Shifting AI controversies:

How do we get from the AI controversies that we have to the controversies that we need?

Workshop report, 17 April 2023

Shaping AI, University of Warwick

Figure 1: Shaping AI controversies during the expert workshop in Friends House (London), 10 March 2023

Executive summary

What features of AI, especially, have triggered controversy in English-language expert debates during the last 10 years? This report discusses insights gathered during a recent research workshop about AI controversies hosted by the ESRC-funded project Shaping AI. The workshop was dedicated to evaluating this project's provisional research results and its main findings are as follows:

- AI controversies during 2012-2022 focused not only on the application of AI in society, such as the use of facial recognition in schools and by the police, but highlighted structural problems with general purpose AI, such as lack of transparency, misinformation, machine bias, data appropriation without consent, worker exploitation and the high environmental costs associated with the large models that define AI today.

- Participation in AI research controversies has been diverse but relatively narrow, with experts from industry, science and activism making notable contributions, but this relative diversity of perspectives appears to be under-utilized in recent media and public policy debates on AI in the UK.

- AI research controversies in the relevant period varied in terms of who participated, the geographic scope of the issues addressed as well as their resolvability, but all controversies under investigation are marked by concern with the concentration of power over critical infrastructure in the tech industry.

1. Introduction

On the 31st of March, Italy became the first country to ban ChatGPT, the AI chatbot, and controversy about this and similar "Generative AI" systems have been extensively reported in the media in recent weeks. However, machine learning — as well as the "deep learning" techniques based on so-called artificial neural networks which underpin Generative AI models—has been the subject of expert, policy and activist debate for at least the last 10 years. These debates have identified a broad range of technological risks, societal harms, and conditions that will need to be put in place for the realization of the benefits of AI. In comparison, today's debates about AI are dominated by a relatively narrow range of industry voices, such as those of Elon Musk and others tech industry representatives who attracted widespread attention with their open letter calling for a temporary pause in AI development. However, many of the controversial features of today's Generative AI were already identified in AI debates during the period 2012 and 2022. In the report, we therefore provide an overview of the views offered by UK-based AI experts during a recent ESRC-funded workshop as to what features of AI proved especially problematic during the last 10 years, as well as their reflections on the role of public debate in identifying and addressing controversial aspects of AI.

2. Workshop topic and methodology: shaping and re-shaping AI controversies

Hosted by the international research project, Shaping AI, the participatory research workshop "Shifting AI Controversies" used data-led participatory design methods (Kimbell, 2019) to evaluate selected AI controversies from a UK perspective. Around 35 UK AI experts from science, government, industry, activism and the arts worked together in small groups to review Shaping AI's on-going analysis of AI controversies, and specifically the UK team's study of English-language research controversies in the areas of AI and AI and society.

To structure this evaluation of AI controversies, Shaping AI researchers had designed an evaluative inquiry (Marres and De Rijcke, 2021) consisting of two activities, that of "shaping" and "re-shaping" AI controversies. Participants worked in small groups to determine the "shape" of selected AI controversies based on controversy dossiers provided by the organisers (Figure 1 and Figure 6). Are recent English-language controversies about AI in good or bad shape when it comes to participation, relevance, engagement with social context, and the role of power? Based on their diagnoses, the groups then worked together to "re-shape" the selected disputes and formulate desired features of AI controversies to come.

To inform this collaborative work in evaluating AI controversies, University of Warwick researchers presented the provisional results of the Shaping AI project at the start of the workshop. Below we will summarise these provisional research results, with a special focus on the UK team's study of English-language AI research controversies for the period 2012-2022, which combined online consultation, social media analysis and interviews. In the second half of this report we present the results of the design-led inquiry into AI controversies during the workshop.

3. Mapping AI controversies across research, media and policy (2012-2022)

the Shaping AI team are currently conducting comparative social research on AI controversies for the period 2012-2022 in four countries: Germany, France, UK and Canada. This on-going study has highlighted the strong influence of promotional industry discourse on AI debates in research, media, and public policy domains in all of these countries. Perhaps inevitably, discussions across these domains have been dominated by sensational claims regarding the unprecedented predictive, diagnostic, and communicative capacities of contemporary AI (Castelle and Roberge, 2020; McKelvey et al., forthcoming). A related insight arising from Shaping AI's on-going research concerns the significant, and relatively sudden, expansion of public policy debates in the area of AI and society across the world since around 2016. This AI & Society discourse is generating an abundance of "problematizations" of AI, definitions of real and potential problems, threats, and harms (misinformation,

exploitation, racism, bias, physical harm, privacy invasion) associated with AI. In part due to this abundance of problem definitions, the dynamics of public policy discourse on technological risk in the case of AI appear to diverge from the standard cycle that academic policy research has identified in other cases, such as genetically modified foods, where the risks identified by experts directly inform problem definition, prioritization and stabilization in public policy (Stirling, 2008). Instead, governments in the four countries appear to favour flexible solutions such as regulatory sandboxes (Ranchordas, 2021), but such experimental approaches leave to a degree unresolved how risks identified by experts and publics feed into problem prioritization and regulatory framings in public policy.

These two observations highlight significant challenges to the development of robust and comprehensive understandings of the new capabilities, benefits, risks and harms generated by AI as a strategic area of research and innovation. Yet, precisely because of the emergence of unprecedented scientific and technological capacities and associated societal risks, benefits and harms, the stabilization of problem definitions and critical intervention in relation to AI are needed more than ever.

4. Findings from the 2021 UK expert consultation: what makes AI controversial?

To address the two above features of recent AI controversies noted above, the strong influence of promotional discourse and the abundance of problem definitions, the UK Shaping AI team undertook an expert consultation in the Autumn of 2021, in which we invited UK-based experts in AI and Society to identify the most important, and possibly overlooked, AI controversies in the last 10 years. The consultation identified facial recognition as a major area of expert concern, alongside corporate research culture, data and machine bias (see Figure 1).

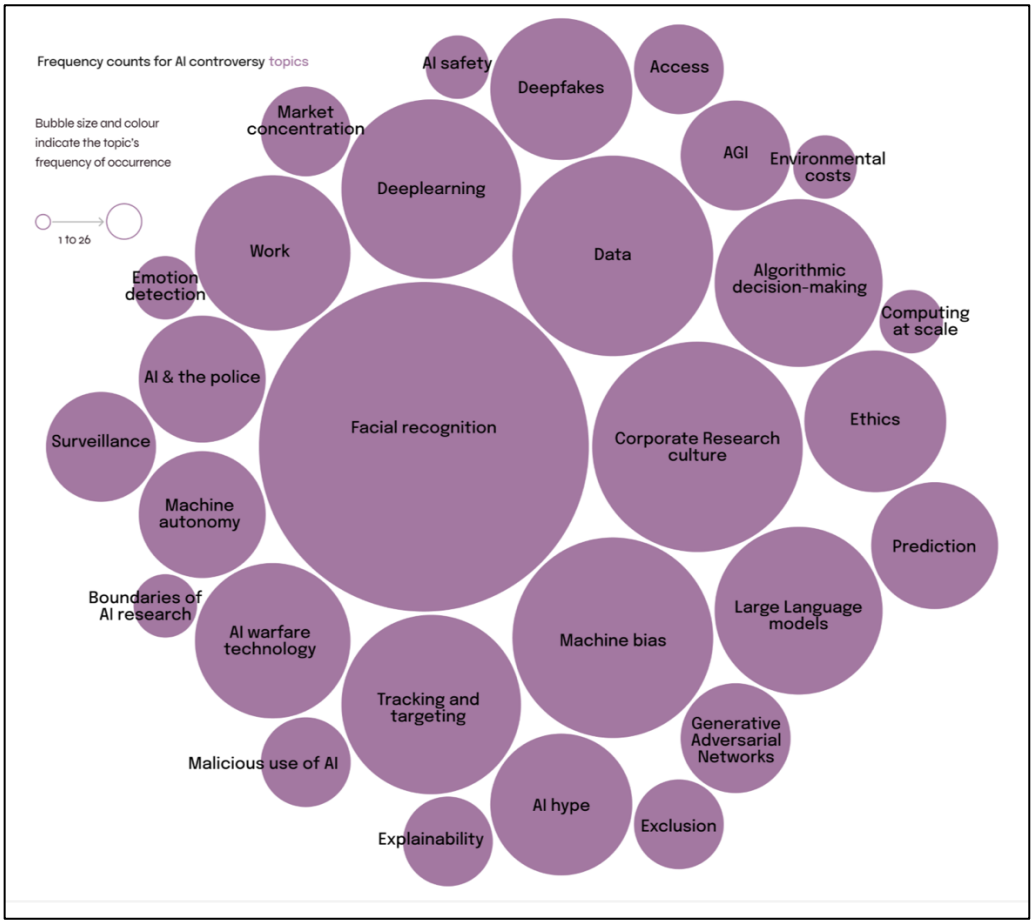

Figure 2: Frequency counts for AI controversy topics identified in the UK expert consultation, Autumn 2021.

Note that this most frequently mentioned topic concerns the application of AI in society, highlighting the use of facial recognition in schools, streets, and transport environments. However, many of the controversial developments identified by UK experts transcend the application of AI in society, and concern the socio-technical architectures and economic arrangements underpinning contemporary AI as a general purpose technology: lack of transparency, machine bias, data appropriation without consent, worker exploitation and corporate control over infrastructure and research culture.[1] UK experts do not overall believe that AI only becomes controversial when it is applied "downstream" in societal domains like health, mobility and education, many of its problems have to do with the architecture of AI itself. However, at the same time, the consultation responses identified a broad range of specific AI applications, systems and in one case, an individual (Timnit Gebru) as subjects of AI controversy, with Large Language Models (GPT; BERT) and discriminatory algorithmic systems (COMPAS; Amazon's hiring engine) especially prominent among these (see Figure 3). The consultation results, then, also highlight that the technological and societal risks and harms associated with the transformer model GPT specifically, and large models, more generally, have been a focus of expert concern since 2021, and indeed, the controversy about GPT has been on-going since 2019.

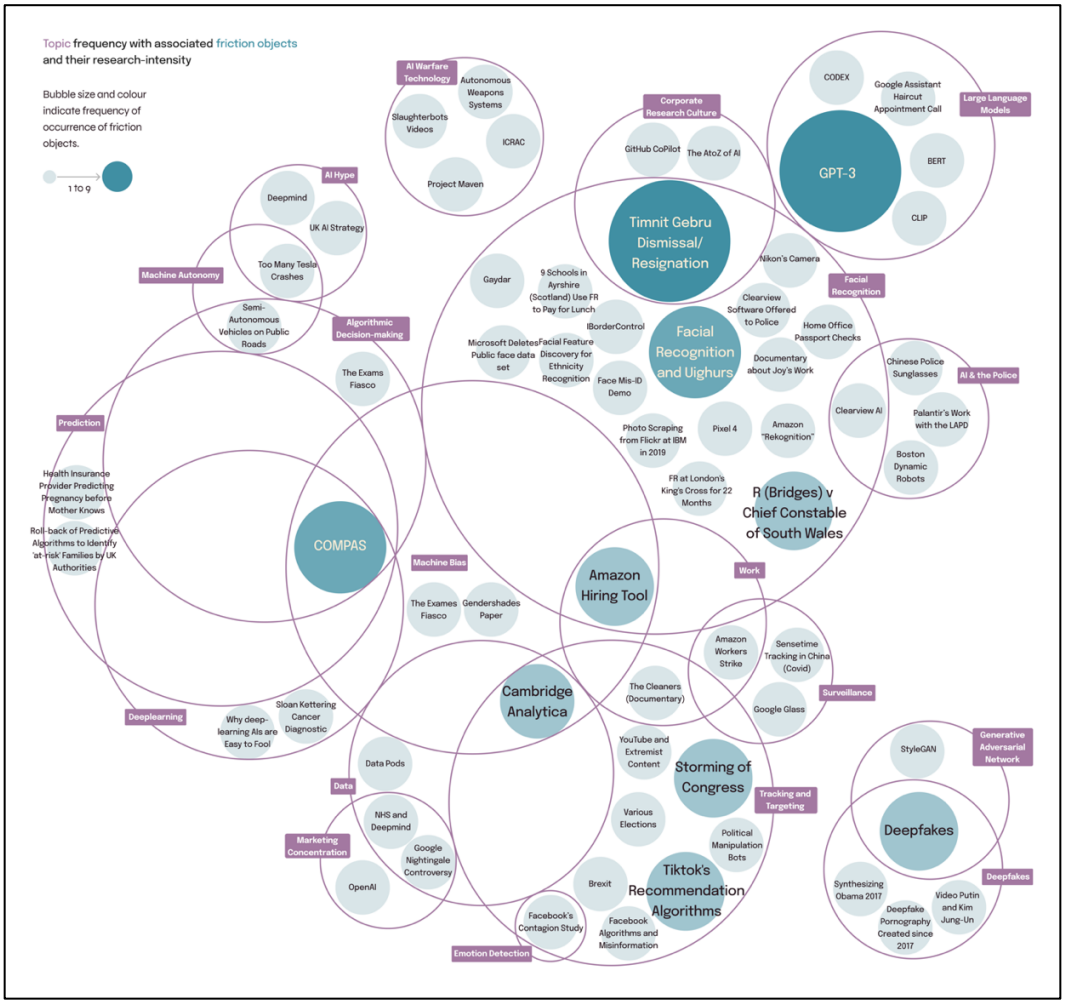

Figure 3: Overview of consultation responses: topic frequency with associated friction objects; and their research-intensity (blue) (UK Shaping AI expert consultation (2021))

1 Other problems identified by UK AI experts in 2021 include: media-based deception using AI; capture and use of personal and public data - online profile photos, medical records, social media posts - by private companies without consent; use of commercial software in the public sector without risk assessment; bias against women in the tech sector; lack of or insufficient checks on data quality and robustness of methods; consolidation of corporate power over research infrastructure and threats to academic freedom.

5. Social media analysis of AI research controversies: what are the main topics of disagreement, who participates, and how?

Building on the consultation, the UK team of Shaping AI identified a set of controversies about AI for the period 2012-2022 for further analysis. In order to delve deeper into the "research layer" of AI controversies, the UK team of Shaping AI identified 5 controversies from the consultation that stood out for their "knowledge-intensity."[2] The selected AI controversies are:

- COMPAS: a controversy about Algorithmic discrimination and the use of algorithmic recommender systems in US courts sparked by the ProPublica report “Machine Bias” (2016)

- NHS+Deepmind: a controversy about data sharing between UK public sector organisations and big tech sparked by the Powles and Hodson (2017) paper " Google DeepMind and healthcare in an age of algorithms"

- Predicting sexual orientation with AI (Gaydar): A controversy about the use of neural networks in predictive social research sparked by the Wang and Kosinkski (2017) paper “Deep neural networks are more accurate than humans at detecting sexual orientation from facial images."

- Large Language Models (Stochastic Parrots): A controversy about Large Models sparked by the Bender et al. (2020) paper “On the Dangers of Stochastic Parrots"

- Machine Learning as a solution for AI: an extended controversy about the capacity of Machine Learning — and specifically the use of trained multilayer neural networks with large numbers of parameters — to solve the problem of AGI (Artificial General Intelligence).

To analyse these research controversies, we turned to Twitter as this media platform was identified by our UK experts as a prominent debate forum (alongside Reddit and Discord), and is well suited to controversy analysis (Housley et al, 2019; Marres, 2015; Madsen and Munk, 2019). For each of the 5 controversies we constructed an English-language Twitter data set,[3] and conducted a controversy analysis focused on Twitter conversations: for all controversies we analysed the topics of disagreement, their actor composition, and the overall "style of engagement".[4] The provisional results of this analysis are presented below.

---

2 We selected controversies for which a comparatively high number of research publications had been suggested by consultaiton respondents. We defined research publications broadly as including think thank reports and investigative journalism, as well as research papers, in accordance with Funtowitz and Ravetz (1997) theory of post-normal science and Whatmore (2009) notion of knowledge controversy.

3 For documentation of our methods of Twitter data capture, analysis and visualisation, see: https://warwickcim.github.io/shifting-ai-controversies-workshop/

4 Style of engagement refers to the length and width of reply chains the conversations in each of the five data sets: the breadth of reaction elicited by a given tweet and the length of chains of successive replies for each conversation (Housley et al, 2018; Marres et al, forthcoming)

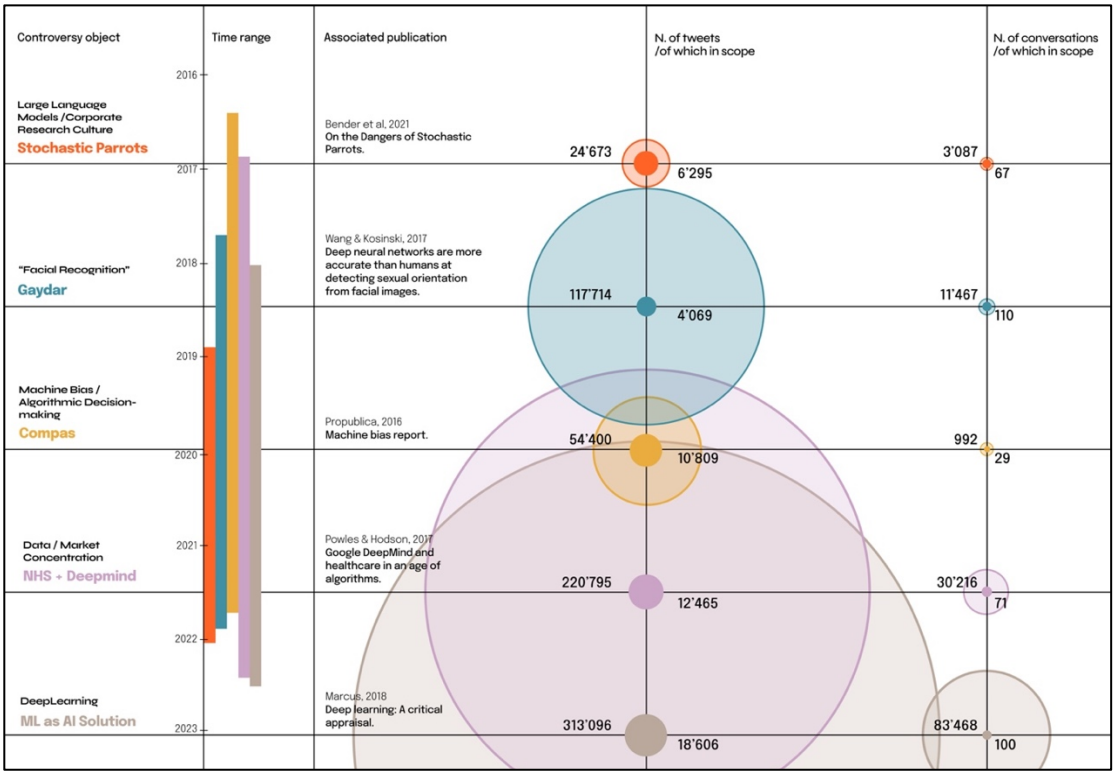

Figure 4: Twitter data for the AI research controversies selected for further analysis

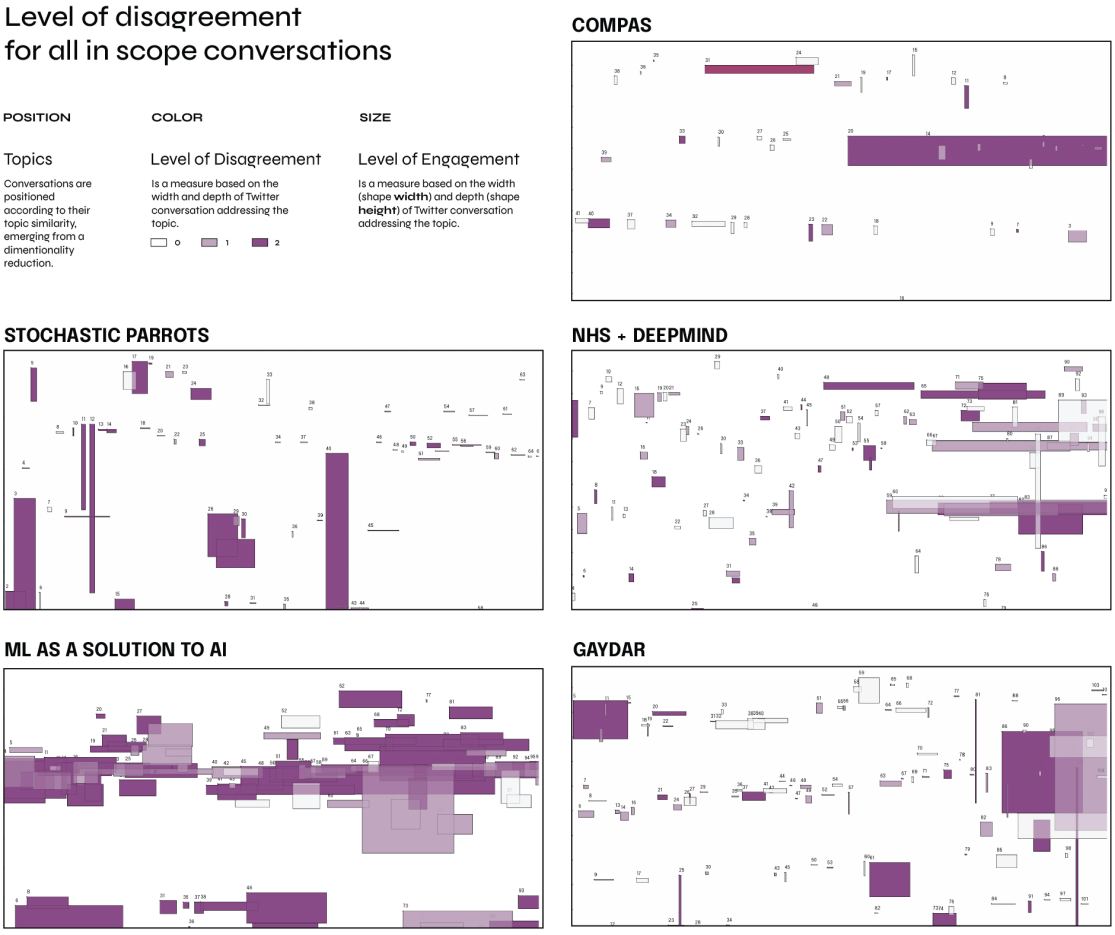

Figure 5: Styles of engagement; level of disagreement for all in scope conversations

Our analysis shows that AI research controversies on Twitter, like those identified in the expert consultation, are primarily concerned with the underlying architectures of AI research and with structural transformations of science, economy, and society through contemporary AI. Engaging with research literature, the selected controversies highlight significant societal risks, harms and problems of AI: disinformation, discriminatory impacts of the use of AI in the public sector, lack of transparency of data sets and methods, growing corporate control over research, the environmental impacts of large language models, and the transfer of personal and public data to private actors. All of the research controversies under scrutiny identified problems with the consolidation of power in and lack of oversight over the technological economy (Slater and Barry, 2005) that underpins AI research and innovation.

CONTROVERSY	MAIN TOPICS	PROMINENT PARTICIPANTS	STYLE OF ENGAGEMENT[5]
COMPAS	Bias; Datasets; Racism; Transparency	Media; Activism	Broadest
NHS+Deepmind	Data protection; Consent and Trust; Data sharing purpose;	Activism; Professions	Broad and deep
AI prediction of sexual orientation	Harms; Validity; Pseudo-science; Bias	Research; Activism	Deep and broad
Machine Learning as solution for AI	Artificial General Intelligence; Human vs Machine learning; Deep learning hitting a wall	Tech Industry; Research	Broad and deep
Large Language Models	Corporate research culture; Women in Tech; Environmental Impact;	Tech industry; Research	Deepest

Table 1: Main topics of disagreement, main actors and overall style of engagement in selected AI research controversies on Twitter.

Also of note is that **participation** in the social media discourse on the AI controversies under scrutiny was **diverse but relatively narrow**. Most controversies were dominated by actors with direct knowledge of AI research and of the societal impact of technology, with participants from the tech industry as well as activism especially dominant. Academic researchers are a notable presence too, but policymakers and professions play a less prominent role. The controversy about the NHS-DeepMind Data sharing agreement was the only debate in which policymakers and legal and medical professionals played a prominent role on Twitter. Twitter of course provides a *highly partial* perspective on what makes AI controversial, and we therefore invited UK-based experts from the 2021 consultation to review the 5 AI controversies from a UK perspective during our "Shifting Controversies" expert workshop in March 2023.

5 The style of engagement refers to the average length and width of the reply chains that constitute the conversations in the Twitter data set for the controversy in question. This is represented by the length and width of blocks in Figure 4, where each block represents one conversations.

Figure 6: The "controversy shape shifter" in use: shaping and re-shaping AI controversies with design methods

6. Expert workshop on AI controversies: what shape are AI controversies in?

The expert workshop[6] that took place in London on March 10 had as its main objective to evaluate from a UK perspective the AI research controversies identified through the consultation and analysed using Twitter. After having been introduced to the controversies, participants worked with a diagnostic tool designed specifically for this purpose - dubbed the "controversy shape shifter" - to determine the shape, and then to re-shape, the selected AI research controversies. During the first, diagnostic, activity of "shaping", participants determined together what is the "state" of the controversy in question: was the controversy in good or bad shape, given the relevance of the issues addressed, the degree of participation, the scope of the issues addressed, and the allocation of responsibility for the problem? To support this work, the Shaping AI team had created a "dossier" for each controversy, which presented a time line of events, actor list, key documents and the Twitter analysis (for an example, see Figure 7). Consulting their dossiers, the groups determined the shape of the selected AI research controversies by producing its outline according to an evaluative grid, which presents a set of axes of interpretation (Figure 8). The controversy " shapes" agreed by the workshop participants are as follows:

CONTROVERSY	RELEVANCE	PARTICIPATION	SITUATEDNESS	POWER	SOLVABILITY
COMPAS	Major	Experts + Policymakers	Local	Few	Easy+Hard
NHS+Deepmind	Minor+Major	Expert + Publics	Local	Relatively Few	Easy+Hard
AI prediction of sexual orientation	Major	Experts + Everybody	Local + Global	Few	Hard
ML as solution for AI	Minor+Major	Experts	Local + Global	Few	(Mostly) Easy
Large Language Models	Major	Experts	Local + Global	Few	Hard

Table 2: The "shapes" of selected research controversies about AI (10 March 2023).

6 A more detailed workshop description can be found here: https://warwick.ac.uk/fac/cross_fac/cim/events/shai/

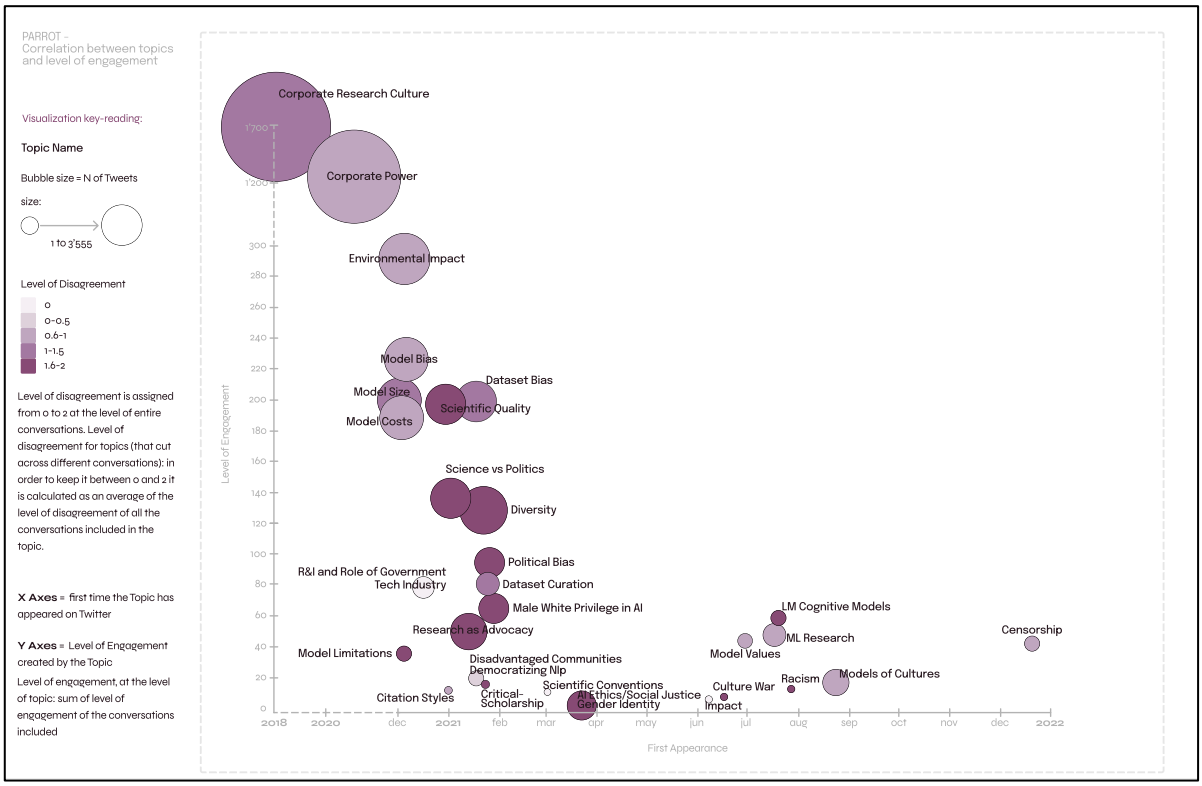

Figure 7: An example of Twitter research results included in the dossier for the Stochastic Parrots controversy: Twitter topics by level of disagreement.

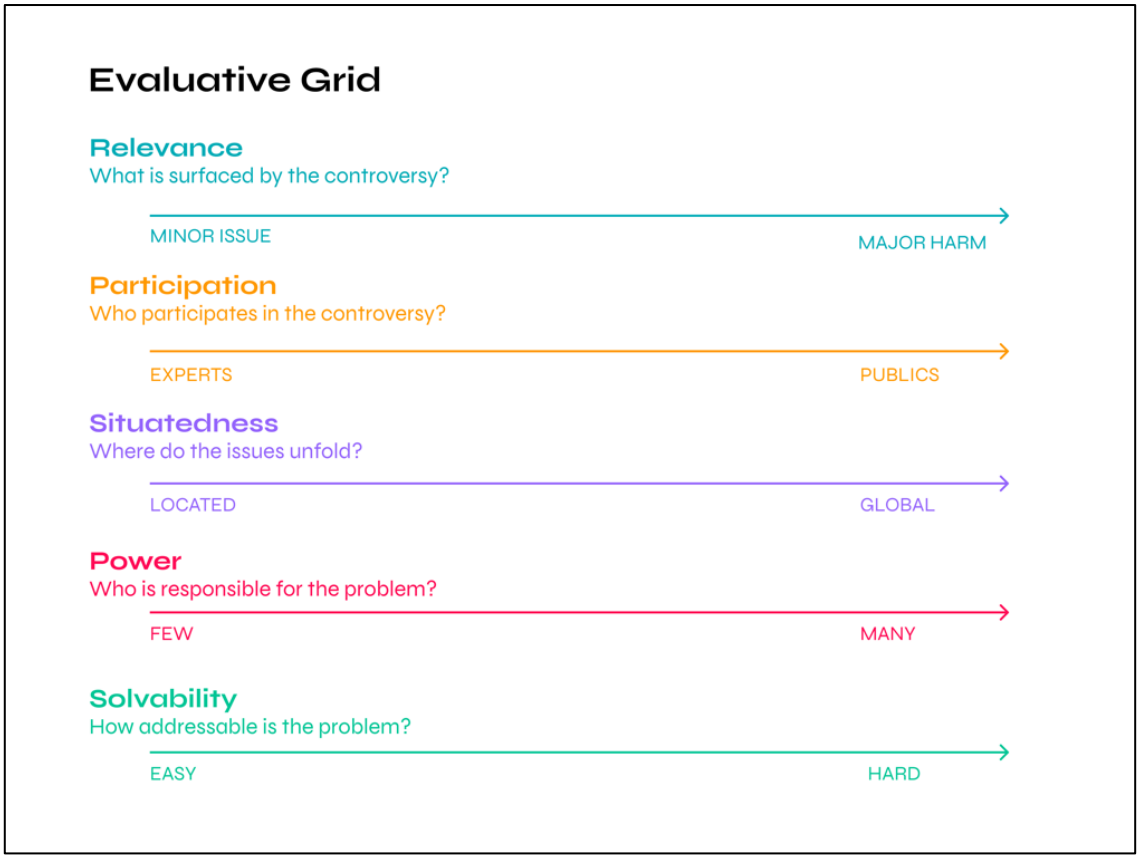

Figure 8: The evaluative grid

Workshop participants provided the following motivations for their evaluations:

COMPAS

This controversy, about the recidivism scoring algorithm used in the US courts in the 2010s, had a narrow focus on the US. Nevertheless, the case received a lot of attention in policy circles in France, the UK and elsewhere because of the link it established between the use of algorithmic systems in the public sector and intersectional biases in institutional settings. While the COMPAS controversy focused on the racial bias and unfairness of the scoring system and underlying data against ethnic groups, it should be clear that the issue of intersectional bias is much larger than just this technology. The controversy thus highlights how societal problematics, such as structural inequality in the US and its large prison population, become materialised in technology. In this respect, the COMPAS controversy is just the tip of the iceberg of a broad societal problem, while at the same time, from a technician's point of view, the problem of algorithmic bias can appear to be a minor issue because it would simply be a matter of fitting the model better or having a larger dataset. As to participation, the controversy was dominated by experts and marked by white middle class bias. Overall, the case underscores the mutual imbrication of technological harm and societal problems in the case of AI, as well as the political issue of who gets to have a voice in AI controversies.

NHS+Deepmind

What is striking about the controversy arising from the collaboration between the NHS and Deepmind in 2015, is that standards on data sharing agreements were in place at the time, but were not adhered to (minor issue, local controversy, relatively easy to solve). Nevertheless, this controversy played an important role in flagging a structural challenge of AI to both national and international public policy communities, namely the appropriation of public data by private companies and the commercial benefit this brings, as well as growing Big Tech's control over public sector data infrastructures (major issue; hard to solve). As regards participation, in its early stages this controversy was dominated by experts with a strong public good framing of their expertise. Later it garnered professional and public attention,[7] enabling patient involvement in the wider issue of NHS contracts with Silicon Valley companies. As such, this controversy also helped empower regulators, though it should be noted it was the NHS and not the company Deepmind that was found at fault in the regulator's ruling (relatively few actors have power).

AI prediction of sexual orientation (Gaydar)

This controversy about the use of neural networks to predict sexual orientation raised a major issue of societal harm, as it demonstrated that machine learning-based analytics could be used to expose people's vulnerabilities and leverage these against them (major harm). The controversy also highlighted the risks that come with people's personal data being used for other than intended purposes, as the users of the dating app where the data was collected had no awareness of what was being inferred from their data. It also flagged exclusionary effects of predictive methodologies, for example, in the context of discrimination against the facial difference community. Overall, the potential of future misuse means that major harm is likely given this predicative capacities of AI. In terms of participation, the controversy was deemed to be closely aligned with media hype, as assumptions are made by the public that neural networks can do certain things that it can't. What is this research really for? The research evinced a simplistic belief in the power of categorization, perhaps even a resurgence of physiognomy, which most experts would reject. It may be the case that the study was not written for scientists, but rather for journalists and commentators to create hype around the revolutionary potential of AI. It is easy to be deceived by the controversy's (un)solvability, as publics are likely to

7 The Information Commissioner's Office ruling on the NHS-Deepmind data sharing, 3 July 2017 https://ico.org.uk/media/action-weve-taken/undertakings/2014353/undertaking-cover-letter-revised-04072017-to-first-person.pdf

overestimate what this technology can do.

Machine learning as a solution to AI

Many of the issues arising from machine learning have to do with the challenge of assessing the societal consequences of a general purpose technology; one that has application across an open-ended set of societal domains. The impressive capabilities of machine learning — and especially the use of large multilayer neural network models in deep learning — currently depend on the appropriation of personal and public data by private companies on a massive scale. The problem in the regulation of machine learning is that the regulator always seems to be one step behind the innovation: ideally government should take a more pro-active approach, but this stands in tension with risks. As to participation, the debate about machine learning has long been dominated by experts but ChatGPT is bringing many more people into this debate. Arguably, indeed, this controversy has always had global reach, as the implementation of machine learning across societal domains means that virtually everyone on the planet is now part of the training data for machine learning applications. Solvability will depend on wide levels of public literacy in the population and the efficacy of monitoring arrangements. The public is currently not engaged actively enough with the harmful consequences of machine learning.

Large Language Models (Stochastic Parrots)

This controversy highlighted that the risks posed by large language models (also referred to as Transformer models, and which are also used in Generative AI) are not only technical but also arise from the concentration of power in big tech companies who control both the research on and societal implementation of AI. This controversy did most to surface several structural challenges of AI, such as the environmental impact of large models, the marginalisation of women in AI research, and the lack of transparency of model data and development (major issues). It also put on display the relative invisibility of governmental actors and regulators in the public debate about the societal risks of AI. It was noted that not only regulators but also societal actors have been largely reactive rather than proactive in specifying requirements on AI development, with the exception of the positive case being made for an inclusive research culture (no innovation without representation). As a consequence of this controversy, big tech firms may have become less willing to support research into ethical and responsible AI, and as such it further underscores the importance of regulation in setting standards in AI research and innovation. But the political economy of big tech makes this controversy hard to solve.

The discussions in the different expert groups surfaced several **common themes**: the appropriation of personal and public data for commercial purposes, in online and public sector settings; the concentration of power in the Silicon Valley companies that is consolidated through the development and implementation of AI across societal domains; the relative invisibility of policy-makers and governments in the debates, suggesting agenda-setting capacity resides mostly with industry and activist organisations; the relative lack of public literacy in most of the areas of controversy, and of public awareness of the societal and environmental risks posed by AI.

7. How do we get from the AI controversies we have to the controversies we need?

Once workshop participants had determined the "shapes" of the selected AI controversies, the small groups were asked to project the shape of the AI controversy that society in their view needs to have (projective activity). Working with plasticine, they projected the desired features and formulated requirements for AI controversies to come. This discussion yielded the following insights and recommendations:

- *Solvability of AI controversies depends on public literacy*

Without increased public understanding of technological, societal and environmental risks posed by AI, it will be much more difficult to regulate these risks.

- *Technical definitions dominate societal understandings.*
In recent AI controversies, the emphasis on technical aspects has been used to avoid going into the societal impacts of AI.

- *Controversy is a Pandora's box that sparks yet more controversy*
If a controversy does not get settled, it just connects to and overspills into other controversies and generates more controversies.

- *There is a lack of infrastructure to facilitate public participation in AI debates*
As a consequence, AI debates are dominated by tech industry and professional experts and activists.

- *The regulatory sense of the state's duty of care is entangled with controversy*
There is a close connection between controversy and regulation. Regulators care about controversies, which serves as a context in which legislation can be brought in.

- *The importance of timing in public consultation*
When issues of technological risk first arise, they tend to be highly specialist, while later on the conflict is likely to have hardened into consolidated positions, public consultation therefore ideally takes place in the middle, in between these two phases in a controversy's development.

- *Many problems with AI are technically solvable, but the political economy of big tech makes them seemingly impossible to solve.*
Mechanisms of data protection, transparency, informed consent, and risk and impact assessments are readily available, but in many domains remain unenforced due to constraints in techno-economical architectures underpinning the development of AI.

Conclusion: will AI controversies be able to inform AI's future?

Our social research on AI controversies is on-going, but one of the main findings to emerge from the "Shifting controversies" workshop is that much expert concern and disagreement is focused on the underpinning socio-technical architectures and techno-economic arrangements of contemporary AI itself, which enable opaque systems of data capture, the widespread use of public and personal data for private gain, as well as favouring an innovation model that carries significant environmental and societal costs (energy use, worker exploitation, discrimination). This contrasts with the current focus of regulatory agenda's on domain-specific implementations of AI,[8] and suggests that the role of expert controversy as a mechanism of policy learning and public awareness raising - a key element in the responsible innovation paradigm - has been relatively limited to date in the case of AI (Coad et al, 2021). However, our research on AI controversies also shows that to date participation in these controversies to date has been diverse even if relatively narrow. Extended expert communities across industry, science and activism have engaged extensively in problem definition during the last 10 years, with much of the debate directing attention to risks, harms and potential benefits arising from the methodological architecture of general purpose AI. As such, AI research controversies of recent years may be understood as offering a research-centric preview of Generative AI controversies that might develop in the coming years, as text and code generators like ChatGPT are marked by precisely the structural features of AI that have been identified as requiring attention and intervention through these controversies.

8 AI regulation: a pro-innovation approach, Policy Paper, Department of Science, Innovation and Technology and the Government's Office for AI, 29 March 2023, https://www.gov.uk/government/publications/ai-regulation-a-pro-innovation-approach

We are currently preparing several journal publications in which the provisional results reported here will presented in more detail. If you would like to stay informed about these publications and our on-going study of AI controversies, please contact Dr. Chiara Poletti at chiara.poletti@warwick.ac.uk.

References

Angwin, J., Larson, J., Mattu, S., & Kirchner, L. (2016). Machine bias, Propublica report.

Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021, March). On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?. In *Proceedings of the 2021 ACM conference on fairness, accountability, and transparency* (pp. 610-623).

Coad, A., Nightingale, P., Stilgoe, J., & Vezzani, A. (2021). The dark side of innovation. *Industry and Innovation*, *28*(1), 102-112.

Funtowicz, S., & Ravetz, J. (1997). Environmental problems, post-normal science, and extended peer communities. *Études et Recherches sur les Systèmes Agraires et le Développement*, 169-175.

Housley, W., Webb, H., Williams, M., Procter, R., Edwards, A., Jirotka, M. & Williams, M. (2018). Interaction and transformation on social media: the case of Twitter campaigns. *Social Media+ Society*, *4*(1), 2056305117750721.

Kimbell, L. (2019). What if there were more policy futures studios?. *Journal of Futures Studies*, *23*(4), 129-136.

Ranchordas, S. (2021). Experimental regulations for AI: sandboxes for morals and mores. *University of Groningen Faculty of Law Research Paper*, (7).

Madsen, A. K., & Munk, A. K. (2019). Experiments with a data-public: Moving digital methods into critical proximity with political practice. *Big Data & Society*, *6*(1), 2053951718825357.

Marcus, G. (2018). Deep learning: A critical appraisal. *arXiv preprint arXiv:1801.00631*.

Marres, N., & de Rijcke, S. (2020). From indicators to indicating interdisciplinarity: A participatory mapping methodology for research communities in-the-making. *Quantitative Science Studies*, *1*(3), 1041-1055.

Marres, N. (2015). Why map issues? On controversy analysis as a digital method. *Science, Technology, & Human Values*, *40*(5), 655-686.

Powles, J., & Hodson, H. (2017). Google DeepMind and healthcare in an age of algorithms. *Health and technology*, *7*(4), 351-367.

Roberge, J., & Castelle, M. (Eds.). (2021). *The cultural life of machine learning: an incursion into critical AI studies*. Palgrave Macmillan.

Slater, D., & Barry, A. (Eds.). (2005). *Technological Economy*. Routledge.

Stirling, A. (2008). "Opening up" and "closing down" power, participation, and pluralism in the social appraisal of technology. *Science, Technology, & Human Values*, *33*(2), 262-294.

Wang, Y., & Kosinski, M. (2018). Deep neural networks are more accurate than humans at detecting sexual orientation from facial images. *Journal of personality and social psychology*, *114*(2), 246.

Whatmore, S. J. (2009). Mapping knowledge controversies: science, democracy and the redistribution of expertise. *Progress in Human Geography*, *33*(5), 587-598.