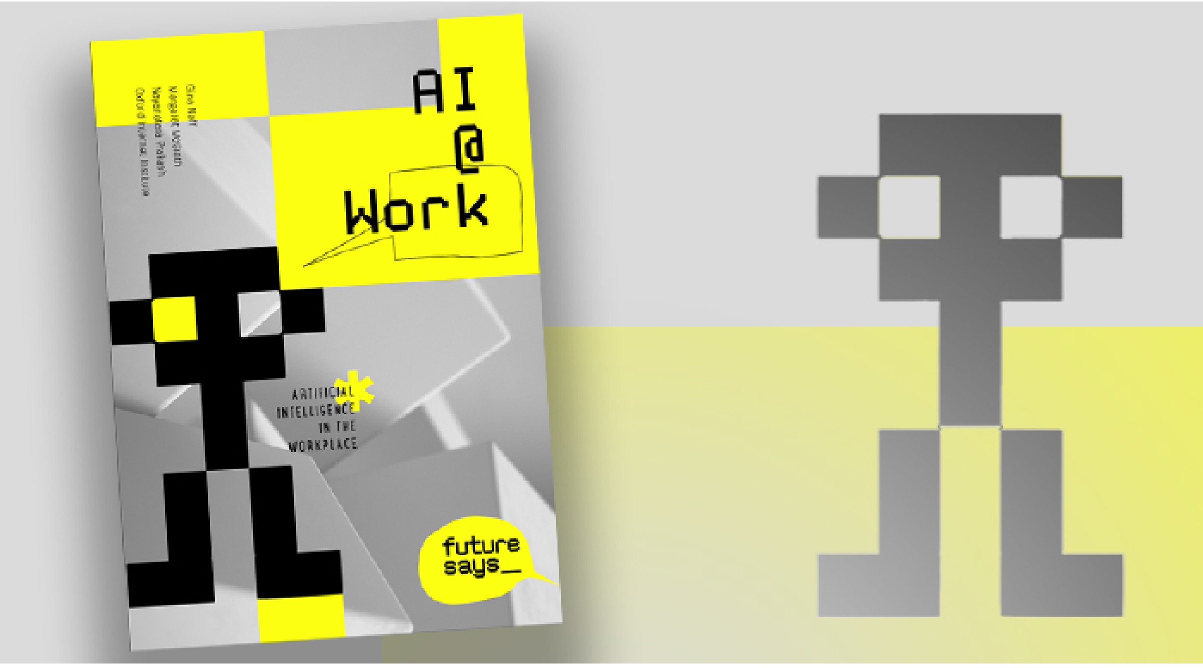
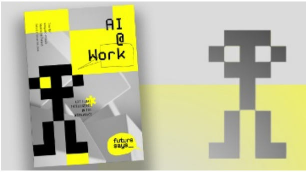
AI @ Work
Artificial Intelligence in the Workplace



By Gina Neff, Maggie McGrath & Nayana Prakash
Oxford Internet Institute



Introduction

This report examines how artificial intelligence (AI) is being used in workplaces. Artificial intelligence technologies are composites of many different kinds of data and technologies and depend on how they are integrated into everyday practices—at work, with workers, in workplaces. Funding for AI ventures last year topped a record US$ 9 billion. As AI moves from the technology sector to more areas of our economy, it is time to take stock critically and comprehensively of its impact on workplaces and workers.

The aim of this report is to inform a more comprehensive dialogue around the use of AI as more workplaces roll out new kinds of AI-enabled systems by looking at the challenges of integrating new systems into existing workplaces.

We analysed themes in over 400 news, academic and industry reports from January 2019 through May 2020, focusing on how they covered AI in workplaces in a wide range of settings. We specifically sought reports on the challenges or failures in the gap between AI technologies and the environments where people use them. We find evidence of this gap, especially in how AI tools used and how people talk about what they are supposed to do.

As we discovered in our thematic gap analysis, there are broadly three major ways that AI fails workers and workplaces. 1) **Integration** challenges happen when settings are not yet primed for AI use, or when these technologies operate at a disjoint between workers and their employers. 2) **Reliance** challenges stem from over and under reliance on AI in workplace systems. 3) **Transparency** challenges, as we define them in this report, arise when the work required by these systems—and where that work is done—is not transparent to users.

From perfumers to oil rigs, AI is now being used outside of the large tech companies that “exist to capture and use digital data. . . That’s different than the rank and file of most enterprise companies.”[1] AI requires global supply chains and a wide range of workers, many in the Global South who increasingly do routine and routinized work to ensure that AI systems function.

Overall, the stories about AI outside of the tech industry show there is much more work to be done in ensuring safe, fair and effective systems that function for workers and in workplaces.

OXFORD INTERNET INSTITUTE

UNIVERSITY OF OXFORD

SUMMARY

AI in Workplaces

- AI is frequently *ineffective*: scaling and transferring AI from one workplace to another presents challenges.
- AI in organizational use is still *simplistic*: it makes binary decisions in complex environments.
- AI is operationally *opaque*: in a micro sense, it often masks human labour. In a macro sense a traditional colonial supply chain remains invisible.

AI and Workers

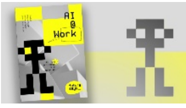
- AI is generating new *problems* at work, including extra work and lessened agency.
- AI is *obscuring* work: Invisible labour is embedded in the AI supply chain.
- AI is mining workers' *process* for data, with often grievous consequence.

A note: we began this project in early March 2020, right before the COVID-19 crisis hit the United Kingdom. Without the benefit of analytical hindsight, we have decided not to include research on AI and COVID-19, although as we emerge from this crisis surely the AI used in the pandemic will warrant future inquiry. Furthermore, this report does not include details on AI usage in and around the Black Lives Matter protests worldwide, though this is a valuable area for further research.

Key Themes

	INTEGRATION	RELIANCE	TRANSPARENCY
WORKPLACES	*INEFFECTIVE* AI	*SIMPLISTIC* AI	*OPAQUE* AI
	Integration Challenges	**People do the Work of AI**	**AI Supply Chain's Colonial Legacy**
	AI Costs Still Outweigh Benefits in Many Places	**AI Makes Binary Decisions in Complex Situations**	
	Challenges of Scale in Multiple Sites of Work		
WORKERS	*PROBLEMS* CAUSED BY AI	PROFESSIONAL *PROCESS* AND AI	AI *OBSCURING* WORKERS
	AI as Workplace Disciplinarian	**AI and Worker's Agency in Decision Making**	**AI Obscuring Human Labour**
	AI Creating New Work	**AI Skills Gap**	
	AI Makes Daily Work a Data Set		

Table of Contents



2. WORKPLACES

2.1 CHALLENGES OF INTEGRATION OF AI: INEFFECTIVE AI

The challenge of ineffective AI covers various aspects of AI not deploying, scaling, or doing a better job than human labour.

2.1.1 INTEGRATION CHALLENGES WITH AI DEPLOYMENT

Description: Integrating AI systems into existing workplaces presents many challenges. Infrastructures are often not ready to provide the requisite data for AI systems, process or manipulate this data, and lack the money and staff required to present a return on investment. AI systems require new kinds of data, new firm-specific skills and know-how and new ways of working.[2]

Case Study: Google's Medical AI was tested in clinics across Thailand to screen patients with diabetic retinopathy, which can cause blindness. A deep learning system was set up to spot signs of eye disease in patients with diabetes. When it worked well, the system sped up the process. But the chaotic reality of workplaces was a world away from the lab. The model had been trained on high quality images, and it rejected images below a certain image quality threshold. For nurses, working quickly and in poor lighting conditions, this meant a great many rejections, and they were often frustrated by their inability to input or sway decisions.

Analysis: In the lab, Google's Health AI scanner was accurate 90% of the time. But medical workplace environments are busy, fast-paced and high-stress. This meant that a successful AI system needs to be flexible enough to handle this environment. Michael Abramoff, an eye doctor at the University of Iowa, notes, real doctors have disagreements all the time. AI-enabled diagnostic systems need to fit into a process where disagreements are routinely discussed. Workers should also be trained to use their own judgment in borderline cases.

Read more:
- "Deloitte Survey: Companies Need Both Data Modernization and Cloud Migration Strategies to Enable Successful AI Initiatives"
- Venturebeat, "Why do 87% of data science projects never make it into production?"
- Will Douglas Heaven, "Google's medical AI was super accurate in a lab. Real life was a different story" *MIT Technology Review*, April 27, 2020.

2.1.2 DESPITE THE COST, AI OFTEN NOT AS EFFECTIVE AS HUMAN LABOUR

Description: At this point, many AI tools and technologies remain relatively basic, and even with heavy investment in AI-related technology, we still aren't at a point where AI is extremely fine-tuned or advanced, or can completely replace human labour. Some estimates note lags of 5-15 years between AI investments and productivity gains.[3] Furthermore, complementary investments in workforce training, hiring and complementary technologies are necessary. Even when companies implement AI systems, results are not always quick. Arvind Krishna, IBM's senior vice-president of cloud and cognitive software, notes that 80% of the work with an AI project is collecting and preparing data, a process which often takes a year or more, and he noted that often halt projects in frustration with the time it takes to get data ready.

Case Study: There are countless examples of expensive and time-consuming AI projects. Take the case of Facebook's chatbot, Blender, supposedly the most human chatbot ever. This incredibly advanced chatbot uses 9 billion variables, but Facebook cautions it still has "many weaknesses compared to humans." Blender can repeat—or even completely make up—information the longer a conversation goes on.

Analysis: Even the most advanced conversational AI at the moment can make mistakes which are incredibly basic. Fully functioning AI data systems can take years, and much of the early work is often simply collecting and cleaning data to get it ready for analysis.

Read more:

- Jared Council, "Data Challenges Are Halting AI Projects, IBM Executive Says", *The Wall Street Journal*, May 28, 2019.
- Jeremy Kahn, "Facebook creates the most 'human' chatbot yet", *Fortune*, April 29, 2020.

2.1.3 SCALING AI ACROSS MULTIPLE SITES OF WORK

Description: Scaling AI across organizations—from say one hospital to another—is often a challenge. Significant differences in how data are collected, managed and used across organisations—or even across divisions within the same organisation—are often to blame. Data has context beyond the collected data set that give shape to many of the issues that make scaling AI systems challenging.

Case Study: At New York's Mount Sinai Hospital, doctors used a machine learning system on chest X-ray images to help them predict which patients had higher risk of pneumonia. The system worked well at Mount Sinai but struggled when applied in other hospitals. The model turned out to spot the difference between the hospital's portable chest X-rays and those taken in the radiology department. Since doctors at Mount Sinai use portable X-rays for more severely ill patients, it was detecting the initial severity of illness based on factors that had little to do with the actual health data or patient outcomes.

Analysis: Unsupervised Machine Learning techniques can pick up on specific factors about the organisation. In this case, differences in where x-rays were taken served as proxies for patient severity, but these proxies were specific to the hospital. Factors in machine learning training can pick up organizational specificities, making it challenging to move AI tools trained on one set of data into other settings.

Read more:

- Liz Szabo, Artificial Intelligence Is Rushing Into Patient Care—And Could Raise Risks, *Kaiser Health News*, December 24, 2019.

2.2 CHALLENGES OF RELIANCE ON AI: SIMPLISTIC AI

2.2.1 PEOPLE STILL DO THE WORK OF AI

Description: AI is marketed as a solution to many problems, from healthcare diagnosis to replacing white-collar workers. However, much of the work that is attributed to AI is currently still done by human labourers—often in countries such as India and China where labour costs are lower. Weak data privacy laws and cheap labour in other countries mean the time-consuming work of data cleansing and data cataloguing can be outsourced overseas. In many cases, businesses market their products and services as being driven by AI, when cheap labour—somewhere else—is actually responsible for getting the job done.

Case Studies: Indian start-up Engineer.ai advertised that their AI technology would help people build 80% of an app from scratch in about an hour. In reality human engineers were behind the service. Company leaders hoped that a marketing bump would help Engineer.ai attract investment to support their work developing 'AI building apps'.

At Google, the team responsible for the data sets that make voice-activated Assistant work Pygmalion, rely on "painstaking" labour to annotate data sets by hand. Much of this vital work behind one of Google's flagship products is done by temporary contract labourers without access to the salaries and benefits of Google employees.

Analysis: These cases show that AI is marketed as automatic, but often based on tedious behind-the-scene work. The Engineer.ai example, along with many others, shows how simply using the phrase 'AI' is appealing for investors and companies, yet it falls to human labourers to fulfil these demands while the capacity for AI-driven work is built. While the phrase AI attracts funds, the work in building and ensuring that the systems work is relegated to lower paid workers.

Read more:

- Nick Statt, 'This AI startup claims to automate app making but actually just uses humans', *The Verge*, August 19, 2019.
- Li Yuan, 'How Cheap Labour Drives China's A.I Ambitions' , New York Times, November 25, 2018.
- Julia Carrie Wong, 'A White Collar Sweatshop: Google Assistant contractors allege Wage Theft,' *The Guardian*, June 25, 2019.

2.2.2 AI MAKES BINARY DECISIONS IN COMPLEX SITUATIONS

Description*:* The AI systems currently in use are often simplistic in nature, allowing for binary decisions. In a number of professional and organizational contexts, nuanced human judgement is still essential to many decision-making processes. To account for this, the "human in the loop" is still part of many AI systems. However, many predictive AI systems are oversimplifying complex situations by using biased training data to make binary decisions, further compounding the effects of algorithmic bias. The focus here is not on the quality or biases embedded in the data, the "garbage in" argument. Instead this case shows focuses on the decision-making process AI systems use to arrive at "garbage out". When AI systems make simplistic decisions, they exacerbate biases and errors embedded in training data.

Case Study: Amazon's AI based hiring system was abandoned after Amazon determined the system was discriminating against female candidates. Because the system was trained on resumes from the past ten years, it prioritised the qualities in the mostly male workforce. The focus of criticism on this case has been primarily on the problems with biases in the training data. However, some Amazon engineers said the system was designed to assign ratings from 1 to 5, like their online product review system. The goal of this design was to narrow a complex and diverse applicant pool down to the top 5 resumes for consideration. Thus, the system design arbitrarily created a binary rejection point at 95 applications.

Analysis: The focus on optimizing and efficiency led candidates to either be accepted or rejected through a simplistic decision making process. While human recruiters are vulnerable to many forms of bias, they might arrive at a pool of 5 candidates through a more nuanced, complex set of metrics than an AI system allows for, especially one designed to mimic Amazon's product review system.

Read more:
- Isobel Asher Hamilton, 'Amazon's AI based hiring tool exhibits gender biases,' *Business Insider*, October 13, 2018
- Chris Baynes, Government 'deported 7,000 foreign students after falsely accusing them of cheating in English language tests' *The Independent* May 2, 2018
- Anna Merlan and Dhruv Mehrotra, Amazon's Facial Analysis Program Is Building A Dystopic Future For Trans And Nonbinary People, *Jezebel*, June 27, 2019

2.3 CHALLENGES OF TRANSPARENCY: OPAQUE AI

2.3.1 AI SUPPLY CHAIN'S COLONIAL LEGACY

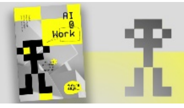
Description: The global supply chain for AI has largely been mapped onto existing postcolonial supply chains. Beyond the disparities between largely Western, well paid data scientists in the Global North and the poorly paid data labellers in the Global South, the infrastructure of postcolonial, transnational capitalism demonstrates similarities with earlier outsourcing booms.

Case Study: Countries like China, India and the Philippines are becoming sources of much of the data labelling labour required to make AI systems function. In China, Yi Yake, co-founder of a data-labelling factory in Henan said, "We're the assembly lines 10 years ago." It is the perfect analogy for a business model that looks alarmingly similar to earlier colonial economic systems.

Analysis: The push to extract cheap labour and goods (data) from a Global South that is enriching the Global North echoes earlier forms of colonial exploitation. This extraction remains largely hidden from most end-users. The public perception of AI systems as automated does not account for the exhaustive human labour required to create enough data to train these systems. As writer Sidney Fussell put it "Each vector of human involvement comes with a way to keep those humans from knowing what's going on." It is this similarity to earlier systems of economic power coupled with opacity about their lineage that makes AI's supply chain particularly alarming.

Read more:

- Sidney Fussell, The AI Supply Chain Runs on Ignorance, *Atlantic Online,* May 14, 2019.
- Li Yuan, How Cheap Labor Drives China's A.I. Ambitions , *New York Times*, November 25, 2018.



3. WORKERS

3.1 PROBLEMS CAUSED BY AI

3.1.1 AI AS DISCIPLINARIAN AT WORK

Description: There is an emergence of the use of AI as a disciplinary tool: facial recognition software to identify criminals and surveillance software. Using AI in this way is posited as a way to improve safety and predict threats, but it does raise the question: should AI be used to discipline humans, rather than workplaces investing in training for their workforce so that they understand the dangers of certain behaviours? Could similar investment be made into training the workforce to address the systemic issues which lead to such problems?

Case Study: Chicago-based company NexLP is developing an AI 'bot' to detect bullying and harassment in workplace emails. The platform is already being used by more than 50 corporate clients worldwide. The bot uses an algorithm trained to identify potential bullying, including sexual harassment, in company documents, emails and chat. Data is analysed for various indicators that determine how likely it is to be a problem, with anything the AI reads as being potentially problematic then sent to a lawyer or HR manager to investigate. The 'red flags' themselves are a company secret, but Leib said the bot looked for anomalies in the language, frequency or timing of communication patterns across weeks, while constantly learning how to spot harassment.

Analysis: There are justified security concerns here, but further issues besides. The bots' capabilities are limited; it is likely to be too sensitive or leave too many gaps. AI is taught to look for specific triggers—it cannot pick up nuance or interpersonal dynamics. Furthermore, if employees feel they are not trusted, they are more likely to work harder to trick the AI or use other means of communication that are not under surveillance. Using disciplinary measures runs the risk of losing employee confidence and encourages workers to find other creative ways to cheat the system. These measures do not target the real root of the problem: why is sexual harassment and bullying rife in workplaces? Using AI rather than involving employees in their own workplace solutions risks alienation and may inhibit long-term solutions.

Read more:

- Rachel Moss, A #MeToo Bot Shouldn't be Necessary, *Huffington Post UK*, January 3, 2020.
- Sylvia L. Mendez et al. The Use of Chatbots in Future Faculty Mentoring: A Case of the Engineering Professoriate, Paper presented at 2019 ASEE Annual Conference.

3.1.2 AI CREATING NEW WORK STREAMS

Description: In creative fields, AI is essentially defining new forms of extra work and impediments to completing work. For creative professionals, as self-branding merges with professional roles, updating design-focused social profiles on sites like Pinterest and Instagram requires extra hours of time outside normal job responsibilities. As the algorithms becomes more complex and opaque, the challenges to remaining relevant increase.

Case Study: On YouTube, users are finding posts flagged and deleted without any transparency. Prominent YouTubers flagged with content that explicitly falls outside guidelines are often not hurt too much because they are such high earners. Yet channels with smaller subscriptions that are more vulnerable. When they are flagged and deleted their revenue stream is blocked, creating significant revenue loss to YouTuber's reliant on that income.

With Pinterest, creative professionals find their "personal brand" and professional work are merging, requiring designers to constantly maintain their digital profiles to stay relevant. This is work that simply didn't exist before platform sites like Pinterest. In addition to blurring the lines between personal and professional lives, they require creatives to be "working" for additional hours each week. This sits counter to the narrative that AI is optimizing efficiency.

Analysis: Both of these examples suggest algorithmic AI is emerging as a new sort of middleman for creative professionals. The complex, opaque ranking algorithms that content creators must interact with to stay professionally relevant require extensive, multi-channel interventions. Far from simplifying or making work easier, these algorithms are complicating professions that already faced many challenges. YouTube content is now mediated by algorithms that often pick up on inaccurate data or misinterpret cues. Because of the nature of YouTube's business model—and the business model of the creators who rely on YouTube for their income—these algorithms are essentially mediating pay and their ability to develop their businesses.

Read more:

- Benjamin Goggin and Kat Tenbarge , 'Like you've been fired from your job': YouTubers have lost thousands of dollars after their channels were mistakenly demonetized for months, *Business Insider*, August 24, 2019
- Leah Scolere and Lee Humphreys, "Pinning Design: The Curatorial Labor of Creative Professionals" *Social Media + Society*, February 24, 2016.

3.1.3 AI IS TETHERING WORKERS BY MAKING DAILY TASKS A DATA SET

Description: Data extraction of customer behaviour is paramount to the development of AI systems. Companies extract data from their customers' browser history, credit card transactions, and television viewing habits to target advertisements. Customers driving new Volvos are training Google's self-driving technology. Data is so valuable at this point that it is not just being extracted from customers, it is also being extracted from employees. Through this, AI is tethering workers by making their daily tasks a new stream of valuable data.

Case studies: WeWork's acquisition of Euclid, a company that focuses on spatial analytics in the workplace, marked a new focus on AI to optimize workplaces. It tracks how space is utilized, and how employees move around physical space. Using WIFI technology and sensors, it allows WeWork to make money off data capturing how people move and operate within an office. Other examples include how workers train the systems that will automate them out of a position.

Analysis: Euclid has been clear they are looking at data on aggregate levels, they're not concerned if your lunch break ran over. However, it feels like an extraordinary violation of privacy. In addition to Euclid, which works from your phone, WeWork was also planning on installing thermal and motion detectors and Bluetooth check-ins. Collectively, the aim of the aggregate is to optimize not just workspaces but workers. It aims to make workers more efficient, using workers themselves to train the algorithms.

Read more:

- Ellen Huet, "Every Move You Make, WeWork Will Be Watching You" *Bloomberg Business Week*, March 15, 2019.
- Hemangini Gupta. "Testing the Future: Gender and Technocapitalism in Start-Up India." *Feminist Review* 123, no. 1 (November 2019): 74–88.

3.2 AI AND PROFESSIONAL PROCESS

3.2.1 AI AND WORKER'S AGENCY IN DECISION MAKING

Description*:* Professionals sometimes have to decide whether to follow AI recommendations or their own judgement. Because of the widespread cultural support for AI, many are being asked to use and trust AI systems. On the ground, in cases like medical diagnoses, doctors often need to override an AI decision. coAsking doctors to choose between either training the AI system better for the future or considering the welfare of their current patient puts them in a bind. "Automation bias" may creep in when predictive systems are right in less critical situations, lulling people into missing important or crucial mistakes.

Case Study: In an analysis of Machine Learning in clinical decision support systems, one study noted the impact of diagnostic decision ML systems on clinical practice. Can systems accommodate late-breaking changes to clinical information and practice when trained on existing data? Can existing patterns of doctors' behaviour be inadvertently reinforced in setting up systems that encourage doctors' reliance on them?

Analysis: Diagnoses carry full legal, professional and moral responsibility. Doctors face extraordinary pressures from a healthcare system geared towards optimizing efficiency. Currently, there are many shortcomings in terms of scaling AI systems across discrete groups. In this context, asking Doctors to weigh their own decisions against AI systems places their agency in the decision-making central to the balance of implementing an AI system and their professional responsibilities. We are likely to see this same challenge emerge with other professionals asked to rely on automated decision tools.

Read more:

- Challen R, Denny J, Pitt M, et al. Artificial intelligence, bias and clinical safety, *BMJ Quality & Safety* 2019; 28:231-237.

3.2.2 AI SKILLS GAP

Description: Even as AI exists, there may be a dearth of workers who have the necessary skill set to work with AI. Reskilling is necessary in order to prevent enormous job losses, but many firms invest in technology rather than complementary investments (such as retraining). The fear of humans losing their jobs to machines still looms large, but there is a difference between humans simply lacking training and AI being inherently better than human labour; the former is true, while the latter is rarely a reality.

Case Study: In India, the AI workforce doubled in 2019, but there are still numerous vacant positions. In the UK, less than half of the workforce are getting enough help and support from their employer to develop the workplace skills they will need in the future. The upskilling gap is even more pronounced among older workers.

Analysis: With all the investment put into developing new AI technologies, if workers are not adequately trained, there will be massive challenges to implementing new systems. Focusing on training workers is an essential complement to the development of AI systems. Skills gaps in both the UK and India suggest the prioritizing of AI development instead of training workers to use systems is endemic of a global AI ecosystem prioritizing big innovative breaks over more mundane tasks like training workers.

Read more:

- ET Bureau. India doubles its AI workforce in 2019, but faces talent shortage. *The Economic Times*. December 27, 2019.
- Aphrodite Papadatou, Skills crisis as UK employers fail to upskill their workers, *HR Review,* Friday, February 1, 2019

3.3 AI OBSCURING WORKERS

3.3.1 AI OBSCURING HUMAN LABOUR

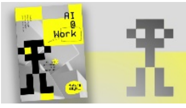
Description: This focuses on the downsides of invisible human labour for both clients/customers and the workers themselves. For workers, being hidden from sight can lead to poor pay, bad working conditions and monotonous work (see Graham and Anwar, 2020). For customers, invisible labour may mean a lack of understanding about who is handling and viewing one's data; in one case, Amazon failed to tell customers that human workers were training the algorithms behind motion detection software. Concerns that such invisible human labour could breach customer privacy have been rife, specifically with Amazon products such as the Alexa.

Case Study: Amazon's Cloud Cam home security promises to monitor your home day and night. However, dozens of Amazon workers in India and Romania were selected to review footage in order to train AI algorithms to do a better job identifying threats. The terms and conditions for Cloud Cam does not mention that human workers have access to this footage, and in some rare cases the footage may be intimate, which suggests it isn't always obtained voluntarily. There are examples of workers sharing footage (and thus, data) with each other in the same workplace.

Analysis: Amazon's Cloud Cam workers are an example of how workers are behind AI but kept 'invisible' so that consumers do not worry about privacy and who has access to their data. Keeping workers invisible in this way may mean that customers share more with the Cloud Cam than they otherwise would, but it also leads to ethical dilemmas about privacy and surveillance. Human annotation is a fairly routine part of AI, but being opaque about the use of humans in this AI loop can lead to a lack of trust between consumer and company.

Read more:

- Natalia Drozdiak, Giles Turner and Matt Day Amazon Workers May Be Watching Your Cloud Cam Home Footage *Bloomberg,* October10, 2019.
- Rebecca Heilwell 'Facebook is flagging some coronavirus news posts as spam' *Vox,* March 17, 2020.
- Tse, T., M. Esposito, T. Mizuno & D. Goh. The Dumb Reason Your AI Project Will Fail. *Harvard Business Review*. June 8, 2020.



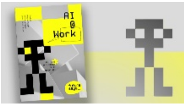
4. COMMON CHALLENGES

AI doesn't always fail due to technical issues. Rather, the technical issues occur because of a misunderstanding of what consumers need and how best to integrate these solutions in specific workplaces and with workers. Broadly, the errors we noticed fit into three categories: transparency, integration, and reliance. While algorithmic transparency has referred to the technical inner workings of AI and algorithmic systems, we advocate here for an expanded definition of transparency to refer to an open dialogue between companies, clients and workers about what AI can do, who does the work of AI systems, and where this work happen in addition to how AI works.

TRANSPARENCY: There is a lack of transparency between companies and consumers about what AI can do, how long it takes, and where humans are involved in the loop. Companies tout AI as automating solutions rather than being open about the amount of money and human labour needed to produce and sustain these systems. This leads to problems down the line as consumers feel deceived by privacy promises. Meanwhile, companies may feel frustrated by the decidedly unglamorous ways AI can work, and how long the process may take.

INTEGRATION: There is a huge gap between the conditions under which AI is trained and the real-life environments it is used in. AI may function best in orderly workplaces with 'perfect' data; unfortunately, real life work settings are unlikely to be so organised. AI without humans trained to use it effectively can exacerbate these issues. Companies may also struggle to scale AI so that it can work across a broad range of systems, presenting a problem for business expansion.

RELIANCE: Companies clearly rely heavily on AI (or rather, the *idea* of AI) and invest a great deal in it, rather than investing in training for workers alongside AI. This means that workers may not be comfortable with decision-making and using AI in their workplaces, leading to integration issues (see above). This over-reliance may also alienate workers, who feel they are being replaced, and lead to a lack of trust in their employer and their contribution to the workplace. Many companies are also still heavily reliant on invisible human labour, often based offshore in India or China. AI global supply chains often trace the maps of colonial power dynamics. Such patterns entrench systems where Western companies take credit for work done abroad by labourers in the Global South.



5. RECOMMENDATIONS

Unfortunately, there are no easy quick fixes to AI challenges. Like the technology itself, solutions to AI issues can be costly and time-consuming. However, companies that can ensure better integration of AI in the workplace are sure to reap the benefits. It is our hope that these recommendations can shape the way AI tools are “ultimately adopted and the organizational value they create,” (Lebovitz, 2020).

As with AI errors, we’ve categorised these recommendations under three key headings.

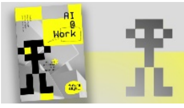
TRANSPARENCY: Companies should be ***clear and open*** with their customers about the use of consumer data and how exactly it is used. This may not erase all customer fears about AI, but it will certainly go a long way towards ***managing expectations***. This transparency should also be applied in earlier stages to talks with any clients who purchase AI solutions. Make it clear from the outset what AI does, and how long it takes to implement. Companies should encourage and ***allow clients to ask questions*** and work with them to ***shape expectations*** about what AI can do.

We need a wider dialogue about ***how we talk about AI in society***. We are used to words like ‘innovative’, ‘transformative’ and ‘‘revolutionary’ to describe AI. These are exciting descriptors, but they mask the more mundane uses and benefits of technology. AI is perhaps currently most effective in automating routine tasks, which can be a huge benefit for companies. Let’s think instead about how we can sell AI as ‘helpful’, ‘essential’ and ‘efficient’, without overstating what the technology can deliver.

The outsourcing of manufacturing was a more visible process, rendered through things like “Made in China” tags in clothing. The AI supply chain ought to have a similar **“Made in” attribution scheme** in order to better understand the global assembly of a technology often considered to be purely technical in nature.

INTEGRATION: The best way to harness AI’s potential is by ensuring workers are comfortable with it and know how to use it, too. ***Upskilling*** is an important way to make sure workers have the abilities necessary to work with and alongside AI; it is a complementary investment to investment in technology. ***Giving workers opportunities*** to use these skills also ensures training is applicable and relevant. Some skills, such as management, may be more valuable than ever in the future, as AI is unlikely to replace core judgment ability.
AI also works best when it works as part of a broad system, where it can get more quality data, rather than being siloed from other business procedures. ***Train employees to think with AI*** by allowing them the ability to critically question systems and interrogate how they fail at the point of intersecting with workers in their workplaces.

RELIANCE: Many of the above suggestions are relevant here, too: companies need to ***level with their customers and users*** about when and where they use human labour in the making of their AI systems and tool, and to clearly detail the potential impact on privacy and security that this work may have. Companies must also ***invest in their workforce*** as well as in AI and ***involve workers in decisions about their data and how it is being used in their jobs and job futures***. This includes workforces abroad, which are crucial to the shared success of the company. Such ***investment in labour*** also helps to prevent exploitative one-sided relationships with offshore workers. ***Training workers to anticipate and handle AI mistakes***, and to use their own judgement when necessary, prevents costly errors and builds a stronger future pipeline of labour. AI itself is only as strong as the team that knows how best to deploy it, which helps build stronger companies. The heavy emphasis on biased training data tends to obscure the simplistic binary decisions AI makes with data. AI systems in-use must be ***frank about their limitations***.



6. CONCLUSION

Our overview of reports on AI in workplaces is meant to be a starting point to shift conversations about agency and autonomy in everyday AI. If we are to build AI systems that are safe, fair and effective, such systems must be able to integrate into workplaces and function with workers. Our review shows the challenges of AI at work. Serious gaps remain between the social and technical infrastructures required for functional AI in many workplace settings. Until these gaps of integration, transparency and reliance can be solved by workers and in workplaces, AI tools and technologies will continue to demonstrate shortcomings in practice.

METHODS ANNEX

This research employs qualitative media analysis (Hodgetts & Chamberlain, 2014). We analysed over 400 news media and scholarly journals articles for stories covering artificial intelligence (AI) in the workplace. We focused on stories where the narratives around the social and technical did not fit. The focus of our searches were AI failures: AI failing to integrate, AI failing to scale, AI failing to launch, and even AI failing to actually be AI.

We focused our search initially on articles from 2019–2020 and on the topic of labour. We included articles from 2016 onwards to capture examples of AI's implementation and its impacts. The scope of our analysis was global, although in part due to the limitations of the researchers, more stories emerged from the United States, the EU and India. We used LexisNexis and Proquest databases for our search queries. After an initial broad sweep of the news, we began to focus on articles loosely centred around the theme of 'labour'.

To organize articles, we constructed a narrative grid with plot synopses. This allowed key themes to emerge from an inductive process, allowing us to construct narratives from the articles themselves, using methods developed by Hodgetts & Chamberlain[5]. This also allowed us to see what was missing from many news articles about AI. Often the information we were looking for was not explicitly written about as an "AI failure". The grid allowed us to iterate with a more focused strategy and expand upon the themes that were emerging.

The limitations of this study included the following: first, we were deliberately looking for counternarratives on AI, which narrowed our focus. Much existing literature on AI in the news takes the form of company releases or positive briefs on the benefit of AI. Second, we were only pulling from published accounts of AI in use. Third, despite our goals of reaching global stories of AI in use, we focused on stories in English which is a limitation of the international reach of this work.

About the Authors

Maggie McGrath is a Doctoral Student at the Oxford Internet Institute researching AI and the design process. In particular, the relationship between everyday algorithmic AI used for image sourcing in ideation and the acceleration of an aesthetic monoculture, focusing on whose agency and values are legitimized. Her research questions emerged from practical experience working as a knitwear designer. Along with Deepak Mallya she recently founded AI Yesterday, a zine that critically engages with AI histories.

Nayana Prakash is a PhD candidate at the Oxford Internet Institute, where she researches Indian women's storytelling online and the Internet as a site of neo-colonialism. In a past life, she was a literary scholar and management consultant. Outside of academia, she has a keen interest in decolonising academia and anti-racism work and organises at a local level around these beliefs. For Future Says_ she co-authored the AI@Work report.

Gina Neff is a professor at the Oxford Internet Institute and the Department of Sociology at the University of Oxford where she studies the future of work and organisations. She is author of *Venture Labor* (MIT Press 2012) and co-author of both *Self-Tracking* (MIT Press 2016) and the forthcoming textbook *Human-Centered Data Science*.

Acknowledgements

This report was authored by Professor Gina Neff, Maggie McGrath and Nayana Prakash from the Oxford Internet Institute (OII). The authors undertook this research in partnership with the Minderoo Foundation, who provided support in the layout and design of the report. This research represents the views and expertise of the authors and does not necessarily reflect the views of the OII or the University of Oxford.

Cite as: Gina Neff, Maggie McGrath and Nayana Prakash. *AI @ Work: Artificial Intelligence in the Workplace*. Oxford Internet Institute,18 August 2020.

References

[1] Bergstein, B. 2019. This is why AI has yet to reshape most businesses, *MIT Technology Review* February 13. https://www.technologyreview.com/2019/02/13/137047/this-is-why-ai-has-yet-to-reshape-most-businesses/

[2] Brynjolfsson, E., D. Rock., & C. Syverson, (2017). Artificial Intelligence and the Modern Productivity Paradox: A Clash of Expectations and Statistics. National Bureau of Economic Research Working Paper No.24001.

[3] Berger, T., and C. Frey (2016). "Structural Transformation in the OECD: Digitalisation, Deindustrialisation and the Future of Work", OECD Social, Employment and Migration Working Papers, No. 193, OECD Publishing, Paris. http://dx.doi.org/10.1787/5jlr068802f7-en

[4] Hodgetts, D. & Chamberlain, K. (2014). Analysing news media. In U. Flick *The SAGE handbook of qualitative data analysis* (pp. 380-393). 55 City Road, London: SAGE Publications, Inc. doi: 10.4135/9781446282243.n26