Generative AI Considered Harmful

Joel E. Fischer
Mixed Reality Laboratory, School of Computer Science
University of Nottingham
United Kingdom
joel.fischer@nottingham.ac.uk

ABSTRACT

The recent months have seen an explosion of interest, hype, and concern about generative AI, driven by the release of ChatGPT. In this article I seek to explicate some potential and actual harms of the engineering and use of generative AI such as ChatGPT. With this I also suggest a reframing for researchers with an interest in interaction. With this reframing I seek to provoke researchers to consider studying the settings of ChatGPT development and use as active sites of production. Research should focus on the organisational, technological and interactional practices and contexts in and through which generative AI and its outputs—harmful and otherwise—are produced, by whom, to what end, and with what consequences on societies.

CCS CONCEPTS

• **Human-centered computing** → **Natural language interfaces**.

KEYWORDS

Large Language Models, LLM, GPT-3, GPT-4, ChatGPT, generative AI, text generation, natural language, NLP, NLG

ACM Reference Format:
Joel E. Fischer. 2023. Generative AI Considered Harmful. In *ACM conference on Conversational User Interfaces (CUI '23), July 19–21, 2023, Eindhoven, Netherlands.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3571884.3603756

1 INTRODUCTION

ChatGPT and the Large Language Models (LLMs) it is based on (e.g, GPT-3/4) have contributed to perhaps never-seen-before hype[1] and panic[2] about generative AI; the latest at the time of writing is that Italy has just banned access to ChatGPT over privacy concerns[3], and a range of illustrious academics and industrialists including Musk and Wozniak have signed an open letter calling on "AI labs to immediately pause for at least 6 months the training of AI systems more powerful than GPT-4"[4]. But the letter has been criticised by some of the researchers it cited, accusing the authoring Future of Life Institute of prioritising imagined apocalyptic scenarios over more immediate concerns about AI, such as racist or sexist bias in exsisting LLMs[5].

From concerns around false authorship becoming commonplace[6], to researchers demonstrating getting a ChatGPT-written paper accepted for publication [7], I believe the discourse around ChatGPT has probably already done a good job "provoking" you—emphatically. That's good because it allows me to focus on a different project in this "provocation paper"[7]. What I want to do is provoke your more level-headed approaches to critique and reason. I want to get past the hype and the panic, and unpack some of what is wrapped up in this generative AI/LLM/ChatGPT phenomenon. To do this, I am going to try to answer two questions: First, what are the potential or actual harms and threats that ChatGPT poses? And second, what can we—a community of researchers interested in CUIs—do about these threats? And with this, I will argue that there are a range of opportunities to study the adoption and use of ChatGPT and similar tools.

In the tradition of articles titled '[x] considered harmful' [8, 17, 20], I want to reflect on both the phenomenon and the concept used to describe it, namely that of 'generative AI', and propose a reframing that hopefully makes it more productive for us researchers interested in conversation and interaction.

Note, I will mostly talk about and refer to "ChatGPT" in this article, but I really mean it as a shorthand for a specific class of generative 'conversational' AI tools that make use of large language models to generate 'human-like' text in response to human prompts (input); currently this seems a more apt shorthand than for instance 'chatbots' or 'dialogue systems', which to my mind have connotations of more limited capabilities (and different technical architectures) than ChatGPT. And while ChatGPT and the GPT-3/4 LLMs it's based on are created by OpenAI, it's worth noting that there are a range of competing companies that are creating generative AI tools and models such as Bard (Google) and Bing (Microsoft), and others that focus on image generation rather than text, such as Stable Diffusion (Hugging Face). And while my focus is on ChatGPT, I think that some of what I cover in this article will also apply to other generative AI tools more broadly.

2 THREATS AND HARMS

When considering the potential threats and harms that ChatGPT may pose I will consider two distinct areas: harms caused by/through the usage of ChatGPT, and harms of ChatGPT introduced in its design and development. These issues are not equal, the burden of

[1] https://www.cnbc.com/2023/03/29/with-chatgpt-hype-swirling-uk-government-urges-regulators-to-come-up-with-rules-for-ai.html
[2] https://www.theguardian.com/commentisfree/2023/jan/07/chatgpt-bot-excel-ai-chatbot-tech
[3] https://www.bbc.co.uk/news/technology-65139406
[4] https://futureoflife.org/open-letter/pause-giant-ai-experiments/
[5] https://www.theguardian.com/technology/2023/mar/31/ai-research-pause-elon-musk-chatgpt
[6] https://studyfinds.org/chatgpt-plagiarize-schools/
[7] This article is written for the CUI 2023 "Provocations" track that "should have the potential to spark debate and discussion at the conference".

CUI '23, July 19–21, 2023, Eindhoven, Netherlands

responsibility lies with the companies developing generative AI, not its users.

2.1 Harmful usage of ChatGPT

The aim here is not to 'blame the users', but to survey the kinds of harms that can stem from the usage of generative AI.

2.1.1 False authorship and plagiarism. This is the issue perhaps most closest to home for those of us working in academic research and teaching. Authorship is linked to income and employment ('publish or perish'), and originality is a key assessment criteria in research funding and publications. The higher education sector, perhaps more than others, relies on the assessment of written texts, from personal statements, CVs and proposals by prospective students and research staff in admissions and hiring, to essays, courseworks and dissertations in assessment of learning outcomes affecting marks and degree classifications. Using ChatGPT to generate text and then pass it off as original could critically undermine the principles and norms around authenticity and authorship. Academics, teachers (and probably many others!) are scratching their heads how to safeguard the work they set students against this kind of 'cheating'. And there's already technical work to auto-detect GPT-generated text (more on that later).

In the meantime, universities are busy updating their rules and definitions around Academic Misconduct to include the use of generative AI, but the devil will be in the detail. Can we really forbid the use of a tool such as ChatGPT completely, or do we instead have to set rules around what is 'acceptable use' and what is not? The latter seems more likely, but just how will be the subject of debates for some time.

2.1.2 Hallucinations and misinformation. It is apt that the verb 'hallucinate' has been coined to describe what generative AI can be observed doing. This is at the heart of what I understand Bender et al. meant when they titled their influential paper "On the Dangers of Stochastic Parrots". According to the authors "an LM is a system for haphazardly stitching together sequences of linguistic forms it has observed in its vast training data, according to probabilistic information about how they combine, but without any reference to meaning" [ibid., p. 617].

The risk that this brings may be acceptable for many use cases, but in the area of legal advice, false information can result in penalties, convictions, and even imprisonment. While ChatGPT has a disclaimer that it should not be used to give legal advice, it will happily write a letter to contest a parking ticket, a speeding ticket, and tell you whether you can be expected to attend court over a breached contract, and so on. And while ChatGPT may have the aforementioned disclaimer, other downstream services that use LLMs and ChatGPT to drive their offering will not be so coy. For example, the service donotpay.com advertises to be the "First Robot Lawyer" that has as a strap line "Fight corporations, beat bureaucracy and sue anyone at the press of a button.". While this service is probably just a "plug-and-chug document wizard"[8], 'real' services that provide legal advice in earnest will probably already be springing up.

2.1.3 Threats to jobs. The moniker 'knowledge economy' encapsulates the sectors of our societies that are built around ideas and intellectual capital. These sectors largely rely on the production of 'the written word', including education and research, human resources, journalism, advertising/PR, legal services, creative industries, the publication industry (academic and popular), and so on. This is an important part of our economies, for instance it has been estimated to be worth £95bn to the UK economy[9]. Aside from the hype and panic, it is probably safe to say that there is a legitimate fear that ChatGPT and other generative AI will threaten people's jobs in these sectors. However, we are also seeing these tools being embraced and used to support existing practices; both the uptake and broader effects on societies and economies will be important areas for empirical research to explore in future.

2.2 Harmful design of ChatGPT

In understanding the potential/actual harms of ChatGPT, we need to consider the 'upstream' design (development, engineering, etc.) of the system, not just the 'downstream' application and usage of ChatGPT and its derivatives.

2.2.1 LLMs replicate biases in society. In the aforementioned paper, Bender et al. [3] set out a range of general concerns about LLMs, including the origins of the data the models are trained on. The authors point out that "large datasets based on texts from the Internet overrepresent hegemonic viewpoints and encode biases potentially damaging to marginalized populations" [ibid., p. 610]. And thus, that "size doesn't guarantee diversity", the point is that LLM-based text has a tendency to 'flatten out' controversial topics and diverse viewpoints towards a societally accepted mean. Moreover there are "risks of harms, including encountering derogatory language and experiencing discrimination at the hands of others who reproduce racist, sexist, ableist, extremist or other harmful ideologies" [ibid., p. 611].

In their influential paper "Man is to computer programmer as woman is to homemaker? Debiasing word embeddings." Bolukbasi et al. shown that vector-based language models (as per their training data) encode biases in societies, such as gender stereotypes. Vector-based representations (i.e., word embeddings) can be used by algorithms to determine 'semantic closeness', which while useful for ontological relationships between concepts such as 'car' and 'tyre' the same leads to harmful stereotyping when relationships between gender and occupation are assumed, for instance.

2.2.2 Copyright, ownership, intellectual property, and attribution. ChatGPT, by its own account, is trained on a massive dataset of text data, including the entire contents of Wikipedia, the Common Crawl archive, books, websites, and "other sources... including articles from scientific journals". The issue is that, by and large, the sources are not cited or otherwise attributed in the output ChatGPT generates. While this may be technically challenging due to the probabilistic transformer-based architecture, it is not inconceivable that for every output it could produce a list of sources that reveals its provenance.

[8] see excellent analysis at https://www.techdirt.com/2023/01/24/the-worlds-first-robot-lawyer-isnt-a-lawyer-and-im-not-sure-its-even-a-robot/

[9] https://www.telegraph.co.uk/business/tips-for-the-future/the-knowledge-economy/

It's worth noting that while Common Crawl and similar internet archives include attribution in the form of URIs, the authorship and ownership of this largely human-written content is effectively erased in the process of training language models. Instead of a heterogeneous collection of web sites (i.e., the WWW), it is treated as just one homogeneous and anonymous "dataset". With these recent developments it may also be incumbent upon the various internet archivists to examine their fair use policies, and at the very least encourage better practices regarding attribution of its sources.

In sum, it appears that current practices of training data collection are predatory. I am not a legal expert, but I assume that court cases will test whether there is copyright and intellectual property infringement inherent in these practices. For instance, Creative Commons CC BY licensing requires attribution: "You must give appropriate credit, provide a link to the license, and indicate if changes were made."[10]

Lacking acknowledgements are particularly bad for artists, creatives, and content creators whose income is directly linked to attribution. While in its current maturity generative AI is limited to text and images, can you imagine a future streaming service that generates music or even video based on an artist or genre? Surely this is already in the works.

2.2.3 The human cost. As has been shown time and again, there is often an element of hidden labour making the AI work behind the scenes [9, 13]. Sadly, the ugly underbelly of AI has now also got a stain with OpenAI written on it. Time Magazine published an article that claims that OpenAI hired Kenyan workers and paid them less than $2 an hour to make ChatGPT less toxic [11], but this is likely to have exposed the workers horrific, mentally scarring content[12]. It will be up to the global tech community to take on more responsible practices that avoid undervalued, exploitative, hidden labour and other harmful practices that amount to a globalised outsourcing and off-shoring culture. What we can do is call out and hold to account these companies for their practices, and develop alternatives that support justice and well-being for workers [14].

2.2.4 Environmental cost. LLMs need a lot of compute to run, and particularly to train. And compute needs electricity. Estimates to train GPT-3 vary from 1.29 gigawatt hours[13] to between 4 tons of CO2 (equivalent to the emissions of one non-US person in a year) to 200 tons of CO2 (equivalent to 50 non-US persons' emissions in a year)—depending on how carbon-intensive the electricity generation is regionally where the data centers are located[14]. Kasper Groes Albin Ludvigsen calculates the carbon footprint of running ChatGPT to be 23.04 kg CO2e daily, or the equivalent of roughly 76 percent of the annual carbon footprint of a Dane[15]. More research is needed to say with more certainty what the environmental impact of ChatGPT is in context of its usage, and how it compares to other information tools we use on a daily basis (e.g., search engines) to understand its real environmental cost.

[10]https://creativecommons.org/licenses/by/4.0/
[11]https://time.com/6247678/openai-chatgpt-kenya-workers/
[12]https://twitter.com/parismarx/status/1615691934432894977
[13]https://www.bloomberg.com/news/articles/2023-03-09/how-much-energy-do-ai-and-chatgpt-use-no-one-knows-for-sure
[14]https://shrinkthatfootprint.com/carbon-footprint-of-training-gpt-3-and-large-language-models/
[15]https://towardsdatascience.com/the-carbon-footprint-of-chatgpt-66932314627d

Whether you are required to go to court if you breach a contract depends on the specific circumstances of the breach and the terms of the contract. In many cases, a breach of contract can be resolved through negotiation or mediation without the need for a court hearing. However, in some cases, a court may be necessary to resolve a dispute over a breach of contract.	If you decide to breach a contract, whether or not you go to court depends on the other party. Occasionally, the other party won't pursue a breach of contract; however, the option to take you to court is totally up to them. If you consult with a lawyer before you breach a contract, you will be well informed about the consequences if the other party chooses to pursue you for the breach of contract.
ChatGPT	https://www.requestlegalservices.com/

Table 1: Are you required to go to court if you breach a contract? ChatGPT response (left) and expert response (right). The majority of an audience (ca. 50) mistakenly believed that the ChatGPT response was written by a human.

3 WHAT CAN CUI RESEARCHERS DO?

Given these potential and actual threats and harms caused by ChatGPT and similar LLM-based tools, what can researchers in CUI, HCI more broadly, and other adjacent disciplines such as digital humanities, science and technology studies, informatics, computer science, data science and so on do in response?

3.1 Countering the harms

Researchers and developers are already working on a suite of tools to detect ChatGPT-written texts. OpenAI itself released a model to distinguish between AI and human-written text[16], but this has been shown to be tricked easily[17]. Others have also developed similar AI-detection tools[18], and these are already being integrated into 'plagiarism detection' tools. These technical solutions to a technical problem strike me to be of the ilk of 'fighting fire with fire' more likely to lead to an 'arms race' of AI vs. AI. This tinkering around the edges of the real issues isn't doing anything about the other issues I have identified above.

There is a role to play for innovations around justice and well-being for AI workers; for instance tools and practices to provide counseling for AI workers, such as content moderators exposed to harmful content [1], and tools such as 'Turkopticon' that give a platform for AI workers' voices where wasn't one before [14].

There will also be technical computer science and engineering work to remedy issues around attribution and citation of sources within the training data. So-called 'Explainable AI' approaches may provide some starting points into how this might be achieved.

Furthermore, there is a role for HCI to 'understand the threats better', by studying the adoption, use and usability of such tools.

3.2 Understanding the harms better

As a community of researchers, we're good at constructing "user studies" through which we can answer research questions such as "how susceptible are people to harm by reliance on false advice?" The majority of an audience (ca. 50 people) of a talk I gave recently mistakenly believed a text answering a legal question (see table 1) that was written by a human was generated by ChatGPT and vice versa.

[16]https://platform.openai.com/ai-text-classifier
[17]https://twitter.com/AlphaSignalAI/status/1620546796236996608
[18]https://www.bbc.co.uk/news/world-us-canada-64252570

This alone does not say whether people would be willing to *rely* on such advice, however, the literature is full of work on trust and reliance [2, 10, 11] that we can draw on to construct studies to test the effects of auto-generated texts on people's likelihood to rely on it, which provide answers ton questions such as "what are the factors that make texts trustworthy and reliable?" Studies are needed that allows us to assess the potential for people to rely on ChatGPT-generated advice in everyday life.

3.2.1 Adoption studies. The study of adoption of AI in everyday life is a longstanding topic for HCI. Often drawing on ethnographic and observational empirical approaches, such studies show *just how* technologies-in-use are interactionally embedded into everyday human conduct, including mobile phones [6], voice interfaces [19], and self-driving cars [5].

People are already using ChatGPT and GPT-3 driven apps in their everyday lives. But how are students, professionals, and creatives using this technology as part and parcel of their studies, work, and creations? Self-reports and 'confessions' are emerging, for instance of artists using image generation as part of their ideation process and text generation to help them structure content for their next YouTube video[19]. *Just what* are people doing to embed this new technology get into their work flows, practices, and tool stack? *Just how* did I make use of ChatGPT to write this paper, you might ask? I will give an insight in the following section.

3.2.2 Interaction studies. In how 'generative AI' is framed (and AI more broadly), the ways in which the outcomes are generated, and the role of the human(s) in it, is arguably either ignored or taken for granted. Actually using generative AI rarely follows the one-shot model it is often made out to do. 'Generative AI' does not generate anything without someone engaging in an interaction with it, in some specifc context and for some specific purpose. Indeed, Cotton et al. who 'used ChatGPT to write their research paper' did so by writing several prompts and structuring the output for the paper, and replacing the bogus references for genuine ones. This certainly looks a lot like the authors followed a structured interactional sequence to get to the outcome they desired.

As a community we have a methodological toolbox to show *just how* things get done in interaction, for instance through studies that draw on conversation analysis [12, 21], Wizard of Oz studies [18], speech acts [16], and embodied interaction [15], to name but a few.

4 EXAMPLE INTERACTION ANALYSIS

Herein, I briefly show how the turn-based exchange you have with ChatGPT could be analysed, showing how the user draws on details of the generated output to shape their next turn to progress the interaction [12]. Consider the example in Figure 1 showing an excerpt of my use of it to find out how ChatGPT works, technically, occasioned by my need to prepare a talk on the topic. After my initial prompt to "explain the GPT-3 LLM" I was curious to learn more about the transformer based architecture mentioned in the response; thus, shaping my subsequent turn. The following response in turn included a mention of 'human feedback' which reminded me of a previous mention of a 'human evaluation' in a prior exchange with ChatGPT. This demonstrates some features

[19]MKBHD on creative AI: https://www.youtube.com/watch?v=0gNauGdOkro

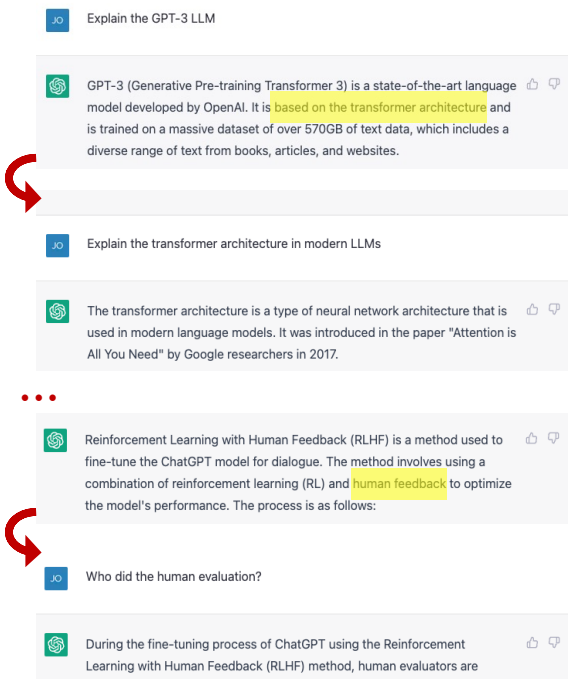

Figure 1: Example of a turn-based interaction with ChatGPT, with the relevant output that shaped my next turns highlighted in yellow.

of interaction analysis: *just what* I paid attention to to shape my prompts in the 'conversation' with ChatGPT to help with writing this talk; accounts for my use of ChatGPT (e.g., to reference the sources of my talk); and the provision of 'evidence' of it 'in action' (in much the same way as I might do with transcriptions of someone using a voice interface).

This brief example already illustrates a rich set of research topics, including the nature of the turn-based interaction through which a sequence of responses and subsequent prompts are generated, how the interaction is shaped by prior knowledge and intentions, and the broader context and intended purpose of the generated outcome that embeds it in the interactant's human endeavours (e.g., 'writing a talk', 'doing research') more broadly [Cf. 19]. How others use it in their work/creative/artistic practices is for us as a community of HCI researchers to find out.

5 CONCLUSIONS

I sought to explicate some potential and actual harms of the engineering and use of LLM-based generative AI systems such as ChatGPT. These harms pertain to the usage of the systems, which includes risks to enable false authorship and plagiarism, following hallucinated and false advice, and threatening jobs of those depending on copy-writing. Pertaining to the development and engineering of these LLM-based systems, harms can stem from replicating the biases and stereotypes in societies from which the training data is created, erasing authorship and ownership, and

nonexistent attribution of the sources that make up the training data, exploited human labour, and unknown environmental cost.

It is probably fair to say that companies are bad at self-regulation (or they would have done a better job already). Standards, regulation and policies may be needed, such as the AI risk management standard ISO/IEC 23894, which provides "guidance on how organizations that develop, produce, deploy or use products, systems and services that utilize artificial intelligence (AI) can manage risk specifically related to AI"[20]. Contributing to initiatives looking to shape best practices, guidelines, regulation and standards in industry is one way in which we can contribute.

As I sought to show, we can also develop countermeasures and innovations that safeguards and mitigate the potential harms. We can also do studies to understand the adoption of ChatGPT in everyday life, including its usage in context and the issues that arise there.

And here's the rub, 'generative AI considered harmful' is both about its potential to cause harms through its development and use, and the potential to stifle research and innovation by its conceptual framing as an autonomous AI-driven black box, which ignores the *interactional* role human(s) have both in the development and use.

The opportunity is to study the settings of ChatGPT as sites and practices of its production. Following this (re-)framing, we may study the probable machine learning-and-performance-focused culture in which developers create ChatGPT, and how users are shaping the broader context of computational tools, which in turn, shapes their adoption practices of ChatGPT.

Providing just one exemplar of how we might approach ChatGPT studies from an interactional perspective, I sought to show that ChatGPT-generated texts can be understood as an 'interactional achievement', the outcome of a turn-based interaction that is thus 'conversational' in nature; and that this doesn't happen in a vacuum. ChatGPT usage is part and parcel of a human endeavour, done in a particular context and with a specific purpose; creating settings that lend themselves well to all sorts of empirical studies.

There are thus ample opportunities to study adoption, use, and the ways in which LLM-based apps hinder and benefit people's work, education, and creative practices, and where areas for innovation, critique and improvement are. There's plenty to do then, for researchers like us.

ACKNOWLEDGMENTS

This work was supported by the Engineering and Physical Sciences Research Council [grant number EP/V00784X/1].

REFERENCES

[1] Andrew Arsht and Daniel Etcovitch. 2018. The human cost of online content moderation. *Harvard Journal of Law and Technology* (2018).

[2] Gagan Bansal, Alison Marie Smith-Renner, Zana Buçinca, Tongshuang Wu, Kenneth Holstein, Jessica Hullman, and Simone Stumpf. 2022. Workshop on Trust and Reliance in AI-Human Teams (TRAIT). In *Extended Abstracts of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI EA '22)*. Association for Computing Machinery, New York, NY, USA, Article 116, 6 pages. https://doi.org/10.1145/3491101.3503704

[3] Emily M. Bender, Timnit Gebru, Angelina McMillan-Major, and Shmargaret Shmitchell. 2021. On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?. In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (Virtual Event, Canada) *(FAccT '21)*. Association for Computing Machinery, New York, NY, USA, 610–623. https://doi.org/10.1145/3442188.3445922

[4] Tolga Bolukbasi, Kai-Wei Chang, James Y Zou, Venkatesh Saligrama, and Adam T Kalai. 2016. Man is to computer programmer as woman is to homemaker? debiasing word embeddings. *Advances in neural information processing systems* 29 (2016).

[5] Barry Brown and Eric Laurier. 2017. The trouble with autopilots: Assisted and autonomous driving on the social road. In *Proceedings of the 2017 CHI conference on human factors in computing systems*. 416–429.

[6] Barry Brown, Moira McGregor, and Eric Laurier. 2013. iPhone in vivo: video analysis of mobile device use. In *Proceedings of the SIGCHI conference on Human Factors in computing systems*. 1031–1040.

[7] Debby R. E. Cotton, Peter A. Cotton, and J. Reuben Shipway. 2023. Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in Education and Teaching International* 0, 0 (2023), 1–12. https://doi.org/10.1080/14703297.2023.2190148 arXiv:https://doi.org/10.1080/14703297.2023.2190148

[8] Andrew Crabtree, Tom Rodden, Peter Tolmie, and Graham Button. 2009. Ethnography Considered Harmful. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Boston, MA, USA) *(CHI '09)*. Association for Computing Machinery, New York, NY, USA, 879–888. https://doi.org/10.1145/1518701.1518835

[9] Kate Crawford and Vladan Joler. 2018. Anatomy of an AI system: the Amazon Echo as an anatomical map of human labor, data and planetary resources. AI now institute and share lab. https://anatomyof.ai/

[10] Mateusz Dubiel, Sylvain Daronnat, and Luis A. Leiva. 2022. Conversational Agents Trust Calibration: A User-Centred Perspective to Design. In *Proceedings of the 4th Conference on Conversational User Interfaces* (Glasgow, United Kingdom) *(CUI '22)*. Association for Computing Machinery, New York, NY, USA, Article 30, 6 pages. https://doi.org/10.1145/3543829.3544518

[11] Justin Edwards and Elaheh Sanoubari. 2019. A Need for Trust in Conversational Interface Research. In *Proceedings of the 1st International Conference on Conversational User Interfaces* (Dublin, Ireland) *(CUI '19)*. Association for Computing Machinery, New York, NY, USA, Article 21, 3 pages. https://doi.org/10.1145/3342775.3342809

[12] Joel E. Fischer, Stuart Reeves, Martin Porcheron, and Rein Ove Sikveland. 2019. Progressivity for Voice Interface Design. In *Proceedings of the 1st International Conference on Conversational User Interfaces* (Dublin, Ireland) *(CUI '19)*. Association for Computing Machinery, New York, NY, USA, Article 26, 8 pages. https://doi.org/10.1145/3342775.3342788

[13] Lilly Irani. 2016. The hidden faces of automation. *XRDS: Crossroads, The ACM Magazine for Students* 23, 2 (2016), 34–37.

[14] Lilly C. Irani and M. Six Silberman. 2013. Turkopticon: Interrupting Worker Invisibility in Amazon Mechanical Turk. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Paris, France) *(CHI '13)*. Association for Computing Machinery, New York, NY, USA, 611–620. https://doi.org/10.1145/2470654.2470742

[15] Razan Jaber and Donald McMillan. 2022. Cross-Modal Repair: Gaze and Speech Interaction for List Advancement. In *Proceedings of the 4th Conference on Conversational User Interfaces* (Glasgow, United Kingdom) *(CUI '22)*. Association for Computing Machinery, New York, NY, USA, Article 25, 11 pages. https://doi.org/10.1145/3543829.3543833

[16] Minha Lee. 2020. Speech Acts Redux: Beyond Request-Response Interactions. In *Proceedings of the 2nd Conference on Conversational User Interfaces* (Bilbao, Spain) *(CUI '20)*. Association for Computing Machinery, New York, NY, USA, Article 13, 10 pages. https://doi.org/10.1145/3405755.3406124

[17] Donald A. Norman. 2005. Human-Centered Design Considered Harmful. *Interactions* 12, 4 (jul 2005), 14–19. https://doi.org/10.1145/1070960.1070976

[18] Martin Porcheron, Joel E. Fischer, and Stuart Reeves. 2021. Pulling Back the Curtain on the Wizards of Oz. *Proc. ACM Hum.-Comput. Interact.* 4, CSCW3, Article 243 (jan 2021), 22 pages. https://doi.org/10.1145/3432942

[19] Martin Porcheron, Joel E Fischer, Stuart Reeves, and Sarah Sharples. 2018. Voice Interfaces in Everyday Life. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI '18)*. ACM, New York, NY, USA, Article 640, 12 pages. https://doi.org/10.1145/3173574.3174214

[20] Stuart Reeves. 2019. Conversation Considered Harmful?. In *Proceedings of the 1st International Conference on Conversational User Interfaces* (Dublin, Ireland) *(CUI '19)*. Association for Computing Machinery, New York, NY, USA, Article 10, 3 pages. https://doi.org/10.1145/3342775.3342796

[21] Yunhan Wu, Martin Porcheron, Philip Doyle, Justin Edwards, Daniel Rough, Orla Cooney, Anna Bleakley, Leigh Clark, and Benjamin Cowan. 2022. Comparing Command Construction in Native and Non-Native Speaker IPA Interaction through Conversation Analysis. In *Proceedings of the 4th Conference on Conversational User Interfaces* (Glasgow, United Kingdom) *(CUI '22)*. Association for Computing Machinery, New York, NY, USA, Article 10, 12 pages. https://doi.org/10.1145/3543829.3543839

[20] https://www.iso.org/standard/77304.html