Automated Decision-Making
Towards a People-Centred Approach
Lomborg, Stine; Kaun, Anne; Scott Hansen, Sne

Published in:
Sociology Compass

DOI:
10.1111/soc4.13097

Publication date:
2023

Document version
Publisher’s PDF, also known as Version of record

Document license:
CC BY-NC-ND

Citation for published version (APA):
Automated decision-making: Toward a people-centred approach

Stine Lomborg | Anne Kaun | Sne Scott Hansen

Abstract
Recently, automated decision-making (ADM) has been increasingly introduced in for example, the public sector potentially ensuring efficiency and more just decision-making. The increasing use of ADM has been reflected by a growing interest by scholarly research. While initially mainly researchers within law and computer sciences engaged with ADM, there has also been a growing engagement by social science and humanities-oriented researchers. This article traces the emergence and evolution of ADM research beyond computer sciences and engineering with a specific focus on social sciences and humanities by identifying central concerns and methods while outlining a stable baseline for future research. Based on a systematic mapping of publications, we outline the contours of ADM as an area of research engaging with an emerging empirical phenomenon. Drawing on findings from the mapping, we discuss ways ahead for ADM research as part of the subfield of digital sociology and suggest that sociological media and communication studies have a crucial role to play in developing future research avenues. Drawing on advances made in audience research, we suggest a radically contextualized and people-centered approach to ADM. Such an approach would help to develop ADM and ground it alongside people's divergent capabilities and contextual arrangements.
1 | INTRODUCTION

Automated decision-making (ADM) with the help of algorithms (ADM) is one of the latest technological developments with far reaching social, political, and economic implications. While there is an increased use of ADM in all kinds of societal domains, the legal frameworks for fostering and regulating the use and implementation are still emerging. The public and legal discourse is often centered around transparency, accountability and explainability, which are also flagged as central concerns in the recent European Digital Services Act. The growing interest is also reflected in several attempts to map ADM across domains. This includes the pioneering work of AlgorithmWatch in Berlin and their Atlas of Automation database as well as several reports. Similarly, there are local registries of AI applications that have been set up by for example, the City of Helsinki and the City of Amsterdam with the aim to make algorithm-based automation more transparent. These two initiatives that are coordinated by the Finnish platform Saidot and similar projects highlight the strong interest in ADM.

The growing public interest in ADM is also reflected in research and a growing academic engagement with ADM systems and practices both in its classic strands of computer science and engineering, law and social science and humanities-oriented fields. However, the question remains if we are seeing the emergence and formation of a field focusing primarily on ADM. This article is an attempt to answer this question through a systematic literature review of academic publications engaging with ADM. The aim is to develop an understanding of the growing body of literature on ADM, and through offering a conceptual map to open avenues for ADM research from the ground up.

This article provides a review of publications that engage with ADM across disciplines. It examines how the area has evolved as well as the disciplinary cross-references and boundaries. By providing a basic mapping of existing research, the aim is to help catapult research on ADM into more interdisciplinary conversations that may enrich public understanding of the prospects and pitfalls of implementing ADM systems in public and private organizations. Furthermore, we highlight the relevance of ADM to the field of digital sociology. Specifically, we underline a need to engage empirically with ADM from the perspective of the people it implicates. The current landscape of ADM research is dominated by high-level guidelines as well as studies of systems that foreground important issues of transparency and data ethics. What is less prominent is the study of ADM systems from the perspective of people as users, producers, and maintainers of these technologies, even if some research does exist (for example Helberger et al., 2020; Ranerup & Henriksen, 2022). We ask for more empirical research that can support the aim to build knowledge on ADM from the ground up; namely as situated and in need of radical contextualization. In addition, we argue that there is a need for theory development in ADM research. Although most papers present themselves as conceptual, the notion of ADM itself remains vague and invites more in-depth conceptual contributions to understand what it is and what it does. We outline one such conceptual intervention derived from sociological media and communication research to suggest that conceptual work can help integrate discussions across disciplines engaging with ADM.

2 | DEFINING ADM

Before we delve into the systematic review of ADM literature, we need to answer the question what do we talk about when we talk about ADM. There is still conceptual confusion regarding what is included in the understanding of ADM, although the term has been suggested as being more specific than for example, artificial intelligence. Thouvenin
et al. (2019) even suggest the term is an oxymoron, because deciding something requires the kind of leeway and contextual judgment that automation crowds out of the process. There are currently different ways of defining or otherwise specifying ADM circulating in public debate and in scholarly articles. For example, the algo:aware report published in 2018 procured by the European Commission’s Directorate for Communications Networks, Content and Technology defines “automated decision-making as a software system—including its testing, training and input data, as well as associated governance processes—that autonomously or with human involvement, takes decisions or applies measures relating to social or physical systems on the basis of personal or non-personal data, with impacts either at the individual or collective level” (p. 5, for a critical discussion of the notion of decision-making in the context of automation see Thouvenin et al., 2019). This definition includes fully automated systems as well as decision-support systems that combine automated aspects with the involvement of human decision-making. Similarly, the advocacy organization AlgorithmWatch defines ADM in its first report as algorithmically controlled, automated decision-making or decision support systems are procedures in which decisions are initially–partially or completely–delegated to another person or corporate entity, who then in turn use automatically executed decision-making models to perform an action. This delegation—not of the decision itself, but of the execution—to a data-driven, algorithmically controlled system, is what needs our attention. (AlgorithmWatch, 2019, p. 9).

The explicit aim of AlgorithmWatch is to introduce a more specific term than artificial intelligence that has become quite fuzzy and a buzzword itself in recent years. At the same time, they aim to include systems that are less complex than machine learning or neural networks (such as scoring systems for risk assessment or other simple rule-based procedures), but nonetheless have far reaching implications for citizens being exposed to them. Accordingly, ADM concerns the process of implementing and delegating tasks to digital, data-based systems and includes both rule- and knowledge-based algorithms. Following this definition, several of the examples of ADM that have been mapped in reports and databases mentioned above include ADM such as comparably simple decision trees. In Sweden there is the example of the so-called Trelleborg model that includes fully automated decisions on recurring social benefit applications since 2017 (Kaun, 2021). Furthermore, there are sorting procedures, for example, a sorting algorithm developed by the Austrian Employment Services that automatically categorizes job seekers into three categories based on a set of criteria (Allhutter et al., 2020; Kayser-Bril, 2019). Third, there are matching algorithms as well as predictive algorithms including predictive risk scoring estimating the likelihood of child abuse developed, but never implemented in Denmark (Alfter, 2019).

The definition and examples provided above highlight the importance of people involved in automation processes especially those affected by automated decisions. Therefore, it is urgent to develop systematic research approaches that take these perspectives seriously and re-center the people in ADM. In the review that follows, we engage with existing literature and their definitions and approaches to ADM to identify important gaps including the continued need to develop such people-centered avenues toward ADM.

3 | APPROACH AND METHODOLOGY

To map existing research on ADM, we queried the Scopus and Web of Science databases for articles that included in their abstracts one or several of the following terms: ADM, algorithmic decision-support, or algorithmic decision-making. We delimit the search to these general keywords derived from AlgorithmWatch’s practical definition above to make the search as inclusive as possible but still specific to the topic of ADM. These search terms are also motivated by the focus on ADM rather than the broader field of automation or artificial intelligence following the aim to engage with the question of how the emerging topic of ADM is reflected in research. By querying the two databases and combining the search results, we sought to mitigate the risk that not all relevant publication outlets are
indexed in the specific database at hand. We did test additional keywords for search, such as algorithmic governance and anticipatory governance, but found that the few additional publications this yielded point in a different direction, namely the by now increasingly consolidated field of (critical) algorithm studies that is adjacent to ADM research.

Our search was performed on 21 February 2022, identifying 1509 texts in Scopus and 1005 in the Web of Science. The identified texts were merged into one database, and duplicates were removed using Zotero. We removed articles with incomplete data, for example, missing abstracts, journal, or metadata such as keywords to make the database coherent. The final database included a total of 432 texts, which is the empirical dataset that forms the basis of this paper.

The dataset was analyzed using content analysis in combination with citation network analysis to describe the overall patterns and trajectories of ADM research. A coding manual was prepared, which was tested and revised in Excel through two pilot coding rounds, where all three authors coded the same randomly chosen 20 texts for each round of pilot coding. Based on the revised coding manual, the 432 texts were coded manually in Excel. Each coding, both in the pilot tests and in the final coding, was based on a reading of the articles’ abstracts. Basing the coding on abstracts only is a common practice in systematic literature reviews and is motivated by the formal character of abstracts highlighting central terminology, methodology and findings in a compressed manner. Where relevant, such as in the case of definitional articles, full articles were read.

The coding manual included the following categories: the research field in question, the type of article (theoretical, empirical, methodological, or other), and for empirical articles which methods were used. We also initially coded for the sector (e.g. public, private or NGO) and if possible context of analysis (e.g. health care, policing), unit of analysis of ADM (e.g. system, designers, workers, citizens), but since—as will be evident below—the main portion of publications address ADM at a general level, it was difficult to glean and validate information from abstracts for these more fine-grained categories. This would have required a close reading of all 432 articles. We thus decided not to proceed with them at the analysis stage. Finally, we coded for the existence of a definition of ADM in the abstract.

The dataset was furthermore analyzed with the help of VOSviewer (Van Eck & Waltman, 2010), which is a network analysis tool for sorting scientific publications according to a range of parameters, including co-citation links, keywords, and journals. We used VOSviewer to identify references shared across the articles and to find clusters of citations. Additionally, we conducted an analysis of the self-assigned article keywords to identify any possible central concepts and terms shared within or across fields and disciplines in an accessible manner. The keyword analysis provides an indication of central concerns to the ADM field and how it is developing.

The analysis, discussion and suggestions for advancing the field of ADM research in the following is informed by the results of our empirical mapping of ADM research and is organized in three sections. The first section reports our empirical observations from mapping the ADM literature: we identify a lack of dialog and integration across disciplines that becomes apparent in the literature. In the second part of the article, we discuss the emergence and consolidation of ADM definitions that are characterizing the field at the moment. Based on this engagement, we propose a conceptual means to bridging disciplines to cross-pollinate ADM research and offer future directions for research grounding ADM studies in empirical research centered around people that are implicated in automated decisions. Accordingly, we propose in conclusion a people-centered, contextual, and empirical research agenda with a view to impact that moves beyond the current focus on general ethical guidelines for ADM technologies and applications.

4 | FINDINGS FROM THE SYSTEMATIC MAPPING OF ADM LITERATURE

4.1 | Disciplinary and field-specific approaches to ADM

We coded each item for disciplinary and field affiliation, approximated by the journal in which the article was published. Most of the articles in our dataset were published within computer science, engineering, or other technical outlets (163 in total), followed by law and legal studies (67), medical science (38), and social science
LOMBORG et al.

journals (38). Humanities (23) and business and marketing studies (22) stand for a comparatively small share of the articles, and 81 articles were coded as interdisciplinary or other. These were published in journals (such as AI & Society) and conference proceedings (ACM AIES conference on Artificial Intelligence, Ethics and Society) that describe themselves as primarily interdisciplinary outlets with focus on digital technologies, their uses and social and political consequences.

When looking at the specific outlets that are most salient for ADM scholarship, the scholarly engagement with ADM appears fragmented in terms of journal outlets. Even with engineering and computer science articles and conference proceedings excluded, the literature is scattered across scientific journals, most of which include only one or two articles about ADM. The journals most often appearing in the data set are AI & Society (17 articles), Computer Law and Security Review (15 articles), Contemporary Readings in Law and Social Justice (8 articles), Information, Communication & Society (8 articles), and Information Polity (8 articles). It seems that there is no clearly designated home to ADM research when it comes to publication avenues.

In terms of chronology, the first articles with references to ADM appeared in the 1970s and 1980s and were mainly published in the field of robotics. Following the evolution of ADM publications over time, it becomes apparent that research of and inspired by ADM exploded since 2018 and stagnates in 2021. As Table 1 below shows there is a steep increase in publications over the years 2015–2019, where the number of articles being published on ADM almost doubles every year. The largest increase can be noted in law between 2018 (8 publications) and 2019 (26 publications), which can be explained with the General Data Protection Regulation (GDPR) entering into force in May 2018. The GDPR includes an explicit reference to ADM. Computer science remains overall dominant in ADM research with a steady increase in published articles over the years and moreover what seems to be a spike from 2018 (21 publications), to 2019 (30 publications), and 2020 (27 publications). This indicates an increasing interest across disciplines in the area, and perhaps particularly so with links to topics such as explainable AI, bias and fairness in machine learning among others, where computer scientists are engaging with themes that are also of great theoretical interest across law and the social sciences more broadly. Finally, we also see an increase of articles in social sciences and humanities covering a wide array of topics from applied analysis of specific sectoral implementations of ADM to theoretical or conceptual articles.

To characterize the orientations of ADM research, we coded the items in the dataset according to the types of articles in question, again excluding computer science and engineering research. Table 2 below presents the findings from this analysis.

Looking closer at the 250 articles coded as social sciences, humanities, law, medical science, business and marketing and interdisciplinary, we can see that a major part of these (117) are theoretical contributions that conceptualize

**Table 1** Development in the number of automated decision-making publications 1974–2021.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of articles</td>
<td>110</td>
<td>12</td>
<td>17</td>
<td>36</td>
<td>56</td>
<td>112</td>
<td>85</td>
<td>4</td>
</tr>
</tbody>
</table>

**Table 2** Article types, excluding publications in computer science and engineering.

<table>
<thead>
<tr>
<th>Article type</th>
<th>Number of articles</th>
</tr>
</thead>
<tbody>
<tr>
<td>Theoretical</td>
<td>117</td>
</tr>
<tr>
<td>Methodological (scientific methods)</td>
<td>29</td>
</tr>
<tr>
<td>Development of applied automated decision-making use case</td>
<td>39</td>
</tr>
<tr>
<td>Empirical</td>
<td>45</td>
</tr>
<tr>
<td>Systematic literature review</td>
<td>7</td>
</tr>
<tr>
<td>Other/unclear</td>
<td>19</td>
</tr>
<tr>
<td>Total</td>
<td>250</td>
</tr>
</tbody>
</table>
ADM, sometimes using examples to illustrate the arguments. The articles labeled as theoretical address, for example, the alleged social and political implications of ADM, or offer conceptual contributions that address specific values such as transparency, fairness, privacy or justice (e.g., Binns, 2018; Borgesius, 2020; Coudert, 2010; Wagner, 2019). Furthermore, many of the theoretical articles are concerned with the many high-level guidelines, on, for example, transparency and data ethics that have been developed over the past couple of years at both organizational, national, and international levels, as well as close-readings and interpretations of the GDPR (e.g. Veale & Edwards, 2018). The theoretical literature appears to have particularly spanned over the past 5 years. When we coded the abstracts for references to specific domains or sectors, for example, policing, health or social services, it became clear that most of the theoretical papers were not specific about their domain of inquiry, but typically offer general discussions of a specific theoretical concept or the consequences of ADM with reference to justice, fairness, transparency, harm, diversity among others, or in the case of law studies, theoretical analyses of legal text, for example, the GDPR.

Another share of articles is dedicated to empirical investigations (45), for example, regarding the implementation of ADM in health care, social services, or other welfare provision in the public sector as well as in policing and criminal justice (e.g. Brown et al., 2019; Dencik et al., 2018). There is little attention paid to specific societal crises, such as the covid-19 pandemic, which has generated some public debate about ADM for instance in relation to contact tracing applications. Given the societal concern caused by the covid-19 crisis, we might have expected more research on this coming out in 2021, although it cannot be ruled out that it will emerge as a more prominent topic in the research pipeline of the coming years. There is also not a lot of what we— with reference to Pink et al. (2022)— might dub as "everyday ADM" research—including research on self-tracking, content moderation systems, chatbots. The dataset includes 29 methodological articles, which deal with big data analytics (e.g., Ibrahim et al., 2020) and data visualization (e.g., Hunt & McKelvey, 2019), as well as 39 articles that describe the development of applied methods for ADM (e.g., Andreeva & Matuszyk, 2019). The remaining articles were either review articles (7), or articles that could not be classified from the abstract (32).

Finally, we analyzed if specific article types are especially common in specific disciplines. Theoretical articles appear particularly prominent in law studies of ADM, where we find almost exclusively theoretical papers, but also in the social sciences, humanities and interdisciplinary research. "Theoretical" means here that the articles engage with questions of how different legal domains are impacted by ADM or how the legal framework needs to be adjusted. In contrast, empirical studies and tests of applied methods for ADM in concrete contexts are predominant in medical science as well as business and marketing studies, where theoretical contributions appear sparse. Finally, tests of applied methods for ADM are also quite common in the interdisciplinary literature. For these studies, naturally, the domain of inquiry is typically clearly indicated, and thus the analysis of ADM is grounded and unpacked within a specific social context.

### 4.2 Common ground in ADM research?

To qualify whether a common ground can be detected across ADM research, we explored our dataset with different visualization tools: a frequency analysis of keywords, and an analysis of citation networks. First, we counted the frequency of the keywords provided for each publication in our data set and used a simple word cloud tool to visualize the relative prominence of keywords (https://www.wordclouds.com). We did a counting of keywords for the entire dataset to look for common patterns across the literature, as well as separate ones for publications in the humanities, law and social sciences to see if there are indications of a specific vocabulary within disciplines (law) or among adjacent fields (humanities, social sciences). For all keyword analyses, keywords directly related to the search terms (automat*, algorithm*, decision-making, support) as well as "big data" and stop words were removed. A word cloud of all the keywords appearing more than 20 times in the dataset is presented in Figure 1, below.

The full dataset holds more than one thousand keywords, the majority of which are just mentioned once or twice. Figure 1 displays a wordcloud of the 126 keywords that appear 20 times or more across the full dataset. Of these, 12
have more than 100 mentions: systems (517), artificial (201), intelligence (187), learning (165), human (152), information (135), machine (131), models (122), computer (112), health (111), clinical (110), and networks (110). These are all very generic top keywords that indicate little about the field beyond the technological topic of study. But a little further down the list are keywords such as care (77), control (77), fairness (58), and privacy (57) that indicate specific and relatively prominent conceptual and topical trajectories in the ADM literature, with “care” typically associated with health care studies, and the others often associated with high-level guidelines for and principal discussions of ethical AI, governance, and data ethics.

If we home in on the keywords provided in the articles published within the social sciences, law and humanities, the frequency analysis indicates that ADM research does not have a tight and clear vocabulary. When cleaned for keywords that also guided the search to identify relevant articles, the 78 social science items in the empirical material contain 453 different keywords, with only 39 of these appearing five times or more across the dataset, and only 13 keywords appear more than 10 times. These are labor (24), public (22), digital (20), artificial (19), intelligence (18), policy (17), learning (13), machine (12), systems (12), technology (12), transparency (12), injury (11), and social (11), again mostly quite generic keywords.

For the 95 law and legal studies items, 293 unique keywords were found, and only 38 of these appear five times or more in the material and 17 keywords with 10 or more mentions. These are: protection (32), law (30), intelligence (27), right (25), artificial (25), privacy (18), regulation (15), human (14), administrative (14), robot (13), public (13), GDPR (13), explanation (12), transparency (11), justice (11), automation (11), and process (10). Again, the keywords appear generic and technical, although in this case unsurprisingly legal terminology also figures prominently in the top list of keywords. Finally, for the 36 humanities items, 206 unique keywords were mentioned, with only 15 of these appearing at least five times, and just 6 keywords with 10 or more mentions: ethics (14), social (13), artificial (11), intelligence (10), learning (10), and machine (10).

From the frequency analysis of keywords in the whole empirical dataset as well as of the disciplinary subsets, the pattern is clear: Most keywords have very low incidences, mainly appearing only once or twice in the dataset.
This indicates that ADM is a highly diverse field of study at least when it comes to terminology. Furthermore, with very few exceptions, the top keywords are so generic and technical across the full ADM literature, and the individual subsets that they do not convey any specific information about disciplinary orientations to the study of ADM, albeit law research stands out with distinct legal terminology in a mix with the more technical and generic keywords.

In the next step we analyzed co-occurrence between the keywords, using VOSviewer. The results are presented in Figure 2 below.

When we look at keyword co-occurrences displayed in Figure 2, we can identify four distinct clusters that indicate core orientations in the current ADM literature. The biggest cluster is the red one, centering on artificial intelligence and comprising legal research (about GDPR, among others) and research into issues of explainability, transparency, privacy, and accountability, which is often of an interdisciplinary nature. Finally, the red cluster includes studies about social welfare, optimization, risk management and efficiency, as well as analysis of computational systems themselves. Hence, the red cluster is quite diverse and cannot be said to clearly follow along the lines of any specific discipline. The other clusters, on the other hand, appear more coherent in terms of focus. The green cluster, for instance, is characterized by various keywords relating to machine learning (e.g., neural networks, convolution, deep learning and diagnostic imaging). The blue cluster predominantly contains keywords associated with informatics (e.g., software, medical informatics, computer), and the yellow cluster is connected by a common focus on health, although the latter two clusters to some extent overlap (e.g., medical informatics and computer-assisted diagnosis in the blue cluster).

We suggest the frequency and keyword co-occurrence analyses indicate that although there might be a common general vocabulary around the kinds of technologies under study in ADM research, and applied research, it seems there is little common language in the theoretical or conceptual sense—that is shared across the articles even within the same or adjacent disciplines. This again, points us toward understanding ADM as a still emerging topic.

Finally, to qualify these observations about a lack of common ground further, we analyzed how the literature is interlinked. Turning to the question if there are specific clusters of ADM research that share common grounds, that is, in terms of key references, we engaged with an analysis of citation networks. Citation networks make visible how intensely articles cite each other or share common references. We took all the all in Scopus listed articles, using the VOS viewer tool to visualize citation networks, we can see that the literature is somewhat clustered.
Figure 3 presents a co-citation analysis showing links between authors that are cited a minimum of 20 times in a sub-dataset. The figure shows the ADM literature is somewhat clustered according to disciplinary orientations.

The red cluster in Figure 3 designates research in so-called critical data studies, with some high-profile contributions around data ethics (e.g., Floridi & Taddeo, 2016), critical algorithm studies (Bucher, 2018; Eubanks, 2018; Gillespie, 2014), digital rights (e.g., Veale & Binns, 2017), and fairness and accountability in ADM (e.g., Barocas et al., 2018). The latter links to the blue cluster, which shows a mix of research on investigative journalism (e.g., Angwin, 2014), public policy and applied computer science research (e.g., Kleinberg et al., 2018). Hence, the blue cluster does not appear specifically associated with a particular topic or discipline, but rather an eclectic collection of texts that cluster around some prominent shared references.

A yellow cluster is formed around computer science with contributions about, for example, explainable AI models (e.g., Guidotti et al., 2018). The yellow cluster to some extent forms a bridge between the red and blue clusters, indicating that computer science contributions to a certain degree engage in cross-disciplinary conversation with social science, law, and humanities research. Finally, the green and somewhat loosely knit cluster designates various types of engineering research that is not connected to the rest of the ADM research.

The most notable finding from the citation analysis, however, is that there are quite few scholars that are cited across the literature. A few individual authors take a somewhat central role (for example the pioneering investigative journalist Julia Angwin or the machine learning ethics scholar Solon Barocas), but they appear to be cited eclectically, and there are no central sources that are shared across the clusters. Out of the 44,265 authors cited by the publications in our database 197 of the authors are cited at least 20 times. Thirty-five authors are cited more than 50 times, and only five authors are cited more than a hundred times (Solon Barocas, Kate Crawford, Luciano Floridi, Frank Pasquale, and Sandra Wachter). These five are all prominent scholars in critical data studies, which could suggest that a core canon of ADM scholarship is consolidated close to critical data studies. The five scholars also share a focus on ethics and AI from socio-legal perspectives. Research environments that aim to develop studies that have relevance...
for the further technological and legal development of ADM and AI. The central position of these five scholars also reflects more generally where overlaps in the literature across disciplines are emerging, namely in the discussion of ethics and policies of AI. The centrality of these scholars may also suggest to their research topics as good candidates for ferment for interdisciplinary ADM research. Altogether, the sparse number of well-cited scholars and publications leads to a confirmation of the assumption that there is very little common ground in terms of citations. While there is an explosion of publications, there is little work that integrates and consolidates ADM research, a tendency that is characteristic of emergent fields and topics (Flensburg & Lomborg, 2021).

5 | DEFINING AND UNDEFINING AUTOMATED DECISION-MAKING

It has been observed before that it is rarely made explicit what we talk about when we talk about ADM both in basic academic research (Pink et al., 2022), but also in practical use cases as Kaun has shown in her study of the Trelleborg model (Kaun, 2021). We examined the dataset for explicit definitions and definitional struggles over ADM. We coded all articles for the inclusion of definitions and conceptualizations in ADM as such, and while we found articles that defined what was meant by ADM in the concrete context (40 in total), we did not find any self-declared conceptual frameworks for ADM as such. This might be related to the very recent upswing of studies on ADM since the implementation of the GDPR in 2018 as well. Nonetheless, given the high number of theoretical papers in the dataset, the lack of clear-cut conceptualizations and definitions of ADM is surprising.

Taking a closer look at those 40 articles that actually included definitions of ADM, articles often cite article 22 of the GDPR that refers to ADM and, in that sense, rely implicitly on a specific ADM definition put forward in practical legal terms. The dominant topics of articles that include an implicit or explicit definition of ADM are engaging with questions of transparency or opacity of decision-making as well as fairness, justice, and accountability, that is, engaged with defining and framing the discussion of the alleged consequences of ADM with reference to core values and human rights rather than concrete people and lived life. Furthermore, the implicit definitions follow along disciplinary divides between computer science, law, and social sciences.

In law studies, a legal terminology is put forward where definitions are discussed in terms of thresholds. The threshold discussion centers around the question “when do we speak of ADM”; should we for example, use a broad definition of ADM that includes all processes that involve computational processing as suggested by Swedish National Audit Office in a report from 2020, or should we apply a definition that includes only fully automated processes that do not involve any human decision-making at all? Similarly, in computer science, a field that has seen a steady flow and growth of ADM literature, ADM is hardly problematized in terms of its definition. It is rather considered as a technical term similar to algorithm. Social sciences and humanities-oriented studies are in the position to contribute to developing a deeper conceptual understanding of ADM and its implications together with other disciplines and partners who develop technological solutions as well as regulators. These fields provide situated and empirical as well as theoretical approaches engaging with the entanglement of human and non-human agencies as well as historical shifts in how societies are organized and experiences.

Having outlined the emerging literature of ADM research that is still eclectic and scattered across disciplines largely lacking common grounds and common language, the findings beg the question why and how the term entered into the different research fields in the first place. The explosion of publications since 2018 can clearly be linked to the implementation of the GDPR in 2016 and article 22 on the human in the loop provision. However, already the EC Directive on data protection from 1995 included an article—Article 15—which grants persons the right not to be subject to fully ADM (Bygrave, 2001). Although the directive was considered ground-breaking for data protection and the regulation of profiling by commentators (see Bygrave, 2001), this directive did not cause a similar upswing in publications on ADM as the GDPR. Why is that so?

In 1995, the discussion of automation with the help of artificial intelligence was slightly different from contemporary debates that strongly focus on the gains of AI and ADM systems as well as the future with general AI (Rahm
& Kaun, 2022). Instead, the discussion was focused on rule-based algorithms including decision trees and robotic process automation. The term ADM accommodated both for technical specificity and openness for further technological development.

In comparison to the year of implementing the GDPR in 2016, public discourse has shifted more firmly toward the question of automation with the help of AI. While AI increasingly returned as a buzzword both in academic and political circles, it also became fuzzier and more controversial. Civil society organizations advocating around algorithmic automation had to define their political targets more clearly. In that context, ADM appears less controversial and rather technical compared to AI that has also become a theme in popular culture with implications for how we make sense of the technology on an everyday level (Goode, 2018). Hence ADM (re-)appeared as a term in the public debate, while social sciences and humanities consequently started to engage with ADM more intensively illustrating a double hermeneutic (Giddens, 1987). Double hermeneutic here means that lay concepts circulating in the public sphere and theories in social sciences have a mutual relationship shaping each other, which, as Anthony Giddens has argued, marks one of the most important differences between the human social sciences and natural sciences.

A second trope in the material are the disciplinary divides. These are emerging along well-known demarcation lines between computer sciences, law and the social sciences and humanities. As Moats (2021, see also Moats & Seaver, 2019) has shown, bridging disciplinary divides requires creative practices and explorative approaches.

Where does this leave us? We have in a sense confirmed our initial impression of a research area in significant growth, attracting interdisciplinary interest but showing little consolidation. But now—arguably—we can document that a shared vocabulary is not established. We do not want to use this finding to argue for a neat definition nor do we subscribe to one overall conceptual framework. But we want to suggest that when sticking with the loose and flexible understanding of ADM, we must at least try to be specific when applying the ADM term. AlgorithmWatch’s practical and specific definition of ADM might still be the closest we come to a common reference point to start conversations about ADM across scholarly fields.

Starting from our citation networks, we suggest bridging the divides displayed by identifying and thinking with the concept of communication. More specifically, we can see that the four central clusters within the citation network are revolving around artificial intelligence (with a focus on explainability, transparency and accountability), machine learning (including neural networks, convolution), informatics and health. The four clusters can meaningfully be linked together through a focus of communication and datafication that helps to understand processes of (automated) data production that are a precondition for artificial intelligence and machine learning to function.

Furthermore, our own experience running interdisciplinary research networks around ADM and hence going beyond the systematic literature review demonstrates what can be achieved by a sustained conversation between disciplines to find useful bridges. Here media and communication research makes an important contribution to the study of ADM and the possible impacts of such an endeavor for democratic processes in datafied societies. While the study of technological systems in the broad sense has historically been outside the scope of media and communication research, media and communication studies has much to offer—both in terms of concepts for illuminating relations between technological systems and the people they implicate (“what media do to people”), and in terms of an empirical sensitivity to the contextual appropriation of said systems (“what people still do with media”) to fit organizational and social agendas and pursuits. Hence, we suggest a communication-based conceptualization of ADM, which also implies an urgently needed rehumanization of ADM (Pink et al., 2022), which reflects how humans and ADM systems shape one another. At the very basic level, then, ADM can be understood as a communicative sequence with feedback loops: it emerges around encoding of data as information in the system, interpretation of these data by processing and socially ordering the information, and output decisions, which must then be decoded, or made sense of, by case workers and other stakeholders in context, with potential to modify future data input and interpretations (Lomborg & Kapsch, 2020). As Jensen and Helles (2017) suggest, all digital communication—however minimally—contributes to reconfiguring the systems in question. Importantly, people, through the interpretations and ensuing actions based on ADM are part of this sequence in different ways, contributing to the ongoing shaping of ADM systems. Following this logic it becomes an empirical task to
explore the different ways people engage with and act through ADM systems as communicative sequences. This, in turn, means to radically contextualize and re-center the people in ADM research as we nuance in the concluding section.

Based on practical and specific definitions of ADM, along the lines laid out above, we can, as suggested by David Moats (2021), develop disciplinary bridges through playfulness and open-ended inquiry into possible concepts that will help us move forward as well as inquiry into the situated uses of ADM. By way of concluding this article, we suggest two trajectories along which to develop ADM research: one conceptual and communication-centered, and one empirical and people-centered.

6 CONCLUSION: CONTEXTUALIZING AND RE-CENTERING THE PEOPLE IN ADM RESEARCH

As we have shown above, the current research landscape is dominated by high-level guidelines, and studies of systems, transparency, and data ethics. Largely, such perspectives start with the ADM system, rather than the people and purposes it is meant to serve. Shifting perspective opens a space for acknowledging that even if ADM systems are highly standardized and often rule-based, they are not entirely fixed. From the vantage point of communication, then, we argue for the need to study ADM from the peoples’ perspective—from the ground up and radically contextualized which means to consider the specific contexts within which people (professionals as well as private citizens) interface with ADM systems, the ways in which they make sense or reject them and ultimately develop frameworks for approaching technological continuity and change. This could for example, include organizational ethnographies of workplaces that have introduced ADM to capture experiences, negotiations, meaning making and power relations around ADM systems. Furthermore, a people-centered approach might mean to engage creative methods such as mind scripting and future vision workshops to highlight user and citizen perspectives (Allhutter, 2012). A people-centered analysis of ADM within digital sociology is slowly evolving but needs further attention.

This includes a research agenda for digital sociologists that considers the practices and care work that people conduct in relation to ADM systems not only in developing, but also maintaining, repairing, abandoning, and repurposing them (Schabacher, 2021). Here care could mean “everything that we do to maintain, continue, and repair our world so that we can live in it as well as possible” (Tronto & Fisher, 1990) in relation to ADM systems that emerge as a practice and as a disposition. Care in relation to ADM systems concerns both the care for technology as well as the role of technology for care (Lindén & Lydahl, 2021; Mol et al., 2010; Puig de la Bellacasa, 2010).

Sociological media and communication studies offer approaches to do just that, namely, to highlight ground up experiences and practices relating to technologies while taking into account the specific context within which they emerge. Especially the strand of audience studies (Das & Ytre-Arne, 2018; Picone et al., 2019) and phenomenologically based investigations of media experiences (Bengtsson & Johansson, 2021; Hill, 2018) seem promising here. These investigations have suggested engaging with the agency of audiences and mundane media users in a nuanced way that allows for both a fundamental critique of media systems and content production as well as acknowledging the mundane acts of actively engaging with media content. Applying these nuanced ways of considering both the constraints and freedom that media users face to ADM means embedding decision-making systems in the everyday world of mundane users including civil servants, managers as well as citizens as for example, welfare recipients. Audience centered research has also highlighted that audiences develop “new and often ambivalent strategies for coping, negotiation, or resistance, but are nevertheless left with uncertainty about the possibilities and consequences of their own communicative actions” (Ytre-Arne & Das, 2021, p. 781). Ytre-Arne and Das (2021) have underlined that communicative agency of people should be understood as interpretative and relational. Similarly, we argue that ADM should be understood within its context and interpretive structures of meaning that emerge in relationships between human users and machinic interfaces. Conceptualizing ADM as fundamentally communicative will bring us closer in this endeavor and will make a crucial addition to the sub-field of digital sociology.
ACKNOWLEDGMENTS
We would like to thank the reviewers, and the members of the ADM Nordic network for their constructive feedback. We would also like to thank Claudia Wladimiro and Iheanyi Amaraizu, who supported the authors in the coding process. Anne Kaun’s research for this article was part of the Automating Welfare in the Baltic Sea Region Project funded by the Baltic Sea Foundation Grant No. S2-20-0007_OS. Sne Scott Hansen’s research for this article was part of the ADM—Nordic Perspectives funded by the Independent Research Fund Denmark Grant No. 9055-00060B, and Stine Lomborg’s research for this article was part of Datafied Living, funded by the ERC Grant No. 947735.

CONFLICT OF INTEREST STATEMENT
The authors have no conflict of interest.

DATA AVAILABILITY STATEMENT
The data that support the findings of this study are available from the corresponding author, SL, upon reasonable request.

ORCID
Stine Lomborg https://orcid.org/0000-0002-5672-7298
Anne Kaun https://orcid.org/0000-0002-5879-2130
Sne Scott Hansen https://orcid.org/0000-0002-0107-6336

ENDNOTES
1 The large reduction in the number of articles is based on the aim for coherence and comparability in the information extracted. Initially, we tested complementing the information, which proved difficult and time consuming. Data quality was validated higher than quantity. Incomplete search returns that did not include keywords, metadata, journal or author information as well as the abstract itself were removed.

2 Initially, we coded the disciplinary affiliation of the authors as well. While this provided further granularity within the social sciences and humanities, particularly, for the present purposes of identifying overall patterns, it was judged superfluous.

REFERENCES


**AUTHOR BIOGRAPHIES**

**Stine Lomborg** is associate professor at the Department of Communication, University of Copenhagen, and the PI of the Datafied Living project, which studies infrastructures and experiences of tracking of the self across personal, work, and institutional life in the welfare state.

**Anne Kaun** is Professor at the Department of Media and Communication Studies, Södertörn University, and an Wallenberg Academy Fellow studying the democratic implications of automated decisions making in the welfare sector.

**Sne Scott Hansen** is a PhD candidate at the Department of Communication, University of Copenhagen. Her PhD project is entitled “Self-tracking the communicating body.”