a difference between those who want to be employed and those who are deemed lazy or tired. These classifications go together as part of membership categorisation 'devices' where the 'lazy' require a different approach than those who 'want to be employed', but a fundamental commonality exists as they are both considered 'able'.

Table 1: Adjectives used by caseworkers to classify, and make sense of citizens.

#	Adjectives to match groups	Membership categories
1	The 'light' or 'good' 2.2's	Those who can work 225 hours a year, the lazy, the tired, those who want to be employed, those who (try to) cheat, those who have not yet been clarified etc.
2	The 'heavy' or 'bad' 2.2's	Those who don't fit into the match group, those who have been here for long, those who are difficult to help, those who do not want to work, those who smell of alcohol etc.

3	The ‘best’ 2.2’s	Those who are closest to employment, those who have a good chance of getting employed etc.
4	The ‘better’ 2.2’s	Those between the light/good and the best 2.2’s, those who might get closer to employment in the future etc.
5	The ‘coming’ 2.3’s	Those who are waiting for the documentation required to become 2.3’s etc.
6	The ‘best’ 2.3’s	Those who are closest to becoming 2.2’s, those who are willing to work, those who can work etc.
7	The ‘heavy’ or ‘permanent’ 2.3’s	Those who complain, the difficult, the mentally ill, those with addictions, the psychopaths, those who are violently aggressive, those with a violent background etc.

From the summarised list above, it becomes clear that caseworkers’ classifications have a moral dimension as they refer not only to vertical distinctions but also normative hierarchies. As it turns out; it is not so much a question of whether this information is ‘informal’, but of how caseworkers define relationships between people and their circumstances. Given that everything is ‘good’ or ‘normal’, a person should be able to work. However, the capacity to be ‘normal’ and fit into the normative models of citizenship is based on presumptions about employability that are often constrained by several other categories. For example, those who have been in the system for a long time may also be someone who will find it difficult to get a job and someone dealing with personal issues may not have the energy to apply for jobs. These relationships are the rich concerns that help caseworkers to inform and constitute decision-making. It renders a visible relationship between values, practical action, and the social organisation of work through caseworkers’ language when reasoning about their everyday routine practices and the premises from which they form valid inferences.

Since the classifications are informal, sharing them is also informal and communicated through word-of-mouth, internal meetings, and daily interactions with colleagues. That is, they are ‘passed on’ to other colleagues as cases are discussed. When working on cases together or when we needed help with a case, the caseworkers told us how the classifications help them to understand the situation at hand better, since, *‘a 2.3 can be many different types of people.’* We also found that most of the classifications were often not individually constructed but taken over from others, such as when a case is handed over to a colleague. The caseworkers know what ‘types’ of 2.2’s and 2.3’s their colleagues are dealing with. When a case is taken over from a colleague who no longer works in the job centre, there are still ways to retrieve relevant information about the citizen. As one caseworker told us: *“I can look at the profile of the citizen in the system and see a list of his previous caseworkers”*. Thus, the informal sharing of classifications helps ensure the taxonomy is not phased out, even as caseworkers leave or are replaced, which happens quite often.

Nevertheless, as we will turn to in the next section, it is also the informal sharing of classifications that helps keep their flexible nature and give them meaning in use. These dynamics of classifications are crucial to the caseworkers. In their view, the knowledge within classifications is often about human character and personality traits. Although documentable, and hence tractable to various ways of formalising, in its essence, it is judgmental and value-laden. Though these judgements are made with considerable thought and concern, with professional elan and care, caseworkers are morally against any outside scrutiny and feel them ill-suited to any formal representation.

5.3 Moral reasons for keeping information 'off-the-record'

We experienced a 'great divide' between the internalities and externalities of caseworkers' classifications from our observations. The 'invisible' nature of much of what is assumed within their community is intentionally kept invisible to others. The classifications of citizens are intended only for the caseworkers who are members of the 'community of practice' who form, use, and maintain them. The caseworkers explained that they are very leery of providing their evaluations of people to the formal systems or any external stakeholder as they find them difficult to articulate in bureaucratic forms. It is a moral judgement what to write down and what not to write down, and one has to be a competent member of a community to make that judgement:

> "In our team, we use a not so nice language, internally (referring to the adjectives made to match groups), but it's something we would never mention to anyone outside these walls, at all. And that's something we've talked about internally. That it's only us. We don't talk about them with the management consultants and consultants in that way, and it's also something we would never mention to anyone else, and we would never write it down."

A related concern they shared with us is that data becomes even more recalcitrant to 'proper understanding' when it is viewed 'from afar', without reference to the real-world character of actual decision-making:

> "We would never write that down. Everything we document the citizen has the right of access to. How would you feel, if you were being talked about as a 'heavy' or 'light' citizen? We have to think about that because the citizen can read everything we write down."

Caseworkers' motivations or motives for not wanting to record and share information outside their community of practice is essential to understanding their practice. Judgement of character cannot - and should not - be summarised in a bullet list, for example. Yet, as we have shown, it is citizens' character traits that caseworkers judge will be key to whether they will succeed in moving away from welfare. This provides both their justification and motive for creating, using, and maintaining the information internally, but, for the reasons provided, they draw a moral boundary along institutional lines between what is right and wrong in terms of recording the information and thus, making it traceable to AI and available for predictive purposes. When asking themselves what information they should make 'public', the caseworkers entertain the thought of what *others* may think of the information when observing it from a distance. For example, *others* may perceive the information differently and only see a person as 'mentally ill'.

To the caseworkers, the meaning of classifications depends on the context. As different people might interpret them differently, they might assimilate an idea different from what was intended and expressed. When caseworkers' classifications get tied down to moments of action, they fulfil particular purposes. Perceptions from afar, from the view of citizens or management consultants, or even system developers, may be quite contrary. From afar, other people can look at these classifications and complain that they should not be in some way. However, they do not know how they are being used in practice. To see the credibility the uses might have, they need to be understood in action and with their particular purpose in mind - at the time and place they are being used. The way classifications are selected, used, combined and configured by caseworkers is oriented to the topic at hand. Their use in any particular situation is purposeful or practical for that topic, rather than abstract references and predictive equations.

Mixed with concerns of understanding the context from afar, the caseworkers also expressed concern about turning classifications into standards by writing them down. They think that classifications, including match groups, tend to have a greater significance for citizens the longer they are used to describe them, based on the assumption that you ought to know more about the citizen. Therefore, citizens do not get informal classifications attached to them 'in a formal sense' and for good reasons:

> "In terms of match groups, I would be, the way I think, I wouldn't like it if we, like, had some subcategories. Let's say; we have someone who's totally 'smart'. And then we have someone who never does what he or she is supposed to do. It quickly becomes a label, I think, because then citizens may have a bad label attached to them for some time, and then change. Because we see this many times, that they change. The behaviour they used to have, it slowly changes, and what you could be afraid of, is that they have that label stuck with them for *too* long [emphasis added]."

The temporality and the idiosyncrasy of the judgement are at stake if written down. Some things the caseworkers allow themselves to think and talk about, but the virtues are not available outside a tightly woven social context. The adjectives become 'dangerous' when formalised as they involve reasoning that carries certain weights for other people as much as for the caseworkers. Once categories become standards, they are given an 'inertia', such that changing them or ignoring them may be difficult and costly, and this type of standardisation enhances the risks. Classifications, such as character judgments, are flexible in nature, and caseworkers are cautious when generalising from one instance to make claims about other places and times. This is exemplified by a caseworker who once handled a case with a citizen that a colleague had previously described as 'aggressive'. But, as she explained to us, she found him to be nothing of that sort and was able to judge him differently. Keeping information 'confidential' thus allows the caseworkers to share, use, reuse, and change sensitive information without the fear that it will end up in the hands of outsiders or as a formal representation.

The caseworkers' fear also pertains directly to their scepticism of introducing AI for predictive purposes. They know AI is supposed to support them in their decision-making activities. They also know the system does not have all the adequate information on how they make decisions. Part of this has to do with information not currently 'fitting' into the system since, as previously mentioned, there are only two match groups to choose from. 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.

The caseworkers know that if they write down their classifications, they might be used by others, in different situations. This also contributes to fear-associated feelings of how that information may be interpreted and used, if known by others. What if citizens get labels stuck on them for 'too' long? The effects of this formalisation are something they are very aware of and cautious about. Therefore, it is not surprising that they also shared their reservations with us about the benefits AI will offer, given that it is implemented for predictive purposes. These reservations are imperative as it is not about whether AI tools can 'do' predictions. It becomes a question about whether the information made available to such tools would be sufficiently comprehensive. Caseworkers' practices make it not so.

6 DISCUSSION

The findings presented in this paper provide implications that are not only relevant to practice studies in CSCW but to all those interested in implementing and evaluating AI-type systems in complex and sensitive decision-making contexts. Our results show that members of work settings make moral judgements about the scope of evidence available to AI. This scope involves data that AI could process, but these members do not find them suited to be 'data-rised' to coin a contemporary phrase. They are ill-suited not just for AI but, as our results show; for any bureaucratic recording system. In this discussion, we relate our findings to previous research and reflect on three implications for AI that emerged as a result. These relate to issues of visibility and predictability and suggestions for future design practices, including the role of AI.

6.1 People in shades of grey

We begin our discussion by considering the nature of classifications and the different shapes they may take. Categories, such as match groups, are often imposed by outside forces but as our findings show, it is through dialogue and interactional work that the task of classifying citizens gets done. Match groups, while defined in legal regulations, are implemented by the caseworkers, and as it turns out, it is in this 'interface zone' that citizens are ultimately defined.

In line with previous research on discretion [11, 66, 89], we found that caseworkers need to interpret and modify formal rules before making decisions about citizens and that they can do so, because of the vagueness of match groups. Following Garfinkel (Suchman [85]), it is their lack of concrete detail that practically makes them useable in this regard. But we must also look to the match groups for important information on how classification 'work'. Caseworkers refer to match groups as revealing nothing about the people within them. Still, they use them in their judgement of people's character. We know that categories are difficult to 'unthink' once they are created, but in a job centre context, they are further backed by government forces. They may, therefore, create particular strong incentives for accommodation [82]. As Douglas rhetorically asks, *"How can we possibly think of ourselves in society except by using the classifications established in our institutions?"* [30]. Although recognising 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'.

If classifications like 'heavy' and 'mentally ill' become visible to others, previous research finds that it works to reinforce and produce heavy and mentally ill people [42]. The caseworkers' problem is that they know that their descriptions of people are not stable [28, 29]. To know the meaning of classifications is to see to the actual use. When asking themselves what information they should make 'public', the caseworkers ruminate on how other people might interpret things differently and assimilate ideas different from what was intended in a moment of time and use. For example, would a citizen who at some point was labelled 'lazy', automatically be perceived as suspicious or blameworthy if read by an algorithm? When classifying people, the caseworkers know that they are also producing a 'type' of person, such as a 'mentally ill' person. These typifications of people are created, used, and reused, in combination, but people can and *do* change. Keeping information 'confidential' allows the caseworkers not only to use but also change their classifications.

Classifications are there for practical purposes. They are needed to make sense of citizens and are used to legitimise organisational action, but they can become 'dangerous' when used in the wrong context. They may not fit into the stable categories assumed by bureaucracies – and it is these stable ones that are typically used in AI. This finding is often ignored in research on AI, which takes stables explanation as given, and it goes beyond the standard critique of technologies as merely 'incomplete' ordering devices.

6.2 'Bad' predictions for 'good' moral reasons

Crucial to our findings are the caseworkers' reasoning for not providing or disclosing information. We discovered that these are rooted in their moral objections. In many respects, this discovery resonates with Garfinkel's classic work [38] on the complex relationship between organisational action and organisational records. In our case, it shows that information used by caseworkers to classify citizens combines both standards and conscious judgements of people's character and the latter are not entered into the bureaucratic record since they, for the reasons provided, are unsuited to that form. The records are, as Garfinkel would say, 'bad' for 'good' organisational reasons. Our findings are not simply echoing Garfinkel's though; these findings point to contemporary issues. Our research shows that these professionals, for their own 'moral reasons', choose how AI should function as part of their decision-making activities. They are the ones who are deciding where AI should not assist. This is not because AI cannot make decisions and predictions, as previous research often suggests. It is that the information that such tools would have as a resource may be inadequate – and for 'good reasons'. 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.

We know from previous research that a dataset cannot encompass the full complexity of the individual it represents [5]. Our findings show that the caseworkers are indeed aware of this. 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. The caseworkers fear the role that AI might play in the future, and because of this, they withhold information. 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. They are sceptical of AI, but also of prior judgements made by themselves and their colleagues. One caseworker exemplified this with the 'aggressive' case, where she was able to reevaluate and judge differently.

For caseworkers, their intermediate judgements should never be seen as either prediction of someone's future or inherent to their character – which is just a roundabout way of predicting futures, if character never changes. Caseworkers believe people can and will change – and changing people's lives is at the very core of their job. Everything from the job centre's official laws, standard procedures and new AI initiative seem to be premised on the idea that caseworkers can and should be predicting people's future employability, except the caseworkers, who take this to be a very contingent task.

6.3 Data visibility and the role of AI

Research in CSCW has long pointed out the limits of technologies in making certain information visible [e.g. 13, 18, 19, 47, 61, 80, 81, 84]. However, if AI can handle complexity and combine criteria in ways that were inconceivable only years ago, how do issues of data and 'visibility' relate to the technologies of today? As was mentioned in the related work section, ML tools can now extract patterns from a vast amount of data, including data with no pre-existing labels. These and other advances change how we might perceive previous issues of, for instance, 'residual' categories. While these might be nowhere else classified [18], their meaning may still be derived from other data patterns and used for predictive purposes. In the context of our research one could argue that, as long as AI developers can detect 'good enough' labels of who is 'really' a 2.2 versus a 2.3, they do not need the intermediate reasoning processes involved in caseworkers' decision-making. Even if caseworkers choose not to write down their classifications, there could be other information from the databases, that can reveal something about 'light' or 'heavy' 2.2's and 2.3's. For example, our findings show that a person may be considered 'heavy' if they have been off work for many years. We might view this attribute as an objective kind and something that an AI system could extract from other datasets. Yet, our study shows that the decision of who is 'really' a 2.2 or 2.3 is not stable. We do not know if all long-term unemployed are considered heavy, or how or in what way this is deemed relevant by the caseworkers, or citizens, across cases. It is not the kind of objective, attribute-like quality that may be more appropriate to formally represent someone's 'real' age [18] or bodyweight [61].

Related concerns addressed in this paper are with the permanence of classifications. Arguably, there is a risk that caseworkers may unwittingly become reliant on specific classifications if they become part of the training data used for decision-support [7]. In any case, it may be easier to follow the recommendation made by the system than to go against it. Eubanks noted that even when humans are 'in the loop' regarding AI decision-making, they still tend to defer to machine-powered recommendations [33]. Besides, our findings show that caseworkers obtain sensory impressions in their interactions with citizens, like *"he smells of alcohol"*. As also noted by previous research [27], this poses a challenge for any AI type system as they cannot obtain these data unless they are recorded. As a general point, we must consider the difference between the possibilities of making certain aspects of work visible and the desirability of doing so [12]. In any case, information is harder to capture 'accurately' in the wild, leaving questions about what should be measured in the first place [64]. There are no guarantees that informative features to machines will produce explanations that are useful to humans. Along similar lines, Bucher [22] asks *"if the data is not reflecting the world, how can it predict what will happen?"* If we believe the caseworkers, it would be unwise to predict people's future and employment, even if AI is never introduced. This uncovering match those observed in earlier studies where caseworkers wanted to turn predictions of citizens towards inefficiencies within their organisation or change the focus from negative outcomes towards positive ones.

Our findings suggest yet another solution, one that takes focus away from prediction. If we do not think of the goal as predicting the future and replacing caseworker judgements, it opens up a space for alternative possibilities. Predictions are not obvious, or necessary. AI may be used for other purposes and support caseworkers in mundane, but crucial tasks of retrieving, organising and consolidating case records. This could help ensure that cases are processed with the right information and at the right time, which was also an issue raised by the caseworkers in our analysis. At the very least, it should be possible to resist the desire to put prediction instruments into practice, particularly in sensitive contexts such as welfare allocation. As noted by previous work, this is not an easy task but rather a mutual learning process, where caseworkers are invited to influence design decisions [62]. Explaining what professionals do is a necessary step in this direction. It relieves us from the idea that citizens, in contexts such as these, can be judged along 'commensurate' metrics calculated via machine learning [64]. 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.

6.4 Directions for future research

There is still ample room for improvement in determining AI's role in sensitive contexts such as welfare allocation. We hope that we have contributed to a better understanding of street-level bureaucrats' practices and values in this context. Further studies on classification, which takes the citizen's perspective, may also be undertaken. This is particularly relevant in a job centre context, where welfare seekers mainly get defined by others. From previous research and our experience in the job centre, we know that caseworkers' phrases to describe citizens might not always be the same as citizens would use to describe themselves. We also know that it may lead to new problems if the role of self-knowledge is ignored in the process. While this is beyond this research's scope, we believe this is an equally important issue to investigate as part of future research.

7 CONCLUSIONS

Based on ethnographic field studies of classification work and its implications for AI in a Danish job centre, this paper shows how caseworkers categorise citizens by adding informal classifications to the binary vocabulary offered by their municipal job centre. The classifications are made on negotiated and situated judgments of people's character that allow caseworkers to operate within a formal framework - yet with sufficient looseness to properly process each individual case and to do so in ways that other caseworkers understand and can act on. While these judgements have an intersubjective character, caseworkers choose not to write them down. Their reasoning is rooted in moral concerns about perceptions of information once it becomes documented and thus viewable away from its context of use. According to the caseworkers, the meaning of classifications depends on the context of use and documenting can thereby lose the real-world character of actual decision-making. In continuation of this, they express deep concern about how information might be interpreted and perpetuated when stored by algorithms and 'stick' with people for too long. For these reasons, while documentable and traceable to AI, the official records are left without a fundamental understanding of how decisions are actually made. This study contributes to CSCW and HCI research in sensitive contexts, such as welfare allocation, by offering empirical evidence for this and by showing that the problem of implementing AI might not only be to do with the technology itself, as previous research suggests, but also human questions about what data is (and should be) made available for AI.

ACKNOWLEDGMENTS

We would like to express our sincere gratitude to the caseworkers, who generously shared their time and experience for this study. Earlier versions of this paper were brought to the CHI 2019 workshop *"Where is the Human?: Bridging the Gap Between AI and HCI"* and the CSCW 2020 workshop *"Better Supporting Workers in ML Workplaces"*. We thank the workshop organisers and participants for the fruitful discussions that contributed to improving this manuscript. We also thank Dave Randall, Peter Garraghan and our colleagues in the EcoKnow.org research project, Naja Holten Møller and Asbjørn Ammitzbøll Flügge, for their encouraging feedback and suggestions for further improvements. Finally, we thank the anonymous reviewers for their helpful insights and valuable input in revising the final manuscript. This work has been supported by a grant from the Innovation Fund Denmark (7050-00034A).

REFERENCES

[1] Alkhatib, A. and M. Bernstein. 2019. Street–Level Algorithms: A Theory at the Gaps Between Policy and Decision. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing* Systems, 1-13.
[2] Andersen, K. B. 2018. Regeringen vil erklære 20.700 danskere på kontanthjælp klar til arbejde. from https://nyheder.tv2.dk/politik/2018-05-29-regeringen-vil-erklaere-20700-danskere-paa-kontanthjaelp-klar-til-arbejde
[3] Bannister, F. and R. Connolly. 2014. ICT, public values and transformative government: A framework and programme for research. Government Information Quarterly, 31(1), 119-128.
[4] Bannon, L. 2011. Reimagining HCI: Toward a More Human-Centered Perspective. Interactions, 50-57.
[5] Barocas, S. and A. D. Selbst. 2016. Big Data's Disparate Impact. California Law Review, Inc., 671-732.
[6] Beich, A. 2019. Jobcentre gør mennesker syge – politikerne må handle nu. from https://jyllands-posten.dk/debat/kronik/ECE11390084/jobcentre-goer-mennesker-syge-politikerne-maa-handle-nu/
[7] Benjamin, R. 2019. Race After Technology: Abolitionist Tools for the New Jim Code. Polity Press.
[8] Berg, M. 1998. The Politics of Technology: On Bringing Social Theory into Technological Design. Science, Technology & Human Values, 23, 456-490.
[9] Bergstrom, C. T. and J. D. West. 2020. Calling Bullshit: The Art of Scepticism in a Data-Driven World: Random House
[10] Beskæftigelsesministeriet. 2016. Bekendtgørelse af lov om en aktiv beskæftigelsesindsats. from https://www.retsinformation.dk/forms/r0710.aspx?id=184891
[11] Black, J. 2001. Managing Discretion. ARLC Conference Papers on Penalties: Policy, Principles and Practice in Government Regulation
[12] Blomberg, J. and H. Karasti. 2013. Reflections on 25 years of ethnography in CSCW. Computer Supported Cooperative Work (CSCW), 22(4-6), 373-423. https://doi.org/10.1007/s10606-012-9183-1
[13] Boulus-Rødje, N. 2018. In Search for the Perfect Pathway: Supporting Knowledge Work of Welfare Workers. Computer Supported Cooperative Work (CSCW), 27, 841-874. https://doi.org/10.1007/s10606-018-9318-0
[14] Bourdieu, P. 1999. Site Effects. In P. Bourdieu, A. Accardo, G. Balazs, S. Beaud and F. Bonvin (Eds.), The Weight of the World (Vol. 2, pp. 123-180). Polity Press.
[15] Bovens, M. and S. Zouridis. 2002. From Street-Level to System-Level Bureaucracies: How Information and Communication Technology is Transforming Administrative Discretion and Constitutional Control. Public Administration Review, 62(2), 174-184.
[16] Bowers, J., G. Button, and W. Sharrock. 1995. Workflow From Within and Without: Technology and Cooperative Work on the Print Industry Shopfloor. In *Proceedings of the European Conference on Computer-Supported Cooperative Work*, 51-66.
[17] Bowker, G. C. 2005. Memory Practices in the Sciences. MIT Press.
[18] Bowker, G. C. and S. L. Star. 1999. Sorting Things Out: Classification and Its Consequences. MIT Press.
[19] Bowker, G. C., S. L. Star, W. Turner, and L. Gasser. 2014. Social Science, Technical Systems, And Cooperative Work: Beyond the Great Divide (2 ed.). Psychology Press.
[20] boyd, d.and K. Crawford. 2012. CRITICAL QUESTIONS FOR BIG DATA. Information, Communication & Society, 15(5), 662-679.
[21] Brown, A., A. Chouldechova, E. Putnam-Hornstein, A. Tobin, and R. Vaithianathan. 2019. Toward Algorithmic Accountability in Public Services: a Qualitative Study of Affected Community Perspectives on Algorithmic Decision-making in Child Welfare Services. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 1-12.
[22] Bucher, T. 2018. If…then: Algorithmic Power and Politics (Vol. 1). Oxford University Press.
[23] Campolo, A.and K. Crawford. 2020. Enchanted Determinism: Power without Responsibility in Artificial Intelligence. Engaging Science, Technology, and Society, 6, 1-19.
[24] Chiusi, F., S. Fischer, N. Kayser-Bril, and M. Spielkamp. 2020. Automating Society Report 2020. AlgorithmWatch.
[25] Crabtree, A., D. M. Nichols, J. O'Brien, M. Rouncefield, and M. B. Twidale. 2000. Ethnomethodologically Informed Ethnography and Information System Design. Journal of the American Society for Information Science, 51(7), 666-682.
[26] Danish Agency for Labour Market and Recruitment. 2018. Active labour market policy measures. from https://www.star.dk/en/active-labour-market-policy-measures/
[27] Dolata, M., B. Schenk, J. Fuhrer, A. Marti, and G. Schwabe. 2020. When the System does not Fit: Coping Strategies of Employment Consultants. Computer Supported Cooperative Work (CSCW), 1-40.
[28] Douglas, J. D. 1967. Social Meanings of Suicide (Vol. 1). Princeton University Press.
[29] Douglas, M. 1966. Purity and Danger: An analysis of the concepts of pollution and taboo (Vol. 1). Routledge
[30] Douglas, M. 1986. How Institutions Think (1 ed.). Syracuse University Press.

[31] Durkheim, É. and M. Mauss. 1903. De Quelques Formes Primitives de Classification. Contribution à l'étude des Représentations Collectives. L'Année Sociologique, 6, 1-72.

[32] Elish, M. C. 2018. The Stakes of Uncertainty: Developing and Integrating Machine Learning in Clinical Care. In *Ethnographic Praxis in Industry Conference Proceedings*, 364-380.

[33] Eubanks, V. 2017. Automating Inequality: How High-tech Tools Profile, Police, and Punish the Poor. St. Martin's Press.

[34] Fitzgerald, R. and W. Housley. 2015. Advances in Membership Categorisation Analysis. Sage Publications Ltd.

[35] Flügge, A. A., T. Hildebrandt, and N. H. Møller. 2021. Street-Level Algorithms and AI in Bureaucratic Decision-Making: A Caseworker Perspective. in *Proceedings of the ACM on Human-Computer Interaction*, 5(CSCW1), Article 40, 22. https://doi.org/10.1145/3449114

[36] Flyverbom, M. 2019. The Digital Prism: Transparancy and Manageed Visibilities in a Datafied World. Cambridge University Press.

[37] Friedman, B., P. H. Kahn, and A. Borning. 2008. Value sensitive design and information systems. The handbook of information and computer ethics, 69-101.

[38] Garfinkel, H. 1967. Good organizational reasons for 'bad' clinic records. In H. Garfinkel (Ed.), Studies on Ethnomethodology (pp. 186-207). Prentice-Hall, Inc.

[39] Garfinkel, H. 1967. Studies in Etnomethodology (Vol. 1). Blackwell Publishers Inc.

[40] Garsten, C. and K. Jacobsson. 2013. Sorting people in and out: The plasticity of the categories of employability, work capacity and disability as technologies of government. ephemera: theory & politics in organization, 13(4), 825-850.

[41] Gillespie, T. 2014. The Relevance of Algorithms. Media technologies: Essays on communication, materiality, and society, 1-32.

[42] Hacking, I. 1985. Making Up People. In T. L. Heller, M. Sosna andD. E. Wellbery (Eds.), Reconstructing Individualism. Stanford University Press.

[43] Hammersley, M. and P. Atkinson. 2007. Ethnography: Principles in practice (Vol. 3). Taylor & Francis: Routledge.

[44] Harper, R. 2019. The Role of HCI in the Age of AI. International Journal of Human–Computer Interaction, 35(15), 1331-1344.

[45] Harper, R., D. Randall, and W. Sharrock. 2016. Choice: The Sciences of Reason in the 21st Century: A Critical Assessment. Polity Press.

[46] Henriksen, L. R. 2019. Jobcentret gør sygemeldte endnu mere syge. from https://www.fyens.dk/danmark/Jobcentret-goer-sygemeldte-endnu-mere-syge/artikel/3359664

[47] Høybye-Mortensen, M. and P. Ejbye-Ernst. 2018. The long way to data-driven decision-making: How do casework registrations become management information? STS Encounters, 10(2.2), 5-36.

[48] Hughes, J. A., D. Randall, and D. Shapiro. 1993. From Ethnographic Record to System Design: Some experiences from the field. Computer Supported Cooperative Work (CSCW), 123-141.

[49] Keddell, E. 2011. Reasoning Processes in Child Protection Decision Making: Negotiating Moral Minefields and Risky Relationships. The British Journal of Social Work, 41(7), 1251-1270.

[50] Keyes, O. 2018. The Misgendering Machines: Trans/HCI Implications of Automatic Gender Recognition. Proceedings of the ACM on Human-Computer Interaction, 88(CSCW, Article 88), 1-22.

[51] Keymolen, E. and D. Broeders. 2011. Innocence Lost: Care and Control in Dutch Digital Youth Care. British Journal of Social Work, 43(1), 41-63.

[52] Lampland, M. and S. L. Star. 2009. Standards and their stories: how quantifying, classifying, and formalizing practices shape everyday life. Cornell University Press.

[53] Law, J. and J. Urry. 2013. Enacting the social. Economy and Society, 33(3), 390-410.

[54] Lipsky, M. 1980. Street-Level Bureaucracy: The Dilemmas of the Individual in Public Service. Russell Sage Foundation.

[55] Loxha, A. and M. Morgandi. 2014. Profiling the Unemployed: A Review of OECD Experiences and Implications for Emerging Economies. Social Protection & Labor, 1424, 1-45.

[56] Martin, D., J. O'Neill, D. Randall, and M. Rouncefield. 2007. How Can I Help You? Call Centres, Classification Work and Coordination. Computer Supported Cooperative Work (CSCW), 16, 231-264.

[57] Matthiesen, S. and P. Bjørn. 2017. When Distribution of Tasks and Skills Are Fundamentally Problematic: A Failure Story from Global Software Outsourcing. PACM on Human-Computer Interaction, 1(CSCW), 74-90.

[58] Maynard, D. W. 1984. Inside Plea Bargaining: The Language of Negotiation. Plenum Press.

[59] Maynard-Moody, S. and M. Musheno. 2003. Cops, Teachers, Counselors: Stories from the Front Lines of Public Service (Vol. 1). The University of Michigan Press.

[60] Ministry of Finance. 2019. National Strategy for Artificial Intelligence: The Danish Government.

[61] Møller, N. H. and P. Bjørn. 2011. Layers in Sorting Practices: Sorting out Patiens with Potential Cancer. Computer Supported Cooperative Work (CSCW), 20, 123-153. https://doi.org/10.1007/s10606-011-9133-3

[62] Møller, N. H., I. Shklovski, and T. T. Hildebrandt. 2020. Shifting Concepts of Value: Designing Algorithmic Decision-Support Systems for Public Services. In *Proceedings of the 11th Nordic Conference on Human-Computer Interaction: Shaping Experiences, Shaping Society*, 1-12. https://doi.org/10.1145/3419249.3420149

[63] Paanakker, H. L. 2020. Perceptions of the Frontline Craft: Assessing Value Convergence Between Policy Makers, Managers, and Street-Level Professionals in the Prison Sector. Administration & Society, https://doi.org/10.1177/0095399720933815
[64] Pasquale, F. 2020. New Laws of Robotics: Defending Human Expertise in the Age of AI. Harvard University Press.
[65] Persson, J., A. Reinwald, E. Skorve, and P. Nielsen. 2017. Value positions in e-government strategies: Something is (not) changing in the state of Denmark.
[66] Petersen, A. C. M., L. R. Christensen, and T. T. Hildebrandt. 2020. The Role of Discretion in the Age of Automation. Computer Supported Cooperative Work (CSCW), 29(3), 303–333. https://doi.org/10.1007/s10606-020-09371-3
[67] QSR International. 2019. NVivo 12. from https://www.qsrinternational.com/nvivo/nvivo-products/nvivo-12-mac
[68] Randall, D. 1995. A Comment on Lucy Suchman's "do Categories Have Politics?". Computer Supported Cooperative Work (CSCW), 3, 47-50.
[69] Randall, D., R. Harper, and M. Rouncefield. 2007. Fieldwork for Design: Theory and Practice. Springer.
[70] Ranerup, A. and H. Z. Henriksen. 2019. Value positions viewed through the lens of automated decision-making: The case of social services. Government Information Quarterly, 36(4), 101377.
[71] Rawls, A. W. 2010. Social Order as Moral Order. In S. Hitlin andS. Vaisey (Eds.), Handbook of the Sociology of Morality (pp. 95-122). New York: Springer Science+Business Media.
[72] Rieder, G. 2018. Big Data: Or, the Vision that Would Not Fade. (PhD Dissertation), IT University of Copenhagen, Copenhagen, Denmark.
[73] Rouncefield, M. 2011. Fieldwork, Ethnography and Ethnomethodology LSCITS Socio-Technical Systems Engineering Handbook (pp. 44-48). http://archive.cs.st-andrews.ac.uk/STSE-Handbook/: University of St Andrew.
[74] Sacks, H. 1992. Lectures on Conversation (Vol. I & II). Blackwell Publishing.
[75] Schmidt, K.and I. Wagner. 2003. Ordering Systems: Coordinative Practices in Architectural Design and Planning. In *Proceedings of the 2003 international ACM SIGGROUP conference on Supporting group work*, 274-283.
[76] Schultz. 2018. MØD ASTA – DIN NYE ASSISTENT: Interview med Simon Kohrtz, produktchef, Schultz. from https://schultz.dk/om-schultz/nyt-fra-schultz/moed-asta-din-nye-assistent/
[77] Sendak, M., M. C. Elish, M. Gao, J. Futoma, W. Ratliff, M. Nichols, et al. 2020. "The Human Body is a Black Box": Supporting Clinical Decision-Making with Deep Learning. In *Proceedings of the 2020 conference on fairness, accountability and transparency*, 99-109.
[78] Skolnick, J. H. 1966. Justice Without Trial: Law Enforcement in Democratic Society. New York: John Wiley & Sons.
[79] Smith, R. J. and P. Atkinson. 2015. Method and Measurement in Sociology, fifty years on. International Journal of Social Research Methodology, 19(1), 99-110.
[80] Star, S. L. and G. C. Bowker. 2007. Enacting silence: Residual categories as a challenge for ethics, information systems, and communication. Ethics and Information Technology, 9, 273-280.
[81] Star, S. L. and A. Strauss. 1999. Layers of silence, arenas of voice: The ecology of visible and invisible work. Computer Supported Cooperative Work (CSCW), 8(1-2), 9-30.
[82] Steensland, B. 2010. Moral Classification and Social Policy. In S. Hitlin and S. Vaisey (Eds.), Handbook of the Sociology of Morality (pp. 455-467). Springer.
[83] Suchman, L. 1993. Do Categories Have Politics? The language/action perspective reconsidered. In Proceedings of the Third European Conference on Computer-Supported Cooperative Work, Milan, Italy. ECSCW'93.
[84] Suchman, L. 1995. Making Work Visible. Communication of the ACM, 38(9), 56-64.
[85] Suchman, L. A. 1987. Plans and situated actions: The problem of human-machine communication. Cambridge University Press.
[86] Suchman, L. A. 2007. Human-machine reconfigurations: Plans and situated actions. Cambridge University Press.
[87] Timmermans, S. and M. Berg. 1997. Standardization in action: Achieving local universality through medical protocols. Social Studies of Science, 27, 273-305.
[88] Verne, G. 2015. "The winners are those who have used the old paper form" On citizens and automated public services. (PhD Dissertation), University of Oslo, Oslo, Norway. (No. 1650)
[89] Villesen, K. 2010. Socialrådgivere modarbejder beskæftigelsespolitikken. from https://www.information.dk/indland/2010/01/socialraadgivere-modarbejder-beskaeftigelsespolitikken
[90] Voida, A., L. Dombrowski, G. R. Hayes, and M. Mazmanian. 2014. Shared values/conflicting logics: working around e-government systems. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 3583-3592.