

# **‘Frontier AI,’ Power, and the Public Interest: Who Benefits, Who Decides?**

**David Leslie<sup>1,2</sup> Carolyn Ashurst<sup>1</sup> Natalia Menéndez González<sup>3</sup>  
Frances Griffiths<sup>4,5</sup> Smera Jayadeva<sup>1</sup> Mackenzie Jorgensen<sup>6</sup>  
Michael Katell<sup>1,2</sup> Shyam Krishna<sup>1,2</sup> Doschmund Kwiatkowski<sup>7</sup>  
Carolina Iglésias Martins<sup>8</sup> Sabeegah Mahomed<sup>1</sup> Carlos Mougán<sup>1</sup>  
Shachi Pandit<sup>9</sup> Mark Richey<sup>10</sup> Joseph W. Sakshaug<sup>11</sup> Shannon Vallor<sup>12</sup>  
Luke Vilain<sup>13</sup>**

<sup>1</sup>The Alan Turing Institute, London, England, United Kingdom,

<sup>2</sup>Queen Mary University of London, London, England, United Kingdom,

<sup>3</sup>Centre for a Digital Society, European University Institute, Florence, Italy,

<sup>4</sup>Warwick Medical School, University of Warwick, Coventry, England, United Kingdom,

<sup>5</sup>Centre for Health Policy, University of the Witwatersrand, Johannesburg, South Africa,

<sup>6</sup>Kings College London, London, England, United Kingdom, <sup>7</sup>AIMEE Project, Fleet, United Kingdom,

<sup>8</sup>Independent Scholar, <sup>9</sup>Roche, United Kingdom, <sup>10</sup>University of Dundee, Dundee, Scotland,

<sup>11</sup>University of Warwick, Coventry, England, United Kingdom,

<sup>12</sup>Centre for Technomoral Futures, Edinburgh Futures Institute, University of Edinburgh,  
Edinburgh, Scotland,

<sup>13</sup>UBS, England, United Kingdom

**The MIT Press**

**Published on:** Sep 09, 2024

**DOI:** <https://doi.org/10.1162/99608f92.4a42495c>

**License:** [Creative Commons Attribution 4.0 International License \(CC-BY 4.0\)](#)

## ABSTRACT

As the rapid industrialization of generative AI (GenAI) reached a crescendo in the fall of 2023, a series of international AI policy initiatives, like the UK AI Safety Summit and G7’s Hiroshima Process, cropped up as a response to corresponding global AI governance challenges. The policymakers and government officials who drove these initiatives emphasized that the rise of ‘frontier AI’ technologies was bringing humankind to a historical inflection point—placing humanity at a crossroads and situating the present generation at an axis of choice which would determine whether the evolution of AI innovation moves toward the exponential advancement of the public good or toward possibilities for potentially irreparable harm to people, society, and the planet. And yet, despite the inflationary rhetoric of historical transformation around which these policy initiatives were framed, their actual results (high-level voluntary commitments, non-binding codes of conduct, the formation of light-touch national AI safety institutes, etc.) seem vastly out of sync with the scope and scale of the problem to which such initiatives claimed to respond.

In this article we argue that, if anything, this framing of the GenAI moment as a historical pivot point raises fundamental questions about the thorny relationship of ‘frontier AI,’ power, and the public interest, namely: Who actually has their hands on the wheel? Who defines the agenda of ‘frontier AI’ innovation? Who controls the means of producing it, and thus ultimately its influence on humanity’s broader fate? These questions cut much deeper than those surrounding the risks arising from unforeseen advances in ‘frontier AI’ capability or those around pre-deployment safety testing that took up much of the oxygen at the UK AI Safety Summit and related international AI policy discussions. They have to do with who possesses agenda-setting power, who decides on when, where, and how these technologies are developed and deployed, and who stands to benefit from or be harmed by them. They also raise further questions about how affected members of society can harness this control over the direction of AI to serve the common good and thereby exercise agency over the trajectories of their own collective futures. We will claim that, ultimately, an effective response to these questions demands a fundamental rethinking of the broader political economy of AI and of the global innovation ecosystem which drives its forward progress—a rethinking that recasts this technology as a global public utility subject to democratic control, community-led agenda-setting, and society-centered regulation.

**Keywords:** generative AI, frontier models, AI ethics and governance, public interest, power, public utility

---

## 1. Introduction

In the weeks that followed the AI Safety Summit at Bletchley Park, reports of its overwhelming success flooded ministerial statements and social media feeds. Just after the summit wound to a close, the UK’s Secretary of State for Science, Innovation, and Technology, Michelle Donelan, told Parliament that the agreement of the “Bletchley Declaration,” a position statement on the future of AI signed onto by 28

participant states, made the summit, “a watershed moment in the history of AI safety” ([Donelan, 2023](#)). She went on to claim that the “securing of an agreement on the government-led testing, pre-deployment of [‘frontier AI’ models][...]was truly a game changer,” making the summit “surpass all expectations.”

News analyses, published at the time of the summit, were less sanguine. An opinion column from *The Guardian* called the whole affair a “vanity jamboree” engineered by “a diminished nation...trading on Silicon celebrity” instead of advancing “democratic accountability” ([Behr, 2023](#)). Another from the *Financial Times* called the Bletchley Declaration “worthy but toothless” ([Thornhill, 2023](#)). An analysis from the *New Scientist* similarly concluded that the declaration

does little more than acknowledge that there are risks and pledge to explore them. The only concrete action promised in the wording is to hold more summits in the future. Perhaps this meeting could have been a ChatGPT-generated email, and saved the carbon expenditure of jetting everyone in. ([Sparkes, 2023](#))

While the mixed reception of the summit’s outcomes among these critics seems to display a sensible sobriety, we need to pay close attention to a more consequential but broadly underreported elephant that remains poised in the middle of the room. The unspoken issue has to do with the contradiction between the missionary rhetoric, and sense of historical self-importance, that seemed to largely motivate and frame the summit and its relatively meek and ineffectual results.<sup>1</sup>

The summit had originally been promoted by the Sunak government as a pivotal historical event. Sunak claimed, in his pre-summit speech launching the UK’s AI Safety Institute, that AI “will bring a transformation as far-reaching as the industrial revolution, the coming of electricity or the birth of the internet” ([Sunak, 2023](#)). “Get this wrong,” he continued, “and AI could make it easier to build chemical or biological weapons. Terrorist groups could use AI to spread fear and destruction on an even greater scale. Criminals could exploit AI for cyberattacks, disinformation, fraud, or even child sexual abuse.” Get this right, however, and “nothing in our foreseeable future will be more transformative for our economy, our society, and all our lives than this technology.”<sup>2</sup>

Indeed, the Bletchley Declaration seemed to parrot this Janus-faced framing: On the one hand, it states, “there is potential for serious, even catastrophic, harm, either deliberate or unintentional, stemming from the most significant capabilities of [‘frontier AI’] models” ([Bletchley Declaration, 2023](#)). On the other hand,

AI has the potential to transform and enhance human wellbeing, peace and prosperity[...]We recognize that this is therefore a unique moment to act and affirm the need for the safe development of AI and for the transformative opportunities of AI to be used for good and for all, in an inclusive manner in our countries and globally. This includes for public services such as health and education, food security, in science, clean energy, biodiversity, and climate, to realize the enjoyment of human rights, and to

strengthen efforts towards the achievement of the United Nations Sustainable Development Goals.  
([Bletchley Declaration, 2023](#))

The contrastive narrative that cropped up in both Sunak’s rhetoric and the Bletchley Declaration attempted to situate the summit as a needed intervention at an inflection point in human history—a moment where decisions made now about AI’s present and future direction of travel, and about how to set the rapid development of ‘frontier’ models on the right course, would dramatically shape the fate of humankind and the biosphere.

But, here’s the elephant in the room: If, as in this framing, AI is, in fact, placing humanity at a turning point—if it is situating the present generation at an *axis of choice* that will determine whether the evolution of AI innovation moves toward the exponential advancement of the public good or toward possibilities for potentially irreparable harm to people, society, and the planet—then the most important questions of our time are: Who actually has their hands on the wheel? Who defines the agenda of ‘frontier AI’ innovation? Who controls the means of producing it, and thus ultimately its influence on humanity’s broader fate?

These questions cut much deeper than those surrounding the risks arising from unforeseen advances in ‘frontier AI’ capability or those around pre-deployment safety testing that took up much of the oxygen at the summit. They have to do with who possesses agenda-setting power, who decides on when, where, and how these technologies are developed and deployed, and who stands to benefit from or be harmed by them. They also raise further questions about how affected members of society can harness this control over the direction of AI to serve the common good and thereby exercise agency over the trajectories of their own collective futures.<sup>3</sup>

One of the greatest threats of the current situation is that such critical public interest questions continue to be placed at the fingertips of the same large private sector tech companies whose exertion of brute-force agenda-setting authority, closely guarded control over critical digital infrastructure, and narrow-minded financial thrust have largely engendered the crisis that motivated the summit itself. In fact, the deafening silence around these issues related to the fraught relationship between AI innovation, power, and the public interest could be reasonably seen as the greatest single failing of the summit and related international AI governance initiatives.

In this analysis piece, we will explore how a critical awareness of this thorny relationship of ‘frontier AI’ technologies, power, and the public interest provides occasion for a radical reconceptualization of AI policy and governance futures. We will begin by interrogating how the living heritage of what was first famously called the “Californian ideology” ([Barbrook & Cameron, 1996](#)) in the 1990s gains new ethical and societal significance in the rising age of large language models and generative AI (GenAI). The heritage of the Californian ideology is characterized by a focal tension. The very same Silicon Valley tech corporations who have, for decades, claimed to be serving humanity’s interest through the prosocial bounty of their innovation activities, have, at the same time, relentlessly pursued self-interested and market-oriented goals that have belied this projected beneficence. In many cases, their actions have operated *in actu* only to further entrench social inequity and systemic discrimination, to displace and exploit human labor, and to noxiously consolidate

financial, market, and techno-scientific power. In this way, the innovation practices of such firms have ended up being precisely at cross-purposes with the public interest they claim to champion. We will argue that surfacing the history of harm, duplicity, and public interest washing that has correspondingly typified the rise of big tech's hegemonic command over the direction of travel of 'frontier AI' innovation has become all-the-more vital amid the rapid and historically unprecedented commercialization of GenAI systems. This is the case because the exponentially expanding scale and scope of their impacts has come to affect the public interest in historically unprecedented ways. We will suggest, finally, that this seemingly intractable dynamic demands a fundamental rethinking of the broader political economy of AI and of the global innovation ecosystem that drives its forward progress—a rethinking that recasts this technology as a global public utility subject to democratic control, community-led agenda-setting, and society-centered regulation.

## 2. The Living Heritage of the Californian Ideology in the Rising Age of 'Frontier AI'

Contradictions are nothing new to the Silicon Valley tech ethos. A generation ago, [Richard Barbrook and Andy Cameron \(1996\)](#) puzzled at how a “bizarre fusion of the cultural bohemianism of San Francisco with the hi-tech industries of Silicon Valley” had created what they called the “Californian Ideology”—a utopian vision of market-based digital society that implausibly combined a libertarian zeal for the saving power of heroic entrepreneurs and a solutionist reliance on the emancipatory potential of information technologies with a naïve faith in the creative and communalist possibilities unlocked by the “electronic agora” (p. 48).

For [Barbrook and Cameron \(1996\)](#), the tendency of “West Coast ideologues” to “amalgamate opposites” came at a cost. Claims to advance the common good and to free society through a commitment to market fundamentalism, big finance, founder genius, tech solutionism, and “an impeccably libertarian form of politics” tended to be specious and self-serving. It was a way for Silicon Valley players to charge ahead with their aspirations for rapid growth, market domination, and windfall profits, no-holds-barred, while strategically evading the underlying problems of socioeconomic inequality, institutional racism, poverty, and environmental degradation that their activities were, in fact, exacerbating. In this way, what might have otherwise seemed like an amusing but harmless syncretism of hippie counterculture and hi-tech libertarianism signaled a more pernicious deception. Ambient Silicon Valley narratives peddling “the countercultural dream of empowered individualism, collaborative community, and spiritual communion,” concealed a bleaker bait-and-switch reality more accurately characterized by the relentless pursuit of profit, finance-driven growth at any cost, and entrenched dynamics of human displacement, exploitation, and discrimination, a reality that Malcolm Harris has simply called, “the Palo Alto system” ([M. Harris, 2023](#)).

In the incipient age of foundation models and GenAI, the living heritage of the Californian ideology gains new ethical and societal significance for two closely connected reasons. First, the direction and pace of the GenAI revolution has largely been steered by a handful of big tech companies whose agenda-setting power, control over digital infrastructure, and market-driven behavior have grown directly out of the Palo Alto system.

Over the last 2 decades, despite the public-spirited slogans, do-gooder rhetoric, and claims to openness and democratization<sup>4</sup> that these corporations have often peddled, their relentless drive for economic growth, expanding market share, and optimizing profits have led to a noxious consolidation of financial and techno-scientific power. This has meant that a small oligopolistic group of private firms has essentially come to monopolize the critical data, compute, and skills infrastructures that are necessary preconditions for the development of 'frontier' GenAI technologies.

These technologies have, in effect, emerged as by-products of market-oriented corporate capture and blitzkrieg-style industrial scaling rather than as end-products of inclusive, equitable, and public interest-oriented scientific research: In the hands of a few large tech firms, massive data corpora and brute force computing power have converged with a monopolized labor force of highly paid tech workers, who have applied the fruits of generations of computational and software engineering research to produce commercializable AI products that can be monetized as digital platforms. The rising age of foundation models and GenAI thus stands at the culminating point of cartel-controlled AI industrialization and a corollary drive to "assetize" ([Birch et al., 2020](#); [Fourcade, 2021](#)) AI "as-a-service" ([Mazzucato et al., 2022](#)). This has signified a systemic siphoning off of the incentives for large-scale AI innovation from the tackling of public good problems into the pursuit of ever narrower economic ends ([Drahos, 2020](#); [Klinger et al., 2022](#))—and questions of when, where, and how these technologies should be developed and deployed and of who stands to benefit from, or be harmed by, them have remained questions primarily decided upon by the predominant corporate heirs of the Palo Alto system.

The significant threat posed to people, society, and the planet by placing such crucial *public interest questions* at the fingertips of private sector tech industry elites leads us to the second reason why the living legacy of the Californian ideology gains unique importance in the GenAI era. As the scope, scale, and gravity of the societal impacts of the industrial-scale deployment of GenAI continue to expand, so too do the real-world ramifications of what tech journalist Kate Losse in 2019 styled as "the false promise of Silicon Valley's quest to save the world" ([Losse, 2019](#)). For several years, tech pundits and practitioners alike have widely perceived the self-contradictory nature of the mission-driven humanitarian pronouncements and platitudes tendered by large tech corporations, which have paraded mottos like 'don't be evil,' 'working for the common good,' and 'giving people the power to build community and bring the world closer together,' while simultaneously engaging in a broad spectrum of antithetic, socially destructive behaviors. Such altruistic bluster—what we might think of simply as public interest washing<sup>5</sup>—has tended to conceal the corrosive, downstroke history of predatorial data practices, sociopathic profiteering, data worker exploitation, staff purging, and squashed tech worker revolts that has more accurately represented the march of large-scale AI research and innovation into the GenAI era. From the [crushing of worker protests against multi-billion dollar military AI contracts at Microsoft](#), the [Frances Haugen whistle-blower scandal at Facebook](#), and the dehumanizing big tech scramble to monetize manipulated behavior at scale in the 'attention economy' ([Wu, 2017](#); [Zuboff, 2019](#)) to the [firing of AI ethicists who criticized large language models at Google](#) and the [exploitation and abuse of data workers in the "Global South,"](#) this darker history of AI industrialization has looked more like an unsustainable

race-to-the-bottom on responsible and socially sustainable AI than a mission-led corporate ascent into the light of AI-for-the-common-good.

The surfacing of such a dismal history of harm, duplicity, and public interest washing has become all-the-more vital amid the rapid and historically unprecedented commercialization of GenAI systems. The release and meteoric rise of ChatGPT in late 2022 triggered a fierce struggle for competitive advantage within the global AI research and innovation ecosystem, with tech titans immediately asserting their proprietary and rentiership claims over enterprise ‘frontier’ AI innovation. Big tech companies, like Google, Meta, and Microsoft, moved quickly to integrate these systems into their flagship services and to consolidate their market control over the population- and ecosystem-level distribution of the affordances of GenAI technologies. At the same time, they marketed their activities as “democratizing access” to AI, promoting “optimal flourishing,” “decentralizing power,” and advancing humanity’s march toward the universally beneficial windfalls of so-called “artificial general intelligence” ([Altman, 2023a](#)). The latter expression (often referred to just as ‘AGI’) emerged here as a term of art of the new Californian ideology of the GenAI era—a vague and speculative idiom that critics have decried both as a pseudoscientific tool pressed into the service of AI boosterism and Silicon Valley eschatology ([Goodlad, 2023](#); [Jarvis, 2023](#)) and as a “science-fiction manoeuvre” ([Hanna & Bender, 2023](#)) deployed to steer statutory and policy attention away from the mobilization of the robust regulatory controls needed to redress tangible risks and harms (like algorithmic bias and discrimination or copyright infringement) and to constrain the hapless behavior of these same corporate actors.

The convergence of this dawning of GenAI’s “age of competition” ([Carugati, 2023](#)) with big tech’s public-interest-washing gambit created something of a perfect storm within a global AI governance ecosystem in which these technologies were now increasingly affecting the fundamental rights and freedoms of hundreds of millions of people across veritably every profession and domain of social life. Alongside the blindered struggle for market expansion and control that ensued, the absence of regulatory mechanisms to put checks on the design, development, and use of these technologies coalesced with widespread protestations among researchers and members of the public about their harmful effects. These increasingly vocal objections to the helter-skelter release of GenAI systems into the social world included complaints about their questionable legality, their far-reaching violations of intellectual property, labor, and privacy rights, their capacity to spew disinformation and false, toxic, and discriminatory content, their environmental costs, and their potential negative impacts and transformative effects on society, more generally (for example, [Alba, 2022](#); [Bender et al., 2021](#); [Birhane et al., 2023](#); [Bommasani et al., 2021](#); [Hao, 2019](#); [Hazell, 2023](#); [Klepper, 2023](#); [Kurenkov, 2022](#); [Leslie & Rossi, 2023](#); [Mascellino, 2023](#); [Perrigo, 2021](#); [Shelby et al., 2023](#); [Vock, 2022](#); [Weidinger et al., 2021](#)). The vacuum created by this outsized gap in technology policy and regulation triggered an international AI governance crisis. The urgent questions of who governs “frontier AI” models and their applications and how lay open and undecided.



And yet, though the stage had been set for definitive, indeed pivotal, policy action to protect the public interest, the entrenched power dynamics of the global political economy of AI came quickly to steer the international AI policy and governance agenda. Instead of safeguarding the public good by confronting the range of extant hazards to people, society, and the biosphere produced by the design, development, and deployment of GenAI systems, policymakers, and governments (largely dominated by the amplified voices of countries from the Global North) deferred to the corporate prerogatives of private sector technical ‘experts,’ narrowing the governance discussion to model-centered ‘AI safety’ and zooming in on “technical methods for avoiding hypothetical ‘extreme risks’ that could emerge from the misuse or loss of control of advanced ‘frontier AI’ systems” ([Smakman et al., 2023](#)). This meant that initiatives like the UK AI Safety Summit “over-indexed on hypothetical future harms,” while failing to effectively target the far-reaching dangers of the use of AI “in the context of broader sociotechnical systems in which it is always embedded” ([Lazar & Nelson, 2023](#); cf. [Gebru et al., 2023](#); [Ryan-Mosley, 2023](#); [Vallor & Luger, 2023](#)). Likewise, the blindered focus on technical measures placed the handful of tech companies—who monopolized the means of ‘frontier AI’ production and were thus exceptionally positioned to shape and implement such measures—in a position of overarching epistemic authority. This erroneously allowed for definitions of essential terms like ‘safety,’ ‘risk,’ ‘trust,’ and the ‘public good’ to be controlled by the ‘mono-culture’ of Silicon Valley tech elites rather than being shaped and decided on by the broad spectrum of communities and stakeholders impacted by the swift proliferation of these technologies ([Lazar & Nelson, 2023](#)).

Such a widespread obeisance of governments to the priorities, interests, and epistemic framings of private sector tech companies unsurprisingly resulted in policy actions and initiatives built around voluntary agreements and nonbinding declarations rather than compulsory statutory and regulatory interventions with substantial enforcement mechanisms. The [Bletchley Declaration](#), the G7 Hiroshima Process’s “[International Code of Conduct for Organizations Developing Advanced AI Systems](#),” and the Biden administration’s “[Voluntary Commitments From Leading Artificial Intelligence Companies to Manage the Risks Posed by AI](#)” all suffered from this deferent fate. Taken together, the geopolitical interests of Western industrialized states locked in a technological race in AI with China and other potential adversaries and the anemic technological infrastructures that characterized the innovation capabilities of these same Western countries (in the wake of decades of the neoliberal impoverishment of public sector capacity) ([Benkler, 2022](#)) incentivized them to pander to the impulses of ‘frontier AI’ companies and to enable further power consolidation among their small number.

Meanwhile, these same private sector firms escalated their execution of public interest washing strategies. In his testimony to the US Senate, Sam Altman, the CEO of OpenAI, vigorously insisted on the need for robust regulation of ‘frontier AI’ to protect the public, citing his organization’s overarching nonprofit status and public interest driven mission ([Altman, 2023b](#)), which, as articulated in OpenAI’s Founding Charter, entailed its “primary fiduciary duty [...] to humanity” and its commitment “to avoid enabling uses of AI that harm humanity or unduly concentrate power” (OpenAI, 2018). At the same time, however, Altman actively moved,

behind the scenes, to block a provision of the [EU AI Act](#) that would have placed stringent regulatory oversight on his company's AI products ([Perrigo, 2023](#)). He also pushed aggressively to accelerate the commercialization of ChatGPT, despite widespread consternation both within OpenAI and in the press and civil society about how such a hasty 'move fast and break things' attitude contradicted the organization's mission and increasingly put the public at significant risk ([Hao & Warzel, 2023](#)). In this context, the contemporaneous Sam Altman termination debacle only further solidified the grim reality of the new Californian ideology: While Altman's abrupt firing in November 2023 appeared to validate the oversight power of OpenAI's mission-driven nonprofit board, the ensuing saga that led to his restoration as CEO a few days later only seemed to affirm that, as [Hao and Warzel \(2023\)](#) put it, "despite its charters and lofty credos, OpenAI was just a traditional tech company after all." In the run-up to Altman's rehiring, an overwhelming counterpressure campaign was launched both by private sector investors in OpenAI's capped-for-profit arm—led by its main multi-billion-dollar backer, Microsoft—and by its employees, who, as equity holders, stood to lose dearly in wake of his departure, following a recent 86-billion-dollar valuation of the company and a tender offer to enrich them with a corresponding payout ([Levine, 2023](#)). Altman's return, the dismissal of the nonprofit board members who had initiated the termination, and the creation of a seat for Microsoft on that same board emphatically signaled, in Brian Merchant's critical framing, "a further consolidation of power of one of the biggest tech companies and less accountability for the product than ever" ([Merchant, 2023](#)). Echoing the growing body of researchers and scholars who have stressed the importance of acknowledging how the dynamics of centralized economic, market, and infrastructural power shape the global data and AI innovation ecosystem ([Abdalla & Abdalla, 2021](#); [Khan & Vaheesan, 2017](#); [Veale et al., 2023](#); [Widder et al., 2023](#); [Whittaker, 2021](#)), Merchant concludes:

If anyone still believes a company can steward the development of a product like AI without taking marching orders from Big Tech, I hope they're disabused of this fiction by the Altman debacle. The reality is, no matter whatever other input may be offered to the company behind ChatGPT, the output will be the same: Money talks. ([Merchant, 2023](#))

### 3. Rethinking AI as a Global Public Utility

A failure to address the fundamental disconnect between the agenda-setting control over AI technologies possessed by self-interested big tech players, who possess immense amounts of infrastructural, economic, and market power, and the undeniable global impact on the public interest that the proliferation of these technologies is having could be disastrous for the future of society, as we know it.<sup>6</sup> And yet, on the other side of last year's UK AI Safety Summit and related international AI policy and governance initiatives, we are left to wonder why such a massively consequential issue remains a purloined letter—largely unseen, unspoken, and unaddressed.

If anything, the summit and adjacent activity were typified by a wholesale avoidance of serious thinking about how to confront the structurally entrenched imbalances in access to and control over the research and

innovation capabilities and resources that are vitally determining the development of the global AI ecosystem. This led these initiatives, and their soft-diplomatic outcomes, into a blind spot. Where large-scale multinational companies control critical digital and labor infrastructures as well as goods and services *that substantially affect the public interest*, these organizations should face concrete forms of ex ante legal intervention and regulatory control that constrain market-driven behaviors, tendencies toward monopolization and power centralization, and profit-led tolerance of societally damaging externalities.

Such a demand for the heightened public accountability of corporate actors whose products and services have, for better or worse, come to significantly affect the interests of society is, of course, nothing new. Rather, it points to a longstanding heritage of public utility thinking. Buried deep in the English common law tradition, in a 1670 work entitled *de Portibus Maris*, Sir Matthew Hale first defined public utilities as those social goods “that are affected with a public interest” ([Hamilton, 1930, p. 1090](#)) and thus cease to be justifiably placed under the control of private sector actors alone. By extension, Hale’s idea is that establishing a public utility becomes necessary where the provision of a good or service is, as the legal scholar K. Sabeel Rahman puts it, “of sufficient social value to be a necessity, and where the provision of this necessity is at risk of subversion or corruption if left to private interests or market forces” ([Rahman, 2018, p. 1635](#)).<sup>7</sup>

Over the course of modern times, industries affected with the public interest—especially those supplying and administering essential infrastructure—have been placed under this kind of public utility rubric and designated as ‘common’ or ‘public’ ‘carriers’ and ‘public callings’ ([Novak, 2017](#)). Such industries, whose production of goods and services required market consolidation of the sort that limited ordinary accountability mechanisms, have been subject to special restrictions such as the duty to serve all comers, to charge just and reasonable prices, and to offer reliable and accessible benefits to the public universally, even-handedly, and fairly ([Boyd, 2014](#); [Rahman, 2018](#)).

The basic idea here is simply that society must be empowered to govern industries, and technologies, that unlock access to essential public benefits and that thus shape crucial dimensions of collective life, which have bearing on the common interest, without there being corresponding checks in place on the exercise of that authority. Public oversight and control must be safeguarded, in this respect, so that affected people can be protected against the considerable moral and societal hazards that attend leaving the access key to the provision of those important public goods in the hands of potentially unchecked market power and private influence ([Boyd, 2014](#); [Rahman, 2018](#)). The public utility concept functions, in this respect, as a means for democratically “harnessing the power of private enterprise and directing it toward public ends” ([Boyd, 2014, p. 1636](#))—whether these ends are related to protecting people, society, and biosphere from the proliferation and use of harmful, unreliable, unsafe, and inequitable technology or to pressing responsible innovation into the service of advancing individual flourishing, societal sustainability, and planetary health.

While of obvious necessity, applying this kind of public utility thinking to the production and use of advanced AI systems is, however, not a straightforward endeavor. In the incipient era of AI’s industrial revolution (a

period so clearly evidenced by the breakneck industrialization of GenAI applications), even those who would advocate for the implementation of a public utility framework in the domain of AI face a new and unprecedented challenge that stems from the boundary-bursting character of its transformative effects. What is unique about AI technologies, in this context, and in contradistinction to other, more conventional, public utilities like electricity or public transportation, is that the expanding scale and scope of their societal impacts exceeds any arbitrarily delimited geopolitical boundaries that might circumscribe notions of the ‘common good’ or ‘public interest’ within local or national peripheries. Not only has the range and number of people, communities, and geographies affected by AI technologies exploded, transgressing nation-state borders; so too has the gravity of their expected consequences—both in terms of the seriousness of potential harms (e.g., emergent species-level risks of AI-enabled nuclear-, chemical-, or bioterrorism) and the significance of potential benefits (e.g., emergent possibilities of pressing AI into the service of tackling previously unbroachable biomedical and public health challenges, of combating climate change and biodiversity loss, of bolstering food security and fighting world poverty, etc.<sup>8</sup>).

This explosion of the scale and scope of the impacts of AI technologies triggers a kind of *moral and legal rescaling imperative*: The ethical and statutory demands arising from the design and implementation of AI systems require a due regard of impacted people that echoes outward from individual agent to community to nation-state to species, and to biosphere, implicating the producers and users of those systems at each step of the way. These are global technologies, par excellence, and their consequences must be considered in that covering light. This is especially true of the potential for the asymmetrically distributed benefits and risks of AI innovations to have devastating downstream effects on the vulnerable and digitally disenfranchised individuals and communities most impacted by widening digital divides and historically entrenched patterns of local, regional, and global inequity.

What all this means is that AI must be understood not merely as a public utility, full stop, but rather as a *global public utility*, which demands policy innovations that respond to the global, indeed biospheric, scale of its risks and negative and positive impacts.

Seen as purveyors of global public utilities, AI industrial actors would be subject to the proactive readjustment of regulatory and governance perimeters on a transnational plane. This would require the codification of binding international legal instruments as well as the cooperative formation of an inclusive network of nonpartisan intergovernmental agencies that:

1. set binding and enforceable gold standards for the ethical, rights-preserving, responsible, and equitable design, development, and deployment of AI systems;
2. put in place mechanisms for both ex-ante and ex-post third-party audits of the safety, social, and environmental sustainability, reliability, and equity of such algorithmic systems;
3. limit risks and catastrophic possibilities by setting up global-scale policing and oversight regimes that ensure sufficient transparency and accountability across the entire multiphase GenAI lifecycle ([Leslie & Rossi,](#)

[2023; Domínguez Hernández et al., 2024](#));

4. establish just and accessible processes for actionable recourse that enable affected people to seek redress when things go wrong without responsabilization ([Pyysiäinen et al., 2017](#); [Rose, 1996](#)) or placing undue burden on them;
5. ensure responsible sourcing and work conditions to safeguard labor rights and promote equitable skills development across the GenAI ecosystem ([Dencik, 2021](#); [Le Ludec et al., 2023](#)); and
6. ensure the advancement of AI innovation for the global public benefit, placing affected communities in the driver's seat for democratically determining present and future AI innovation agendas and fostering inclusive and equitable access to state-of-the-art research capabilities and resources.<sup>9</sup>

Crucially, such a recasting of AI as a global public utility would involve transforming the incentivization structures that currently drive AI research and innovation at scale. This would entail reshaping the market prerogatives that currently steer the growth-driven 'move fast and break things' attitudes of large technology firms whose disproportionate size and power enable them to act with reckless disregard and virtual impunity. More essentially, it would involve compelling these firms, as common carriers and trustees of the common benefit, to prioritize the responsible, ethical, and equitable provision and deployment of critical digital capacities, resources, and infrastructure to address the urgent and universally consequential public good problems—like global poverty, public health crises, biodiversity drain, and climate change—that will ultimately decide the fate of humanity and shared planetary life.

More critically still, as the last point in the list above underscores, the task of pinpointing and prioritizing such global public good problems must be continuously driven and shaped by "collective intelligence processes that leverage [the] participation and deliberation" ([Spada & Paulson, 2023, p. 78](#)) of impacted stakeholders. It is only through inclusive and society-led participation and deliberation processes that grand societal challenges—which cut across sectors, geographies, disciplines, socioeconomic stratifications, and national and cultural self-identifications—can be effectively and equitably confronted ([Warnke et al., 2023](#)). Tackling these whole-of-society challenges requires that the directional agency of mission-oriented AI innovation ([Mazzucato, 2018](#); [Mazzucato et al., 2022](#)) is steered by multiple vectors of insight from myriad perspectives and understandings (especially from historically marginalized, minoritized, or vulnerable social groups that would otherwise not be included in decision-making processes). Only in this way can the identification and articulation of the shared goals that can orient AI innovation in the direction of the public interest and the common good be underwritten by democratic legitimacy, social license, and public consent.<sup>10</sup>

In this vein, when integrated into AI innovation and policy lifecycles, community-driven participatory and deliberative processes of making explicit the public values behind implicit notions of the common good provide opportunities for the multiple publics that constitute the greater public sphere to push back on asymmetries of power that have ossified into non-dissolving socioeconomic structures. Such processes, in turn, capacitate affected members of the public to exercise influence over political and industrial decision-making.

In the specific context of the innovation ecosystem of AI and data-driven technologies, this enables collectively-articulated public interest claims to be “forced upon the ecosystem’s architecture” as sites of democratic resistance that challenge the “unprecedented infrastructural, sectoral, and intersectoral powers” of large tech companies that have gained seemingly unyielding predominance through scaled datafication, commodification, and techno-industrial centralization ([Van Dijck et al., 2018, p. 138](#)).

The vital role played by this kind of critical and constructive grassroots democratic practice spells the need for institutional creativity and ingenuity to leverage existing and emerging forms of community-led participation and deliberation to ever more inclusively convene impacted publics in the ends of identifying and articulating the shared goals that can orient AI innovation projects, initiatives, and policy in the direction of the public interest and the common good. Such novel forms of democratic empowerment would be the procedural linchpin of mobilizing AI as a global public utility. They entail conceiving and resourcing new ways of democratically assembling local, regional, and global communities from the bottom-up. They also entail the need to gather, convene, and connect impacted publics cross-nationally and transversally (i.e., in a way that decenters concentrated constellations of socioeconomic, cultural, and geopolitical hegemony) so to build and strengthen a global public sphere that is a sufficient democratic match to the global platform power and infrastructural preponderancy of big tech corporations and their geopolitical supporters and sponsors.

## 4. Conclusion

Discussions around the first wave of international AI governance initiatives this past year laid emphasis on the claim that the rise of foundation models and GenAI was bringing humankind to a socio-historical inflection point, placing humanity at a crossroads. Turn the right and responsible way, and all walks of species life across all parts of the globe could reap the benefits of advanced AI research and innovation, putting sustainable and flourishing global futures within reach. Turn the wrong way, and people, society, and the planet could suffer from potentially irreparable harm. If anything, this framing of the GenAI moment as a historical pivot point raises fundamental issues about the relation of technology, power, and the public interest, namely: Who should govern and decide the trajectories of advanced AI research and innovation that are radically *affected with the public interest*? Who should control the data, compute, and skills infrastructures that enable the progress of such research and innovation practices? And, how should the benefits and risks that derive from corresponding techno-scientific activities be distributed? As we have claimed in this article, such normative questions can be properly addressed only by first answering more basic, critical, and empirical questions about who, in the current political economy of AI, *actually* governs and decides on its direction of travel, who *actually* defines and dictates what is in ‘the public interest,’ who *actually* controls essential, innovation-enabling infrastructure, and who *actually* benefits or is put at risk from all this.

If there is one lesson to take away from the failures of the Bletchley Summit and related international governance initiatives, it is that, despite the lofty rhetoric of historical transformation around which this policy activity was framed, a set of high-level voluntary commitments to pre-deployment model evaluation, non-

binding codes of conduct, the formation of light-touch national AI safety institutes, and agreements to talk more are sorely lacking as coherent and satisfactory responses to the grave issues that humanity now faces. They come nowhere close to setting in motion the fundamental changes that need to be made in the political economy of AI to adequately address the serious risks and the urgent public good problems that have, in no small measure, been spurred and exacerbated by very same irresponsible corporate actors who now stake claim to saving humanity from them and who are driving the agenda of international AI policy discussions about possible solutions. Instead, as we have argued here, a confrontation of the current pathologies of systemic power asymmetry, structural inequity, economistic short-sightedness, and democratic deficit that shape the global AI innovation ecosystem demands a fundamental rethinking of the broader political economy of AI that recasts this technology as a global public utility subject to democratic control, community-led agenda-setting, and robust, society-centered regulation.

---

## Disclosure Statement

David Leslie, Carolyn Ashurst, Natalia Menéndez González, Frances Griffiths, Smera Jayadeva, Mackenzie Jorgensen, Michael Katell, Shyam Krishna, Doschmund Kwiatkowski, Carolina Iglésias Martins, Sabeehah Mahomed, Carlos Mougán, Shachi Pandit, Mark Richey, Joseph W. Sakshaug, Shannon Vallor, and Luke Vilain have no financial or non-financial disclosures to share for this article.

---

## References

- Abdalla, M., & Abdalla, M. (2021). The Grey Hoodie Project: Big tobacco, big tech, and the threat on academic integrity. In M. Fourcade & B. Kuipers (Eds.), *Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society* (pp. 287–297). ACM. <https://doi.org/10.1145/3461702.3462563>
- Aguilar, J., Garces-Jimenez, A., R-Moreno, M. D., & García, R. (2021). A systematic literature review on the use of artificial intelligence in energy self-management in smart buildings. *Renewable and Sustainable Energy Reviews*, 151, Article 111530. <https://doi.org/10.1016/j.rser.2021.111530>
- Alba, D. (2022, December 8). OpenAI Chatbot spits out biased musings, despite guardrails. *Bloomberg*. <https://www.bloomberg.com/news/newsletters/2022-12-08/chatgpt-open-ai-s-chatbot-is-spitting-out-biased-sexist-results?leadSource=uverify%20wall>
- Altman, S. (2023a, February 24). Planning for AGI and beyond. Open AI. <https://openai.com/index/planning-for-agi-and-beyond/>
- Altman, S. (2023b). *Written testimony of Sam Altman Chief Executive Officer OpenAI before the U.S. Senate Committee on the Judiciary Subcommittee on Privacy, Technology, & the Law*. US Senate.



<https://www.judiciary.senate.gov/imo/media/doc/2023-05-16%20-%20Bio%20&%20Testimony%20-%20Altman.pdf>

Antonelli, A., Gorla, S., Sterner, T., & Silvestro, D. (2021). *Optimising biodiversity protection through artificial intelligence*. BioRxiv. <https://doi.org/10.1101/2021.04.13.439752>

Bächtiger, A., Dryzek, J. S., Mansbridge, J., & Warren, M. E. (Eds.). (2018). *The Oxford handbook of deliberative democracy*. Oxford University Press.

Barbrook, R., & Cameron, A. (1996). The Californian ideology. *Science as Culture*, 6(1), 44–72. <https://doi.org/10.1080/09505439609526455>

Behr, R. (2023, November 1). Rishi Sunak's vanity jamboree on AI safety lays bare the UK's Brexit dilemmas. *The Guardian*. <https://www.theguardian.com/commentisfree/2023/nov/01/rishi-sunaks-ai-safety-britains-brexit-dilemmas-elon-musk>

Ben Ayed, R., & Hanana, M. (2021). Artificial intelligence to improve the food and agriculture sector. *Journal of Food Quality*, 2021, Article 5584754. <https://doi.org/10.1155/2021/5584754>

Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (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* (pp. 610–623). ACM. <https://doi.org/10.1145/3442188.3445922>

Benkler, Y. (2022). Power and productivity: Institutions, ideology, and technology in political economy. In D. Allen, Y. Benkler, L. Downey, R. Henderson, & J. Simons (Eds.), *A political economy of justice* (pp. 27–60). University of Chicago Press. <https://dx.doi.org/10.2139/ssrn.3503962>

Berka, J., Balata, J., Jonker, C. M., Mikovec, Z., van Riemsdijk, M. B., & Tielman, M. L. (2022). Misalignment in semantic user model elicitation via conversational agents: A case study in navigation support for visually impaired people. *International Journal of Human–Computer Interaction*, 38(18–20), 1909–1925. <https://doi.org/10.1080/10447318.2022.2059925>

Birch, K., Chiappetta, M., & Artyushina, A. (2020). The problem of innovation in technoscientific capitalism: Data rentiership and the policy implications of turning personal digital data into a private asset. *Policy Studies*, 41(5), 468–487. <https://doi.org/10.1080/01442872.2020.1748264>

Birhane, A., Kasirzadeh, A., Leslie, D., & Wachter, S. (2023). Science in the age of large language models. *Nature Reviews Physics*, 5, 277–280. <https://doi.org/10.1038/s42254-023-00581-4>

Bletchley Declaration. (2023). *The Bletchley Declaration by Countries Attending the AI Safety Summit*, 1–2 November 2023. UK Government. <https://www.gov.uk/government/publications/ai-safety-summit-2023-the->



[bletchley-declaration/the-bletchley-declaration-by-countries-attending-the-ai-safety-summit-1-2-november-2023?ref=nural.cc](https://bletchley-declaration/the-bletchley-declaration-by-countries-attending-the-ai-safety-summit-1-2-november-2023?ref=nural.cc)

Bommasani, R., Hudson, D. A., Adeli, E., Altman, R., Arora, S., von Arx, S., Bernstein, M. S., Bohg, J., Bosselut, A., Brunskill, E., Brynjolfsson, E., Buch, S., Card, D., Castellon, R., Chatterji, N. S., Chen, A. S., Creel, K. A., Davis, J., Demszky, D., ... & Liang, P. (2021). *On the opportunities and risks of foundation models*. ArXiv. <https://doi.org/10.48550/arXiv.2108.07258>

Bonnet, P., Joly, A., & Lee, S. H. (2022, May). Pl@ntNet: Harnessing the power of artificial intelligence for biodiversity conservation. *NCT*. <https://ctsoc.ieee.org/images/CTSOC-NCT-2022-05-FA.pdf>

Boyd, W. (2014). Public utility and the low-carbon future. *UCLA Law Review*, 61, 1614-1710.

Bozeman, B. (2007). *Public values and public interest: Counterbalancing economic individualism*. Georgetown University Press.

Busnatu, Ș., Niculescu, A. G., Bolocan, A., Petrescu, G. E., Păduraru, D. N., Năstasă, I., Lupușoru, M., Geantă, M., Andronic, O., Grumezescu, A. M., & Martins, H. (2022). Clinical applications of artificial intelligence—An updated overview. *Journal of Clinical Medicine*, 11(8), Article 2265. <https://doi.org/10.3390%2Fjcm11082265>

Carugati, C. (2023, May 11). *The age of competition in generative artificial intelligence has begun*. Bruegel. <https://www.bruegel.org/first-glance/age-competition-generative-artificial-intelligence-has-begun>

Curato, N., & Steiner, J. (2018). Deliberative democracy and comparative democratization studies. In A. Bächtiger, J. S. Dryzek, J. Mansbridge, & M. Warren (Eds.), *The Oxford handbook of deliberative democracy* (pp. 490–502). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780198747369.013.35>

Dencik, L. (2021). Towards data justice unionism? A labour perspective on AI governance. In P. Verdegem (Ed.), *AI for everyone?* (pp. 267–284). University of Westminster Press. <https://doi.org/10.16997/book55>

Domínguez Hernández, A., Krishna, S., Perini, A. M., Katell, M., Bennett, S. J., Borda, A., Hashem, Y., Hadjiloizou, S., Mahomed, S., Jayadeva, S., Aitken, M., & Leslie, D. (2024). Mapping the individual, social, and biospheric impacts of Foundation Models. In *FAccT '24: Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency* (pp. 776–796). ACM. <https://doi.org/10.1145/3630106.3658939>

Donelan, M. (2023, November 9). *AI Safety Summit 2023: Statement by the Secretary of State for Science, Innovation, and Technology*. House of Commons Hansard. <https://hansard.parliament.uk/Commons/2023-11-09/debates/D73007F1-646E-471F-A47A-248B4204DB49/ArtificialIntelligenceSafetySummit>

Downs, A. (1962). The public interest: Its meaning in a democracy. *Social Research*, 29(1), 1–36.

Drahos, P. (2020). Responsive science. *Annual Review of Law and Social Science*, 16, 327–342.

<https://doi.org/10.1146/annurev-lawsocsci-040220-065454>

El Himer, S., Ouaisa, M., Ouaisa, M., & Boulouard, Z. (2022). Artificial Intelligence of Things (AIoT) for renewable energies systems. In El Himer, S., Ouaisa, M., Emhemed, A. A. A., Ouaisa, M., & Boulouard, Z. (Eds), *Artificial Intelligence of Things for smart green energy management* (pp. 1–13). Springer, Cham.

Fourcade, M. (2021). Ordinal citizenship. *The British Journal of Sociology*, 72(2), 154–173.

<https://doi.org/10.1111/1468-4446.12839>

Frischmann, B. M. (2012). *Infrastructure: The social value of shared resources*. Oxford University Press.

Geburu, T., Bender, E., McMillan-Major, A., & Mitchell, M. (2023). *Statement from the listed authors of Stochastic Parrots on the “AI pause” letter*. DAIR Institute. <https://www.dair-institute.org/blog/letter-statement-March2023/>

Goodlad, L. M. (2023). Editor's introduction: Humanities in the loop. *Critical AI*, 1(1–2).

<https://doi.org/10.1215/2834703X-10734016>

Habermas, J. (1991). *The structural transformation of the public sphere: An inquiry into a category of bourgeois society*. MIT Press. (Original work published 1962).

Habermas, J. (2001). *Moral consciousness and communicative action*. MIT Press. (Original work published 1990).

Habermas, J. (2003). *On the pragmatics of communication*. Polity. (Original work published 1998).

Hanna, A. & Bender, E. (2023, August 12). AI causes real harm. Let's focus on that over the end-of-humanity hype. *Scientific American*. <https://www.scientificamerican.com/article/we-need-to-focus-on-ais-real-harms-not-imaginary-existential-risks/>

Hao, K. (2019, November 11). The computing power needed to train AI is now rising seven times faster than ever before. *MIT Technology Review*. <https://www.technologyreview.com/2019/11/11/132004/the-computing-power-needed-to-train-ai-is-now-rising-seven-times-faster-than-ever-before/>

Hao, K., & Warzel, C. (2023, November 19). Inside the chaos at OpenAI. *The Atlantic*. <https://www.theatlantic.com/technology/archive/2023/11/sam-altman-open-ai-chatgpt-chaos/676050/>

Hamilton, W. H. (1930). Affectation with public interest. *The Yale Law Journal*, 39(8), 1089–1112.

Harris, K. (2023, November 1). *Remarks by Vice President Harris on the future of artificial intelligence*. The White House. <https://www.whitehouse.gov/briefing-room/speeches-remarks/2023/11/01/remarks-by-vice->

[president-harris-on-the-future-of-artificial-intelligence-london-united-kingdom/](#)

Harris, M. (2023). *Palo Alto: A history of California, capitalism, and the world*. Hachette UK.

Hern, A. (2024, February 6). TechScape: Why is the UK so slow to regulate AI? *The Guardian*. <https://www.theguardian.com/technology/2024/feb/06/techscape-uk-artificial-intelligence-government-regulation>

Hazell, J. (2023). *Spear phishing with large language models*. ArXiv. <https://doi.org/10.48550/arXiv.2305.06972>

Huntingford, C., Jeffers, E. S., Bonsall, M. B., Christensen, H. M., Lees, T., & Yang, H. (2019). Machine learning and artificial intelligence to aid climate change research and preparedness. *Environmental Research Letters*, 14(12), Article 124007. <https://doi.org/10.1088/1748-9326/ab4e55>

Jarvis, J. (2023, November 19). Artificial general bullshit. *Medium*. <https://medium.com/whither-news/artificial-general-bullshit-e480939332f6>

Jung, J., Maeda, M., Chang, A., Bhandari, M., Ashapure, A., & Landivar-Bowles, J. (2021). The potential of remote sensing and artificial intelligence as tools to improve the resilience of agriculture production systems. *Current Opinion in Biotechnology*, 70, 15–22. <https://doi.org/10.1016/j.copbio.2020.09.003>

Kaelin, V. C., Valizadeh, M., Salgado, Z., Parde, N., & Khetani, M. A. (2021). Artificial intelligence in rehabilitation targeting the participation of children and youth with disabilities: Scoping review. *Journal of Medical Internet Research*, 23(11), Article e25745. <https://doi.org/10.2196/25745>

Khan, L. M., & Vaheesan, S. (2017). Market power and inequality: The antitrust counterrevolution and its discontents. *Harvard Law & Policy Review*, 11, 235.

Klepper, D. (2023, January 24). It turns out that ChatGPT is really good at creating online propaganda: 'I think what's clear is that in the wrong hands there's going to be a lot of trouble.' *Fortune*. <https://fortune.com/2023/01/24/chatgpt-open-ai-online-propaganda/>

Klinger, J., Mateos-Garcia, J., & Stathoulopoulos, K. (2022). *A narrowing of AI research?* ArXiv. <https://doi.org/10.48550/arXiv.2009.10385>

Kurenkov, A. (2022, June 12). Lessons from the GPT-4chan controversy. *The Gradient*. <https://thegradient.pub/gpt-4chan-lessons/>

Kwok, R. (2019). AI empowers conservation biology. *Nature*, 567(7746), 133–135. <https://doi.org/10.1038/d41586-019-00746-1>

Lazar, S., & Nelson, A. (2023). AI safety on whose terms? *Science*, 381(6654), 138.

<https://doi.org/10.1126/science.adi8982>

Le Ludec, C., Cornet, M., & Casilli, A. A. (2023). The problem with annotation: Human labour and outsourcing between France and Madagascar. *Big Data & Society*, 10(2).

<https://doi.org/10.1177/20539517231188723>

Leslie, D., & Rossi, F. (2023). *ACM Tech Brief: Generative artificial intelligence* (Issue 8). ACM.

<https://doi.org/10.1145/3626110>

Levine, M. (2023, November 27). OpenAI is still an \$86 billion nonprofit. *Bloomberg*.

<https://www.bloomberg.com/opinion/articles/2023-11-27/openai-is-still-an-86-billion-nonprofit?leadSource=verify%20wall>

Lezoche, M., Hernandez, J. E., Alemany Díaz, M. D. M. E., Panetto, H., & Kacprzyk, J. (2020). Agri-food 4.0: A survey of the supply chains and technologies for the future agriculture. *Computers in Industry*, 117, Article 103187. <https://doi.org/10.1016/j.compind.2020.103187>

Linaza, M. T., Posada, J., Bund, J., Eisert, P., Quartulli, M., Döllner, J., Pagani, A., Olaizola, I. G., Barriguinha, A., Moysiadis, T., & Lucat, L. (2021). Data-driven artificial intelligence applications for sustainable precision agriculture. *Agronomy*, 11(6), Article 1227. <https://doi.org/10.3390/agronomy11061227>

Losse, K. (2019, February 7). The false promise of Silicon Valley's quest to save the world. *The New Republic*.

<https://newrepublic.com/article/153034/false-promise-silicon-valleys-quest-save-world>

Mascellino, A. (2023, January 9). ChatGPT used to develop new malicious tools. *Infosecurity Magazine*.

<https://www.infosecurity-magazine.com/news/chatgpt-develop-malicious-tools/>

Mazzucato, M. (2018). Mission-oriented innovation policies: Challenges and opportunities. *Industrial and corporate change*, 27(5), 803–815. <https://doi.org/10.1093/icc/dty034>

Mazzucato, M., Schaake, M., Krier, S., & Entsminger, J. (2022). *Governing artificial intelligence in the public interest* (Working Paper No. 2022-12). UCL Institute for Innovation and Public Purpose.

<https://www.ucl.ac.uk/bartlett/public-purpose/wp2022-12>

Merchant, B. (2023, November 20). OpenAI's board had safety concerns. Big Tech obliterated them in 48 hours. *The Los Angeles Times*. <https://www.latimes.com/business/technology/story/2023-11-20/column-openais-board-had-safety-concerns-big-tech-obliterated-them-in-48-hours>

Miotto, R., Wang, F., Wang, S., Jiang, X., & Dudley, J. T. (2018). Deep learning for healthcare: Review, opportunities and challenges. *Briefings in Bioinformatics*, 19(6), 1236–1246.

<https://doi.org/10.1093/bib/bbx044>

Munz, J., Gindele, N., & Doluschitz, R. (2020). Exploring the characteristics and utilisation of Farm Management Information Systems (FMIS) in Germany. *Computers and Electronics Agriculture*, 170, Article 105246. <https://doi.org/10.1016/j.compag.2020.105246>

Nordgren, A. (2022). Artificial intelligence and climate change: ethical issues. *Journal of Information, Communication and Ethics in Society*, 21(1), 1–15. <https://doi.org/10.1108/JICES-11-2021-0106>

Novak, W. J. (2017). The public utility idea and the origins of modern business regulation. In N. R. Lamoreaux & W. J. Novak (Eds.), *Corporations and American democracy* (pp. 139–176). Harvard University Press. <https://doi.org/10.4159/9780674977686-006>

Pyysiäinen, J., Halpin, D., & Guilfoyle, A. (2017). Neoliberal governance and ‘responsibilization’ of agents: Reassessing the mechanisms of responsibility-shift in neoliberal discursive environments. *Distinktion: Journal of Social Theory*, 18(2), 215–235. <https://doi.org/10.1080/1600910X.2017.1331858>

OpenAI. (2018). *OpenAI charter*. <https://openai.com/charter>

Panch, T., Szolovits, P., & Atun, R. (2018). Artificial intelligence, machine learning and health systems. *Journal of Global Health*, 8(2), Article 020303. <https://doi.org/10.7189%2Fjogh.08.020303>

Perrigo, B. (2021, August 23). An artificial intelligence helped write this play. It may contain racism. *Time*. <https://time.com/6092078/artificial-intelligence-play/>

Perrigo, B. (2023, June 20). Exclusive: OpenAI lobbied the E.U. to water down AI regulation. *Time*. <https://time.com/6288245/openai-eu-lobbying-ai-act/>

Rahman, K. S. (2018). The new utilities: Private power, social infrastructure, and the revival of the public utility concept. *Cardozo Law Review*, 39(5), 1621–1692. <http://cardozolawreview.com/wp-content/uploads/2018/07/RAHMAN.39.5.2.pdf>

Rahman, K. S., & Teachout, Z. (2020). *From private bads to public goods: Adapting public utility regulation for informational infrastructure*. Knight First Amendment Institute at Columbia University. <https://knightcolumbia.org/content/from-private-bads-to-public-goods-adapting-public-utility-regulation-for-informationalinfrastucture>

Rose, N. (1996). Governing “advanced” liberal democracies. In A. Barry, T. Osborne, & N. Rose (Eds.), *Foucault and political reason* (pp. 37–64). Routledge.

Royal Society. (2020, December 2). *Digital technology and the planet: Harnessing computing to achieve net zero UK*. <https://royalsociety.org/news-resources/projects/digital-technology-and-the-planet/>

Ryan-Mosley, T. (2023, June 12). It's time to talk about the real AI risks. *MIT Technology Review*.  
<https://www.technologyreview.com/2023/06/12/1074449/real-ai-risks/>

Schubert, G. (1961). *The public interest: A critique of the theory of a political concept*. Free Press.

Shelby, R., Rismani, S., Henne, K., Moon, A., Rostamzadeh, N., Nicholas, P., Yilla-Akbari, N., Gallegos, J., Smart, A., Garcia, E. & Virk, G. (2023). Sociotechnical harms of algorithmic systems: Scoping a taxonomy for harm reduction. In F. Rossi, S. Das, J. Davis, K. Firth-Butterfield, & A. John (Eds.), *Proceedings of the 2023 AAAI/ACM Conference on AI, Ethics, and Society* (pp. 723–741). ACM.  
<https://doi.org/10.1145/3600211.3604673>

Silvestro, D., Goria, S., Sterner, T., & Antonelli, A. (2022). Improving biodiversity protection through artificial intelligence. *Nature Sustainability*, 5(5), 415–424. <https://doi.org/10.1038/s41893-022-00851-6>

Simons, J., & Ghosh, D. (2020, August 4). *Utilities for democracy: Why and how the algorithmic infrastructure of Facebook and Google must be regulated*. Brookings Institute.  
<https://www.brookings.edu/articles/utilities-for-democracy-why-and-how-the-algorithmic-infrastructure-of-facebook-and-google-must-be-regulated/>

Smakman, J., Davies, M., & Birtwistle, M. (2023). *Mission critical: Lessons from relevant sectors for AI safety*. The Ada Lovelace Institute. <https://www.adalovelaceinstitute.org/policy-briefing/ai-safety/>

Smith, P., & Smith, L. (2021). Artificial intelligence and disability: Too much promise, yet too little substance? *AI and Ethics*, 1(1), 81–86. <https://doi.org/10.1007/s43681-020-00004-5>

Spada, P., & Paulson, L. (2023). Measuring the effect of collective intelligence processes that leverage participation and deliberation. In S. Boucher, C. A. Hallin, & L. Paulson (Eds.), *The Routledge handbook of collective intelligence for democracy and governance* (pp. 78–107). Routledge.  
<https://doi.org/10.4324/9781003215929-6>

Sparkes, M. (2023, November 2). What did the UK's AI Safety Summit actually achieve? *New Scientist*.  
<https://www.newscientist.com/article/2400834-what-did-the-uks-ai-safety-summit-actually-achieve/>

Stephenson, N., Shane, E., Chase, J., Rowland, J., Ries, D., Justice, N., Zhang, J., Chan, L., & Cao, R. (2019). Survey of machine learning techniques in drug discovery. *Current Drug Metabolism*, 20(3), 185–193.  
<https://doi.org/10.2174/1389200219666180820112457>

Sunak, R. (2023, October 26). *The prime minister's speech on AI*. UK Government.  
<https://www.gov.uk/government/speeches/prime-ministers-speech-on-ai-26-october-2023>

Thornhill, J. (2023, November 2). Step aside world, the US wants to write the AI rules. *The Financial Times*.  
<https://www.ft.com/content/e605dc93-545e-4fe8-b9cf-7cd5a88b49ed>

Tran, K. P. (2021). Artificial intelligence for smart manufacturing: Methods and applications. *Sensors*, 21(16),  
Article 5584. <https://doi.org/10.3390/s21165584>

Vallor, S. & Luger, E. (2023). *A shrinking path to AI safety*. Edinburgh Futures Institute. <https://efi.ed.ac.uk/a-shrinking-path-to-safety-how-a-narrowly-technical-approach-to-align-ai-with-the-public-good-could-fail/>

Van Dijck, J., Poell, T., & De Waal, M. (2018). *The platform society: Public values in a connective world*.  
Oxford University Press.

Veale, M., Matus, K., & Gorwa, R. (2023). AI and global governance: Modalities, rationales, tensions. *Annual  
Review of Law and Social Science*, 19, 255–275. <https://doi.org/10.1146/annurev-lawsocsci-020223-040749>

Vicini, S., Bortolotto, C., Rengo, M., Ballerini, D., Bellini, D., Carbone, I., Preda, L., Laghi, A., Coppola, F., &  
Faggioni, L. (2022). A narrative review on current imaging applications of artificial intelligence and radiomics  
in oncology: Focus on the three most common cancers. *La radiologia medica*, 127(8), 819–836.  
<https://doi.org/10.1007/s11547-022-01512-6>

Vinuesa, R., Azizpour, H., Leite, I., Balaam, M., Dignum, V., Domisch, S., Felländer A, Langhans, S. D.,  
Tegmark, M., & Fuso Nerini, F. (2020). The role of artificial intelligence in achieving the Sustainable  
Development Goals. *Nature Communications*, 11(1), Article 233. <https://doi.org/10.1038/s41467-019-14108-y>

Vock, I. (2022, December 9). ChatGPT proves that AI still has a racism problem. *New Statesman*.  
<https://www.newstatesman.com/thestaggers/2022/12/chatgpt-shows-ai-racism-problem>

Wainberg, M., Merico, D., Delong, A., & Frey, B. J. (2018). Deep learning in biomedicine. *Nature  
Biotechnology*, 36(9), 829–838. <https://doi.org/10.1038/nbt.4233>

Wald, M. (2021). AI data-driven personalisation and disability inclusion. *Frontiers in Artificial Intelligence*, 3,  
Article 571955. <https://doi.org/10.3389/frai.2020.571955>

Warnke, P., Bratan, T., & Wunderle, U. (2023). Public engagement in the tradition of participatory approaches—  
an approximation. In V. Block (Ed.), *Putting responsible research and innovation into practice: A multi-  
stakeholder approach* (pp. 123–146). Springer International. [https://doi.org/10.1007/978-3-031-14710-4\\_7](https://doi.org/10.1007/978-3-031-14710-4_7)

Weidinger, L., Mellor, J., Rauh, M., Griffin, C., Uesato, J., Huang, P.-S., Cheng, M., Glaese, M., Balle, B.,  
Kasirzadeh, A., Kenton, Z., Brown, S., Hawkins, W., Stepleton, T., Biles, C., Birhane, A., Haas, J., Rimell, L.,  
Hendricks, L. A., ... Gabriel, I. (2021). *Ethical and social risks of harm from language models*. ArXiv.  
<https://doi.org/10.48550/arXiv.2112.04359>



Whittaker, M. (2021). The steep cost of capture. *Interactions*, 28(6), 50–55. <https://doi.org/10.1145/3488666>

Widder, D. G., West, S., & Whittaker, M. (2023, August 17). *Open (for business): Big tech, concentrated power, and the political economy of Open AI*. SSRN. <https://dx.doi.org/10.2139/ssrn.4543807>

Wu, T. (2017). *The attention merchants: The epic scramble to get inside our heads*. Vintage.

Yigitcanlar, T., Corchado, J. M., Mehmood, R., Li, R. Y. M., Mossberger, K., & Desouza, K. (2021). Responsible urban innovation with local government artificial intelligence (AI): A conceptual framework and research agenda. *Journal of Open Innovation: Technology, Market, and Complexity*, 7(1), Article 71. <https://doi.org/10.3390/joitmc7010071>

Yost, M. A., Kitchen, N. R., Sudduth, K. A., Sadler, E. J., Drummond, S. T., & Volkmann, M. R. (2017). Long-term impact of a precision agriculture system on grain crop production. *Precision Agriculture*, 18(5), 823–842. <https://doi.org/10.1007/s11119-016-9490-5>

Zuboff, S. (2019). *The age of surveillance capitalism: The fight for a human future at the new frontier of power*. Public Affairs.

---

©2024 David Leslie, Carolyn Ashurst, Natalia Menéndez González, Frances Griffiths, Smera Jayadeva, Mackenzie Jorgensen, Michael Katell, Shyam Krishna, Doschmund Kwiatkowski, Carolina Iglésias Martins, Sabeedah Mahomed, Carlos Mougan, Shachi Pandit, Mark Richey, Joseph W. Sakshaug, Shannon Vallor, and Luke Vilain. This article is licensed under a [Creative Commons Attribution \(CC BY 4.0\) International license](https://creativecommons.org/licenses/by/4.0/), except where otherwise indicated with respect to particular material included in the article.

## Footnotes

1. Indeed, as recent commentary in *The Guardian* notes, the UK government has remained reluctant to make any substantive moves, in the domestic context, toward crafting meaningful regulation, which seems to challenge claims that the United Kingdom intends to be a ‘world leader’ in AI governance (Hern, 2024). The government’s February 2024 response to last year’s Artificial Intelligence White Paper consultation makes clear that there is little appetite (and even less funding) for an AI regulatory strategy. While acknowledging strong support by consultation respondents for legislative action, the government has chosen not to take any. Even an attempt by AI developers and the creative industries to agree on a voluntary code for AI and intellectual property has fallen flat and produced nothing. It all gives the appearance that the summit, and the AI developments that prompted it, have done little to shift the government’s thinking away from its ‘light touch’ approach to regulation. [↵](#)
2. It is helpful to note that Sunak was not alone in painting this contrastive picture. In her summit launch speech at the British Embassy in London, U.S. vice president Kamala Harris also stressed the pivot point:



“AI has the potential to do profound good to develop powerful new medicines to treat and even cure the diseases that have for generations plagued humanity, to dramatically improve agricultural production to help address global food insecurity, and to save countless lives in the fight against the climate crisis. But just as AI has the potential to do profound good, it also has the potential to cause profound harm. From AI-enabled cyberattacks at a scale beyond anything we have seen before to AI-formulated bio-weapons that could endanger the lives of millions, these threats are often referred to as the ‘existential threats of AI’ because, of course, they could endanger the very existence of humanity” (K. Harris, 2023). [↵](#)

3. This is especially important insofar as the benefits and the risks may not be evenly spread. The stakes for different groups are different, and, without intervention, the beneficiaries of the current AI innovation ecosystem, who are also those holding the most agenda-setting power (e.g., larger technology corporations), would define the rules and the trajectory into the future. [↵](#)

4. Critics of the rhetoric of openness and democratization claim that the GenAI ecosystem at large hides behind a tenuous veil of ‘openness’ and scientific research beneficent to all of humanity. The current scaling up of AI capacities and capabilities, and the underlying sociotechnical ecosystem, are notably built on ‘open source’ tenets but lead perceptibly to corporate capture, market influence, and commercial benefit (Widder et al., 2023). [↵](#)

5. One might go further here to trace out what we might think of as the five-step big tech public interest washing playbook: (1) Put forward vacuous platitudes about advancing the common benefit through breakthrough innovations that serve the purpose of virtue signaling; (2) Use corresponding public interest and social good vocabularies without specifying the exact meaning of the relevant terms deployed; (3) Stand behind the prima facie legitimacy of scientific authority and techno-industrial success to justify the epistemic soundness and robustness of the public interest platitudes set forth; (4) Avoid any sort of meaningful public participation, engagement, or input in defining public interest and social good vocabularies and in delimiting how these should be put into practice; (5) In contrast to this outward-facing beneficent image, act behind the curtain to secure private sector interests, optimizing profitability, shareholder value, market share, and commercial frictionlessness. [↵](#)

6. A further concern, in this context, is that ‘frontier AI’ technologies could be abused, misused, or weaponized by state sponsors and procurers who act through what appears to the public as independent commercial organizations. [↵](#)

7. Rahman attempts to rekindle the public utility concept in the context of what he and Zephyr Teachout have characterized as “Adapting Public Utility Regulation for Informational Infrastructure” (Rahman & Teachout, 2020). The idea here is to view information platforms like Google, Amazon, and Meta as public utilities that provide essential infrastructure and that “serve their critical functions as the bedrock of a democratic public sphere” (Rahman & Teachout, 2020, p. 6). The central concern driving this idea that

informational platforms are 'new public utilities' draws on earlier calls to see the Internet as a public utility (which concentrated more on issues around net neutrality), but it focuses regulatory attention on the need to redress power imbalances originating from private sector control of critical infrastructure. Simons & Ghosh (2020) make the explicit connection of informational platforms and algorithmic systems, identifying "algorithms as infrastructure" that are directly affected by the public interest in their ecosystem-level power to form and norm publics. As they write: "Asserting public power over internet platforms...requires a clear understanding of what these algorithms are and how they work. How we conceptualize them will influence the internet regulation we develop [...] [We] should think of these algorithms as a kind of infrastructure, one that shapes how citizens consume advertisements, access news and information, and engage with one another at unprecedented speed and on an unprecedented scale. Designing and operating this algorithmic infrastructure involves unavoidably political choices that benefit the interests of some over others and promote some fundamental values while violating others" (Simons & Ghosh, 2020, p. 3). In this Policy Forum article, we seek to further extend these notions of the 'new public utilities' to capture the way advanced AI systems and GenAI technologies are, on the one hand, continuous with such understandings of essential informational infrastructure and, on the other hand, more far-reaching in their characteristic function as 'intellectual infrastructure,' to borrow a term elaborated on by Brett Frischmann. On Frischmann's account, "intellectual infrastructure, such as basic research, ideas, general purpose technologies, and languages, creates benefits for society primarily by facilitating a wide range of downstream productive activities, including information production, innovation, and the development of products and services, as well as education, community building and interaction, democratic participation, socialization, and many other socially valuable activities. The foundational role of intellectual infrastructures in cumulative, dynamic, and complex systems merits attention. Courts and commentators frequently refer to intellectual infrastructure resources as 'building blocks' to capture their role as basic inputs. But while the 'building blocks' metaphor is evocative, it fails to fully reflect the complex relationships among participants in intellectual systems that derive value from intellectual infrastructures as producers, users, consumers, or incidental beneficiaries" (Frischmann, 2012, p. 253). We believe that Frischmann's notion of "intellectual infrastructure" can be effectively adapted to the purposes, functions, and uses of LLMs and advanced GenAI technologies. [↵](#)

8. Examples of the potential of AI technologies to serve the public good abound across domains such as health and social care, social and biospheric sustainability, and the areas of focus covered by the UN Sustainable Development Goals. (See Aguilar et al., 2021; Antonelli et al., 2021; Ben Ayed & Hanana, 2021; Berka et al., 2022; Bonnet et al., 2022; Busnatu et al., 2022; El Himer et al., 2022; Huntingford et al., 2019; Jung et al., 2021; Kaelin et al., 2021; Kwok, 2019; Lezoche et al., 2020; Linaza et al., 2021; Miotto et al., 2018; Munz et al., 2020; Nordgren, 2022; Panch et al., 2018; Royal Society, 2020; Silvestro et al., 2022; Smith & Smith 2021; Stephenson, 2019; Tran, 2021; Vicini et al., 2022; Vinuesa et al., 2020; Wainberg, 2018; Wald, 2021; Yigitcanlar et al., 2021; Yost et al., 2017). [↵](#)

9. This latter call for governance institutions to foster inclusive and equitable access to state-of-the-art research capabilities and resources should also be extended to the creation of a more level playing field for small- and medium-sized businesses to access research and development capabilities and resources. The political economy of AI needs to be reshaped to better facilitate fair competition and access on the innovation and development side, so that smaller businesses seeking to do basic research and deliver innovative solutions with advanced AI technology are better able to deliver benefits to their local economies and communities. If integrated with society-led participatory co-design, this will also strengthen the democratic governance of the technology by bringing innovation processes closer to impacted communities.

└









10. Admittedly, determining exactly what the 'public interest' or the 'common good' amounts to, at any given time, is often a semantically difficult and politically fraught endeavor, but it is at the same time an ongoing effort that is at the very heart of modern democratic forms of life. While some have been skeptical even of the possibility of reaching a basic collective understanding of what the public interest or the common good is (Downs, 1962; Schubert, 1961), others see a shared orientation to it (and to its cooperative distillation and elaboration) as a progressive force of democratic empowerment and as a precondition of robust deliberation itself (Bächtiger et al., 2018; Habermas, 1962/1991, 1990/2001, 1998/2003). On the skeptical side, critics claim that concepts like the 'public interest' and the 'common good' are inherently vague and indeterminate, forever breeding multiple interpretations that are difficult to reconcile and resolve (Bozeman, 2007). However, proponents of the important role that these concepts play in the moral and political life of the community see this ambiguity as a strength that continually spurs the deepening of democratic will-formation and participatory problem-solving. These latter thinkers point to the history of democratic transitions and development as a proving ground that well evidences that the inclusive and reflective exchange of reasons in the ends of generating shared conceptions of the common good "has transformative power in realizing democratic goals" (Curato & Steiner, 2018, p. 491).

└

## References

- └
- Abdalla, M., & Abdalla, M. (2021). The Grey Hoodie Project: Big tobacco, big tech, and the threat on academic integrity. In M. Fourcade & B. Kuipers (Eds.), *Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society* (pp. 287–297). ACM. <https://doi.org/10.1145/3461702.3462563>
- └
- Alba, D. (2022, December 8). OpenAI Chatbot spits out biased musings, despite guardrails. *Bloomberg*. <https://www.bloomberg.com/news/newsletters/2022-12-08/chatgpt-open-ai-s-chatbot-is-spitting-out-biased-sexist-results?leadSource=uverify%20wall>

└

- Altman, S. (2023a, February 24). Planning for AGI and beyond. Open AI.  
<https://openai.com/index/planning-for-agi-and-beyond/>  

- Altman, S. (2023b). *Written testimony of Sam Altman Chief Executive Officer OpenAI before the U.S. Senate Committee on the Judiciary Subcommittee on Privacy, Technology, & the Law*. US Senate.  
<https://www.judiciary.senate.gov/imo/media/doc/2023-05-16%20-%20Bio%20&%20Testimony%20-%20Altman.pdf>  

- Barbrook, R., & Cameron, A. (1996). The Californian ideology. *Science as Culture*, 6(1), 44–72.  
<https://doi.org/10.1080/09505439609526455>  

- Behr, R. (2023, November 1). Rishi Sunak's vanity jamboree on AI safety lays bare the UK's Brexit dilemmas. *The Guardian*. <https://www.theguardian.com/commentisfree/2023/nov/01/rishi-sunaks-ai-safety-britains-brexite-dilemmas-elon-musk>  

- Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (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* (pp. 610–623). ACM. <https://doi.org/10.1145/3442188.3445922>  

- Benkler, Y. (2022). Power and productivity: Institutions, ideology, and technology in political economy. In D. Allen, Y. Benkler, L. Downey, R. Henderson, & J. Simons (Eds.), *A political economy of justice* (pp. 27–60). University of Chicago Press. <https://dx.doi.org/10.2139/ssrn.3503962>  

- Birch, K., Chiappetta, M., & Artyushina, A. (2020). The problem of innovation in technoscientific capitalism: Data rentiership and the policy implications of turning personal digital data into a private asset. *Policy Studies*, 41(5), 468–487. <https://doi.org/10.1080/01442872.2020.1748264>  

- Birhane, A., Kasirzadeh, A., Leslie, D., & Wachter, S. (2023). Science in the age of large language models. *Nature Reviews Physics*, 5, 277–280. <https://doi.org/10.1038/s42254-023-00581-4>  

- Bletchley Declaration. (2023). *The Bletchley Declaration by Countries Attending the AI Safety Summit, 1–2 November 2023*. UK Government. <https://www.gov.uk/government/publications/ai-safety-summit-2023-the-bletchley-declaration>

↑

- Bommasani, R., Hudson, D. A., Adeli, E., Altman, R., Arora, S., von Arx, S., Bernstein, M. S., Bohg, J., Bosselut, A., Brunskill, E., Brynjolfsson, E., Buch, S., Card, D., Castellon, R., Chatterji, N. S., Chen, A. S., Creel, K. A., Davis, J., Demszky, D., ... & Liang, P. (2021). *On the opportunities and risks of foundation models*. ArXiv. <https://doi.org/10.48550/arXiv.2108.07258>

↑

- Boyd, W. (2014). Public utility and the low-carbon future. *UCLA Law Review*, 61, 1614-1710.

↑

- Carugati, C. (2023, May 11). *The age of competition in generative artificial intelligence has begun*. Bruegel. <https://www.bruegel.org/first-glance/age-competition-generative-artificial-intelligence-has-begun>

↑

- Dencik, L. (2021). Towards data justice unionism? A labour perspective on AI governance. In P. Verdegem (Ed.), *AI for everyone?* (pp. 267–284). University of Westminster Press. <https://doi.org/10.16997/book55>

↑

- Domínguez Hernández, A., Krishna, S., Perini, A. M., Katell, M., Bennett, S. J., Borda, A., Hashem, Y., Hadjiloizou, S., Mahomed, S., Jayadeva, S., Aitken, M., & Leslie, D. (2024). Mapping the individual, social, and biospheric impacts of Foundation Models. In *FAccT '24: Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency* (pp. 776–796). ACM. <https://doi.org/10.1145/3630106.3658939>

↑

- Donelan, M. (2023, November 9). *AI Safety Summit 2023: Statement by the Secretary of State for Science, Innovation, and Technology*. House of Commons Hansard. <https://hansard.parliament.uk/Commons/2023-11-09/debates/D73007F1-646E-471F-A47A-248B4204DB49/ArtificialIntelligenceSafetySummit>

↑

- Drahos, P. (2020). Responsive science. *Annual Review of Law and Social Science*, 16, 327–342. <https://doi.org/10.1146/annurev-lawsocsci-040220-065454>

↑

- Fourcade, M. (2021). Ordinal citizenship. *The British Journal of Sociology*, 72(2), 154–173. <https://doi.org/10.1111/1468-4446.12839>

↑

- Gebru, T., Bender, E., McMillan-Major, A., & Mitchell, M. (2023). *Statement from the listed authors of Stochastic Parrots on the “AI pause” letter*. DAIR Institute. <https://www.dair-institute.org/blog/letter-statement-March2023/>

[↩](#)

- Goodlad, L. M. (2023). Editor's introduction: Humanities in the loop. *Critical AI*, 1(1–2).  
<https://doi.org/10.1215/2834703X-10734016>

[↩](#)

- Hamilton, W. H. (1930). Affectation with public interest. *The Yale Law Journal*, 39(8), 1089–1112.

[↩](#)

- Hanna, A. & Bender, E. (2023, August 12). AI causes real harm. Let's focus on that over the end-of-humanity hype. *Scientific American*. <https://www.scientificamerican.com/article/we-need-to-focus-on-ais-real-harms-not-imaginary-existential-risks/>

[↩](#)

- Hao, K. (2019, November 11). The computing power needed to train AI is now rising seven times faster than ever before. *MIT Technology Review*. <https://www.technologyreview.com/2019/11/11/132004/the-computing-power-needed-to-train-ai-is-now-rising-seven-times-faster-than-ever-before/>

[↩](#)

- Hao, K., & Warzel, C. (2023, November 19). Inside the chaos at OpenAI. *The Atlantic*.  
<https://www.theatlantic.com/technology/archive/2023/11/sam-altman-open-ai-chatgpt-chaos/676050/>

[↩](#)

- Harris, M. (2023). *Palo Alto: A history of California, capitalism, and the world*. Hachette UK.

[↩](#)

- Hazell, J. (2023). *Spear phishing with large language models*. ArXiv.  
<https://doi.org/10.48550/arXiv.2305.06972>

[↩](#)

- Jarvis, J. (2023, November 19). Artificial general bullshit. *Medium*. <https://medium.com/whither-news/artificial-general-bullshit-e480939332f6>

[↩](#)

- Khan, L. M., & Vaheesan, S. (2017). Market power and inequality: The antitrust counterrevolution and its discontents. *Harvard Law & Policy Review*, 11, 235.

[↩](#)

- Klepper, D. (2023, January 24). It turns out that ChatGPT is really good at creating online propaganda: 'I think what's clear is that in the wrong hands there's going to be a lot of trouble.' *Fortune*.  
<https://fortune.com/2023/01/24/chatgpt-open-ai-online-propaganda/>

[↑](#)

- Klinger, J., Mateos-Garcia, J., & Stathoulopoulos, K. (2022). *A narrowing of AI research?* ArXiv. <https://doi.org/10.48550/arXiv.2009.10385>

[↑](#)

- Kurenkov, A. (2022, June 12). Lessons from the GPT-4chan controversy. *The Gradient*. <https://thegradient.pub/gpt-4chan-lessons/>

[↑](#)

- Lazar, S., & Nelson, A. (2023). AI safety on whose terms? *Science*, 381(6654), 138. <https://doi.org/10.1126/science.adi8982>

[↑](#)

- Le Ludec, C., Cornet, M., & Casilli, A. A. (2023). The problem with annotation: Human labour and outsourcing between France and Madagascar. *Big Data & Society*, 10(2). <https://doi.org/10.1177/20539517231188723>

[↑](#)

- Leslie, D., & Rossi, F. (2023). *ACM Tech Brief: Generative artificial intelligence* (Issue 8). ACM. <https://doi.org/10.1145/3626110>

[↑](#)

- Levine, M. (2023, November 27). OpenAI is still an \$86 billion nonprofit. *Bloomberg*. <https://www.bloomberg.com/opinion/articles/2023-11-27/openai-is-still-an-86-billion-nonprofit?leadSource=uverify%20wall>

[↑](#)

- Losse, K. (2019, February 7). The false promise of Silicon Valley's quest to save the world. *The New Republic*. <https://newrepublic.com/article/153034/false-promise-silicon-valleys-quest-save-world>

[↑](#)

- Mascellino, A. (2023, January 9). ChatGPT used to develop new malicious tools. *Infosecurity Magazine*. <https://www.infosecurity-magazine.com/news/chatgpt-develop-malicious-tools/>

[↑](#)

- Mazzucato, M. (2018). Mission-oriented innovation policies: Challenges and opportunities. *Industrial and corporate change*, 27(5), 803–815. <https://doi.org/10.1093/icc/dty034>

[↑](#)

- Mazzucato, M., Schaake, M., Krier, S., & Entsminger, J. (2022). *Governing artificial intelligence in the public interest* (Working Paper No. 2022-12). UCL Institute for Innovation and Public Purpose.

↵

- Merchant, B. (2023, November 20). OpenAI’s board had safety concerns. Big Tech obliterated them in 48 hours. *The Los Angeles Times*. <https://www.latimes.com/business/technology/story/2023-11-20/column-openais-board-had-safety-concerns-big-tech-obliterated-them-in-48-hours>

↵

- Novak, W. J. (2017). The public utility idea and the origins of modern business regulation. In N. R. Lamoreaux & W. J. Novak (Eds.), *Corporations and American democracy* (pp. 139–176). Harvard University Press. <https://doi.org/10.4159/9780674977686-006>

↵

- Perrigo, B. (2021, August 23). An artificial intelligence helped write this play. It may contain racism. *Time*. <https://time.com/6092078/artificial-intelligence-play/>

↵

- Perrigo, B. (2023, June 20). Exclusive: OpenAI lobbied the E.U. to water down AI regulation. *Time*. <https://time.com/6288245/openai-eu-lobbying-ai-act/>

↵

- Pyysiäinen, J., Halpin, D., & Guilfoyle, A. (2017). Neoliberal governance and ‘responsibilization’ of agents: Reassessing the mechanisms of responsibility-shift in neoliberal discursive environments. *Distinktion: Journal of Social Theory*, 18(2), 215–235. <https://doi.org/10.1080/1600910X.2017.1331858>

↵

- Rahman, K. S. (2018). The new utilities: Private power, social infrastructure, and the revival of the public utility concept. *Cardozo Law Review*, 39(5), 1621–1692. <http://cardozolawreview.com/wp-content/uploads/2018/07/RAHMAN.39.5.2.pdf>

↵

- Rose, N. (1996). Governing “advanced” liberal democracies. In A. Barry, T. Osborne, & N. Rose (Eds.), *Foucault and political reason* (pp. 37–64). Routledge.

↵

- Ryan-Mosley, T. (2023, June 12). It’s time to talk about the real AI risks. *MIT Technology Review*. <https://www.technologyreview.com/2023/06/12/1074449/real-ai-risks/>

↵

- Shelby, R., Rismani, S., Henne, K., Moon, A., Rostamzadeh, N., Nicholas, P., Yilla-Akbari, N., Gallegos, J., Smart, A., Garcia, E. & Virk, G. (2023). Sociotechnical harms of algorithmic systems: Scoping a taxonomy for harm reduction. In F. Rossi, S. Das, J. Davis, K. Firth-Butterfield, & A. John (Eds.), *Proceedings of the*



[↩](#)

- Smakman, J., Davies, M., & Birtwistle, M. (2023). *Mission critical: Lessons from relevant sectors for AI safety*. The Ada Lovelace Institute. <https://www.adalovelaceinstitute.org/policy-briefing/ai-safety/>

[↩](#)

- Spada, P., & Paulson, L. (2023). Measuring the effect of collective intelligence processes that leverage participation and deliberation. In S. Boucher, C. A. Hallin, & L. Paulson (Eds.), *The Routledge handbook of collective intelligence for democracy and governance* (pp. 78–107). Routledge. <https://doi.org/10.4324/9781003215929-6>

[↩](#)

- Sparkes, M. (2023, November 2). What did the UK's AI Safety Summit actually achieve? *New Scientist*. <https://www.newscientist.com/article/2400834-what-did-the-uks-ai-safety-summit-actually-achieve/>

[↩](#)

- Sunak, R. (2023, October 26). *The prime minister's speech on AI*. UK Government. <https://www.gov.uk/government/speeches/prime-ministers-speech-on-ai-26-october-2023>

[↩](#)

- Thornhill, J. (2023, November 2). Step aside world, the US wants to write the AI rules. *The Financial Times*. <https://www.ft.com/content/e605dc93-545e-4fe8-b9cf-7cd5a88b49ed>

[↩](#)

- Vallor, S. & Luger, E. (2023). *A shrinking path to AI safety*. Edinburgh Futures Institute. <https://efi.ed.ac.uk/a-shrinking-path-to-safety-how-a-narrowly-technical-approach-to-align-ai-with-the-public-good-could-fail/>

[↩](#)

- Van Dijck, J., Poell, T., & De Waal, M. (2018). *The platform society: Public values in a connective world*. Oxford University Press.

[↩](#)

- Veale, M., Matus, K., & Gorwa, R. (2023). AI and global governance: Modalities, rationales, tensions. *Annual Review of Law and Social Science*, 19, 255–275. <https://doi.org/10.1146/annurev-lawsocsci-020223-040749>

[↩](#)

- Vock, I. (2022, December 9). ChatGPT proves that AI still has a racism problem. *New Statesman*. <https://www.newstatesman.com/thestaggers/2022/12/chatgpt-shows-ai-racism-problem>

[↩](#)

- Warnke, P., Bratan, T., & Wunderle, U. (2023). Public engagement in the tradition of participatory approaches—an approximation. In V. Block (Ed.), *Putting responsible research and innovation into practice: A multi-stakeholder approach* (pp. 123–146). Springer International. [https://doi.org/10.1007/978-3-031-14710-4\\_7](https://doi.org/10.1007/978-3-031-14710-4_7)

↵

- Weidinger, L., Mellor, J., Rauh, M., Griffin, C., Uesato, J., Huang, P.-S., Cheng, M., Glaese, M., Balle, B., Kasirzadeh, A., Kenton, Z., Brown, S., Hawkins, W., Stepleton, T., Biles, C., Birhane, A., Haas, J., Rimell, L., Hendricks, L. A., ... Gabriel, I. (2021). *Ethical and social risks of harm from language models*. ArXiv. <https://doi.org/10.48550/arXiv.2112.04359>

↵

- Whittaker, M. (2021). The steep cost of capture. *Interactions*, 28(6), 50–55. <https://doi.org/10.1145/3488666>

↵

- Widder, D. G., West, S., & Whittaker, M. (2023, August 17). *Open (for business): Big tech, concentrated power, and the political economy of Open AI*. SSRN. <https://dx.doi.org/10.2139/ssrn.4543807>

↵

- Zuboff, S. (2019). *The age of surveillance capitalism: The fight for a human future at the new frontier of power*. Public Affairs.

↵