AI & SOCIETY
https://doi.org/10.1007/s00146-024-02162-0

MAIN PAPER

“Everybody knows what a pothole is”: representations of work and intelligence in AI practice and governance

SJ Bennett[1] · Benedetta Catanzariti[2] · Fabio Tollon[3]

Received: 2 April 2024 / Accepted: 2 December 2024

Abstract

In this paper, we empirically and conceptually examine how distributed human–machine networks of labour comprise a form of underlying intelligence within Artificial Intelligence (AI), considering the implications of this for Responsible Artificial Intelligence (R-AI) innovation. R-AI aims to guide AI research, development and deployment in line with certain normative principles, for example fairness, privacy, and explainability; notions implicitly shaped by comparisons of AI with individualised notions of human intelligence. However, as critical scholarship on AI demonstrates, this is a limited framing of the nature of intelligence, both of humans and AI. Furthermore, it dismisses the skills and labour central to developing AI systems, involving a distributed network of human-directed practices and reasoning. We argue that inequities in the agency and recognition of different types of practitioners across these networks of AI development have implications beyond RAI, with narrow framings concealing considerations which are important within broader discussions of AI intelligence. Drawing from interactive workshops conducted with AI practitioners, we explore practices of data acquisition, cleaning, and annotation, as the point where practitioners interface with domain experts and data annotators. Despite forming a crucial part of AI design and development, this type of data work is frequently framed as a tedious, unskilled, and low-value process. In exploring these practices, we examine the political role of the epistemic framings that underpin AI development and how these framings can shape understandings of distributed intelligence, labour practices, and annotators' agency within data structures. Finally, we reflect on the implications of our findings for developing more participatory and equitable approaches to machine learning applications in the service of R-AI.

Keywords Artificial intelligence · Machine learning · Labour · Automation · Intelligence · Responsible AI

1 Introduction

Artificial Intelligence is often framed as an attempt to recreate biological models of intelligence, overlooking the ways in which specific epistemic legacies and supply chains of labour shape contemporary AI practice. We argue that conceptual framings of AI intelligence structure AI tasks and shape the valuation of different social actors throughout AI development. In our analyses we draw upon alternate framings of AI which account for its history and development by examining how it has embedded and concretised dominant labour logics since its very inception in the 1950s. Bringing findings from our empirical study in dialogue with the rich body of critical scholarship on AI and data work, we argue that the classification of skill and intelligence are central to the maintenance of these labour dynamics. To do so, we look at the practical contexts of this classification and division of labour—the points of contact between practitioners and data workers—and illustrate how practitioners' own framing of skill and intelligence propagate the labour and power structures that underpin it. We examine how narrow representations of intelligence feed into hegemonic power dynamics, shape the nature of tasks and instantiate epistemic inequities, resulting in narrow 'intelligence' by design.

✉ Benedetta Catanzariti
benedetta.catanzariti@ed.ac.uk

SJ Bennett
sj.bennett@durham.ac.uk

Fabio Tollon
ftollon@ed.ac.uk

[1] Durham University, Geography, Durham, UK

[2] Science, Technology and Innovation Studies, University of Edinburgh, Edinburgh, UK

[3] University of Edinburgh, Philosophy, Edinburgh, UK

Our findings explore how epistemic, technical and socio-cultural expectations shape perceptions of data workers and data work tasks. We examine how assumptions about the presumed universality of AI models and associated sub-tasks can impose cultural understandings upon data in the abstract. However, these translate poorly into the reality of the data and associated work. Given this, practitioner assumptions regarding the complexity (or the simplicity) of a task, amount of work and number of iterations required to complete it, may poorly account for the messy and ambiguous nature of data work, in a mismatch between the simplistic, abstracted biological framings of intelligence shaping task design, and the distributed intelligence underpinning its execution. Thus, we borrow Hallam Stevens' use of the image of *network* (Stevens 2013) to propose it as an alternative to that of a linear and frictionless AI pipeline, arguing that this can better account for complexity whilst avoiding implicit hierarchies of labour valuation and indeed inclusion/exclusion of tasks/roles. In his account of bioinformaticians' methods and processes involved in turning DNA into sequence data, Stevens notes how the spatial metaphor of the pipeline obscures the situated judgements and sometimes messy practices necessary to make sense of data. As he notes, "pipelines are directional, the water or the data must flow down the pipe, following the single path laid out for them and never moving side to side or going backward" (p. 103). A "densely connected network" (p. 15), on the other hand, better represents the myriad ways in which scientists interpret and articulate DNA data. Similarly, we argue that, in machine learning, the image of the pipeline serves to conceal the fragmented, iterative, and highly changeable nature of data work. Paraphrasing Stevens, imagining AI development as a linear process normalises the idea that machine learning practice "is an automatic, black-boxable activity" (p. 110). Drawing upon our findings, we then introduce the concept of *representation coils,* or looping effects with downstream implications in which these assumptions shape and concretise labour dynamics. Within these coils, the interaction between abstract representations of task complexity and data worker competency, and their messy translation into practice, subtly reinforce dominant epistemologies and power dynamics.

Drawing upon this study of AI practitioners' perspectives on data workers and practices of data acquisition, cleaning, and annotation—or "data wrangling" (Muller et al. 2019)—we illustrate the importance of broader framings of intelligence, examining the critical role of AI practice and its entanglements with the supply chains and labour which enable the AI algorithms commonly marketed as 'intelligent'. Accounting for these networks of labour and skill helps us identify how high-level framings of AI materially impact the practical construction of sub-tasks in ways which can be narrow and inequitable. Reflecting on the findings of our study, we consider the impacts of the narrow framings of intelligence commonly employed in discussions of AI, particularly regarding the downstream impacts of these deficits in studying and governing AI. We illustrate how sidelining the impact of labour networks creates the inequities for which narrow 'intelligence' is often criticised. We then suggest that research into this area can be useful to inform the field of Responsible Innovation, contributing the conceptual tool of 'representation coils' to facilitate analysis of the complex interactions between AI framings, practices and social actors, and their ethico-political implications. This tool can inform discussions around the 'responsible' development of machine learning systems that look beyond a linear model of responsibility and help to highlight the recursive and situated nature of truly 'responsible' innovation.

2 Background

2.1 Intelligence, skill and automation

Common portrayals of the field of Artificial Intelligence (AI) typically frame it as the attempt to imitate and automate human intelligence, as canonically inaugurated at the 1956 Dartmouth workshop on AI (McCarthy et al. 2006). This research has been split into distinctive camps subscribing to different theories about how to approach the characterisation and operationalisation of intelligence. For many years, the dominant subfield of AI was symbolic learning, which focuses on the modelling of relationships between concepts and problems and programming logical rules to dictate actions. For instance, the rules that describe simple insect models of vision can provide the logic by which a robot attempts to navigate a space (de Croon et al. 2022) Meanwhile, connectionist approaches to AI sidestep the need for such granular prior knowledge, and privilege instead the algorithmic modelling of datasets which identify associations between sets of datapoints, concerned with the relationships between such datapoints rather than concepts.

A rich body of critical and historical work, however, points to the role of automation in shaping understandings of intelligence as they relate to the social classification of skill and labour (Ali et al. 2023; Schaffer 1994) and how the operationalisation of intelligence was and is shaped by the affordances of technology (Dick 2011; 2015; Jones 2016). Histories of calculations, for example, show how automation contributed to the creation of a new form of intelligence in the mid-nineteenth century that was necessary to operate the newly introduced calculating machines (Daston 1994, 2018a, 2018b; Schaffer 1994). This process of automation took advantage of and exacerbated existing inequities in its division of labour and assignation of value to types of expertise. To produce accurate results, these early calculators required in fact the vigilant attention of human operators—unmarried women who could be paid less than their male counterparts—tasked with entering numbers, punching cards, pulling levers, and checking results. Although considered tedious, this type of work could not be performed mindlessly, and instead required a form of "analytical intelligence", or *mindful drudgery*, and it underpinned the large-scale calculations necessary to advance the scientific, technical, and military projects of the nineteenth and twentieth century (Daston 2018a, 2018b). While automation had not significantly reduced human labour, the mechanisation of calculation contributed to its disqualification as an intelligent task. Automation differentiated analytical intelligence—which was associated with the tedious but mindful activity of cheap female staff—from the mental labour of the mathematical (male) genius.

More recently, Matteo Pasquinelli describes the project of AI as a continuation of this social classification of intelligence and skill. Central to this view is the recognition that all forms of labour require specialised cognitive skills learned through practice and training and that "ultimately, *all labour is logic*" (Pasquinelli 2023, 3). In this sense, Pasquinelli posits that the project of AI, especially in its current form of machine learning, has not emerged simply as the result of the automation of biological intelligence, but rather as a method to order forms of intelligence and organise labour accordingly. From Charles Babbage's *engine machine* to Rosenblatt's invention of the first artificial neural network in the 1950s, automation has reinforced the "imposition labour and knowledge hierarchies that reinforce the polarisation of skilled and unskilled workers in the job market" (21). A particularly salient example of this process is cybernetics' efforts to automate perception by demarcating visual labour—here reframed as 'pattern recognition'—from mental and manual labour. This mechanisation of perception (later extended to non-visual pattern recognition practices) has "come to be traded as the mechanisation of cognition or artificial intelligence" (164). Unearthing these historical (re) framings helps us better understand how current classifications of skill and intelligence underpinning AI development came to be. Much-needed ethnographic work has shown how the vast and distributed human intelligence, essential to the creation and maintenance of supervised machine learning models, is in fact typically framed as *unskilled, low level, and uncreative* (Sambasivan and Veeraraghavan 2022; Lilly Irani 2015a; 2015b; Kandel et al. 2012). This 'labour theory of automation' (Pasquinelli 2024) situates empirical studies of data work and data design (DiSalvo et al. 2024; Miceli and Posada 2022; Miceli et al. 2021b, 2020; Feinberg 2017) within the longer history of framings of skill and intelligence as they relate to automation, and further illuminates the power structures that govern current data practice.

The data workers that annotate the ground truth necessary to train machine learning algorithms are seen—like their nineteenth century counterparts—as computational resources, without expertise or specialised skill of their own. This definition and attribution of skill and expertise in data work is reflected in the power asymmetry of machine learning design which posits the creative work of engineers and researchers at the top of the knowledge hierarchy, maintained and reinforced through uneven labour dynamics. Because the annotation of large datasets is often outsourced to subcontracted workers, annotators' feedback and concerns are rarely taken into consideration in the model design process. As scholarship in this area has shown, the perceived intuitive nature of annotation tasks often results in labour structures that posit annotators as interchangeable and disenfranchised from the overall system design. For instance, Emily Denton et al. have traced the norms and assumptions that underpin ImageNet, a large computer vision dataset that includes 1,281,167 training images spanning 1000 object categories, labelled by crowdsourced Amazon Mechanical Turkers. As the authors have noted,

> Framing the label verification as an act that requires little reflective judgement not only suggests that anyone can participate, but that annotators are interchangeable because they share the same innate faculty of seeing objects and because they exercise vision in the same way (Denton et al. 2021, 10).

In contrast, empirical work in this area has surfaced the heterogeneous and complex nature of data annotation, which requires interpretation and translation between different forms of knowledge (Mao et al. 2019; Muller et al. 2019; Sambasivan et al. 2021). These insights have shown how data work, contrary to its framing as an unskilled activity, requires varied forms of discretionary and situated human judgment to ensure trust in algorithmic results and in the overall design process. However, a growing body of work has emphasized how the power structures that govern data practices within machine learning design development can influence and shape data annotation (Miceli et al. 2020, 2021b, 2021a).

Notably, ethnographic research on data work has pointed to the global labour dynamics that underpin many commercial machine learning models: workers, often outsourced from low-income regions or vulnerable populations, clean and label the image, video, text, and sound data used to train and maintain learning algorithms (Irani and Silberman 2013; Irani 2015b; Gray and Suri 2019). The "ghost work" behind much of today's automation (Gray and Suri 2019) is often recruited through crowdsourcing platforms (such as Amazon Mechanical Turk) or impact-sourcing annotation companies that employ workers from the Global South to offer data annotation services at competitive prices to practitioners and researchers largely based in the Global North. Labour dynamics and structures within these platforms and organizations can reinforce the power asymmetries present in the design pipeline: the social and material conditions of data work – the vulnerable status of workers and the complete alienation from their employer (and the purpose of their tasks) – might further shape the annotation process.

Finally, scholarship in this area has noted how annotation standards (the requirement and expectations of clients or managers), multiple layers of power (client, team leaders, reviewers, and annotators), and the naturalization of annotation practices (the idea that labels are "self-evident"), can all shape the interpretation of data. This "imposition of meaning" (Miceli et al. 2020) has the potential to reinforce normative assumptions around the data and, at the same time, forestall opportunities to challenge the overall system design. Once again, data workers (collectors, cleaners, and annotators) are not seen as relevant stakeholders in the design process.

2.2 Representation coils in data work

The ramifications of these "impositions of meaning" go beyond specific, immediate tasks and datasets of focus to more broadly shape future algorithmic developments, by dictating the socio-material landscapes and narratives employed in framing and developing AI projects. Recent scholarship has examined the implications of socio-technical entanglements for emerging social worlds in the 'recursive society' (Beer 2022) and how these are shaped by data coils which arise from the layered feedback loops between data-centric technology and the societal contexts of their implementation (Beer 2022). Data coils occur when data-centric technologies alter the social fabric in some way and the data captured from this augmented existence is again modelled, in recursive loops which become so difficult to separate out from each other that they become concretised in tightly wound coils. In Beer's view, salient socio-political concerns are not just "what algorithms are doing or how data is harvested" but involve how loops are formed by previous loops, "on repeat" (Beer 2022, p.2). This can be seen for example, when We draw on Beer's concept of data coils to propose the concept of representation coils or recursive looping of machine learning and data practices, building on the idea that representation of labor and tasks is subject to this coiling or layering of feedback loops, which can become so entangled as to be difficult to separate from the social fabric which they shape. In proposing this, we draw attention to multiplicity in the framing and representations of machine learning tasks and pipelines and suggest that these be considered and analysed in the study and governance of machine learning models and data work. Representation coils are shaped by power asymmetries regarding who gets to define tasks and roles, who is afforded visibility and attributed expertise, that is, who gets to 'self-represent' (Suchman 1995). In this way, much like data coils, representation coils are recursive processes which [re]produce systemic inequities such as epistemic injustice (Bhakuni and Abimbola 2021; Fricker 2007)—a form of injustice where knowledge contributions are silenced, omitted, or deemed inferior to dominant epistemic framings, due to the identity of the knowledge contributor and the tools available to them, perpetuating and even concretising these inequities (Fricker 2007; Bennett et al. 2023). Representation coils make epistemic injustice in data work and machine learning practice *visible* and can help us better understand how and where classifications of skill and intelligence are formed and maintained.

Critical scholars have indeed noted that the body of literature on data labelling often focuses on the annotators' role in reproducing bias, pointing to issues of data quality and reliability that can arise from inconsistencies in annotation work or annotators' biased perceptions of the data, without questioning the interpretations of the practitioners who define and evaluate these tasks (Miceli et al. 2021a). *Representation coils*, then, are stacked, socio-technically embedded recursive loops where the narrow interpretations that define a task or role go beyond the immediate material structuring of tasks and labor to shape the epistemic framings and space of future tasks and roles, thus illuminating novel sites of epistemic injustice. These coils are not limited to the relationship between machine learning practitioners and data annotators but exist within complex networks of relations between multiple machine learning stakeholders and tasks. For example, researchers and engineers are subject to a similar demeaning of data work when it comes to their own practice, with data work being far less rewarded than other types of development such as 'model work' (Sambasivan et al. 2021) although some practitioners even find the process rewarding (Godsey 2017).

In this way, data work is subject to friction between, on one hand, systemic devaluation and, on the other hand, the growing need for "data expertise" (Hutchinson et al. 2021). As work in critical dataset studies (Thylstrup 2022) has shown, datasets "form a foundational element of machine learning cultures" (656): training data not only shapes model performance, it also reproduces epistemic assumptions about what does and does not count as valid knowledge. The interpretation and categorization of data yields the power to make decisions about the system's outputs and, ultimately, to shape societal outcomes. Framing certain data tasks and work as less valuable or less creative facilitates certain types of epistemic contributions and forecloses upon others, similarly to how AI projects are often dictated by available data sources. In the same way, practitioners choose to tackle concepts which are (perceived as) easily modelled under dominant epistemic hierarchies, reproducing specific ideals of intelligence whilst marginalizing others. This serves to reinforce the power asymmetries present in the social and material conditions of data work and further shrink the representation coils of data work, which serves to narrow the ‘intelligence’ of the eventual models and outputs.

3 Findings

We examine the practical implications introduced by large-scale attempts to automate data work via a case study workshop with AI practitioners, which probed their expectations of data work and data annotators. Our findings illustrate how framings of data work, shaped by assumptions around the nature of AI automation and intelligence, can play a key role in constructing annotation tasks, labour practices, and annotators’ agency within data structures. Key themes within this are the role of assumed universalism, inequitable valuation of expertise and uncertain trajectories of work in influencing how practitioners conceptualize data work, maintaining inequitable representation coils with numerous downstream implications. Researchers and policymakers working in the AI space typically draw upon representations given by AI practitioners and mainstream narratives to conceptualize the nature of AI intelligence and development. However, these representations—situated within the experiences, cultural knowledge and assumptions of the practitioners involved—often overlook or misinterpret the crucial infrastructure exemplified by data work.

3.1 Methods

We investigated representations of annotation work in machine learning practice using qualitative methodology, specifically interactive online workshops. In 2021, the first two authors conducted two workshops with 8 informants, recruited internationally from both academic and industry contexts, through Twitter, LinkedIn, academic mailing lists, and by emailing contacts working at technology companies. Often, qualitative studies of scientific and technical knowledge refer to research participants as “experts” by virtue of their role as informants, differentiating between scientific and technical expertise from other forms of everyday knowledge. However, literature in the sociology of knowledge and expertise has problematized this definition to include forms of non-institutional or specialized knowledge and account for new forms of knowledge production and foregrounded the political role that attributions of expertise and authority have in decision-making processes (Epstein 1995). As argued in the paper, attributions of skill and expertise can indeed play a political role within data structures. For this reason, we use the term ‘practitioners’ to refer to our informants, in the pragmatic and literal sense of the term as “those who practice a profession or an art”. With this definition, we aim to be inclusive with respect to the various forms of knowledge of those involved in AI development. Our informants (Table 1) were based in roles working on industry products, research projects, and annotation and consultancy services, and spanned a range of disciplinary areas, including computational neuroscience, natural language processing, user experience, and digital literacy. Importantly, our informants worked at various stages of the development process, with varying degrees of interaction with data annotators. For instance, Amie, as a Data Lead at an impact-sourcing annotation company, worked closely with both annotators and engineering teams, translating and mediating requests and concerns to both groups. Conversely, machine learning engineers, such as Shawn or Antony, had only interfaced with ‘domain experts’ or had employed anonymous annotators through on-demand platforms. These differences in experiences gave us a better understanding of the various ways in which data work is represented throughout AI development.

Table 1 Table of Informants

Role	Pseudonym
AI Researcher	Abbie, Shawn, Antony, Aurora
AI Researcher and Consultant	Luc
Project Manager	Rohan
Data Lead	Amie
CEO	Lilian

This research was approved by the research ethics board of the second author’s university department. We have pseudonymised the transcripts to protect informants’ identities and use these pseudonyms wherever we employ quotes. Informants were not reimbursed for their participation.

Each workshop was 90 min long and conducted virtually, using Zoom and Miro. We designed a shared Miro board to elicit reflections regarding practitioners’ experience working with annotators. Using the Miro board, we presented informants with a machine learning workflow modelled on the findings of previous literature, prompting discussions about practitioners’ own workflows, the points at which they engaged with data work, and reflection on how collaboration and delegation played into this. We also asked them to leave suggestions for revisions or additions to the workflow. We then used these comments to prompt group discussion in Zoom breakout rooms facilitated by the authors. We recorded and then analysed this data using reflexive thematic analysis, a variant of thematic analysis which facilitates examination of patterns in the data whilst flexible enough to allow engagement with emergent themes and relate findings to "wider socio-cultural contexts" (Braun and Clarke 2012). This analysis was conducted using NVivo and printed-out transcripts, involving regular meetings to discuss the emergent themes. This study is constrained by some methodological limitations and, particularly, by issues related to sampling and scope, as these were greatly influenced by the effect of the COVID-19 pandemic on face-to-face qualitative research. However, findings from this research aim to be illustrative, rather than representative, of common representations and framings of annotation, in line with portrayals of data work as found in critical AI literature. While we observed coherence within our findings—informants offered similar accounts of their practices and assumptions—a potential avenue for future research could tease out more nuanced distinctions between, as well as within, the research and industry sectors.

3.2 Assumptions in task design and interpretation

Practitioners' situated experiences, which informed the specific examples they gave of data work practices, shaped how data work was discussed from task conceptualization to evaluation. The assumptions underpinning their practices included judgements about the sorts of information represented within datasets, and disciplinary norms which devalued certain tasks such as data work. These assumptions had *socio-cultural and technical* facets. Tensions arose when these limited viewpoints were viewed as universal.

Socio-cultural assumptions. Assumptions around the presumed universality of data practices were reflected in the imposition of cultural understandings of data in the abstract, which do not translate into the reality of the data and associated work. This is especially apparent in the assumption that the annotation of everyday objects should be universally consistent, irrespective of the cultural and geo-political context. Amie, who coordinates a team of annotators at an impact-sourcing annotation company, reflected on her experiences of dealing with unanticipated complexity and uncertainty over the meaning of the data. As she recounted, often the engineering team will send her instructions on how to annotate image data that are *"very prescriptive"* and leave no room for interpretation. However, she would often have to moderate conflicting annotations of seemingly mundane objects, such as potholes:

> *"You can think 'oh everybody knows what [a pothole] is. And it's quite simple. I could send that task to anyone. We started to do it, but then we quickly realized the dataset that we had was from Japan. And what a pothole looks like in Japan and the streets in general, was really different from, say, Canada. So, our own annotators had a hard time, and we are talking with each other, being like 'I think this is a pothole. No, I don't think so.'"*

As exemplified by this quote, annotation work is further from menial and straightforward, requiring workers to continuously negotiate and fine-tune their interpretation of the data. Conversely, practitioners viewed machine learning as a process where creative, high-level, qualified stakeholders were distinct and separate from those in roles perceived to be rote and "low-level", not necessarily reflecting how annotators conceptualised their work. However, framing annotation tasks as intuitive or self-evident raises implications for how we conceptualize agency and skill within machine learning projects.

Technical assumptions. Data practices were further complicated by a lack of shared language between technical and non-technical actors, with technical teams assuming their understandings as universal, often disregarding differences in understanding as due to lack of knowledge and expertise, rather than different types of knowledge and degrees of agency within project development. This was illustrated by Antony's reflection that *"you can have a very different idea of what's actually possible given the data",* while Amie responded that communicating these differences in understanding to the machine learning team would elicit a somehow patronising response: *"Flagging this to the machine learning engineers.. This is difficult because I see from sitting with them, when they open up an asset that is difficult, they all laugh"*. Similarly, practitioners' responses depicted a binary world of data work consisting of fundamentally competent or incompetent data workers and requiring machine learning practitioners to sift them out (rather than considering whether the issue is communication or the nature of the task). As Aurora recounted, this is particularly relevant in the case of workers crowdsourced through on-demand platforms:

> *So, if you are referring maybe to AMT [Amazon Mechanical Turk], I think it's really important before you start. For example, surveys to make some qualification tasks for your annotators, and it's really, really important to give some attention, maybe think about something else that will catch the good annotators. Because it's so problematic."*

Amie, who works at the intersection between annotators and machine learning engineers, offered a contrast of expertise and skill in data work, constructing boundaries to distinguish between expert and lay annotators. The boundaries she drew reflected broader societal judgements of how types of skill and domain are perceived.

> *"It's interesting because these are both different types of annotators. If I work for an insurance company and I'm building a model to assess the risk of a client, I'm*

probably going to use my underwriters to label some data for me and get me the knowledge so I can train my models versus if I work in agriculture and I want to map out fields and where they stop.. Well, I can send that task to the gig economy and ask, like, a hundred agents to sort of label the field distance and how to distribute it."

Universalist assumptions strike again here. There is a specialist knowledge underlying what seem intuitively simple tasks easily completed by 'lay' annotators, judged as able to contribute the knowledge required to annotate agricultural data despite the complexities that characterise such tasks. These assumptions reflect the above-mentioned historical classifications of tasks and associated forms (and values) of intelligence, with vision tasks typically perceived as possible and likely to be automated (in this case, the identification of potholes or field borders), as opposed to tasks that rely on expert knowledge. As shown in the next section, assumptions around the motivations behind annotators' work—namely, the lack of integrity and low interest in the quality of the annotation—might contribute to the creation of work structures that treat annotators as replaceable workers, simultaneously overprescribing rules whilst ignoring the true complexity of their work.

3.3 Structural inequities stratify representations of skill and labour

Representational asymmetries between practitioners and data workers seemed to be further complicated by implicit assumptions held about annotators in different geographical and socio-economic contexts, and how these factors impacted their assumed levels of 'data expertise'. Practitioners described the experiences of "expert" annotators (such as physicians annotating medical image data or legal experts coding documents for law firms) who are *"extremely high paid and sought after"*, as radically different to that of annotators recruited in lower income geographies. Antony was keen to draw boundaries between categories of annotators before offering more in-depth views, asking *"are we concerned really with low-level implementation or more high-level stakeholders as well?"* These categorizations serve to implicitly entrench existing structural inequities, by reinforcing associations between the value offered by actors in global minority countries, and vice versa. The power differential between these two annotator groups further stratifies collaboration. At the same time, assumptions and expectations around crowdsourced workers lacking integrity and motivation seemed to be most persistent among practitioners. For Antony, the disregard for data quality was a direct reflection of the labour structures that sustain on-demand platforms such as Amazon Mechanical Turk:

"There seems to be a conflict between what the annotators are doing and what the machine learning people are doing because obviously if you're working on Mechanical Turk and you are paid \$0.25 or \$0.20 or whatever it happens to be per click or per annotation, then you want to do as many as you can. And then sometimes there's a conflict between that and then what Google wants, which is maybe to have high quality base. And so, there's an extra kind of thing where you have to ensure that quality and it becomes much, much harder to do the quality because people don't really care about it. They want to do it as quickly as possible and sort of vice versa. The machine learning engineers don't want to check up on what people are doing to make sure they get the quality."

Of course, this perception does not necessarily reflect the reality of crowdsourced work. Ethnographic research in this area has shown how many on-demand data 'wranglers' depend on crowdsourced work as a primary means of income and view themselves as skilled and creative workers (Gray and Suri 2019). However, the ways in which practitioners and organizations conceptualize this work can contribute to representation of data annotation work as rote and low skill, perhaps in a self-perpetuating culture. These inequitable valuations do not account for the complexity and highly involved nature of tasks deemed more menial due to reduced cultural value. In addition to stemming from disciplinary and identity-based assumptions, this devaluation of tasks reproduces the bifurcation of roles into rote or intelligent, as examined in genealogies of computational labour (Daston 1994, 2018a, 2018b; Pasquinelli 2023). Our findings illustrate how this process of devaluation is underpinned by techno-colonialist structures (Madaio et al. 2020), which "re/produce power/knowledge relations in data and labour" (Miceli and Posada 2022, 1), particularly given the precarity of the conditions of data work, which exacerbate power inequities. Reductive framings of data work can result in representation coils where the narrow framing of labour and tasks simultaneously creates and perpetuates further knowledge asymmetries and underspecified tasks.

3.4 Seamless pipelines or networks of labour?

A hidden factor contributing to this complexity and shaping representations was the unpredictable nature of the work. As discussed, practitioner assumptions regarding the complexity (or the simplicity) of an annotation task often bear little relation to the nature of the work that will be necessary to complete it. This also rings true regarding the amount of work, and number of iterations, required to complete the work. Despite being commonly portrayed as linear and frictionless, machine learning development is fragmented, unsystematic, and highly changeable. In the words of Lilian, *"there is a lot of looping back and forth, and a lot of activities happen more in parallel than sequentially"*. These idiosyncrasies can impact annotators in the course of their own work. As Amie recounted, machine learning workflows often do not allow the time necessary for annotators to build annotation expertise:

> *"The instructions, they change every two weeks with the machine learning engineers discovering new things as they train and test. So, they send updates and they're like, 'oh, now we're going to switch and do this like that and change the annotation to this.' Then, this has to be communicated to annotators, who just learned to master one specific way of things. Like, 'you know what? Forget about it. We're switching to this new way of doing things.' But the gap is really always to understand, why are we changing every two weeks? Why are we constantly iterating on the instructions? And it's often hard to keep it up to pace."*

In addition, negotiations with clients and end-users around model evaluation (*"is the model actually solving a problem?"*) can intervene at any stage of the model development. This, as Lilian described, can sometimes result *"in a complete redo of the entirety of the previous bits of the pipeline!"* This unpredictability, according to Amie, could be addressed by improving transparency and visibility throughout AI development, as well as fostering active collaboration between annotators and engineers. This requires, for instance, that engineers provide annotators with information about problem definitions, model developments, and outputs. Similarly, annotators' feedback on data properties and differences in interpretation (*"What data the annotators have seen in a week. What was hard and what was easy. When did they ping each other to ask, 'how do you do that one?'"*) should be considered *"hard assets"* for the engineering teams.

These very framings of the tools shaping datafication are imbued with power inequities, drawing on the assumptions and values of the practitioners who propose popular representations of machine learning pipelines. These inequities are underpinned by implicit positionings of practitioner knowledge and contexts as universal and more valuable, cascading to minimize the expertise and material contributions of other stakeholders in networks of machine learning development. This partly stems from perceptions of annotators as providing a discrete service, whereas this does not hold true for all annotators, who might view themselves as creative workers, contributing to the overall process and outputs. Stereotyping representations of data workers as unengaged and unmotivated also shape the ways in which tasks and roles are constructed, impacting both the worker and the data work itself.

As suggested by critical scholars, reframing of data work is not simply a matter of making annotators visible to practitioners but, instead, of making "the rest of the machine learning supply chain" visible to annotators (Miceli and Posada 2022, 30). In doing so, researchers and engineers must recognize annotators' contribution to the overall model development, both conceptually and economically, and "regard data workers as tech workers as much as we do when we think of engineers" (31). In this regard, Rohan spoke of processes of engaging with data workers and other non-technical practitioners as a form of translation work:

> *"The first thing that you can't deny is that the individuals who are going to be doing this* [annotation] *are external to your domain knowledge, your workflow, your expertise. So, you end up doing translation work. That's effectively what it is. It's translation work. And that is really, really both challenging but also beneficial, because one of the first things that translation work does is it shows you the extent to which you are removed from, you know, like the thinking and the needs of these other people. So, everything that you've taken for granted in your domain of expertise, just to be able to function every day and talk to other people and be heard and whatever all of that."*

Nonetheless, both machine learning practitioners and annotators are grappling with barriers impacting their goals for work. This is partly due to external factors; it might be financially impossible for a smaller machine learning project to have a more involved annotation process for a massive dataset, and work as an annotator within the gig economy is more common. These factors are not strictly distinct from considerations of the external representation of annotation work, however. Devaluing such work also complicates the process for the machine learning practitioners themselves, when the complexity and related socio-political contexts shaping tasks are under-anticipated. Relying on practitioner intuition is risky business, with implications for machine learning practice and its governance. Similarly, representing data work as an independent input to a linear pipeline, perhaps iterated upon but nonetheless an external material to input, can misrepresent the reality of real-world machine learning practice fundamentally intertwined not only with data, but with the work that shapes it. Indeed, as work in this area has shown, annotators see themselves as creative workers, contributing to the overall outputs of the machine learning system, but are often excluded from decision-making processes (Gray and Suri 2019; Irani 2015b, 2023). These observations about the highly iterative and changeable nature of machine learning practice challenge common portraits of data pipelines or workflows portrayed as a linear and frictionless process. Borrowing from Hallam Stevens (Stevens 2013), the image of the pipeline (an oil pipeline or a water pipeline) obscures the messy, situated, and iterative nature of data work. In contrast, we see the distributed work that takes place within data structures as a network of situated and dynamic practices. We propose this representation of the network as a perspective from which to approach the study and governance of AI practice, particularly when such work addresses specific processes and outputs of machine learning.

4 Discussion

4.1 Intervening with representation coils in machine learning practice

In addition to the socio-cultural norms applied to representations of data work, disciplinary norms and the technical requirements of the aforementioned networks of AI development influence how tasks are framed. Our discussions of tasks and labour highlight ways in which these are shaped by their intended purpose of feeding into an 'intelligent' model, necessarily compressed to fit the assumptions underpinning the envisioned system. In this way, the notion of the 'intelligent' model shapes the task, which shapes the practitioners' representation of the annotator. This can result in feedback loop, or representation coils, where assumptions shape materialities, resulting in lock-in and foreclosing upon other possibilities for contributions from data workers.

Understanding how and why practitioners construct representations of data work improves our abilities to interrogate AI and develop better theorising and governance practices. In order to effectively contest AI practices we need to understand where the key points of contestation happen, but building our representations of technology development solely from the testimonies of powerful actors risks missing some of the most important points for change. In technologies primarily taking the form of knowledge work such as machine learning practice, this is particularly important.

Many of these representations are not malicious but reflect the broader concentration of socio-political and geo-political power shaping the machine learning industry. Practitioners make seemingly innocuous assumptions around universality, complexity, and skill, and rationalize bumps in the road according to these assumptions, neglecting the impact of implicit misrepresentations and mischaracterizations of both data work and data workers on collaborative processes. These representations may be then picked up as the foundation for studying techno-mediation of harm and injustice.

With our work, we aim to tie abstract notions of AI as a form of knowledge production (the reproduction of specific modes of intelligence) to AI as a social practice, enacted by social actors with different capacities for agency, and demonstrate how the two are materially connected. Probing practitioners' assumptions around data work can, therefore, illuminate the local and situated ways in which different actors within the network of machine learning development enact classifications of skill and intelligence. Previous work has considered how the organisation of labour can shape responsible AI practice and how responsibility is distributed accordingly within the AI supply chain, viewed as modular and isolated from broader accountability (Widder and Nafus 2023). We show how this organisation of labour is directly shaped by framings of skill and intelligence, and that the contingencies of AI practice are not limited to the specific contexts of model development but have concrete downstream impacts on future AI projects and outputs. Research in this area has also pointed to the compounding negative effects triggered by poor data practices (such as lack of documentation or the disregard of domain expertise) (Sambasivan et al. 2021; Sambasivan and Veeraraghavan 2022). We add to this concerns arguing that cultural and disciplinary assumptions around data work can further impact project outputs and outcomes.

We see representation coils as a useful analytical tool to locate and map the points of interaction between practitioners where these classifications of skill and intelligence are formed and maintained. We also view them as an opportunity for critical intervention. As our informants highlighted, it is in the tension between the different interpretations of data, tasks, and project objectives across the development network that we can implement translational work. Mediating between these epistemic differences offers an opportunity to reflexively challenge practitioners' assumptions about what constitutes skilled work and knowledge. Improved transparency and communication across the nodes of the network can foreground the messy, ambiguous, and uncertain work that is necessary to maintain machine learning models, and challenge the image of seamlessness and certainty projected by common AI narratives.

Importantly, this work can also contribute to already existing literature on the material components of data work. For example, by contributing to understanding 'best' (or common) practices in (human centered) data science (Kogan et al. 2020), or how the work of annotators is shaped by organizational power structures and incentives (Wang et al. 2022).

Thus, the conceptual tools that we have introduced here facilitate understanding of how AI assumptions are bridged in practice and, therefore, have applications in the domains of AI ethics and Responsible Innovation, offering a locus of study for how AI ethics principles are translated into the material processes of knowledge production, highlighting sites of epistemic injustice and potential for intervention.

4.2 Implications for responsible innovation governance and frameworks

Foregrounding the often contested and messy practices that underpin AI development allows us to bridge gaps in how we understand the different epistemic norms that operate in different spheres of AI development. Viewing AI as a social concept mediated by specific practices has implications for the responsible governance of these systems. In this vein, 'Responsible AI' (R-AI) has become a buzzword capturing attempts to foster AI development that respects human rights and democratic values (Hagendorff 2020).

Most R-AI guidelines focus on the *allocation* of responsibility throughout the network of actors that make up the AI ecosystems through the operationalization of various *principles* (Jobin et al. 2019). This is of course an important task, and there are many cases where this attribution matters, such as recruiting (Dastin 2022), facial recognition (Buolamwini and Gebru 2018), criminal sentencing (Angwin et al. 2016), and voter profiling (Susser et al. 2019). While these broad-strokes approaches are of course significant, we, following recent work in this domain by Widder and Nafus, are interested in articulating local and situated practitioner accounts to better understand what interventions are required to enable and promote R-AI (2023).

While R-AI has traditionally focused on principles and how those principles can be translated into practice, this is insufficient with respect to the given material dynamics that exist between different agents along the network of AI development (in our case, practitioners and data workers). More specifically, our insights here suggest that R-AI ought to take seriously the *relations* and *practices* that underpin AI development, and we find conceptual support for this kind of approach in the literature on Responsible Research and Innovation (RRI). Von Schomberg, for example, notes that RRI is a "process by which societal actors and innovators become mutually responsive to each other with a view on the (ethical) acceptability, sustainability and societal desirability of the innovation process and its marketable products" (Von Schomberg 2011, 50). An important question, then, is how to ensure that these actors remain responsive to one another in cases where we find different epistemic norms in operation.

Significantly, representation coils offer us another way to understand these relations. By highlighting the ways that recursive feedback loops come to shape data practices, we get a better handle on precisely what might go wrong in the process of AI development. The ways that certain tasks are represented by powerful agents (such as machine learning practitioners or organisations) go on to artificially shape and constrain how those tasks are understood. The issue for governance, then, is that principles might be interpreted by these agents in one way, and that specific *assumptions* of what these principles mean will be the ones implemented in practice.

Additionally, understanding AI as a system of networks pushes back against narratives that see AI development as a clean and linear process, or as a 'pipeline'. As outlined in our empirical work above, the reality of this development in practice is often a messy affair. Abstract principles articulated in a top-down manner, therefore, might not be the best approach to deal with the often evolving and contested bodies of practice that come to underpin AI. These principles run the risk of being imposed from above and, therefore, do not fully capture the very real epistemic difficulties involved in the process of AI projects.

The implications of this, then, is that more attention needs to be given to the material practices that come to shape specific understandings of key concepts in AI. The lessons for AI governance are that it is not enough to focus on a responsible *product*, which is often the case in RRI literature, but we also need to be keenly aware of how narrow framings of intelligence can obfuscate our ability to implement proper mechanisms for responsible governance. For example, the role of skills and labour that we have outlined suggest that R-AI research needs to incorporate more empirical work of the way that these are assessed by different agents with different levels of power in the network of AI development, and how these assumptions feed into inequitable design practices. One way to deal with this is to see responsibility in a *proactive* manner. Not as a *response* to crisis, but as a careful *curation* or *maintenance* over time (Browne et al. 2024). What this means, in practice, is paying careful attention to the way that certain types of labour or skill are framed by practitioners themselves, and investigating the ways that this can come to shape and influence societal outcomes.

5 Conclusion

In this paper, we have demonstrated the ways in which framings of intelligence shape AI development. We explored how epistemic, technical and socio-cultural assumptions shape perceptions of data workers, and data work tasks, within the inherently iterative, explorative nature of AI practice. We illustrated how assumptions about presumed universality of AI models and associated sub-tasks can impose cultural understandings upon data in the abstract, and how these translate poorly into the reality of the data and associated work. Borrowing the image of a 'network', borrowed from Stevens (2013), we outlined how linear 'pipeline' models of AI development obscure and obstruct the true nature of data work. The notion of network can better account for the complexity of data work whilst avoiding implicit hierarchies of labour valuation. Significantly, this highlights how relying only on equally linear 'principles' to steer R-AI research and practice is likely to lead to predictable failures in how we understand and describe AI development, as these principles often fail to capture the material granularity of practices and tasks associated with AI.

Similarly, we found that representational asymmetries are exacerbated by implicit assumptions held about annotators in different geographical and socio-economic contexts, and their assumed levels of 'data expertise', tied in with biases that crowdsourced workers lack integrity and motivation. These assumptions around task complexity are exacerbated by the unpredictable nature of the work, particularly regarding the amount of labour and number of iterations required to complete tasks. Given this, practitioner assumptions regarding the complexity (or the simplicity) of a task may bear little relation to the nature of the work that will be necessary to complete it, in a mismatch between the simplistic, abstracted biological framings of intelligence shaping task design, and the distributed intelligence underpinning its execution.

Responding to how practitioner perspectives shape how data work is framed, from task conceptualization to evaluation, we introduce the concept of representation coils. Representation coils are processes by which tasks and roles are iteratively concretised in interactions between abstract representations of task complexity and data worker competency, and their messy translation into practice, subtly reinforcing dominant epistemologies and power dynamics. Significantly, this approach can help render visible novel forms of epistemic injustice. Finally, we reflected on how these conceptual tools come to bear on discussions around R-AI, and what lessons we might draw for the effective governance of AI-systems.

Acknowledgements The authors would like to thank the practitioners who took part in this research, as well as the STIS community at Edinburgh for providing helpful feedback to earlier versions of this manuscript. We would also like to thank the reviewers for their valuable and insightful comments.

Author contributions Study conception and design, funding acquisition, material preparation, data collection and analysis were performed equally by SJ Bennett and Benedetta Catanzariti. All three authors contributed to the first draft of the manuscript and commented on consequent versions of the manuscript. All authors read and approved the final manuscript.

Funding SJ Bennett and Benedetta Catanzariti received partial financial support from the Edinburgh Futures Institute. Fabio Tollon's work on this paper was supported by the Arts and Humanities Research Council grant number AH/X007146/1.

Data availability Raw data from this study has not been made publicly available to preserve individuals' privacy under the European General Data Protection Regulation, and in keeping with the consent forms and ethics procedures of the study.

Declarations

Conflict of interest The authors have no relevant financial or non-financial interests to disclose.

References

Ali SM, Dick S, Dillon S, Jones ML, Penn J, Staley R (2023) Histories of artificial intelligence: a genealogy of power. BJHS Themes 8(January):1–18. https://doi.org/10.1017/bjt.2023.15

Angwin J, Larson J, Mattu S, Kirchner L (2016). 'Machine Bias'. *ProPublica*, 23 May 2016. https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing.

Beer DG (2022) The problem of researching a recursive society : algorithms, data coils and the looping of the social. Big Data Soc. https://doi.org/10.1177/20539517221104997

Bennett SJ, Claisse C, Luger E, Durrant AC. 2023. Unpicking epistemic injustices in digital health: on the implications of designing data-driven technologies for the management of long-term conditions. https://doi.org/10.1145/3600211.3604684.

Bhakuni H, Abimbola S (2021) Epistemic injustice in academic global health. Lancet Glob Health 9(10):e1465–e1470. https://doi.org/10.1016/S2214-109X(21)00301-6

Braun V, Clarke V (2012) 'Thematic Analysis. In: APA handbook of research methods in psychology, Vol 2: Research designs: quantitative, qualitative, neuropsychological, and biological, 57–71. APA Handbooks in psychology®. Washington, DC, US: American Psychological Association. https://doi.org/10.1037/13620-004.

Browne J, Drage E, McInerney K (2024) Tech workers' perspectives on ethical issues in AI development: foregrounding feminist approaches. Big Data Soc 11(1):20539517231221780. https://doi.org/10.1177/20539517231221780

Buolamwini J, Gebru T (2018) Gender shades: intersectional accuracy disparities in commercial gender classification. In: Proceedings of the 1st conference on fairness, accountability and transparency, 77–91. PMLR. https://proceedings.mlr.press/v81/buolamwini18a.html.

Dastin J (2022) Amazon scraps secret AI recruiting tool that showed bias against women*. Auerbach Publications, In Ethics of Data and Analytics

Daston L (1994) Enlightenment calculations. Crit Inq 21(1):182–202

Daston L (2018a). 'Calculation and the division of labor, 1750–1950'. for work / Against work. pure.mpg.de. 2018. https://onwork.edu.au/bibitem/2018-Daston,L-Calculation+and+the+Division+of+Labor,1750-1950/.

Daston L (2018b) Calculation and the division of labor, 1750–1950'. Bull GHI Wash 62(Spring):9–30

de Croon GCHE, Dupeyroux JJG, Fuller SB, Marshall JAR (2022) Insect-inspired AI for autonomous robots. Sci Robot 7(67):6334. https://doi.org/10.1126/scirobotics.abl6334

Denton E, Hanna A, Amironesei R, Smart A, Nicole H (2021) On the genealogy of machine learning datasets: a critical history of ImageNet. Big Data Soc 8(2):20539517211035956. https://doi.org/10.1177/20539517211035955

Dick S (2011) AfterMath: the work of proof in the age of human-machine collaboration. Isis 102(3):494–505. https://doi.org/10.1086/661623

Dick S (2015) Of models and machines: implementing bounded rationality. Isis 106(3):623–634. https://doi.org/10.1086/683527

DiSalvo C, Rothschild A, Schenck LL, Shapiro BR, DiSalvo B (2024) When workers want to say no: a view into critical consciousness and workplace democracy in data work. Proc ACM Hum-Comput Interact. 8(CSCW1):156:1-156:24. https://doi.org/10.1145/3637433

Feinberg M (2017) 'A design perspective on data'. In: Proceedings of the 2017 CHI conference on human factors in computing systems, 2952–63. CHI '17. New York, NY, USA: Association for computing machinery. https://doi.org/10.1145/3025453.3025837.

Fricker, Miranda. 2007. *Epistemic Injustice: Power and the Ethics of Knowing*. Clarendon Press.

Godsey B (2017) Think Like a data scientist: tackle the data science process step-by-step. Simon and Schuster.

Gray ML, Suri S (2019) Ghost work: how to stop silicon valley from building a new global underclass. Houghton Mifflin Harcourt.

Hagendorff T (2020) The ethics of AI ethics – an evaluation of guidelines. Mind Mach 30(1):99–120. https://doi.org/10.1007/s11023-020-09517-8

Hutchinson B, Smart A, Hanna A, Denton E, Greer C, Kjartansson O, Barnes P, Mitchell M (2021) Towards accountability for machine learning datasets: practices from software engineering and infrastructure'. ArXiv:2010.13561 [Cs], January. http://arxiv.org/abs/2010.13561.

Irani L (2015a) Justice for "Data Janitors"'. Public Books, 15 January 2015. https://www.publicbooks.org/justice-for-data-janitors/.

Irani L (2015b) The cultural work of microwork. New Media Soc 17(5):720–739. https://doi.org/10.1177/1461444813511926

Irani L (2023) Algorithms of suspicion: authentication and distrust on the Amazon mechanical turk platform. SSRN Scholarly Paper, Rochester, NY. https://doi.org/10.2139/ssrn.4482508

Irani LC, Six Silberman M (2013) Turkopticon: interrupting worker invisibility in Amazon mechanical turk. In Proceedings of the SIGCHI conference on human factors in computing systems, 611–20. CHI '13. New York, NY, USA: Association for Computing Machinery. https://doi.org/10.1145/2470654.2470742.

Jobin A, Ienca M, Vayena E (2019) The global landscape of AI ethics guidelines. Nat Mach Intell 1(9):389–399. https://doi.org/10.1038/s42256-019-0088-2

Jones ML (2016) Reckoning with Matter: calculating machines, innovation, and thinking about thinking from pascal to babbage. University of Chicago Press, Chicago, IL

Kandel S, Paepcke A, Hellerstein JM, Heer J (2012) Enterprise data analysis and visualization: an interview study. IEEE Trans Visual Comput Gr 18(12):2917–2926. https://doi.org/10.1109/TVCG.2012.219

Kogan M, Halfaker A, Guha S, Aragon C, Muller M, Geiger S (2020) Mapping out human-centered data science: methods, approaches, and best practices. In: Companion proceedings of the 2020 ACM international conference on supporting group work, 151–56. GROUP '20. New York, NY, USA: Association for computing machinery. https://doi.org/10.1145/3323994.3369898.

Madaio MA, Stark L, Wortman Vaughan J, Wallach H (2020) 'Co-designing checklists to understand organizational challenges and opportunities around fairness in AI', 20.

Mao Y, Wang D, Muller M, Varshney KR, Baldini I, Dugan C, Mojsilović A (2019) How data scientists work together with domain experts in scientific collaborations: to find the right answer or to ask the right question?'. Proc ACM Hum-Comput Interaction 3(GROUP):237:1-237:23. https://doi.org/10.1145/3361118

McCarthy J, Minsky ML, Rochester N, Shannon CE (2006) A proposal for the dartmouth summer research project on artificial intelligence, August 31, 1955. AI Mag 27(4):12–12. https://doi.org/10.1609/aimag.v27i4.1904

Miceli M, Posada J (2022) 'The data-production dispositif'. arXiv. https://doi.org/10.48550/arXiv.2205.11963.

Miceli M, Schuessler M, Yang T (2020) Between Subjectivity and imposition: power dynamics in data annotation for computer vision. ArXiv:2007.14886 [Cs], July. http://arxiv.org/abs/2007.14886.

Miceli M, Posada J, Yang T (2021a) Studying up machine learning data: why talk about bias when we mean power?' Proceedings of the ACM on human-computer interaction 6.GROUP. arXiv. https://doi.org/10.48550/arXiv.2109.08131.

Miceli M, Yang T, Naudts L, Schuessler M, Serbanescu D, Hanna A (2021b) Documenting computer vision datasets: an invitation to reflexive data practices, 12.

Muller M, Lange I, Wang D, Piorkowski D, Tsay J, Liao QV, Dugan C, Erickson T (2019) '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, 1–15. CHI '19. New York, NY, USA: Association for Computing Machinery. https://doi.org/10.1145/3290605.3300356.

Pasquinelli M (2023) The eye of the Master: a Social History of Artificial Intelligence. Verso Books.

Pasquinelli M. 2024. 'Beyond the schism of value form & labor form - Notes - e-Flux'. *E-Flux*, 13 June 2024. https://www.e-flux.com/notes/614431/beyond-the-schism-of-value-form-and-labor-form-in-ai-studies-and-the-humanities-a-response-to-critical-inquiry.

Sambasivan N, Veeraraghavan R (2022) The Deskilling of domain expertise in AI development. In: CHI conference on human factors in computing systems, 1–14. CHI '22. New York, NY, USA: Association for Computing Machinery. https://doi.org/10.1145/3491102.3517578.

Sambasivan N, Kapania S, Highfill H, Akrong D, Paritosh P, Aroyo LM (2021) "Everyone Wants to Do the Model Work, Not the Data Work": data cascades in high-stakes aI'.

Schaffer S (1994) Babbage's intelligence: calculating engines and the factory system. Crit Inq 21(1):203–227

Stevens H (2013) Life out of sequence: a data-driven history of bioinformatics. University of Chicago Press, Chicago, IL

Suchman L (1995) Making work visible. Commun ACM 38(9):56–64. https://doi.org/10.1145/223248.223263

Susser D, Roessler B, Nissenbaum H (2019) Technology, autonomy, and manipulation. Int Policy Rev 8 (2). https://policyreview.info/articles/analysis/technology-autonomy-and-manipulation.

Thylstrup NB (2022) The ethics and politics of data sets in the age of machine learning: deleting traces and encountering remains. Media Cult Soc 44(4):655–671. https://doi.org/10.1177/01634437211060226

Von Schomberg R (2011) 'Prospects for technology assessment in a framework of responsible research and innovation'. SSRN Scholarly Paper. Rochester, NY. https://doi.org/10.2139/ssrn.2439112.

Wang D, Prabhat S, Sambasivan N (2022) 'Whose AI Dream? in search of the aspiration in data annotation.' In: Proceedings of the 2022 CHI conference on human factors in computing systems, 1–16. CHI '22. New York, NY, USA: Association for computing machinery. https://doi.org/10.1145/3491102.3502121.

Widder DG, Nafus D (2023) Dislocated accountabilities in the "AI Supply Chain": modularity and developers' notions of responsibility. Big Data Soc 10(1):20539517231177620. https://doi.org/10.1177/20539517231177620