Studying Up Machine Learning Data: Why Talk About Bias When We Mean Power?

MILAGROS MICELI, Technische Universität Berlin, Weizenbaum Institute, Germany
JULIAN POSADA, University of Toronto, Canada
TIANLING YANG, Technische Universität Berlin, Weizenbaum Institute, Germany

Research in machine learning (ML) has largely argued that models trained on incomplete or biased datasets can lead to discriminatory outputs. In this commentary, we propose moving the research focus beyond bias-oriented framings by adopting a power-aware perspective to "study up" ML datasets. This means accounting for historical inequities, labor conditions, and epistemological standpoints inscribed in data. We draw on HCI and CSCW work to support our argument and critically analyze previous research in three ML-related areas: data bias, data work, and data documentation. This way, we point at both co-existing lines of work within our community — one bias-oriented, the other power-aware — and highlight the need for dialogue and cooperation. In the first area, we argue that reducing societal problems to "bias" misses the context-based nature of data. In the second one, we highlight the corporate forces and market imperatives involved in the labor of data workers that subsequently shape ML data. Finally, we propose expanding current transparency-oriented efforts in dataset documentation to reflect the social context of data design and production.

Additional Key Words and Phrases: bias, power, machine learning datasets, training data, data work, dataset documentation

ACM Reference Format:
Milagros Miceli, Julian Posada, and Tianling Yang. 2022. Studying Up Machine Learning Data: Why Talk About Bias When We Mean Power?. *Proc. ACM Hum.-Comput. Interact.* 1, 1 (September 2022), 14 pages. https://doi.org/10.1145/XXXXX

1 INTRODUCTION

In 2015, Facebook's "real name" policy caught some media attention after the platform's algorithm failed to recognize the names of hundreds of North American Indigenous users as "real" and proceeded to cancel their accounts [43, 85]. According to Facebook's algorithm, real names seemed to be defined by Anglo-Western conventions. Thus, the system flagged names composed of several words or with unusual capitalization. Moreover, despite the many contextual factors that determine how a name sounds and looks like, Facebook enforces its policy algorithmically, that is, in a narrow, unquestionable, and predefined way.

At first sight, the issues raised by users whose names were flagged could indicate the presence of biased training data: As Anglo-Western names are dominant and names from other cultures are underrepresented, the unbalanced dataset leads to "unfairness." This approach is not wrong but it is also not sufficient to fully address the issue at stake, i.e., that some worldviews are considered

Authors' addresses: Milagros Miceli, Technische Universität Berlin, Weizenbaum Institute, Berlin, Germany; Julian Posada, University of Toronto, Toronto, Canada; Tianling Yang, Technische Universität Berlin, Weizenbaum Institute, Berlin, Germany.

2573-0142/2022/9-ART $15.00
https://doi.org/10.1145/XXXXX

more valid than others. Framing such issues as "bias" tends to obscure a set of persistent questions behind and beyond the technical domain: What *is* a real name? Who decides over the *realness* of a name? And, do we need a real name policy at all?

In the past decade, injustice and harm produced by data-driven systems has often been addressed under the umbrella term "bias." Research has shown that biases can penetrate ML systems at every layer of the pipeline, including data, design, model, and application [71]. Special attention has been paid to the quality of data, arguing that models trained on incomplete or biased datasets can lead to discriminatory or exclusionary outcomes [15, 29, 71]. Moreover, significant academic focus lies upon bias in data work and crowdsourcing [13, 21, 34, 36, 49]. Because of the interpretative character of tasks such as labeling, rating, and sorting data, abundant research has focused on the individual subjectivities of data workers to account for biases in data, investigating ways of mitigating them by constraining workers' judgment.

With the present commentary, we aim to contribute to the discussion around data bias, data workers' bias, and data transparency by broadening the field of inquiry with a Social Science perspective: *from bias research towards an investigation of power differentials that shape data.* As we will argue in the next sections, the study of biases locates the problem within technical systems, be them data or algorithms, and obscures its root causes. Moreover, the very understanding of bias and debiasing is inscribed with values, interests, and power relations that inform what counts as bias and what does not, what problems debiasing initiatives address, and what goals they aim to achieve. Conversely, the power-oriented perspective looks into technical systems but sets the focus on larger organizational and social contexts. It investigates the set of relations that intervene in data and system production and aims to make visible power asymmetries that inscribe particular values and preferences in them.

Computing has become so widely integrated in society, both influencing and being shaped by it, that a broader understanding of sociotechnical system, in our case machine learning, becomes key to address social concerns surrounding their development and deployment. In this sense, "debiasing" efforts are not sufficient to fully address the questions posed by "real-name" algorithms and other data-driven systems that are deeply engrained in our everyday lives. Such approaches could be expanded by applying a relational view on the power dynamics and the economic imperatives that drive machine learning, i.e., considering that biases do not occur in a vacuum but are deeply entangled with naturalized ways of doing things within the organizations where datasets and systems are developed. This requires an epistemological shift in terms of how to think of these problems, what questions to ask, and what methods to use. Such a shift can only be achieved through more dialogue between Computer Science and disciplines such as Sociology, Anthropology, and Economy. Given the important interdisciplinary tradition in HCI and CSCW, we believe in the key role of these communities in prompting conversations around power and ML systems.

On that basis, we follow the line of previous work [5, 37, 87] that has borrowed the concept of "studying up" from anthropologist Laura Nader. In anthropology, studying up means expanding the field of inquiry to study power, i.e., to interrogate elites that have remained significantly understudied in the anthropological tradition. In their call to study up algorithmic fairness, Barabas and colleagues [5] explain that this endeavor "requires a new set of reflective practices which push the data scientist to examine the political economy of their research and their own positionality as researchers working in broken social systems." In a similar vein, our appeal is to "study up" machine learning data by investigating labor conditions, institutional practices, infrastructures, and epistemological stances encoded into datasets, instead of looking for biases at the individual level and proposing purely technical interventions.

In the following sections, we will zoom into three critical ML-related fields of inquiry: *data, data work, and data documentation.* While our argument is based on previous research, it is worth mentioning that a systematic literature review is not within the scope of this commentary. Here, we critically discuss CSCW and HCI work that revolves around the concept of "bias", while building on previous initiatives within those communities that have striven for a more comprehensive understanding of sociotechnical systems instead. By contrasting both co-existing perspectives, we highlight the importance of fostering more dialogue between them to produce research that expands the investigation of individual biases into a consideration of power asymmetries within organizations and among them. Our argumentation concludes with suggestions as to how to study up machine learning data and why.

2 THE LIMITS OF BIAS

Studies on data and algorithmic biases have demonstrated how data-driven systems can enhance discriminatory practices and result in exclusionary experiences in various domains, including credit [25, 59] and algorithmic filtering [4, 70]. CSCW and HCI research has explored algorithmic bias in the job market [20, 46], advertisement [1], and image search engines [54], among several other domains. Moreover, researchers have shown how algorithms contribute significantly to the visibility of information [58] and how stereotypes are perpetuated by gender recognition systems [44]. The quest for addressing these problems has prompted the development of an area of research that emphasizes the issue of bias, and the values of fairness, accountability, and transparency in mitigating its effects [28]. The fact that research in the technical realms takes issue with social inequities and examines the harmful effects of technology is a significant step. However, work in Critical Data and Algorithmic Studies as well as CSCW and HCI has argued for a shift of perspective from individual cases and individual biases towards the comprehensive analysis of social practices and power relations involved in creating the systems that surround us [5, 11, 27, 52, 62, 86].

Technological development is sociotechnical in nature and data, as an abstraction [28, 56], is not given but created through human discretion [68] and shaped by power dynamics [62]. By focusing on technical solutions for personal subjectivities, bias-oriented approaches are mostly unable to account for the social processes underway that comprise increasing surveillance and privacy intrusion to satisfy the voracious need for more and different data [22] and the important shifts in labor that include the mobilization of largely precarized workforces to process data and make it "readable" for ML systems [78]. Through a power-aware lens, it is possible to interrogate why accurate, efficient, and seemingly "debiased" ML systems are still not *good for everyone*. For example, accurate facial recognition used for surveillance is dangerous in the hands of unscrupulous organizations or oppressive governments. Debiasing efforts sometimes mitigate harm, but if systems remain controlled by powerful organizations that follow their own agendas, machine learning will inevitably perpetuate injustice. In this context, attempts to address and mitigate biases appear as "a tiny technological bandage for a much larger problem" [28]. Research efforts that focus on designing "debiased" systems are not bad. However, the question stands: "debiased" according to whom and for whom? [52]. The bias-oriented approach provides only limited tools to explore this and other important questions.

Moreover, framing sociotechical problems as bias constitutes what Powles and Nissenbaum call "a seductive diversion" [80]: On the one hand, we are told that biases can be fought and mitigated, and that data can be cleaned and systems debiased. On the other hand, it is argued that bias is not a technical but a societal issue; hence, biases are everywhere and nowhere. If society is biased, then biased AI cannot be avoided. This way, the bias framing presents a puzzle that keeps us continually busy because technical fixes are inadequate solutions to societal issues. We are always on the way of identifying and mitigating them in an attempt to build debiased systems while knowing that the ideal of a debiased system can never be achieved. Still, considerable efforts, both within and beyond HCI and CSCW, are invested in technical tools to mitigate data biases, algorithmic biases, and workers' biases in domains where interrogation and reflexivity could be more fruitful. This way, the bias puzzle distracts us from addressing fundamental questions about who owns data and systems, who are the data workers, whose worldviews are imposed onto them, whose biases we are trying to mitigate, and what kind of power datasets perpetuate. "It also denies us the possibility of asking: should we be building these systems at all?" [80]. These questions could contribute to shifting the perspective because they interrogate privilege and naturalized worldviews encoded in data and systems that (re)produce the status quo. Consequently, such questions are more about power than they are about bias.

In the following and to unpack this argument, we will go deeper into discussing problematic aspects of framing power differentials and injustice as "bias" in ML data, data work, and dataset documentation.

2.1 Data is Always Biased

Data bias has been defined as "a systematic distortion in the data" that can be measured by "contrasting a working data sample with reference samples drawn from different sources or contexts."[71] This definition encodes an important premise: *that there is an absolute truth value in data and that bias is a "distortion" from that value.* This key premise broadly motivates approaches to "debias" data and ML systems. However, we argue that the problem with this assumption is that data never represents an absolute truth. Data, just like truth, is the product of subjective and asymmetrical social relations.

In their groundbreaking analysis of three commercial gender classifiers made available by Microsoft, Face++, and IBM, Buolamwini and Gebru [15] show that darker-skinned women are up to 44 times more likely to be misclassified than lighter-skinned men. This work is often cited as a paradigmatic example of how data can contain biases as related to the underrepresentation of certain groups. Looking at this problem from a bias-oriented perspective, the solution seems straightforward: add more and diverse data to training datasets. However, as Gebru also points out in an interview, biased data is only part of the story: "[. . .] not just bias in the training data, but ethics in general — what's okay to do, what's okay not to do, the power dynamics of who has data, who doesn't have data, who has access to certain kinds of models, and who doesn't" [50]. The contextual issues that escape technical fixes also include: where is diverse data — the "missing faces" — harvested? Under which conditions? Who classifies them? Moreover, and considering that Black and Brown populations have historically been subject to surveillance, persecution, and police violence [9, 14], it is worth asking if improving facial-recognition systems so that they can properly "see" dark-skinned faces would further perpetuate such injustice.

Our point is that biased data is undoubtedly one issue to consider when it comes to discriminatory outcomes from machine learning systems, but so are social structures, the definition of social problems to be solved in computational terms, and the widespread assumption that algorithms are neutral where people are not. These factors, as well as data, are deeply political. Machine learning systems are fundamentally trained to cluster and classify data. When these classifications are value-laden and interest-informed, they can result in imposing and promoting the particular set of interpretations and worldviews of some groups, which could reinforce injustice [12]. In other words, ML systems have real effects on real people. Therefore, it is important to consider that their quality cannot be thought of only in terms of accuracy and performance. Some issues do not just get solved by throwing in more data and quantification not always leads to better representation or less harm. In a broader sense, harms produced by ML systems manifest existing power asymmetries: they are about having the power to decide how systems will "see" and classify, what data is worth including, and whose data we can afford to ignore. Those harms are about the power to impose a hegemonic worldview over others possible.

Tracing the links to historical and ongoing asymmetries can be helpful to understand how data comes to be [27] and what kind of political work ML systems perform [45]. This means, of course, acknowledging that the data that fuels machine learning is produced by humans and hence is laden with subjective judgments. However, discussions around human intervention on data ought to consider that the subjective forces that shape data and systems are not just about the personal biases of individual actors. Data is produced within organizations and through practices that "embody specific technical ideals and business values" [74] that also shape the subjectivities of data workers. We are for sure not the first ones to make this statement: Researchers in Human-Centered Data Science (HCDS) [57, 68, 69, 73, 74, 94] and Human-Centered Machine Learning (HCML) [19] have explored data as a "human-influenced entity"[68]. A series of CSCW/HCI workshops on Data Science work practices [66, 67] has fostered interesting conversations on collaboration, meaning making, trust, craft, and power. This line of work has shown that narratives, preferences, and values related to larger socio-economic contexts are embedded in processes of data production [75]. Practices such as the framing of real-world questions as computational problems [10, 72], the choice of training data and data-capturing measurement interfaces [77], the establishment of taxonomies to label data [62], and the selection of data features [68] as well as the design of data to be recognizable, tractable, and analyzable [35, 68], all are decisions that are hardly ever made by individual choice and in a vacuum. Instead, they concern organizational structures and depend on what is possible in terms of time and budgets, and what is expected in terms of computational output and revenue plan.

As the examples in the following section will show, despite the abundant CSCW and HCI initiatives that have argued that "datasets aren't simply raw materials to feed algorithms, but are political interventions" [23], a considerable number of investigations within those research communities still comprise the assumption that data represents an absolute truth value and that bias is just a distortion that can be mitigated. The problem is that framing arbitrary representations in data as bias misses the political character of datasets: there is no neutral data and no apolitical standpoint from where we can call out bias [23]. Datasets are always "a worldview" [26] and, as such, data always remains biased.

2.2 "Mitigating Worker Biases" Should Not Be the Goal

Datasets are conditioned by the networked systems in which they are created, developed, and deployed. The examination of the provenance of data and the work practices involved in dataset production are essential to the investigation of subjectivities embedded in data-driven systems [62, 68, 69, 73]. In formal terms, data work for machine learning involves tasks such as the collection, curation, and cleaning of data, labeling and keywording, and, in the case of image data, it can also involve semantic segmentation (i.e., marking and separating the different objects contained in a picture) [17, 18, 90]. Outsourced data workers perform these tasks through digital labour platforms (crowdsourcing) or business process outsourcing companies (BPOs). In this regard, outsourced data work is part of the broader gig economy landscape, in the case of platforms [92], and other digital service BPOs, like those providing content moderation [82]. In both cases, these types of work are characterized for low- or piece-wages, limited-to-no labor protection, and high levels of control and surveillance.

The tasks that data workers perform are fundamentally about making sense of data [62, 68], that is, about interpreting the information contained in each data point. Because of the subjective character of data-related tasks, bias-oriented research in this space has focused mainly on the individual subjectivities of workers, considering their judgments to be a significant source of biases and data quality errors [13, 21, 40, 49, 91]. For example, abundant research considers labelers' subjectivity in annotation tasks to be one of the main reasons for biased labels. The field of research directed towards the study of crowdworkers and crowdsourcing platforms [13, 21, 34, 36, 49] offers several examples of such an approach. Some of this work argues, for example, that data workers' cognitive biases [30], their own preferences [81], and political stances [93] can negatively affect data. Moreover, research has proposed methods to identify and monitor annotator bias within datasets [3, 39, 49, 91]. In a paper presented at CHI 2019, Hube et al.[49] explore how crowdworkers annotate machine learning data and propose a framework for mitigating their biases. The authors argue that extreme personal opinions of workers can affect data labeling tasks and produce biased data, especially when the tasks involve opinion detection and sentiment analysis. Consequently, they add that "the ability to mitigate biased judgments from workers is crucial in reducing noisy labels and creating higher quality data." This reasearch follows the line of many of the work in crowdsourcing that rests on three premises: (1) that data represents an absolute ground truth and that bias is a deviation from that truth value, (2) that data workers have enough agency to interpret data according to their personal judgment and could, therefore, be prone to deviating from that predefined truth value, (3) and that workers using their own subjectivities to interpret data is *per se* detrimental to the quality of data. Quite often, these approaches to detect and mitigate workers' bias do not consider that data workers constitute automation's "last mile" [41], that is, the bottom end of hierarchical labor structures, and that they collect and label data within organizational structures and according to predefined truth values instructed to them by managers and clients.

Several issues framed by previous research as "workers' bias" are actually manifestations of broader power asymmetries that fundamentally shape data: power asymmetries that are as trivial as being the boss in a tech company with decision-making power or being an underpaid crowdworker who risks being banned from the platform if they do not follow instructions. Socio-technical systems are complex in nature and this also includes the data work that fuels them. We argue that research that focuses on the personal biases of workers and aims at mitigating them could benefit from an interrogation of power differentials, normalized preconceptions, and profit-oriented interests that shape labor conditions in data work.

Let us look at some examples from our on-going research project that focuses on data work in Latin America. These examples should provide an idea of the identity of the workers whose biases research attempts to mitigate. As with many data workers, they are located in Argentina and Venezuela. The Venezuelan economy is currently experiencing the highest levels of inflation in the world and many people look for work with crowdsourcing platforms because they offer a steady income, paid in US dollars. Melba, one of the crowdworkers interviewed by us, is a retired woman. Her monthly pension is the equivalent to USD$1, which, as she puts it, is "not enough to buy half a dozen eggs; not enough to buy a piece of cheese or bread." The payment she receives for doing data work is also meager by international standards. However, in a country experiencing hyperinflation, it allows her to supplement her income. In the case of Juan, another crowdworker from Venezuela, the income from the platform is comparable to what he would receive doing harvest work in the neighboring country, Colombia. However, doing data annotation allows him to stay in Venezuela with his family instead of migrating and being apart. In the case of Argentina, most of the data workers we interviewed live in the impoverished areas that surround Buenos Aires. Despite the meager salaries they receive for data collection and annotation tasks (the equivalent to US$1.80 per hour), and the exhausting nature of the work they perform, all interviewees expressed being proud of their work. For many of them, doing data work means finally having a desk job and breaking with generations of unlicensed cleaning or construction work. Similarly, for many of the Venezuelan crowdworkers, having access to this type of work means avoiding extreme poverty and having a means to circumvent many of the difficulties present in their local labour market.

The cases described above are not extreme or marginal. They represent the reality of an industry that outsources data work to global locations where the lack of better employment opportunities forces workers to be cheap, quiet, and obedient. A growing body of literature in CSCW and HCI has taken crowdworkers' perspective and pointed to the issues of underpayment [47], crowdworkers' growing dependency on performing crowdsourcing tasks to make ends meet [83], the use of parameters and processes (e.g. the rate of previously approved and paid tasks) to select and recruit crowdworkers [6], and the power asymmetries introduced by crowdsourcing platform design and inherent in the relations between service requesters and crowdworkers [51, 61, 84]. Ekbia and Nardi use the term *heteromation* to characterize the shift in technological-mediated work and labor in which human intervention and action are indispensable for technical systems to function [31]. They argue that heteromated systems, like MTurk, are the outcome of socioeconomic forces rather than of the essential attributes of humans and machines, as commonly assumed [31]. The authors not only scrutinize the asymmetrical labor relations in which crowdworkers are put at a significantly disadvantaged position, but also emphasize that crowdworkers are regarded as mere "functionaries" of an algorithmic system and are rendered invisible [31]. Apart from drawing attention to invisible labor and asymmetrical labor relations, a political economic perspective further highlights the profit-driven imperative of capital, the surveillance and social control enabled and reinforced by digital technologies, and the political nature of design choices and technologies that mediate work and labor [32, 33]. These studies are important examples within CSCW and HCI of how shifting researcher's gaze upwards to look into power dynamics can expose fundamentally different issues with sociotechnical systems. However, they unfortunately have not received enough attention from scholars in those very same research communities that investigate bias in data work.

Social and labor conditions affect the dependency of workers on data work, and that dependency has an effect on how datasets are produced, such as restricting workers' ability to raise questions about annotation instructions and tasks. Starting from the question of how this form of heteromated labor affects crowdworkers, broader communities, and polities [31], we propose also asking *how power asymmetries in heteromation inform ML datasets and systems*. Starting from the assumption that such imbalances are the problem, not just bias, leads to fundamentally different research questions and methods of inquiry. We believe that this perspective can significantly contribute to broaden research on data worker and crowdsourcing bias.

2.3 Data Documentation Beyond Bias Mitigation

Several frameworks and tools to document machine learning datasets and models have been proposed and applied. Significant examples are the work of Bender and Friedman with the *Data Statements for Natural Language Processing* [8], Holland et al. with the *Dataset Nutrition Label* [48], and most prominently, Mitchell et al. with *Model Cards for Model Reporting* [64], and Gebru et al. with *Datasheets for Datasets* [38]. In these investigations, data bias appears as a core motivation for developing documentation frameworks. The authors argue that documentation can help "diagnose sources of bias" [48], and has potential to "mitigate unwanted biases in machine learning systems" [38]. In the present subsection, we would like to discuss two ways to complement these data documentation approaches. The first one is to consider expanding the documentation of dataset composition beyond merely listing dataset's elements. The second one is to consider the complex and intricate relationship between dataset creators and dataset consumers. As we will argue, both considerations could allow us to expand this line of research and explore power relations in machine learning through a CSCW-informed perspective, beyond bias-oriented framings.

First, we argue for the inclusion of further information beyond the proposed list of data "ingredients". For instance, one of the questions in *Datasheets for Datasets* asks "does the dataset identify any subpopulations?" (e.g. by race, age, or gender). This way of documenting dataset composition is key but it also brings along what we consider to be a valid question: *Is this information sufficient in itself to explicate unjust outcomes?* Disclosing whether a dataset includes racial categories and listing said categories “does not speak to the problem of such categories’ reductiveness, nor makes the assumptions behind race classifications embedded in datasets explicit” [63]. We believe that documentation can and should tell us more, for instance, about how data collectors and annotators have established the correspondence between data point and category. Moreover, “to impose order onto an undifferentiated mass, to ascribe phenomena to a category — that is, to name a thing — is in turn a means of reifying the existence of that category” [23], as Crawford and Paglen eloquently put it. Similarly, when the documentation of racial categories contained in a dataset is limited to listing them without further reflection, the risk exists that the documentation could contribute to the reification and naturalization of such categories.

Our second idea is to provide more context on the intricate relationship between data workers and requesters. For instance, in their investigation, Gebru et al. [38] argue that *Datasheets for Datasets* would improve communication between dataset creators and dataset consumers. The clear differentiation between dataset creators and consumers surely applies to large open datasets commonly used for benchmarking, such as ImageNet. However, such a clear separation does not correspond with the totality of machine learning datasets or even to most of the ML products that are created for commercial use. For instance, Feinberg [35] unveils "a multilayered set of interlocking design activities" in data infrastructure, collection and aggregation in data production. In many settings of data production, design activities and decisions are shaped, if not determined, by dataset consumers and other external stakeholders rather than data workers, which makes them co-designers of datasets. In such settings, the distinction between consumers and producers is more ambiguous. Previous work [55, 63] has explored companies producing (or outsourcing the production of) tailor-made datasets to train their own ML models. These companies have particular requirements in mind and produce data specifically tailored to the ML product they aim to develop. Many of these organizations do outsource data collection and labelling but, even then, tasks are completed according to the specific instructions provided by the model developers — what Gebru et al. call “dataset consumers.” Once the labeled dataset is sent to the model developers, data is further cleaned and sometimes re-labeled. In a similar vein, Seidelin [88], building on and extending Feinberg’s design perspective of data, situates data work and practices in organizational, cross-organizational, and multi-stakeholder contexts. Her research reveals that data work and data-based services are by nature collaborative and cooperative, and that the design and production of data are rather co-design processes. These perspectives challenge the clear separation between dataset producers and consumers and show that dataset consumers are also dataset co-creators.

With both ideas described above, we seek to expand previous work in data documentation beyond the data-bias motivation. Merely listing the composition of a dataset without interrogating the origins of its categories might be sufficient if the aim of documentation is “mitigating unwanted biases”. However, it is not enough to unveil the political work those categories perform. Similarly, the stiff differentiation between producers and consumers seems to reproduce an analog logic as the studies on data worker’s biases described in the previous section: The responsibility for data quality issues lies with data workers exclusively and requesters (a.k.a model developers) have no control over assumptions encoded in datasets because they are mere “consumers.” We argue that this field of inquiry could be better explored by moving the focus away from the documentation of datasets’ technical features and biases, to highlighting the importance of documenting production contexts, aiming to make visible the dynamics of power and negotiation within and among organizations that shape datasets.

Such an extended perspective could also help to explicate why, despite growing calls for more transparency in machine learning, data documentation practices are still limited in the field. Some factors to take into account are that requesters often regard the information that should be documented as corporate secrecy and that documentation is often perceived as an optional task, in some cases even as a burden, that is time-consuming and expensive [63]. Moreover, the lack of knowledge and training, be they technical or ethical, makes data workers less equipped to reflect on what should go into documentation [55] and, even among informed workers, hierarchical managerial structures in BPOs and the risk of being banned in data work platforms, would probably make workers reluctant to use documentation to reflect upon taken-for-granted practices. To address such difficulties, researchers developing documentation frameworks could benefit from the acknowledgement that data production is a collaborative project, which demands cooperative efforts from actors that hold different (organizational and social) positions and decision-making power to shape data [24, 62, 63]. While the bias-orientation of existing frameworks counteracts documentation's potential to make power explicit and contestable, we believe that CSCW research could significantly contribute to this line of work. More than diagnosing "the source of bias," documentation should aim at interrogating work practices and decision-making hierarchies within and among organizations.

3 CONCLUSION

This commentary has critically explored several implications of framing diverse socio-technical problems as "bias" in machine learning. Through examples related to the study of ML datasets, data work, and dataset documentation, we have argued for a shift of perspective to orient efforts towards considering the effects of power asymmetries on data and systems.

Such reorientation not only concerns privileged groups among machine learning practitioners. It is also about the role of researchers and the intertwined discourses in industry and academia [42]. We need more research that interrogates the relationship between human subjectivities and (inter-)organizational structures in processes of data production. Most importantly, power-oriented investigations could allow researchers to "shift the gaze upward" [5] and move beyond a simplistic view of individual behaviors and interpretations that, in many cases, ends up allocating responsibilities with data workers exclusively. Moreover, it could be helpful to investigate workers' dissent not as a hazard but as an asset that could help flag broader data quality issues, as Aroyo and Welty [2] have argued. A view into corporate work practices and market demands can offer a broader perspective to this line of research [78].

Rather than technically correcting bias, this commentary is a call to "study up" machine learning data, that is, to interrogate the set of power relations that inscribe specific forms of knowledge in machine learning datasets. CSCW and HCI offer good examples of how different power conceptualizations can help broaden the study of socio-technical systems. For instance, scholars have drawn on feminist [7, 28, 65] and postcolonial [51, 76] theories to ask "Who" questions and make visible power dynamics in technoscientific discourses, highlighting their political nature. Further examples are the exploration of data annotation practices and meaning imposition through the Bourdieusian concept of *Symbolic Power* [62], the exploration of how race and gender are constructed in computer vision systems by Scheuerman et al. [86], and the work of Kannabiran and Petersen [53] who use Foucault's notion of power to discuss biases in user interfaces.

Our call also includes considering data workers as allies and assets in the quest for producing better and more just data, instead of portraying them as bias-carrying hazards. It means asking ourselves, "how is AI shifting power?" [52] rather than "how can workers' biases be mitigated?" Practitioners and researchers would do good by reflecting on power asymmetries that are inherent to creating data if the goal is accounting for "biased" data but, most importantly, for unjust socio-technical systems. Despite the abundant work (including several examples cited here) that has shown how power differentials shape data and data work, a number of investigations within our research community still direct their efforts towards mitigating biases in data work and crowdsourcing without considering the experiences and conditions of workers. Therefore, we insist on the need to foster interdisciplinary dialog. Both lines of research — the study of power and the study of bias in ML data production — co-exist in parallel within CSCW and HCI. It is our hope that this commentary will prompt conversations that lead to more collaboration and, ultimately, to the advancement and broadening of this field of inquiry.

3.1 How and Why Study Up Data?

We conclude by proposing a power-oriented research agenda to study ML data along three interrelated lines:

First, we propose conducting more qualitative and ethnographic research on data workers and data work production: Who are data workers? In what contexts do they perform data work? Specifically, what are the workflows, corporate infrastructures and cultures, and intra- and inter-organizational collaboration in data production contexts? How do these contexts affect data workers and dataset production? Restoring and exposing data work settings can further make explicit the assumptions, norms, and values that inform and are inscribed in data work, allowing the "arenas of voice" [16, 89] and ethical considerations of workers [79] to emerge. In this sense, we argue that a deeper investigation into data workers and data work production cannot be achieved through mere quantitative measures and necessitates qualitative and exploratory research as well as the expertise of social scientists.

Second, we propose "shift[ing] the gaze upward" [5] and studying the actors who commission the creation of machine learning datasets: Who are data work requesters? What are their needs and wants? What rationale and priorities do they inscribe in data work tasks? What are the organizational forces driving them to produce and request data in specific ways? How do their needs and demands affect data workers' labor conditions? Investigating the role of ML practitioners commissioning data-related tasks could help to explore the collaborative nature of data work and would see requesters as co-designers of data, and not as mere consumers. Here, too, it is important to look into the organizational settings in which the work of model developers is embedded. Drawing attention to data work requesters and their organizations can therefore reveal the service relationships, market logics, and the resulting power asymmetries that shape data work and, thereby, data.

Finally, we propose expanding data documentating research and existing documentation frameworks: How can data documentation become sensitive to power relations and data production contexts? What would such a data documentation framework look like? How could organizations be incentivized to adopt such a documentation approach? How can we go beyond recognizing the power imbalances inscribed in data work and take action to bridge the power gap? Recognizing and investigating power relations are the initial steps to challenge them [28]. In this sense, a power-oriented data documentation framework can be one of the tools to render power — and its imbalances — visible in data work. In line with previous research [38, 60, 63], we argue that documentation frameworks should be grounded on the needs of workers, be integrated into existing workflows and organizational infrastructure, and have the flexibility to accommodate specific work scenarios.

ACKNOWLEDGMENTS

Funded by the German Federal Ministry of Education and Research (BMBF) – Nr 16DII113, the International Development Research Centre of Canada, and the Schwartz Reisman Institute for Technology and Society. We would like to acknowledge the data workers that have shared their knowledge and experience with us so that we could develop the ideas outlined in these pages. Special thanks to Alex Hanna, Bettina Berendt, and our anonymous reviewers for providing insightful comments that helped us strengthen our argument.

REFERENCES

[1] Ali, M., Sapiezynski, P., Bogen, M., Korolova, A., Mislove, A., and Rieke, A. Discrimination Through Optimization: How Facebook's Ad Delivery Can Lead to Biased Outcomes. *Proc. ACM Hum.-Comput. Interact. 3*, CSCW (Nov. 2019), 199:1–199:30. Article 199.

[2] Aroyo, L., and Welty, C. Truth Is a Lie: Crowd Truth and the Seven Myths of Human Annotation. *AI Magazine 36*, 1 (Mar. 2015), 15.

[3] Artstein, R., and Poesio, M. Bias decreases in proportion to the number of annotators. In *Proceedings of FG-MoL 2005 : the 10th Conference on Formal Grammar and the 9th Meeting on Mathematics of Language, Edinburgh, 5–7 August, 2005* (2005), pp. 139–148.

[4] Baker, P., and Potts, A. 'Why do white people have thin lips?' Google and the perpetuation of stereotypes via auto-complete search forms. *Critical Discourse Studies 10*, 2 (May 2013), 187–204.

[5] Barabas, C., Doyle, C., Rubinovitz, J., and Dinakar, K. Studying up: reorienting the study of algorithmic fairness around issues of power. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency* (Barcelona, Spain, 2020), FAT* '20, Association for Computing Machinery, pp. 167–176.

[6] Barbosa, N. M., and Chen, M. Rehumanized Crowdsourcing: A Labeling Framework Addressing Bias and Ethics in Machine Learning. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow Scotland Uk, May 2019), ACM, pp. 1–12.

[7] Bardzell, S., and Bardzell, J. Towards a feminist HCI methodology: social science, feminism, and HCI. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Vancouver BC Canada, May 2011), ACM, pp. 675–684.

[8] Bender, E. M., and Friedman, B. Data Statements for Natural Language Processing: Toward Mitigating System Bias and Enabling Better Science. *Transactions of the Association for Computational Linguistics 6* (2018), 587–604.

[9] Benjamin, R. *Race After Technology: Abolitionist Tools for the New Jim Code*, 1. edition ed. Polity, Medford, MA, June 2019.

[10] Berendt, B. AI for the Common Good?! Pitfalls, challenges, and ethics pen-testing. *Paladyn, Journal of Behavioral Robotics 10*, 1 (Jan. 2019), 44–65.

[11] Birhane, A., Kalluri, P., Card, D., Agnew, W., Dotan, R., and Bao, M. The Values Encoded in Machine Learning Research. *arXiv:2106.15590 [cs]* (June 2021). arXiv: 2106.15590.

[12] boyd, d. How an Algorithmic World Can Be Undermined, 2018.

[13] Brodley, C. E., and Friedl, M. A. Identifying Mislabeled Training Data. *Journal of Artificial Intelligence Research 11* (Aug. 1999), 131–167.

[14] Browne, S. *Dark Matters: On the Surveillance of Blackness.* Duke University Press, Durham, NC, 2015.

[15] Buolamwini, J., and Gebru, T. Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification. In *Proceedings of the 1st Conference on Fairness, Accountability and Transparency* (2018), vol. 81, PMLR, pp. 77–91.

[16] Casilli, A. A. Digital labor studies go global: Toward a digital decolonial turn. *International Journal of Communication 11* (2017), 3934–3954.

[17] Casilli, A. A., and Posada, J. The Platformisation of Labor and Society. In *Society and the Internet*, M. Graham and W. H. Dutton, Eds., vol. 2 ed. Oxford University Press, Oxford, 2019.

[18] Casilli, A. A., Tubaro, P., Le Ludec, C., Coville, M., Besenval, M., Mouhtare, T., and Wahal, E. *Le Micro-Travail en France. Derrière l'automatisation de nouvelles précarités au travail ?* Projet DiPLab « Digital Platform Labor », Paris, 2019.

[19] Chancellor, S., Baumer, E. P., and De Choudhury, M. Who is the "human" in human-centered machine learning: The case of predicting mental health from social media. *Proceedings of the ACM on Human-Computer Interaction 3*, CSCW (2019).

[20] Chen, L., Ma, R., Hannák, A., and Wilson, C. Investigating the Impact of Gender on Rank in Resume Search Engines. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC Canada, Apr. 2018), ACM, pp. 1–14.

[21] Cheng, J., and Cosley, D. How annotation styles influence content and preferences. In *Proceedings of the 24th ACM Conference on Hypertext and Social Media - HT '13* (Paris, France, 2013), Association for Computing Machinery, pp. 214–218. tex.ids: cheng2013a.

[22] Couldry, N., and Mejias, U. A. Data Colonialism: Rethinking Big Data's Relation to the Contemporary Subject. *Television & New Media 20*, 4 (May 2019), 336–349.

[23] Crawford, K., and Paglen, T. Excavating AI: The Politics of Images in Machine Learning Training Sets, Sept. 2019. tex.ids: zotero-3263.

[24] Dafoe, A., Bachrach, Y., Hadfield, G., Horvitz, E., Larson, K., and Graepel, T. Cooperative AI: machines must learn to find common ground. *Nature 593*, 7857 (may 2021), 33–36.

[25] Danielle K. Citron. The Scored Society: Due Process for Automated Predictions. *Washington Law Review 89*, 1 (Mar.

2014), 1–33.
[26] Davis, H. A Dataset is a Worldview, Mar. 2020. Library Catalog: towardsdatascience.com.
[27] Denton, E., Hanna, A., Amironesei, R., Smart, A., Nicole, H., and Scheuerman, M. K. Bringing the People Back In: Contesting Benchmark Machine Learning Datasets. *arXiv:2007.07399 [cs]* (July 2020). arXiv: 2007.07399.
[28] D'Ignazio, C., and Klein, L. F. *Data feminism.* Strong ideas series. The MIT Press, Cambridge, Massachusetts, 2020.
[29] Dixon, L., Li, J., Sorensen, J., Thain, N., and Vasserman, L. Measuring and Mitigating Unintended Bias in Text Classification. In *Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society - AIES '18* (New Orleans, LA, USA, 2018), ACM Press, pp. 67–73.
[30] Eickhoff, C. Cognitive Biases in Crowdsourcing. In *Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining* (Marina Del Rey CA USA, Feb. 2018), ACM, pp. 162–170.
[31] Ekbia, H., and Nardi, B. Heteromation and its (dis)contents: The invisible division of labor between humans and machines. *First Monday* (May 2014).
[32] Ekbia, H., and Nardi, B. The political economy of computing: the elephant in the HCI room. *Interactions 22*, 6 (Oct. 2015), 46–49.
[33] Ekbia, H., and Nardi, B. Social Inequality and HCI: The View from Political Economy. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (San Jose California USA, May 2016), ACM, pp. 4997–5002.
[34] Fan, S., Gadiraju, U., Checco, A., and Demartini, G. CrowdCO-OP: Sharing Risks and Rewards in Crowdsourcing. *Proceedings of the ACM on Human-Computer Interaction 4*, CSCW2 (Oct. 2020), 1–24.
[35] Feinberg, M. A Design Perspective on Data. In *CHI '17: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA, 2017), CHI '17, Association for Computing Machinery, pp. 2952–2963. tex.ids: feinberg2017a.
[36] Finin, T., Murnane, W., Karandikar, A., Keller, N., Martineau, J., and Dredze, M. Annotating named entities in Twitter data with crowdsourcing. In *Proceedings of the NAACL HLT 2010 Workshop on Creating Speech and Language Data with Amazon's Mechanical Turk* (Los Angeles, California, June 2010), CSLDAMT '10, Association for Computational Linguistics, pp. 80–88.
[37] Forsythe, D. E. *Studying Those Who Study Us: An Anthropologist in the World of Artificial Intelligence.* Stanford University Press, Stanford, 2001.
[38] Gebru, T., Morgenstern, J., Vecchione, B., Vaughan, J. W., Wallach, H., Daumé III, H., and Crawford, K. Datasheets for Datasets. *arXiv:1803.09010 [cs]* (Mar. 2020). arXiv: 1803.09010.
[39] Geva, M., Goldberg, Y., and Berant, J. Are We Modeling the Task or the Annotator? An Investigation of Annotator Bias in Natural Language Understanding Datasets. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)* (Hong Kong, China, 2019), Association for Computational Linguistics, pp. 1161–1166.
[40] Ghai, B., Liao, Q. V., Zhang, Y., and Mueller, K. Measuring Social Biases of Crowd Workers using Counterfactual Queries.
[41] Gray, M. L., and Suri, S. *Ghost Work: How to Stop Silicon Valley from Building a New Global Underclass.* Houghton Mifflin Harcourt, Boston, May 2019.
[42] Green, B. Data Science as Political Action: Grounding Data Science in a Politics of Justice. SSRN Scholarly Paper ID 3658431, Social Science Research Network, Rochester, NY, July 2020.
[43] Haimson, O. L., and Hoffmann, A. L. Constructing and enforcing "authentic" identity online: Facebook, real names, and non-normative identities. *First Monday* (June 2016).
[44] Hamidi, F., Scheuerman, M. K., and Branham, S. M. Gender Recognition or Gender Reductionism?: The Social Implications of Embedded Gender Recognition Systems. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC Canada, Apr. 2018), ACM, pp. 1–13.
[45] Hanna, A., Denton, E., Smart, A., and Smith-Loud, J. Towards a Critical Race Methodology in Algorithmic Fairness. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency* (Barcelona, Spain, 2020), FAT* '20, Association for Computing Machinery, pp. 501–512. tex.ids: hanna2020a.
[46] Hannák, A., Wagner, C., Garcia, D., Mislove, A., Strohmaier, M., and Wilson, C. Bias in Online Freelance Marketplaces: Evidence from TaskRabbit and Fiverr. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing* (Portland Oregon USA, Feb. 2017), ACM, pp. 1914–1933.
[47] Hara, K., Adams, A., Milland, K., Savage, S., Callison-Burch, C., and Bigham, J. P. A Data-Driven Analysis of Workers' Earnings on Amazon Mechanical Turk. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC Canada, Apr. 2018), ACM, pp. 1–14.
[48] Holland, S., Hosny, A., Newman, S., Joseph, J., and Chmielinski, K. The Dataset Nutrition Label: A Framework To Drive Higher Data Quality Standards. *arXiv:1805.03677* (2018).
[49] Hube, C., Fetahu, B., and Gadiraju, U. Understanding and Mitigating Worker Biases in the Crowdsourced Collection of Subjective Judgments. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (New York,

NY, USA, 2019), CHI '19, Association for Computing Machinery, pp. 1–12. tex.ids: hube2019a event-place: Glasgow, Scotland Uk.

[50] Hunter-Syed, A., and Gebru, T. Timnit Gebru on Algorithmic Bias & Data Mining Ethics, Apr. 2020.

[51] Irani, L. C., and Silberman, M. S. Turkopticon: interrupting worker invisibility in amazon mechanical turk. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Paris, France, Apr. 2013), CHI '13, Association for Computing Machinery, pp. 611–620.

[52] Kalluri, P. Don't ask if artificial intelligence is good or fair, ask how it shifts power. *Nature 583*, 7815 (July 2020), 169–169. Number: 7815 Publisher: Nature Publishing Group.

[53] Kannabiran, G., and Petersen, M. G. Politics at the interface: a Foucauldian power analysis. In *Proceedings of the 6th Nordic Conference on Human-Computer Interaction Extending Boundaries - NordiCHI '10* (Reykjavik, Iceland, 2010), ACM Press, p. 695.

[54] Kay, M., Matuszek, C., and Munson, S. A. Unequal Representation and Gender Stereotypes in Image Search Results for Occupations. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (Seoul Republic of Korea, Apr. 2015), ACM, pp. 3819–3828.

[55] Kazimzade, G., and Miceli, M. Biased Priorities, Biased Outcomes: Three Recommendations for Ethics-oriented Data Annotation Practices. In *Proceedings of the AAAI/ACM Conference on Artificial Intelligence, Ethics, and Society.* (New York, NY, USA, Feb. 2020), AIES '20, Association for Computing Machinery, pp. 1–7. tex.ids: kazimzade2020a, kazimzade2020b.

[56] Kitchin, R. *The data revolution: big data, open data, data infrastructures & their consequences.* SAGE Publications, Los Angeles, California, 2014. OCLC: ocn871211376 tex.ids: kitchin2014a.

[57] Kogan, M., Halfaker, A., Guha, S., Aragon, C., Muller, M., and Geiger, S. Mapping Out Human-Centered Data Science: Methods, Approaches, and Best Practices. In *Companion of the 2020 ACM International Conference on Supporting Group Work* (Sanibel Island Florida USA, Jan. 2020), ACM, pp. 151–156.

[58] Kulshrestha, J., Eslami, M., Messias, J., Zafar, M. B., Ghosh, S., Gummadi, K. P., and Karahalios, K. Quantifying Search Bias: Investigating Sources of Bias for Political Searches in Social Media. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing* (New York, NY, USA, 2017), CSCW '17, Association for Computing Machinery, pp. 417–432. tex.ids: kulshrestha2017a event-place: Portland, Oregon, USA.

[59] Lee, M. S. A., and Floridi, L. Algorithmic Fairness in Mortgage Lending: from Absolute Conditions to Relational Trade-offs. *Minds and Machines 31*, 1 (Mar. 2021), 165–191.

[60] Madaio, M. A., Stark, L., Wortman Vaughan, J., and Wallach, H. Co-Designing Checklists to Understand Organizational Challenges and Opportunities around Fairness in AI. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA, Apr. 2020), CHI '20, Association for Computing Machinery, pp. 1–14. tex.ids: madaio2020a, madaio2020b.

[61] Martin, D., Hanrahan, B. V., O'Neill, J., and Gupta, N. Being a turker. In *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing* (Baltimore Maryland USA, Feb. 2014), ACM, pp. 224–235.

[62] Miceli, M., Schuessler, M., and Yang, T. Between Subjectivity and Imposition: Power Dynamics in Data Annotation for Computer Vision. *Proceedings of the ACM on Human-Computer Interaction 4*, CSCW2 (Oct. 2020), 1–25.

[63] Miceli, M., Yang, T., Naudts, L., Schuessler, M., Serbanescu, D., and Hanna, A. Documenting Computer Vision Datasets: An Invitation to Reflexive Data Practices. In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (Virtual Event Canada, Mar. 2021), ACM, pp. 161–172.

[64] Mitchell, M., Wu, S., Zaldivar, A., Barnes, P., Vasserman, L., Hutchinson, B., Spitzer, E., Raji, I. D., and Gebru, T. Model Cards for Model Reporting. In *Proceedings of the Conference on Fairness, Accountability, and Transparency* (2019), FAT* '19, Association for Computing Machinery, pp. 220–229.

[65] Muller, M. Feminism asks the "Who" questions in HCI. *Interacting with Computers 23*, 5 (Sept. 2011), 447–449.

[66] Muller, M., Aragon, C., Guha, S., Kogan, M., Neff, G., Seidelin, C., Shilton, K., and Tanweer, A. Interrogating Data Science. In *Conference Companion Publication of the 2020 on Computer Supported Cooperative Work and Social Computing* (Virtual Event USA, Oct. 2020), ACM, pp. 467–473.

[67] Muller, M., Feinberg, M., George, T., Jackson, S. J., John, B. E., Kery, M. B., and Passi, S. Human-Centered Study of Data Science Work Practices. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow Scotland Uk, May 2019), ACM, pp. 1–8.

[68] Muller, M., Lange, I., Wang, D., Piorkowski, D., Tsay, J., Liao, Q. V., Dugan, C., and Erickson, T. How Data Science Workers Work with Data: Discovery, Capture, Curation, Design, Creation. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk, 2019), CHI '19, Association for Computing Machinery, pp. 1–15.

[69] Muller, M., Wolf, C. T., Andres, J., Ashktorab, Z., Joshi, N. N., Desmond, M., Sharma, A., Brimijoin, K., Pan, Q., Duesterwald, E., and Dugan, C. Designing Ground Truth and the Social Life of Labels. 17.

[70] Noble, S. U. *Algorithms of Oppression: How Search Engines Reinforce Racism.* NYU Press, New York, 2018.

[71] Olteanu, A., Castillo, C., Diaz, F., and Kiciman, E. Social Data: Biases, Methodological Pitfalls, and Ethical Boundaries. *SSRN Electronic Journal* (2016). tex.ids: olteanu, olteanu2016a, olteanu2016b, olteanua.

[72] Passi, S., and Barocas, S. Problem Formulation and Fairness. In *Proceedings of the Conference on Fairness, Accountability, and Transparency* (Atlanta, GA, USA, 2019), FAT* '19, Association for Computing Machinery, pp. 39–48.

[73] Passi, S., and Jackson, S. Data Vision: Learning to See Through Algorithmic Abstraction. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing* (Portland, Oregon, USA, 2017), CSCW '17, Association for Computing Machinery, pp. 2436–2447.

[74] Passi, S., and Jackson, S. J. Trust in Data Science: Collaboration, Translation, and Accountability in Corporate Data Science Projects. *Proc. ACM Hum.-Comput. Interact. 2*, CSCW (Nov. 2018), 1–28.

[75] Paullada, A., Raji, I. D., Bender, E. M., Denton, E., and Hanna, A. Data and its (dis)contents: A survey of dataset development and use in machine learning research. *arXiv:2012.05345 [cs]* (Dec. 2020). arXiv: 2012.05345.

[76] Philip, K., Irani, L., and Dourish, P. Postcolonial Computing: A Tactical Survey. *Science, Technology, & Human Values 37*, 1 (Jan. 2012), 3–29.

[77] Pine, K. H., and Liboiron, M. The Politics of Measurement and Action. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (New York, NY, USA, 2015), CHI '15, Association for Computing Machinery, pp. 3147–3156. event-place: Seoul, Republic of Korea.

[78] Posada, J. The Future of Work Is Here: Toward a Comprehensive Approach to Artificial Intelligence and Labour. *Ethics of AI in Context* (2020).

[79] Posada, J. Unbiased: AI Needs Ethics from Below. In *New AI Lexicon*, N. Raval, A. Kak, and L. Strathman, Eds. AI Now Institute, New York, NY, 2021.

[80] Powles, J., and Nissenbaum, H. The Seductive Diversion of 'Solving' Bias in Artificial Intelligence, Dec. 2018.

[81] Ramanath, R., Choudhury, M., Bali, K., and Roy, R. S. Crowd Prefers the Middle Path: A New IAA Metric for Crowdsourcing Reveals Turker Biases in Query Segmentation. In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)* (Sofia, Bulgaria, 2013), Association for Computational Linguistics, pp. 1713–1722.

[82] Roberts, S. T. *Behind the Screen: Content Moderation in the Shadows of Social Media.* Yale University Press, New Haven, CT, 2019.

[83] Ross, J., Irani, L., Silberman, M. S., Zaldivar, A., and Tomlinson, B. Who are the crowdworkers?: shifting demographics in mechanical turk. In *CHI '10 Extended Abstracts on Human Factors in Computing Systems* (Atlanta Georgia USA, Apr. 2010), ACM, pp. 2863–2872.

[84] Salehi, N., Irani, L. C., Bernstein, M. S., Alkhatib, A., Ogbe, E., Milland, K., and Clickhappier. We Are Dynamo: Overcoming Stalling and Friction in Collective Action for Crowd Workers. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (Seoul Republic of Korea, Apr. 2015), ACM, pp. 1621–1630.

[85] Sampat, R. Protesters target Facebook's 'real name' policy, June 2015.

[86] Scheuerman, M. K., Wade, K., Lustig, C., and Brubaker, J. R. How We've Taught Algorithms to See Identity: Constructing Race and Gender in Image Databases for Facial Analysis. *Proc. ACM Hum.-Comput. Interact. 4*, CSCW1 (2020). Article 058.

[87] Seaver, N. Studying Up: The Ethnography of Technologists, Mar. 2014.

[88] Seidelin, C. *Towards a Co-design Perspective on Data : Foregrounding Data in the Design and Innovation of Data-based Services.* Ph.D. thesis, IT-Universitetet i København, 2020.

[89] Star, S. L., and Strauss, A. Layers of Silence, Arenas of Voice: The Ecology of Visible and Invisible Work. *Computer Supported Cooperative Work 8*, 1-2 (Mar. 1999), 9–30.

[90] Tubaro, P., and Casilli, A. A. Micro-work, artificial intelligence and the automotive industry. *Journal of Industrial and Business Economics* (2019).

[91] Wauthier, F. L., and Jordan, M. I. Bayesian Bias Mitigation for Crowdsourcing. In *Proceedings of the 24th International Conference on Neural Information Processing Systems* (Granada, Spain, 2011), NIPS'11, Curran Associates Inc., pp. 1800–1808.

[92] Woodcock, J., and Graham, M. *The Gig Economy: A Critical Introduction.* Polity Press, London, 2020.

[93] Yano, T., Resnik, P., and Smith, N. A. Shedding (a Thousand Points of) Light on Biased Language. In *Proceedings of the NAACL HLT 2010 Workshop on Creating Speech and Language Data with Amazon's Mechanical Turk* (Los Angeles, 2010), Association for Computational Linguistics, pp. 152–158.

[94] Zhang, e. a. How do data science workers collaborate? roles, workflows, and tools.