OPENNESS HYPE AND OPEN WASHING:

A CRITICAL ANALYSIS OF OPENNESS DISCOURSES IN GENERATIVE AI

By

ELISABETH GREVE, B.A. Hons

Supervisor: Dr. Andrea Zeffiro

A Major Research Paper

Submitted to the Department of Communication Studies and Media Arts

in Partial Fulfillment of the Requirements

for the Degree

Master of Arts

in Communication and New Media

McMaster University

© Copyright by Elisabeth Greve, August 2024

Table of Contents

Introduction
Literature Review
Contextualizing Transparency and Openness
Openness in Technology10
From Ethics Washing to Open Washing12
Openness as Hype16
Theoretical Framework
Methods
Analysis
<i>Case 1: Meta</i> 24
Case 2: IBM
Case 3: OpenAI
Discussion
Conclusion
References
Appendix

Introduction

In December 2023, Meta and IBM launched the AI Alliance in collaboration with over fifty other founding members and collaborators, including startups, enterprises, science organizations, non-profits, and select universities. The group proclaims itself a "community of technology creators, developers and adopters collaborating to advance safe, responsible AI rooted in open innovation" (The AI Alliance, n.d.). Simultaneously, many key industry players in artificial intelligence (AI) development have made broad attempts to market their AI endeavours as "transparent" and "open." As MIT Technology Review reports, "[0]pen source' is the latest buzzword in AI circles" (Gent, 2024). This emphasis on openness is also reflected in the words of Meta's chief AI scientist, Yann LeCun, who claims that "[o]penness is the only way to make AI platforms reflect the entirety of human knowledge and culture" (Associated Press, 2023). Even OpenAI, a technology company that builds closed AI systems, including the much-lauded ChatGPT project, places the term directly in its name. Closed AI systems can broadly be understood as those which do not share details about the structure of their models, including the training processes, training data, code, and other proprietary information. In contrast to this, openness has emerged as a salient idea which has become popular in the discourses surrounding AI, particularly in discussions of ethics, safety and responsibility. However, this idea of openness, along with the many ways in which it is being implemented and mobilized, is certainly worth closer examination.

Openness is a term that seems to defy a clear definition; however, it broadly can be understood as a lack of restriction and/or secrecy and thus describes the quality of being accessible and/or honest (Merriam-Webster, n.d.; Oxford University Press, n.d.). Beyond these definitions, the word has adopted various meanings and significances throughout its long history (Schlagwein et al., 2017; Jones, 2015). In many contexts, openness and the closely related idea of transparency, operate as strong metaphors which have come to represent cultural values and inform ethics discourses. For example, transparency, which describes the quality of being see-through, has been popularized to represent an institutional value and a public good (Adams, 2020). Similarly, the adoption of openness in AI discourse has seemingly resulted in the conflation of being 'open' with being ethical and responsible. In this project, I investigate openness as it appears in AI discourse, particularly in the public-facing documents of major technology companies that are developing generative AI.

On November 30, 2022, technology company OpenAI released their Large Language Model (LLM)-based chatbot ChatGPT and instantly it was hailed as world-changing (Leaver & Srdarov, 2023). While ChatGPT is perhaps the most widely acclaimed, it exists in a larger pool of generative AI projects which work to generate seemingly novel outputs, such as text and image, based on user prompts. As these tools have continued to be released to the public, they have inspired equal parts fear and excitement, whether over the potentially transformative effects or the threats posed by AI to education, creative industries, the labour market and more. Meanwhile, the considerable hype advanced by major industry players has contributed to a blurry distinction between what is real and what is imaginary. As Meredith Broussard (2018) asserts, artificial general intelligence (AGI) and super-intelligence remains a not-yet-actualized fantasy, whereas the AI that does exist, "narrow AI," is purely mathematical. Narrow AI, including generative AI, works by analyzing a dataset, identifying patterns and probabilities based on this data, and then "codifying these pattens and probabilities into a computational construct called a *model*" (Broussard, 2018, p. 32). Subsequently, as Leaver & Srdarov (2023) suggest, "[e]scaping the hype and hypocrisy deployed by AI companies is vital for repositioning generative AI not as magical, not as a saviour, and not as a destroyer, but rather as a new technology that needs to be critically and ethically understood" (para. 12).

Studying the discourses surrounding AI ethics and safety is immensely important, particularly in light of ongoing work that reveals AI's vast potential for harm. Indeed, a significant array of scholarship has continued to demonstrate how AI holds the capacity to reinforce historical injustices and biases (Broussard, 2023; Noble, 2018), including but not limited to the perpetuation of racial inequalities (Buolamwini, 2023; Benjamin, 2019) and gendered inequalities (Criado-Perez, 2019; Adam, 1998). Importantly, these biases are not simply the result of computational systems being fed biased data to produce biased results, a process often referred to as "garbage in, garbage out." Instead, scholars have continued to demonstrate how biases exist structurally (Broussard, 2023) and how these biases are made durable, invisible and inaccessible to human audit within computational systems (Mullaney et al., 2021).

Additionally, AI holds the potential to significantly amplify labour precarity (Munn, 2023) while also maintaining labour relations entrenched in patriarchy and racial capitalism (Atanasoski & Vora, 2019). This can be witnessed in the dominant technological imaginaries which envision a future where AI technologies take on the dull, repetitive and 'dirty' work which is associated with racialized, gendered and colonized labour populations (Atanasoski & Vora, 2019). Troublingly, this reinforces notions of devalued work, but it also serves to disguise the human labour that remains necessary to prop up AI systems. Training and maintaining AI

systems requires a significant amount of labour which often takes the shape of "precarious and exploitative piecework carried out by a digital underclass" (Munn, 2022, p. 6). This ofteninvisible workforce contends with limited job security, demanding conditions, and intrusive surveillance (Kakalic & Hancock, 2023). One of the ways in which this labour demand is being supported is through outsourcing labour to the Global South in ways that allow troubling colonial legacies to persist (ElGhadban, 2023). For example, the roughly 200 workers in Nairobi, Kenya who perform the critical work of content moderation for Sama, an AI outsourcing company based in California that has data-labelling contracts with Google, Meta and Microsoft, are paid as little as \$1.50 an hour (ElGhadban, 2023).

Ultimately, as many scholars identify, AI has shown great potential for harm and can significantly contribute to the "deepening of systemic inequalities globally" (Ricaurte, 2022a, para. 1). In response to growing concerns, major technology corporations have begun to develop ethical frameworks that promise to reduce the potential harm of these technologies (Ricaurte, 2022a). However, in many cases, these frameworks are merely part of broader 'ethics washing' strategies. Because of AI's potential consequences and harms, ethics frameworks are necessary. Yet it remains difficult to discern the extent to which these frameworks impact AI development in the tech sector. Do ethics discourses allow companies to deflect scrutiny and maintain dominant positions in the industry? How do we gauge the sincerity of these approaches and parse when these attempts are shallow, performative, inadequate or insufficient?

Conversations surrounding the implications of AI technologies and discussions surrounding appropriate safeguards have been abundant over the past few years. Amidst these conversations, openness has seemingly emerged as a highly favoured talking point. From using 'open' as a buzzword to signal an array of unspecified yet inferred values, to promoting open source models, advocating for open science and open innovation approaches and more, the idea of openness has become conspicuous. While openness, in its many various interpretations and implementations, certainly offers an array of benefits, the leveraging of this rhetoric and the heralding of openness as an ultimate value and ideal must be examined.

In this project, I investigate how and in what ways openness has taken root in and shaped the conversations surrounding AI ethics, safety and responsibility. Furthermore, this work considers the popularity of openness in AI discourses in conjunction with the subsequent praise of open source and open innovation approaches. To this end, I examine how discourses of openness steer the development of AI products while simultaneously projecting particular visions of desirable futures.

The central questions guiding this study are as follows: 1) How has the technology industry stabilized a particular vision of 'open' AI over the past two years? 2) In what ways have industry players advanced the idea of openness alongside efforts to position AI products as safe and responsible? 3) How is openness positioned and articulated? What does this rhetoric overwrite, neglect, and/or leave out? Broadly, this project investigates the promotion of openness, paying particular attention to how the idea is deployed, while considering who it benefits and what it obscures.

In the paper to follow, I examine openness amongst the public-facing documents of three major technology companies that are producing generative AI products, namely Meta, IBM, and OpenAI. Through critical discourse analysis, along with the theoretical framework of sociotechnical imaginaries and carceral imaginaries, I examine the openness discourse as it is

deployed by major industry players. After outlining the relevant literature and articulating the theory and methods that structure the project, I present the analysis case-by-case. I then bring these cases together and discuss the insights gained across the three cases as a whole. Ultimately, the project endeavours to advance an understanding of how openness discourses are wielded by major players in the generative AI industry, while working to evaluate the extent to which these discourses can be understood as open washing and even hype.

Literature Review

The array of existing scholarship this project draws upon and works to be in conversation with is set out thematically in the literature review to follow. I begin by tracing the history and exploring the contexts of openness and transparency paradigms to consider how openness has been taken up in the technology sector. Situating the concept further in the work of critical technology and data scholars, I link this scholarship to academic work that critically evaluates ethics discourses and ethics washing. Finally, the literature review concludes with scholarship that theorizes and offers a critical assessment of 'hype' narratives operationalized by big tech.

Contextualizing Transparency and Openness

While I primarily focus on openness in the study to follow, transparency, a closely interrelated value and ideal often used as a synonym for openness (Meijer, 2009), is also important to set out and understand. Both transparency and openness exist as strong metaphors in AI discourses, lending the evocation of their literal meanings to represent cultural values and inform ethics frameworks. Metaphors, a common instrument of human language and cognition, are particularly prominent in technological discourse due to their ability to aid understanding of the novel and abstract (Puschmann & Burgess, 2014). Explaining new technology is often a difficult endeavour, particularly due to the technical and abstract nature of emerging technologies, and the instability or uncertainty accompanying them (Wyatt, 2004; Puschmann & Burgess, 2014; Chown & Nascimento, 2023). In light of this, metaphors have become a popular instrument in dominant discourses surrounding digital technology (Wyatt, 2021).

The metaphors of transparency and openness are prominent not only in the realm of technology ethics but have long shaped ethics discourses broadly. Indeed, the term 'transparency,' has long been displaced from its literal meaning to become "a metaphor for political openness" and has subsequently been effectively entrenched in modern discourse (Adams, 2020, p. 17). Similarly, openness borrows from its literal definition to reflect larger ideas of access, participation, and more (Schlagwein et al., 2017). While both ideals have gained immense popularity in the present, they also have long traditions in ideas such as freedom of information (Adams, 2020) and 'open government' (Meijer, 2009). Indeed, the rise of transparency and openness has been cataloged and traced by various scholars, some of which I will recount briefly to offer some useful context.

Various scholars examining transparency, including Meijer (2013) and Adams (2020), establish the concept as arising from the Enlightenment ideals of the 18th century. As Schlagwein et al. (2017) document, the closed paradigm of science and academia was challenged and overcome in the Enlightenment. As a result, 'open science,' where scientists and experts published their methods and findings for others, would become the default (Schlagwein et al., 2017). Beyond this, Adams (2020) describes that the importance of transparency would be established in the aftermath of World War II as the United Nations (UN) sought to assert the importance of freedom of information. The UN General Assembly Resolution states, "[f]reedom of information is a fundamental human right and is the touchstone of all the freedoms to which the United Nations is consecrated" (UN General Assembly Resolution, 1947, as cited in Adams, 2020). Subsequently, freedom of information, and transparency more broadly, would become enshrined as a moral obligation and as Adams (2020) argues, would be central to the new world order that international human rights sought to bring about. Interestingly, as Adams writes, the freedom-seeking narratives of both the Enlightenment and human rights claim a "universalism and morality that transcend both law and politics" (2020, p. 21). Therefore, transparency may effectively render itself outside of history and ideology, thus masking its political rationalities (Adams, 2020, p. 21). However, it is important to recognize that the Enlightenment itself was a distinctly Western historical moment which centred the West and was premised upon a Western epistemology. It is this Western epistemology that Adams (2020) argues transparency arises directly from. Moreover, transparency's alignment with the ideas of freedom and rights has enabled the concept itself to become naturalized and subsequently, it is often interpreted as a neutral and universal idea (Adams, 2020).

Schudson (2015) also historically locates the the rise of transparency rhetoric and a cultural 'right to know,' finding that "openness was a key element in the transformation of politics, society, and culture from the late 1950s through the 1970s" (p. 5). This shift towards "expanded disclosure practices," which Schudson (2015) identifies, would allow transparency to be increasingly institutionalized and widely accepted to the extent that "both the experience of being human and the experience of constructing a democracy have changed in response to a new

transparency imperative and a new embrace of a right to know" (p. 15-16). Transparency itself would effectively come to be identified with self-government and good citizenship while openness would be heralded as a public and private virtue (Schudson, 2015, p. 3). Furthermore, Schudson (2015) echoes Adams's (2020) characterization of transparency's insistence on its own disconnect from politics, noting the way that both transparency and openness are often supported across the political spectrum, linking various political parties and nongovernmental organizations under their broad banners of support (Schudson, 2015).

In order to further develop an understanding of openness, I will turn to a range of work which investigates how the idea is defined and implemented. As Jones (2015) writes, both 'open' and 'openness' have taken on an array of meanings while also carrying "powerful associations with broad social discourses" (p. 331). In fact, as Jones (2015) notes, openness resists simple definition because "it is a complex assemblage of social, political and technological elements developed over time" (p. 335). Accordingly, Schlagwein et al. (2017) set out the various ways in which 'open' and 'openness' are described across openness literature. First, open terminology can be seen referring to higher-order concepts such as a "philosophy of openness." Beyond this, open terminology can refer to the nature of resources, particularly information resources, such as open data or open-source code. In these cases, openness means accessibility and transparency (Schlagwein et al., 2017). In other situations, open terminology may refer to the nature of a process, such as crowdsourcing or open innovation. In these instances, openness means "wide participation in a process" (Schlagwein et al., 2017, p. 299). Lastly, open terminology can refer to the intended effects on specific domains such as "open government," "open education," or "open science." In these instances, openness implies an 'opening' of the domain or area oriented

towards "democracy, equality and liberalism" (Schlagwein et al., 2017, p. 299). Furthermore, even when occurring within a specific domain, openness often carries several meanings. Looking at academic literature on government, Meijer et al. (2012) discover two dimensions of openness which are often considered separately, namely transparency and participation. These interpretations of openness offer insight into the term's vast array of meanings and applications. Building on this, I will also set out how scholars have identified openness in relation to technology.

Openness in Technology

Openness manifests in a myriad of ways in relation to technology. Schlagwein et al. (2017) assert that many of the general principles of openness, including transparency, access, participation and democracy, are made more efficient and are fundamentally shaped by information technologies. Notably, one of the principal ways in which openness was established in the digital technology realm is through the open source movement. The term "open source" was first adopted after a discussion at the O'Reilly Freeware Summit in 1998, which would later be known as the Open Source Summit (Schlagwein et al., 2017, p. 301). Truscello (2005) examines the discourses of Free and Open Source Software (FOSS) movements, noting how they have become "central social developments in the computing world and beyond" (p. 346). As Truscello (2005) describes, advocates of FOSS promote both the "accessibility and distribution of software source code for operating systems and software applications" (p. 347). Truscello (2005) notes that Free Software activists "equate the transparency of source code with an egalitarian society" whereas Open Source advocates "see software licenses as a pragmatic

engineering choice, not an ethical one" (p. 347). Another movement which has developed in tandem with the open source movement is that of the open data movement (Kitchin, 2014). This movement is built on the principles of "openness, participation and collaboration" and advocates that through "sharing and working together the value of data for society can be realized" (Kitchin, 2014, p. 2).

Another key way in which openness is taken up in regard to technology is through 'open innovation.' Chesbrough (2003) defines open innovation as a "paradigm that assumes that firms can and should use external ideas as well as internal ideas, and internal and external paths to market, as firms look to advance their technology" (p. XXIV). Interestingly, Dahlander & Gann (2010) examine open innovation literature to study the idea and definition of openness. After conducting a content analysis on the relevant literature published in Thomson's ISI Web of Knowledge, Dahlander & Gann (2010) discovered significantly varying interpretations of the "openness construct." Moreover, they found that it was often unclear "exactly what type of openness" the authors were referring to (Dahlander & Gann, 2010, p. 702).

The ways in which openness manifests in AI discourses are yet to be extensively studied. Some research on this topic exists; however, it is not comprehensive. Some notable work includes a study by Robinson (2020) that examines Nordic national AI policy documents to ascertain the influence of cultural values of trust, transparency and openness. Additionally, research by Liesenfeld & Dingemanse (2024) and Liesenfeld et al. (2023) examines open washing and evaluates the openness of numerous generative AI projects. While these studies examine openness in the context of AI, they do not specifically examine how technology companies leverage this discourse. Also, this existing research emerges out of the Netherlands and Nordic countries. The lack of research on the ways in which the generative AI industry coopts openness, particularly within the context of North America, thus appears to be a notable gap in the literature to which this project seeks to contribute.

From Ethics Washing to Open Washing

At the beginning of the 2010s, digital technology production was characterized by the "brash spirit of Facebook's motto 'move fast and break things' and the superficial assurances of Google's motto 'do not be evil'"(Green, 2021, p. 209). However, as Green (2021) notes, by the end of the decade, there was a considerable "crisis of conscience" following multiple technology controversies, such as disinformation, algorithmic bias, and the Cambridge Analytica scandal. In response, academics, technology companies and governments embraced 'tech ethics,' which often took shape in statements of ethical principles and oversight bodies (Green, 2021). This practice has remained prominent in recent years across the technology industry as tech ethics and its subsequent ethics codes and frameworks have continued to garner popularity and endorsement (Bietti, 2021). It is within this longer trajectory of the rise of tech ethics over the past decade or so that openness and open washing must be understood in relation to.

In light of rising concerns, Hoffmann (2021) argues that those who are invested in a datadriven technological future have required alternative discourses, "specifically discourses that acknowledge the harms perpetrated by and through data science and technology without also conceding the social, political, or economic power they afford" (p. 3544). This often takes shape in new and emergent commitments to technology ethics such as 'data ethics,' 'AI ethics,' and 'social responsibility' (Hoffmann, 2021, p. 3544). In fact, Stark & Hoffmann (2019) note that there has been an increased focus on developing codes of ethics within the field of data science and AI. While codes of ethics, as well as broader commitments to ethical practices, are longstanding mechanisms through which professionals have sought to organize and define themselves and their priorities (Stark & Hoffmann, 2019), the extent to which they are being proliferated is critiqued by Bietti (2021) who argues that "AI ethics is becoming a business" (p. 274).

Subsequently, a practice which can be understood as 'ethics washing' has become increasingly prevalent in the efforts of technology companies and the industry at large (Bietti, 2021). Ethics washing generally refers to the shallow appearance of ethical behaviour often weaponized by technologists as justification for deregulation, self-regulation or market-driven governance (Bietti, 2021, p. 267). Indeed, as Bietti (2021) argues, focusing on AI 'ethics' instead of AI 'politics' works to depoliticize and normalize the impacts of these technologies while simultaneously limiting the opportunity for real debate or change. Concerningly, when companies 'ethics wash' they are able to perform ethics rather than making the necessary structural changes which would truly achieve particular social values (Bietti, 2021). While efforts to adopt ethical frameworks, guidelines, or values may appear proactive and be interpreted positively, Bietti (2021) warns against how these practices are "driven by market incentives and techno-centric perspectives" as well as "motivated primarily by the need to avoid financial and other company risk" (p. 275). Accordingly, corporate ethics discourses, particularly those which emerge from countries with immense control over AI governance, can be interpreted as what Ricaurte (2022b) terms "Western-ethical-white-corporate-washing" (p. 729).

A range of scholarship has investigated and critiqued tech ethics discourses and the practice of ethics washing. In examining antidiscrimination discourse, Hoffmann (2019) critiques the overemphasis on 'bad actors,' the tendency towards single-axis thinking, which neglects intersectional perspectives, and a limited understanding and focus on modes of disadvantage. Through this critique, Hoffmann (2019) demonstrates how these antidiscrimination discourses are unable to combat algorithmic discrimination because they focus on producing "reactionary technical solutions" that fail to address the "underlying logic that produce unjust hierarchies" in the first place (p. 911). In another study, Hoffmann (2021) investigates discourses of inclusion in data ethics. Inclusion discourses "gesture toward social transformation as a way to diffuse resistance to structural conditions of domination and subordination" (Hoffmann, 2021, p. 3540). These discourses can neutralize critical calls "to not collect certain kinds of data or build and deploy certain technologies" by reframing the issue itself as solvable by "doing things more inclusively" (Hoffmann, 2021, p. 3548). It also works to frame data ethics problems as best solved by technical intervention from those with expertise in areas like machine learning or AI (Hoffmann, 2021). Thus, Hoffmann (2021) argues that these inclusion discourses reify the exclusive nature of technical intervention and reproduce the legitimating power of dominant actors. Those who produce and build data technologies can also benefit economically and politically as "splashy appeals to inclusion and 'doing better' can help stave off external scrutiny" (Hoffmann, 2021, p. 3549). As Hoffman's work demonstrates, a range of particular ethics discourses can adopt the performance of ethics while failing to deliver adequate solutions or action.

The ethics discourses which surround AI have also been studied by scholars. One dominant approach to studying ethics discourses is through identifying themes and patterns. For example, Jobin et al. (2019) mapped and analyzed the global landscape of principles and guidelines for ethical AI, finding a convergence emerging around the principles of transparency, justice and fairness, non-maleficence, responsibility, and privacy. However, Jobin et al. (2019) note that there is "substantive divergence in relation to how these principles are interpreted; why they are deemed important; what issue, domain or actors they pertain to; and how they should be implemented" (p. 1). Similarly, Fjeld et al. (2020) analyzed 30 prominent AI principles documents, discovering an array of thematic trends. These themes include privacy, accountability, safety and security, transparency and explainability, fairness and non-discrimination, human control of technology, professional responsibility and promotion of human values (Fjeld et al., 2020).

Scholarship has also sought to criticize the trajectories of AI ethics discourses. Korn (2021) is critical of how 'ethics' is applied broadly across many technologies and simultaneously fails to explicitly address how technologies have contributed to the violence and marginalization of racialized communities. Korn (2021) advocates for creating "critical tech ethics" which would prioritize connecting race to technology ethics. Moreover, as Ricaurte (2022b) writes, ethics frameworks must be investigated as a strategic discursive mechanism that primarily serves the interests of countries and corporations developing AI. The discourse of 'ethical AI,' Ricaurte (2022b) argues, is strategically aligned so that AI developers may avoid restrictions on their technologies while obfuscating the "historical imperial, colonial, neoliberal and patriarchal infrastructures" through which their AI products are developed (p. 729).

Ultimately, the practices of ethics washing in AI, as well as across the tech industry, are documented and analyzed extensively by scholars. This project builds on this work while engaging particularly with the idea of open washing. Open washing is set out by Weller (2014) who uses the term to describe the ways in which market capital can be gained by "proclaiming open credentials" (p. 155). Interestingly, Weller (2014) notes that open washing is facilitated by the ambiguity around the term, allowing its misuse for commercial reasons. In this project, I adopt this understanding of open washing, which describes the practice of brandishing openness in ways that are performative and even mercenary.

Openness as Hype

This project seeks to approach openness through an understanding of hype. Therefore, I will briefly set out some scholarship that has studied technological discourses through the prominent belief and hype it often engenders. Scholars have offered many various terms to describe the hype that surrounds technologies, including, but not limited to, technological benevolence, technochauvinism, and even "honeymoon objectivity." These terms help to describe how overblown and misplaced belief and faith, often upheld and communicated discursively, work to shape the dominant perception of technological development.

As Ruha Benjamin (2019) writes, technological benevolence often takes shape in the buzzwords and "do-good sound bites" which work to characterize the "*revolutionary* culture of innovation" (p. 102). This is because, technological benevolence promises that technology alone can offer solutions to contemporary problems (Benjamin, 2019). However, this idea of "revolution" or "safety," "efficiency," and "diversity" are often merely marketing catchphrases

which are used to entice investors and social justice advocates alike (Benjamin, 2019). Similarly, Meredith Broussard (2018) describes technochauvinism as the belief that technology is "always the solution" (p. 8). It is misplaced technochauvinist optimism that, as Broussard (2018) argues, allows marketers to promise that digitization will be revolutionary while these promises rarely come to pass or produce worse digital realities. This idea is also documented by Sun-Ha Hong (2020) who sets out the idea of "honeymoon objectivity" to describe the recurring hope that contemporary technological innovations will deliver certainty and objectivity. This hope is sustained largely through discourse, particularly in the discussion of unfulfilled prototypes and "never-quite-actualized promises" which, in turn, operate as a "panoply of fantasies" (Hong, 2020, p. 15). These reductive and ever-recycled fantasies are found to be significantly lacking in Hong's (2020) view as they promise and envision machine-optimized futures that are fuelled by naïve liberal individualism and technocracy rather than new visions of distributed resources or collectives. These beliefs are created and sustained through particular discourses, and as Mullaney et al. (2021) argue, it is the dominant narratives of technological neutrality that, in turn, have lulled the public into complacency.

Interestingly, scholarship has also investigated how this takes shape in the hype which surfaces in relation to technological developments. In particular, Leaver & Srdarov (2023) investigate the hype surrounding ChatGPT and generative AI over the first six months of 2023, finding that these technologies were consistently positioned as either utopian or dystopian, "always seemingly magical, but never banal" (para. 2). They critique this utopian/dystopian binary positioning of generative AI, arguing that the focus on these extremes displaces conversations about the "more grounded and immediate challenges" which are posed by generative AI (Leaver & Srdarov, 2023, para. 2). Similarly, work by Hockenhull & Cohn (2021) set out to analyze the role of hype in performing and translating socio-technical imaginaries of digital technologies. Drawing on fieldwork in Denmark, Hockenhull & Cohn (2021) propose the idea of 'hot air' as a concept which can describe how "hype for the future performs these imaginaries" (p. 302). This range of work sets out the ways that collective belief is mobilized and harnessed despite technology's harms or deficiencies. In what follows, openness will be examined in relation to hype and, to a certain extent, I will suggest that openness can be understood as hype.

Theoretical Framework

The intersecting concepts of sociotechnical imaginaries and carceral imaginaries inform the methodological approach of this project. Both develop out of scholarship involved with conceptualizing imagination as a social practice, such as the accounts advanced by Benedict Anderson in *Imagined Communities* (1983) and Charles Taylor in *Modern Social Imaginaries* (2004). While Anderson employed the concept of imaginaries to describe nations as imagined political communities, Taylor described a "social imaginary" as that which "enables, through making sense of, the practices of a society" (2004, p. 2). These theorizations of imaginaries would be taken up further by Science and Technology Studies (STS) scholars. This project is primarily informed and shaped by the work of Sheila Jasanoff and Ruha Benjamin, both of whom offer their contributions to the theorization of imagination through their concepts of sociotechnical imaginaries and carceral imaginaries. This theoretical framework supports the project by lending crucial insight into the excavation of openness as an imaginary and as hype.

Sheila Jasanoff's concept of sociotechnical imaginaries builds on the work of Anderson and Taylor by occupying the theoretically undeveloped space that lies between "the idealistic collective imaginations identified by social and political theorists" and the "politically neutered networks or assemblages with which STS scholars often describe reality" (2015, p. 19). Jasanoff (2015) defines sociotechnical imaginaries as "collectively held, institutionally stabilized, and publicly performed visions of desirable futures, animated by shared understandings of forms of social life and social order attainable through, and supportive of, advances in science and technology" (p. 4). As Jasanoff explains, sociotechnical imaginaries are "collective, durable, capable of being performed" while also being "situated and culturally particular" (2015, p. 19). Jasanoff also clarifies that these 'imaginaries' are products as well as instruments of science, technology and society. Through engaging with the concept of sociotechnical imaginaries, I question how the AI industry has stabilized openness within particular visions of the future. As Jasanoff writes, imaginaries "encode not only visions of what is attainable through science and technology but also of how life ought, or ought not, to be lived" (2015, p. 4). This framework is essential to studying how certain agendas and ideas about the future, in this case about openness and AI, are set, performed, and stabilized. It also supports an understanding of how imaginaries, including those advanced by the generative AI industry, hold the power to direct development and shape the future.

Ruha Benjamin's (2016) concept of 'carceral imaginaries' builds on Jasanoff's conceptualization of collective imaginations of the future to similarly train scholarly attention towards the particular ways that science and technology is imagined and created. Benjamin (2016) draws together STS and critical race theory to advocate for examining the relationship between innovation and containment, urging scholars to consider "who and what are fixed in *place*—classified, corralled, and/or coerced—to enable technoscientific development?" (p. 145). As Benjamin (2016) elaborates, "carceral imaginaries seek to contain individual bodies *and* collective visions of the future" (p. 151). For example, Benjamin (2016) writes that "newfangled forms of classification and control" are often hidden under noble aims, "such as 'heath' and 'safety'" (p. 150). Benjamin's insistence on recognizing and studying that which remains unchanged underneath veneers of scientific or technological progress directs the focus of this study. Indeed, engaging with this framework supports an examination of openness through questioning what is locked in place as well as who and what are left out of these visions. It also allows critical reflection on who has access to openness and who defines what makes AI ethical, responsible, and safe.

Methods

This project adopted a case studies approach. I examined the public-facing documents of three generative AI companies: Meta, IBM and OpenAI. Meta and IBM were selected due to their involvement in advancing openness discourse. Meanwhile, OpenAI was selected for different reasons. Firstly, OpenAI places 'open' in its name which intersects with openness discourses in spite of their notable lack of openness rhetoric otherwise. Also, OpenAI is perhaps one of the most widely popular generative AI companies and thus contributes greatly to shaping the discourses surrounding AI. Examining each of these companies through the case study approach was deemed most suitable for the project due to the way that each case was able to be assessed separately before examining the cases together. Moreover, the case study approach

enabled in-depth examination of the way in which these companies deploy openness. As opposed to a more broad-view approach which might have looked at more companies, limiting the analysis to three specific cases allowed for more thorough contemplation of the ways that particular companies choose to engage with openness. This prioritization of depth rather than breadth is also reflected in the number of texts with which the analysis is conducted.

For each of the three companies (Meta, IBM, OpenAI) ten public-facing company documents were selected. This produced a collection of 30 documents in total. The documents themselves include webpages, corporate blogposts and company newsletters. This corpus size proved ideal because it was constrained enough to allow for more in-depth analysis of the texts while still remaining large enough to examine themes and patterns across the texts. These texts were selected using a non-probability purposive sampling approach, whereby I selected documents according to their coverage of openness and/or inclusion of key ideas or keywords, such as open, open source, open innovation, open science, transparency, safety, responsibility, and ethics. For Meta and IBM, priority was given to documents which contained more discussion of openness or use of 'open.' For OpenAI, which does not champion openness rhetoric, documents were selected on the basis of references to safety, responsibility and ethics. Additionally, I concentrated on a two-year time frame, limiting the focus to 2022-2024. As such, the documents were selected in accordance with this date range. This timeline aligns with the surfacing of hype surrounding generative AI projects, which was prominent after the release of ChatGPT in November 2022.

After selecting the collection of texts, I closely analyzed the documents, locating themes and patterns. I extracted relevant excerpts and organized them thematically, allowing me to study the ways in which each of the three companies discussed openness and/or safety. When analyzing Meta and IBM, I sought to study how openness was situated, positioned, and articulated across the texts, noting patterns and discrepancies. With OpenAI, I sought to observe how the absence of openness was filled with alternative discourses. Across all three cases, I worked to decipher what values and beliefs seemed to animate the idea of openness while remaining attentive to what was left out, neglected and/or locked in place by the focus on openness. When analyzing each case, I also examined how the future was invoked amongst these discourses. For Meta and IBM, I examined how the discourse of openness was promoted in tandem with future visioning. For OpenAI, I observed how the company envisioned and articulated a future without openness.

More specifically, the cases were examined using critical discourse analysis (CDA). Following Michel Foucault's theorization, discourse is produced by and for power (1980). Discourse itself can (re)produce realties, including assumptions, beliefs and inequalities. In regard to technology, Ian Roderick (2016) asserts that how technology is talked about and represented effectively shapes how technology is experienced and how it comes to evolve. Broadly, discourse analysis concerns itself with studying texts in ways that understand their meanings, contexts and implications (Berger, 2020). In particular, CDA allows scholars to investigate how messages are communicated and how discourse can act as a means of social construction (Fairclough et al., 2011), with particular attention on how power relations are produced, maintained, and/or challenged through discourse (Locke, 2004). Norman Fairclough, one of the founders of CDA, outlines an approach to studying discourse that focuses on three dimensions of discursive practice, namely its linguistic form, its instantiation of social practice, and its social constructions (1992). This project examined openness discourses at the levels of these three dimensions to trace the uses of language therein while also examining the practices and beliefs the discourses advance.

Although the 'critical' in CDA can be interpreted in various ways, this project follows an understanding that 'critical' can be broadly understood as "having distance to the data, embedding the data in the social, taking a political stance explicitly" and focusing on self-reflection throughout the research process (Locke, 2004, p. 26). The CDA approach this research employed is the practice of "de-naturalizing" discourse. As Locke (2004) describes, discourses are often naturalized and thus understood as 'common sense,' however CDA allows for these 'common sense' positions to be demystified, thus exposing them as discursive constructions (p. 32). Broadly, critical discourse studies challenge the "self-evident character and acceptability of a given set of practices" (Roderick, 2016, p. 58). The analysis within this project therefore aspires to participate in denaturalizing the openness discourses advanced by Meta, IBM and OpenAI, revealing them to be strategic constructions.

Analysis

The analysis undertaken was done on the basis of case-by-case investigations. As such, the findings for each case will be presented separately before they are discussed together in the discussion section.

Case 1: Meta

Meta is a key proponent of openness discourse in the generative AI industry and thus provides an ideal case study. In addition to co-launching the AI Alliance with IBM, Meta advocates for open innovation and loudly advertises its large language model (LLM), Llama, as open source. As researchers Liesenfeld & Dingemanse (2024) suggest, perhaps the strongest claims to the open source label in the realm of generative AI "come from Meta and its Llama 2 and Llama 3 models" (p. 1776). For example, the corporate blog post introducing Llama 2 proclaims, "we're introducing the availability of Llama 2, the next generation of our open source large language model" (Meta, 2023a). Similarly, when Llama 3 was introduced, Meta proclaimed it to be "the next generation of our state-of-the-art open source large language model" (Meta, 2024a). However, Meta's promotion of openness extends beyond marketing the Llama models as open source. It is clear that the use of 'open' and 'openness' across Meta's documents is not exclusively tied to open source or to the open innovation and open science approaches that the AI Alliance advocates for.

Findings

To illustrate the findings, I will recount the many ways in which openness appeared in the documents of analysis. One of the most obvious ways that Meta engages with the idea of openness is through promoting the Llama models as open source. Indeed, one document reads, "[o]ur open source approach promotes transparency and access" (Meta, 2023a). This statement conveys that open sourcing Meta's models is meant to bolster broader claims of being committed to the values of transparency and access. Beyond this, Meta's documents also frequently reference its support of "open innovation and open science in AI" (Meta, 2023d), while also throwing around the term "open community" (Meta, 2023d). While the idea of "open

community" was left undefined within the documents of analysis, the term was often used in association with open innovation and open science and thus appears to reflect the desire for these approaches to be conflated with ideas of community and collectivity. In addition to these uses of 'open,' a frequent deployment of the idea appeared in the promotion of an "open ecosystem" (Meta, 2024a). One company blogpost reads, "we're committed to the continued growth and development of an open AI ecosystem" (Meta, 2024b). Elsewhere, Meta vows to "[r]esponsibly advance the ecosystem of open foundation models" (Meta, 2023d).

In addition to these more concrete uses of openness, the term was also found to be applied in ways that betrayed its usage as a buzzword. One of the ways in which this occurred most frequently was by using 'open' as a descriptive word. This ranged from Meta claiming to "share [its] work openly" (Pineau, 2023) and "developed openly" (Meta, 2023d) to professing to do research "in an open and responsible way" (Pineau, 2023). Another frequent form of usage was explicitly referencing openness, whether as a key principle to which Meta subscribes or as a key fixture of the company's culture. For example, a company blog post writes, "our key principles of openness, collaboration, excellence, and scale" (Meta, 2024b). This is echoed in statements such as "our culture of responsibility, excellence, and openness!" (Pineau, 2023). It is also observable in proclamations such as, "[w]e have long believed that openness leads to better, safer products, faster innovation, and a healthier overall market" (Meta, 2024a). In the excerpt that is included below, the many different uses of openness are observable.

Figure 1

Meta Llama's Open Approach

Open Approach Our investment in open trust and safety

We are furthering our commitment to an open approach to build generative AI with trust and safety in mind. We believe in transparency, open science, and cross-collaboration, and to date, we've released over a thousand open-source libraries, models, datasets, and more. Our trust and safety models and tools mark additional contributions towards creating an open AI ecosystem, involving academics, policymakers, industry professionals, and society at large in the responsible development of generative AI.

Note. Excerpt from a Meta webpage (n.d.-b).

In this short paragraph, openness is referred to innumerable times, ranging from references to open source and open science to mentions of an open AI ecosystem and the undefined idea of "open trust and safety." On the same webpage, a heading reads "[w]hy Meta Llama" after which the subtitle declares "[a]n approach to open trust and safety in the era of generative AI" (Meta, n.d.-b). One is meant to infer that Meta Llama is superior simply because of its purported openness.

Openness as Universal.

In addition to the wide-ranging applications of openness, two core themes were also observable amongst the texts of analysis one of which was the way that openness was positioned as universally applicable and beneficial. This appeared potently in phrases such as "benefits to all" (Meta, 2023d) and "building AI that works well for everyone" (Meta, 2023c). In a statement of support for Meta's open approach it reads, "[o]pening today's Llama models will let everyone benefit from this technology" (Meta, 2023b). Similarly, upon releasing the Llama 2 model, Meta proclaimed, "[b]y making AI models available openly, they can benefit everyone" (Meta, 2023a). Not only are these claims highly exaggerated and disproportionate in their promises, but they also suggest a universality of openness by claiming that openness ensures safety and wide benefit. This suggests that openness alone can enable AI to be beneficial, even though it is the inequalities and harms perpetrated by AI systems, rather than lack of access, that makes them unsafe. Crucially, it also assumes that anyone can interface with these systems, when, in truth, it is only those with the required expertise who will be able to do so. Therefore, it remains a limited 'expert' group, albeit perhaps external to Meta, that the system is open to. Notably, it is not accessible to those most harmed by these systems. Moreover, claims such as these advocate that openness is a universal approach to developing and deploying AI, which not only neglects other similarly important values but seems to exclude and overwrite alternative perspectives and approaches.

Openness as the Future.

Another key theme observed in the documents was the referencing of the future. Whether appearing in company blog posts where the future is explicitly discussed or in more covert references to future development, Meta's documents appeared to prioritize future thinking and envisioning. Importantly, the future which is speculated and performed by Meta is one which insists that openness and open approaches will shape the future of AI. More specifically, Meta argues that open source foundation models, conveniently the very products which they are releasing, will come to define the industry. On a webpage for Llama, Meta claims that the model "serves as a bedrock for innovation in the global community" (Meta, n.d.-a). An explicit discussion of the future can be observed below in an excerpt from a featured article by Joelle Pineau, the VP of AI Research at Meta.

Figure 2

The future

While much of the progress in AI over the last decade was achieved by a divide and conquer approach, breaking up the problem into separate well-defined tasks, in the next decade, we are increasingly looking at ways of putting the puzzle pieces together to advance AI. The rise in Foundation models is just the beginning of this: large models with increasingly general abilities, which we can flexibly adapt to our specific needs and values. World models, which can be used to reason and plan, will be increasingly common, allowing us to overcome limitations of current AI models. Rather than a single AGI, we expect the future to feature numerous and diverse populations of AIs deployed across platforms, which will transform how we work, how we play, how we connect, how we create, how we live.

Pursuing this path requires that we also have a deep understanding of how to build AI models responsibly, from beginning to end. We remain committed to doing that work safely and responsibly. Our commitment to open science is a key part of this, and will continue to be part of FAIR's DNA. When we aim to share our work openly, whether it be our papers, code, models, demos or responsible use guides, it helps us set the highest standards of quality and responsibility, which is the best way for us to help the community build better AI solutions. This also directly helps Meta build AI solutions that are safer, more robust, equitable and transparent, and can benefit the many different people that use our products around the world.

Note. Excerpt from Pineau's (2023) article celebrating the 10-year anniversary of Meta's Fundamental AI Research (FAIR) team.

As can be observed throughout this excerpt and the majority of the documents of analysis, there is a clear vision of the future of AI, which Meta constructs. This future will be built on open foundation models, which, unsurprisingly, is the exact product Meta offers. Moreover, open source, open science and even openness are all positioned as key to this vision. Elsewhere in Pineau's article it is suggested that "Meta is uniquely poised to solve AI's biggest problems" due to their "culture of openness" (Pineau, 2023). The promotion of openness within Meta's future visioning therefore operates as a strategic move to not only position Meta favourably as ethical and open but also to position Meta as an alleged industry leader which will usher in the future of AI.

Analysis / Critique

Upon critical reflection, Meta's mobilization of openness rhetoric appears performative. This is made increasingly apparent in instances where 'open' acts as a buzzword. It is also evidenced in the self-serving ways open labels, such as open source and open science, are brandished. Liesenfeld & Dingemanse (2024) evaluated the openness of over 45 generative AI systems. In their ranking of the openness of these systems, Meta's Llama models are among some of the lowest ranked (Liesenfeld & Dingemanse, 2024, p. 1779). As these researchers note, open source has been increasingly co-opted by generative AI, resulting in an observable rise in open washing. As Liesenfeld & Dingemanse (2024) describe, there is ample evidence that open washing is a highly effective communication strategy. Indeed, the launches of Meta's Llama models were met with excitement, while the open source claim was uncritically praised as a key selling point. For example, one *Wired* headline announced, "Meta's Open Source Llama 3 Is Already Nipping at OpenAI's Heels" (Knight, 2024). It is therefore clear that Meta's openness discourse aligns with the practice of open washing, and that this endeavour has been largely successful.

Further scrutiny of Meta's openness, particularly as a purported safety-advancing approach, leaves much to be desired. A blogpost that introduced Llama 2 reads, "[o]pening access to today's AI models means a generation of developers and researchers can stress test them, identifying and solving problems fast, as a community" (Meta, 2023a). As is suggested here, the open source approach seemingly offloads the work of producing safety onto other actors. Meta positions the granting of access to their models as celebratory, meanwhile it is left to other developers and researchers, perhaps even the public at large, to truly identify the problems. Therein lies a clear issue; making a model open does not make it safe and does not reflect responsible practice. Therefore, Meta's deployment of openness rhetoric to gesture at an array of inferred values while supporting their claims of working towards safety and responsibility is certainly troublesome.

Despite some of the problems associated with Meta's openness, its contributions to openness discourse surrounding generative AI not only takes up immense space in the dialogue at large but also, as a result, steers the conversation. The themes of universalism and future envisioning, rampant throughout the documents, work to advance the belief that Meta's openness will be massively beneficial for everyone, everywhere and that it will also be a driver of the future of AI. Crucially, this future envisioning hopes to steer generative AI's future with Meta at the helm.

Case 2: IBM

IBM provides another interesting case study. As previously mentioned, IBM co-founded the AI Alliance with Meta and thus advocates for open science and open innovation. IBM has also focused on releasing open source foundation models, namely through its Watsonx platform. Watsonx is tailored for businesses and is described as a platform that allows enterprises to build their own generative AI models and traditional machine learning systems (Yusuf, 2023). The corporate blog post which announced its release declared, "IBM is making available a wide selection of open-source models through watsonx.ai" (Yusuf, 2023). In addition to IBM's participation in advocating for open innovation and open source, IBM promotes openness rather loudly, thus offering a particularly potent case of analysis.

Findings

Examining the company's documents revealed the ways in which IBM engages with openness discourse. Notably, much of the discussions of openness focused on either open source or open innovation. Along with this more constrained focus on certain types of openness, IBM's rhetoric was considerably heavy-handed. This manifested in statements such as, "[a]t IBM, we take open source seriously" (IBM, n.d.-b), and "[a]t our core, we believe that open source is the bedrock of modern computing" (Maximilien et al., 2023). Furthermore, an entire webpage showcases all of IBM's "Open Projects" (IBM, n.d.-d). Open innovation is similarly advertised rather forcefully. An IBM Institute for Business Value article declares that "[i]n the current ecosystem economy that runs on partnerships, the clear choice for growing one's business is open innovation" and "open innovation has become a prevailing aspiration" (Marshall et al., 2023). The praise of open innovation is also echoed by Arvind Krishna, IBM Chairman and CEO, who comments, "[t]he progress we continue to witness in AI is a testament to open innovation" (IBM, 2023). These consistent references to the benefits of open innovation seem to reveal IBM's desire not only for open innovation, but also for 'open ecosystems,' to be widely adopted by the industry. Interestingly, this echoes Meta's similar promotion of open ecosystems.

In addition to these manifestations of openness, other usages of 'open' were also found throughout IBM's texts. These included references to open culture, such as calls to "[c]reate an open culture" (IBM, n.d.-e) or broad references to fostering a "culture of open" (IBM, n.d.-e). Openness was also elicited in references to "open source communities" and "open communities" (IBM, n.d.-f). Once again, this mirrors Meta's similar use of 'open communities' language. Lastly, 'open' was present in references to open governance and open standards. For example, an article suggests that "[o]pen governance ensures the long-term success and viability of open projects" (IBM, n.d.-e). These various references to openness all work to contribute to the openness discourse that was found strongly throughout IBM's documents.

IBM as an Openness Leader.

A prominent theme observed across the texts was IBM's consistent effort to position itself as a leader of open source and openness. For example, an IBM Developer article dedicated to discussing IBM's open approach to technology declared "IBM is one of the most prolific contributors to open source" (Maximilien et al., 2023). The article also claims that "IBM's commitment and contribution to open source is unrivaled in the industry" (Maximilien et al., 2023). These brazen claims are echoed elsewhere as well. One webpage reads, "IBM is unmatched in the breadth of our open source involvement." (IBM, n.d.-f). Not only does IBM claim to be a central and prolific contributor to open source, but they also claim to have done so before other industry players, as is observable in the excerpt below.

Figure 3

Before Open Source was Cool



Note. Excerpt from "A strong history and commitment to open source" (IBM, n.d.-a).

As is clear, IBM is eager to proclaim their early involvement in open source projects.

This is a position that they quickly leverage, professing that "IBM is recognized by many within the open source community as a leader in open source" (Maximilien et al., 2023). An entire webpage is dedicated to sharing their open source expertise, claiming that "IBM Open Source Support can improve developer productivity" (IBM, n.d.-e). Thus, IBM positions itself as a key figure in the open source community and asserts itself as the expert in openness.

Openness as the Future.

Another theme in the texts was references to the future, ranging from covert to explicit. For example, one webpage exclaims, "[b]uild the future of tech with us" after which it elaborates, "IBM is unmatched in the breadth of our open source involvement" (IBM, n.d.-b). Across the documents, IBM defines the future of technology as one built through open source and open innovation. Moreover, IBM seems to prioritize situating itself as the expert and leader of this future. This rhetoric is easily observable in the excerpt below.

Figure 4

The Open Future (According to IBM)

Envisioning an open future for the enterprise

Hopefully you've gathered that our goal when we get involved in open source is to bring a project under open governance so that it can quickly reap the benefits of an active, diverse developer ecosystem.

A trend we're seeing now is that organizations themselves are becoming open source in the way that they operate as a development organization. As organizations change the way they develop code, infusing open source principles and practices into the development process, the technology that's created is better, more innovative, and more secure.

Part of this openness is the way that teams across different communities are coming together. For instance, the Node.js and JavaScript communities recently joined their foundations so that they can create better, more closely aligned and scalable projects.

We'll be in the fray, continuing to lead by example in the way we infuse open source into our own products and technology, the open way we lead our teams, and how we interact transparently with our clients. We are excited to be a part of this renaissance and to continue to push this trend through our leadership and our commitment to open governance and standards.

Note. Excerpt from "A strong history and commitment to open source" (IBM, n.d.-a).

In phrases like "continuing to lead by example" as seen above, it becomes clear that IBM situates itself as a leader not only of open source but of the future. Moreover, near the end of the excerpt IBM describes the open future as a force of change, a "renaissance," which they are excited to participate in. This rhetoric, which positions openness as a movement of change, is also echoed elsewhere. Writing in support of the AI Alliance, IBM's Chairman and CEO, Arvind Krishna, wrote that "[t]his is a pivotal moment in defining the future of AI" (IBM, 2023). Conveniently, this allegedly important moment is seemingly ushered in by IBM's openness endeavours.

Analysis / Critique

When examining these findings, it is blatant that IBM leverages openness as a key selling point of both the Watsonx platform and the company at large. While this openness discourse is certainly grounded in IBM's various open source endeavours, it is also clear that it benefits IBM to advertise it loudly. Like Meta, IBM promotes their products through their openness. For example, one webpage reads "[e]nterprises turn to Watsonx because it is: Open" (IBM, n.d.-c). Here, openness is suggested as what makes IBM's products superior. This deliberate marketing of openness can be interpreted as open washing, particularly when considering that IBM announced Watsonx in a corporate blog post which touted its openness. This approach, which Liesenfeld & Dingemanse (2024) refer to as a "release by blog post" strategy, has been central to the growth of open washing in generative AI as it allows companies to market their products through claims of openness that are left unchecked. In addition to the consistent leveraging of openness, IBM also uses open source rhetoric to reassert its power in the industry. Throughout almost all the discussions of openness, IBM eagerly asserts its authority, consistently reminding the public that IBM was the earliest, the most prolific and the most dedicated.

While IBM generally supports open source projects, and openness broadly, Watsonx does not appear to be entirely open. In fact, while Watsonx source codes and licences may be open, its full extent and use exists behind a paywall. IBM offers a free trial, but accessing the full capabilities of the platform requires payment. Although open source does not always also refer to software being free —recall the distinction between free software and open source advocates (Truscello, 2005)— they are generally understood in conjunction with one another. I am uninterested in debating the particulars of these movements; instead, what is important here is that *discursively* openness is often understood as also referring to access. However, in the case of Watsonx, despite being purportedly open, wide accessibly is certainly not prioritized. While the platform approach aligns well with Watsonx being tailored towards businesses and enterprises, it certainly cannot be claimed to contribute to opening access to those beyond these limited circles.

This suggests that some of the rampant openness discourse is perhaps relatively shallow. As a result, IBM's touting of openness may, in turn, undermine and dilute the understanding and meaning of openness.

Case 3: OpenAI

OpenAI is a case study vastly unlike the previous two. This is because OpenAI is literally open in name only. They offer a markedly closed and proprietary generative AI product to the public, namely the ChatGPT models. Not only is ChatGPT not open source, but OpenAI also largely avoids engaging in openness rhetoric. For example, OpenAI is not a member of the AI Alliance. Therefore, it may act as a benchmark against which the prior two cases can be better evaluated. This approach is like that of Liesenfeld & Dingemanse (2024) who used ChatGPT as a "closed reference point" in their study, and subsequent ranking, of the openness of generative AI projects. Arguably, Meta and IBM, in part, hope to use openness as a way to differentiate their projects from that of OpenAI's immensely popular ChatGPT.

Findings

In contrast to Meta and IBM, OpenAI does not advance or participate in openness rhetoric. In fact, across the documents examined, there were hardly any usages of the word 'open' apart from that which appears in OpenAI's name. Indeed, the only substantial discussion of openness in the texts of analysis was found in the footnotes of an article written by OpenAI's CEO, Sam Altman, which reads as follows:

[W]e now believe we were wrong in our original thinking about openness, and have pivoted from thinking we should release everything (though we open source some things, and expect to open source more exciting things in the future!) to thinking that we should

figure out how to safely share access to and benefits of the systems. (Altman, 2023) In this excerpt, Altman suggests that openness, while perhaps initially part of the company's aspirations, has since been abandoned. Altman even goes so far as to say they were "wrong" in this thinking. Notably, he goes on to broadly cite 'safety' and 'benefits' as the reasoning for not prioritizing openness.

Safety Instead of Openness.

Across the texts of analysis, safety was discussed in ways that suggest it is being offered as the antithesis of openness. In its charter, OpenAI discusses the publishing of its AI research, explaining that "we expect that safety and security concerns will reduce our traditional publishing in the future" (OpenAI, n.d.-c). Similarly, in a safety update, OpenAI discloses that its security efforts prioritize "restricting access to training environments and high-value algorithmic secrets on a need-to-know basis" (OpenAI, 2024b). The update goes on to say, "[w]e believe that protecting advanced AI systems will benefit from an evolution of infrastructure security" (OpenAI, 2024b). As is clear in these examples, OpenAI seems to suggest that openness threatens safety and security. In the excerpt below, OpenAI discusses its approach to safety, citing "rigorous testing" and "feedback."

Figure 5

Building Safe AI Systems (According to OpenAI)

Building increasingly safe AI systems

Prior to releasing any new system we conduct rigorous testing, engage external experts for feedback, work to improve the model's behavior with techniques like reinforcement learning with human feedback, and build broad safety and monitoring systems.

For example, after our latest model, GPT-4, finished training, we spent more than 6 months working across the organization to make it safer and more aligned prior to releasing it publicly.

We believe that powerful AI systems should be subject to rigorous safety evaluations. Regulation is needed to ensure that such practices are adopted, and we actively engage with governments on the best form such regulation could take.

Note. Excerpt from "Our approach to AI safety" (OpenAI, 2023a).

In other texts similarly focused on safety, various approaches and values are referenced instead of openness. One document pledges that OpenAI will reinforce "safety, security and trustworthiness" (OpenAI, 2023b). Another suggests that addressing safety requires "extensive debate, experimentation, and engagement" (OpenAI, 2023a). One of the only mentions of 'open' appears as follows: "[w]e have and will continue to foster collaboration and open dialogue among stakeholders to create a safe AI ecosystem" (OpenAI, 2023a). Notably, this is merely a reference to "open dialogue" with stakeholders. Across these texts, OpenAI appears to deliberately place emphasis on safety and security, while doing so in a way that stands in opposition with openness. Instead of discussing openness, OpenAI describes its efforts to rigorously enforce security and consult with experts, rather than opening up the conversation or its source code. Moreover, this emphasis on safety is leveraged to advance OpenAI's bold claims, such as assertions that its products are "used for the benefit of all" and that its "primary fiduciary duty is to humanity" (OpenAI, n.d.-c).

The Future without Openness.

In addition to discussing safety without featuring openness, the future that OpenAI signals across its content also lacks mention of openness. Instead, the future that OpenAI consistently and rather fervently predicts is one that will be defined by artificial general intelligence (AGI). In an article written by CEO Sam Altman, the future is predicted to be transformed by the AGI that OpenAI will supposedly develop.

Figure 6

The Long Term (According to OpenAI)

The long term

We believe that the future of humanity should be determined by humanity, and that it's important to share information about progress with the public. There should be great scrutiny of all efforts attempting to build AGI and public consultation for major decisions.

Note. Excerpts from "Planning for AGI and beyond" (Altman, 2023).

While this excerpt mentions sharing information and consulting the public, it does not explicitly engage with the idea of openness. Later in the above article, Altman writes, "[w]e can imagine a world in which humanity flourishes to a degree that is probably impossible for any of us to fully visualize yet. We hope to contribute to the world an AGI aligned with such flourishing" (2023). This future envisioning is present throughout much of OpenAI's texts. For example, a link to "view careers" is accompanied by large text proclaiming, "[j]oin us in shaping the future of technology" (OpenAI, n.d.-a). This, once again, works to assert that the future of technology will be shaped by OpenAI.

Analysis / Critique

In March 2024, OpenAI was brought under scrutiny because of Elon Musk filing a lawsuit against the company (which has since been dropped) for their lack of openness. The suit stated that, "OpenAI Inc has been transformed into a closed-source, de facto subsidiary of the largest technology company in the world: Microsoft" (Robins-Early, 2024). To further illustrate the grievance of his lawsuit, Musk posted a meme in which OpenAI was renamed as "ClosedAI" (Robins-Early, 2024). OpenAI subsequently responded to these allegations in a corporate blog post, explaining that "[w]e're making our technology broadly usable in ways that empower people and improve their daily lives, including via open-source contributions" (Brockman et al., 2024). OpenAI also added that "[w]e provide broad access to today's most powerful AI, including a free version that hundreds of millions of people use every day" (Brockman et al., 2024). Here, OpenAI appears to suggest that it is reasonably open due to the widespread usage of ChatGPT. However, as Liesenfeld & Dingemanse (2024) argue, "ubiquity and free availability are not equal to openness and transparency" (p. 1774). Although these ideas are often conflated, ChatGPT's wide usage and free access do not mean the model is open source or that the company is prioritizing openness. Later in the post, documentation of an email chain includes a clarification of the company's name, explaining that the "[o]pen in openAI means that everyone should benefit from the fruits of AI after its built" (Brockman et al., 2024). Interestingly, this explanation seems to once again point to the way that 'open' and 'openness' are consistently conceived of and applied in unfixed ways. In the case of OpenAI's name, openness is purportedly merely meant to gesture towards widespread benefit.

As was discovered, OpenAI implies that openness must be sacrificed to advance safety. However, the use of safety as a reason or rationale for rejecting openness certainly falls short, particularly in light of the ways that public-facing proprietary models often produce evident harms (Liesenfeld et al., 2023). Instead, the decision to pivot away from openness likely has more to do with serving the interests of the company than safety. An email included in the response to Musk's lawsuit mentions that as OpenAI approaches its goal "it will make sense to start being less open" (Brockman et al., 2024). Across the texts, OpenAI promotes a version of safety primarily concerned with bad actors and model safeguards. For example, a safety update explains that OpenAI hopes to "increase our security posture to ultimately be resilient to sophisticated state actor attacks and to ensure that we introduce additional time for safety testing before major launches" (OpenAI, 2023c). In addition to this, OpenAI reveals who and what it is concerned with keeping safe as they proclaim, "[w]e prioritize protecting our customers, intellectual property, and data" (OpenAI, 2024b). This seems to indicate that, perhaps above anything else, what OpenAI hopes to hold safe and secure is its own trade secrets.

Discussion

In reflecting on these case studies, it might be helpful to ask: what does openness mean in the generative AI landscape? However, the answer is neither clear nor simple. Taken together, the case studies work to reveal the manifold dimensions and tensions of openness discourse. On the one hand, Meta and IBM deploy openness through the use of the open source label, open innovation, communities and ecosystems language, and broad applications of open rhetoric. This rhetoric ranges from claims of doing things "openly" or in an "open way" (Meta, 2023d; Pineau, 2023) to proclaiming cultures of open and openness (IBM, n.d.-e; Pineau, 2023). As surmised, it is not an accident that both Meta and IBM engage in this discourse and do so similarly because

both are co-founding members of the AI Alliance and thus seek to advance a similar openness agenda. Conversely, OpenAI was found only to deploy openness insofar as it reasserts its commitment to safety and security. For example, an OpenAI spokesperson claimed, "[w]e only open-source powerful AI models once we have carefully weighed the benefits and risks, including misuse and acceleration" (Gent, 2024). As this tension demonstrates, there is not the unequivocal consensus on openness that Meta and IBM would have the public believe. These claims, such as Yann LeCun declaring that "[o]penness is the only way" (Associated Press, 2023), therefore demonstrate how openness is deployed strategically to shed favourable light on AI projects, to bolster claims of ethical practice or to advance a company's position, in one way or another, in the generative AI market at large. This is perhaps most potently observed when Meta and IBM position their products as desirable simply due to purported openness. Several documents pose the question: why choose Meta Llama or Watsonx? To which Meta and IBM respond: because its 'open' (Meta, n.d.-b; IBM, n.d.-c).

As a result of these findings, it is reasonable to suggest that Meta and IBM's openness discourse appears to reflect the practice of open washing. As Liesenfeld & Dingemanse (2024) find, many key industry players do not exhibit systems that are significantly more open than OpenAI's ChatGPT. And yet, these same companies will bill their projects as open source when they are "at best open weight" (Liesenfeld & Dingemanse, 2024, p. 1781). This is certainly the case with Meta Llama, which was among the lowest ranked for openness in Liesenfeld & Dingemanse's (2024) study and perhaps may also be the case with Watsonx. Therefore, it is clear that not only broader promotions of openness but also the touting of the open source label contributes to open washing. Concerningly, Meta and IBM's loud proclamations of the open

source label in spite of potentially insufficient follow-through, as well as a broader lack of industry consensus over what truly makes a generative AI project open source (Gent, 2024), works to undermine and dilute the label.

As can also be observed, openness exists as a loaded concept which carries immense rhetorical baggage that, as scholars have set out, harkens back to the Enlightenment and even human rights. As was discovered in the cases, Meta and IBM employ openness in ways that draw on the concept's inferred values and connoted ideals. Indeed, simply calling something 'open' would perhaps not be read as beneficial without drawing on associated ideas of access, participation, transparency, democracy and more. This can be observed explicitly when Meta claims, "[o]ur open source approach promotes transparency and access" (Meta, 2023a). Openness is thus being extrapolated to signal an array of values, including, as is stated, transparency and access. This is witnessed throughout the discourses analyzed as open rhetoric was found to frequently signal towards often unspecified but nonetheless inferred values and ideals.

Beyond simply understanding and tracing the term's unfixed definition and array of applications, it is important to reflect on how openness is wielded across the case studies. Building on an understanding of imaginaries supported by Jasanoff (2015) and Benjamin (2016), I will suggest that openness might be understood as an imaginary which is being formulated in various ways by major industry players. Therefore, I would like to briefly unpack the imaginary and subsequent future vision that is advanced and, crucially, who is included in it.

One of the ways in which openness can be analyzed as an imaginary is through examining the insistence on universalism which was found prominently in the cases of Meta and IBM. Interestingly, this ties back to the idea that openness and transparency have long been positioned as apolitical. As has already been set out, both Adams (2020) and Schudson (2015) identify that throughout history, openness and transparency have been notably unterhered to specific social and/or political convictions and they are often conceived of as geographically and ideologically neutral. Through advancing ideals of access and participation, openness and transparency gesture towards ideas like equality and democracy without being expressly prioritized with power and power dynamics. I will suggest that we interpret the way that generative AI companies advance an assumed universalism of openness as a continuation of this. For example, this is particularly observable in Meta's claims that openness will allow "benefits to all" (Meta, 2023d) and that it will "benefit everyone" (Meta, 2023a). In these instances, and others, openness is positioned as universally beneficial while it remains unclear who the "all" or "everyone" really refers to. Following Benjamin's (2016) assertions, it is crucial to remain attentive to who and what is left out. In this case, the collective being signalled appears to be more limited and select than advertised which becomes evident when considering the groups which do not necessarily stand to gain from openness. For example, the Global Indigenous Data Alliance (2019) have indicated that openness does not adequately protect Indigenous peoples. As they advocate, prioritizing visibility and access does not account for the power differentials and historical contexts which produce harm.

To further unpack the openness imaginary, it is also necessary to examine the futures which are envisioned and projected across the case studies. Both Meta and IBM insist upon the future of the generative AI industry being 'open' and fundamentally transformed by openness. In this imagined future, the industry is dominated by open foundation models, shared in the spirit of open science and open innovation, to foster open ecosystems and open communities. However, the communities which Meta and IBM invoke are comprised primarily of technology experts and enterprises, not the everyday person. Therefore, this future vision does not challenge or meaningfully redistribute power and arguably postulates a future in which nothing fundamentally changes. This can also be observed in the future vision which OpenAI passionately set out. In this vision, OpenAI produces AGI. When setting out this future Sam Altman writes, "[w]e can imagine" (2023). Building once more on Benjamin's (2016) work, it is crucial to question who the "we" doing the imagining refers to. In this case, it can be assumed that it is OpenAI, and perhaps Altman specifically, who is declaring this 'desirable' future vision. Across the cases, the future visions that are fabricated, whether defined by openness or not, are ones that are imagined by and stand to serve the companies that advance them.

To conclude this discussion, I will return to the idea that openness might be read as hype. In observing the cases together, it becomes clear that there are rather polarizing ways in which openness is deployed. As discussed, Meta and IBM position openness as widely beneficial and integral to their imagined futures. Conversely, OpenAI rejects and dismisses openness on account of safety and security, relying on emphasizing these priorities instead. These opposing views can be explained and rationalized by recognizing that how openness is postured has to do with hype. As a tool through which major industry players have long sought to advance their technologies, hype serves to direct attention and shape public conversation. In many ways, openness discourse might be understood as an emergent pathway for this to occur. For Meta and IBM, openness is leveraged to hype their generative AI projects and situate themselves as the leaders and experts. For example, upon the release of Llama, Meta was lavished with praise for its open approach. *MIT Technology Review* described that the "news [was] welcomed by many," citing several statements which applaud Meta (Heaven, 2022). Therefore, openness can even be seen feeding into the hype surrounding the release of generative AI products. Openness can also enable companies to sidestep scrutiny while appealing to reduced regulation in favour of self-regulation. Indeed, loudly proclaiming commitments to openness allows industry players to appear to be engaging in appropriate self-regulation, subsequently making external regulation seem less necessary. The AI Alliance might be understood in this light as a form of self-regulation that endeavours to keep concern and criticism at bay. Although the case of OpenAI stands in contrast to this, it also serves to demonstrate how the hyping up or down of openness is wielded deliberately by industry players.

Before continuing, I want to clarify that I am not claiming that openness is always and exclusively hype. As has been elaborated prior, openness manifests in many different areas, such as government, education and more, in important ways. In fact, the open source movement has been integral to the development of software and the internet at large (Heaven, 2023). However, the ways in which openness has been taken up by the generative AI industry is a different situation entirely. In this space, not even the open source label has a clear definition. As has been demonstrated throughout the case studies, lack of definition has not stopped major industry players from adopting the open source label while also broadly brandishing openness. As a result, it may be suggested that openness operates as and for hype. Ultimately, this discussion sought to trace through and reflect on the many nuances and tensions which were observed across the case studies.

Conclusion

Considering the immense attention and awe which has surrounded AI in the past few years as well as the documented "steep rise in generative AI systems that claim to be open" (Liesenfeld & Dingemanse, 2024, p. 1774), this project set out to investigate how openness discourse has manifested and proliferated within the generative AI industry. Relying on critical discourse analysis and the theoretical framework of sociotechnical and carceral imaginaries, this project investigated openness discourse while working to uncover the extent to which this discourse can be understood as open washing and even hype. The paper began by thematically setting out the array of existing literature which the project draws from and builds upon. Next, the conceptual framework of imaginaries, specifically that which Jasanoff (2015) and Benjamin (2016) have theorized, was elaborated. This theory, along with the methodology of critical discourse analysis, informed the methods which structured the project as a whole. These methods were then deployed by analyzing three case studies focusing on Meta, IBM and OpenAI. The findings and analysis of each case were presented separately, after which the discussion sought to bring the cases together and reflect on the discoveries as a whole.

As became evident throughout the analysis and discussion, openness discourse is complex and nebulous but is often wielded deliberately by major players in the generative AI industry. As the findings illuminate, this strategic and often performative deployment of open rhetoric reflects the practice of open washing, whereby companies leverage claims of openness and open credentials to gain market capital. Furthermore, building on the theoretical framework, openness could be unpacked as an imaginary. One which advances a particular agenda and subsequent future vision, particularly in the cases of Meta and IBM. However, this future vision is one which, despite an insistence on universalism, reflects a limited view of the future of AI in which very little is fundamentally changed. Considering the various ways in which each company sought to use openness, it was suggested that openness, in this context, could be understood and read as hype.

While this project is only able to reflect and capture a particular moment in what is sure to be a continuously evolving conversation, it nonetheless offers a view into the specific ways in which three industry players (Meta, IBM, OpenAI) are participating, or not, in openness discourses at large. However, it is necessary to outline the limitations which have constrained the project. Most notably, the limited amount of data sources and focused case study approach, while determined to be the most suitable for the project, did indeed present its own limitations. In fact, this focused approach is unable to present and investigate the entirety of how openness discourse is playing out in the generative AI industry. Further, in focusing deliberately on analyzing publicfacing company documents, the study could not reflect and examine how these discourses are also taken up in technology journalism and beyond. Lastly, in focusing on how the discourse operates, the project is not suited to critiquing the actual implementation of openness and, therefore, was not able to extensively explore how openness may be insufficient as an ethics framework in the AI realm.

In light of these limitations, there are many directions for future research. As is suggested throughout, openness discourse may allow technology companies to perform ethical behaviour without adequate change or accountability. Accordingly, future research may endeavour to specifically investigate what the performance of openness enables industry players to do and to what extent. How effective is openness discourse at staving off scrutiny, supporting moves for

de-regulation and more? Further, as is set out in the project, openness discourse appears to be a potent way in which major players hope to dominate and steer the AI ethics conversations. As Ricaurte (2022a) insists, when ethical frameworks materialize it is vital to remember that "AI is a matter of power," a tool that allows wealth, knowledge and control to be concentrated (para. 2). Therefore, it remains important to evaluate and question: what are the conditions, values and frameworks that would meaningfully create a safe and accountable AI sector? This remains unanswered within this project but presents a viable and important avenue for ongoing and future scholarship.

References

- Adam, A. (1998). Artificial knowing: Gender and the thinking machine. Routledge. https:// doi.org/10.4324/9780203005057
- Adams, R. (2020). *Transparency: New Trajectories in Law* (1st ed.). Routledge. https://doiorg.libaccess.lib.mcmaster.ca/10.4324/9780429340819
- Altman, S. (2023, February 24). *Planning for AGI and beyond*. OpenAI. https://openai.com/ index/planning-for-agi-and-beyond/
- Anderson, B. (1983). Imagined Communities: Reflections on the Origin and Spread of Nationalism. Verso Editions and NLB.
- Associated Press. (2023, December 5). *Meta and IBM launch 'AI Alliance' to promote opensource AI development*. The Guardian. https://www.theguardian.com/technology/ 2023/dec/05/open-source-ai-meta-ibm
- Atanasoski, N., & Vora, K. (2019). The Surrogate Human Effects of Technoliberalism. In Surrogate Humanity: Race, Robots, and the Politics of Technological Futures (1st ed., pp.1-26). Duke University Press. https://doi.org/10.1515/9781478004455
- Benjamin, R. (2016). Catching our breath: Critical Race STS and the Carceral Imagination. Engaging Science, Technology, and Society, 2. https://doi.org/10.17351/ests2016.70
- Benjamin, R. (2019). Technological Benevolence: Do Fixes Fix Us? In *Race After Technology* (pp. 137-159). Polity Press.
- Berger, A. A. (2020). *Media and Communication Research Methods: An Introduction to Qualitative and Quantitative Approaches* (Fifth Edition). Sage Publications.

- Bietti, E. (2021). From Ethics Washing to Ethics Bashing: A Moral Philosophy View on Tech Ethics. *Journal of Social Computing*, 2(3), 266–283. https://doi.org/10.23919/ JSC.2021.0031
- Brockman, G., Sutskever, I., Schulman, J., Altman, S. & Zaremba, W. (2024, March 5). *OpenAI* and Elon Musk. OpenAI. Retrieved July 20, 2024, from https://openai.com/index/ openai-elon-musk/
- Broussard, M. (2018). Hello, Reader. In Artificial unintelligence: How computers misunderstand the world (pp. 3-12). MIT Press.
- Broussard, M. (2023). *More Than a Glitch: Confronting Race, Gender, and Ability Bias in Tech* (1st ed.). MIT Press. https://doi.org/10.7551/mitpress/14234.001.0001
- Buolamwini, J. (2023). Unmasking AI: My mission to protect what is human in a world of machines. Random House Publishing Group.
- Chesbrough, H. W. (2003). *Open innovation: The new imperative for creating and profiting from technology*. Harvard Business School Press.
- Chown, E., & Nascimento, F. (2023). Digital Metaphors. In *Meaningful Technologies: How Digital Metaphors Change the Way We Think and Live* (pp. 51–72). Lever Press. http://www.jstor.org/stable/10.3998/mpub.12668201.8

Criado-Perez, C. (2019). Invisible women: Data bias in a world designed for men. Abrams Press.

Dahlander, L., & Gann, D. M. (2010). How open is innovation? *Research Policy*, *39*(6), 699–709. https://doi.org/10.1016/j.respol.2010.01.013

- ElGhadban, A. (2023, January 30). *The return of East India companies: AI, Africa and the new (digital) colonialism.* https://datapopalliance.org/the-return-of-east-india-companies-ai-africa-and-the-new-digital-colonialism/
- Fairclough, N. (1992). Discourse and Social Change. Polity Press.
- Fairclough, N., Mulderrig, J., & Wodak, R. (2011). Critical discourse analysis. In (T. A. Van Dijk, Ed.) *Discourse Studies: A Multidisciplinary Introduction* (2 ed., pp. 357-378).
 SAGE Publications Ltd, https://doi.org/10.4135/9781446289068
- Fjeld, J., Achten, N., Hilligoss, H., Nagy, A., Srikumar, M. (2020, January 15). Principled Artificial Intelligence: Mapping Consensus in Ethical and Rights-Based Approaches to Principles for AI. Berkman Klein Center Research Publication No. 2020-1. http:// dx.doi.org/10.2139/ssrn.3518482
- Foucault, M. (1980). *Power/Knowledge: Selected Interviews and Other Writings* 1972–1977, Trans. C Gordon. New York: Pantheon Books.
- Gent, E. (2024, March 25). The tech industry can't agree on what open-source AI means. That's a problem. MIT Technology Review. https://www.technologyreview.com/ 2024/03/25/1090111/tech-industry-open-source-ai-definition-problem/
- Global Indigenous Data Alliance. (2019). CARE Principles for Indigenous Data Governance. GIDA-global.org
- Green, B. (2021). The Contestation of Tech Ethics: A Sociotechnical Approach to Technology Ethics in Practice. *Journal of Social Computing*, 2(3), 209–225. https://doi.org/ 10.23919/JSC.2021.0018

- Heaven, W. D. (2022, May 3). Meta has built a massive new language AI—and it's giving it away for free. MIT Technology Review. https://www.technologyreview.com/ 2022/05/03/1051691/meta-ai-large-language-model-gpt3-ethics-huggingfacetransparency/
- Heaven, W. D. (2023, May 12). The open-source AI boom is built on Big Tech's handouts. How long will it last? MIT Technology Review. https://www.technologyreview.com/ 2023/05/12/1072950/open-source-ai-google-openai-eleuther-meta/
- Hockenhull, M., & Cohn, M. L. (2021). Hot air and corporate sociotechnical imaginaries:
 Performing and translating digital futures in the Danish tech scene. *New Media & Society.*, 23(2), 302–321. https://doi.org/10.1177/1461444820929319
- Hoffmann, A. L. (2019). Where fairness fails: data, algorithms, and the limits of antidiscrimination discourse. *Information, Communication & Society, 22*(7), 900–915. https://doi.org/10.1080/1369118X.2019.1573912
- Hoffmann, A. L. (2021). Terms of inclusion: Data, discourse, violence. *New Media & Society.,* 23(12), 3539–3556. https://doi.org/10.1177/1461444820958725
- Hoffmann, A. L., Proferes, N., & Zimmer, M. (2018). "Making the world more open and connected": Mark Zuckerberg and the discursive construction of Facebook and its users. New Media & Society., 20(1), 199–218. https://doi.org/ 10.1177/1461444816660784
- Hong, S.-H. (2020). Honeymoon Objectivity. In *Technologies of Speculation: The Limits of Knowledge in a Data-Driven Society* (pp. 13-29). NYU Press. https://doi.org/ 10.2307/j.ctv17db2hd

- Hong, S.-H. (2021, February 18). Why Transparency Won't Save Us. Centre for International Governance Innovation. https://www.cigionline.org/articles/why-transparency-wontsave-us/
- IBM. (2023, December 5). AI Alliance Launches as an International Community of Leading Technology Developers, Researchers, and Adopters Collaborating Together to Advance Open, Safe, Responsible AI. IBM Newsroom. https://newsroom.ibm.com/AI-Alliance-Launches-as-an-International-Community-of-Leading-Technology-Developers,-Researchers,-and-Adopters-Collaborating-Together-to-Advance-Open,-Safe,-Responsible-AI
- IBM. (n.d.-a). A strong history and commitment to open source. Open Source @ IBM. Retrieved July 19, 2024, from https://www.ibm.com/opensource/story/
- IBM. (n.d.-b). *Innovating in the open*. Open Source @ IBM. Retrieved July 19, 2024, from https://www.ibm.com/opensource/?mhsrc=ibmsearch_a&mhq=open
- IBM. (n.d.-c). *Meet Watsonx*. Retrieved July 19, 2024, from https://www.ibm.com/watsonx
- IBM. (n.d.-d). Open Projects. IBM Developer. Retrieved July 19, 2024, from https:// developer.ibm.com/open/projects/
- IBM. (n.d.-e). Sharing our expertise in open source. Open Source @ IBM. Retrieved July 19, 2024, from https://www.ibm.com/opensource/enterprise/
- IBM. (n.d.-f). *Take your skills to the next level with generative AI*. IBM Developer. Retrieved July 19, 2024, from https://developer.ibm.com/generative-ai-for-developers/

- Jasanoff, S. (2015). Future Imperfect: Science, Technology, and the Imaginations of Modernity. In S. Jasanoff & S.-H. Kim (Eds.), *Dreamscapes of Modernity: Sociotechnical Imaginaries and the Fabrication of Power* (pp. 1-33). University of Chicago Press.
- Jobin, A., Ienca, M., and Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence 1*(9): 389–399. https://doi.org/10.1038/s42256-019-0088-2
- Jones, C. (2015). Openness, technologies, business models and austerity. *Learning, Media and Technology*, 40(3), 328–349. https://doi.org/10.1080/17439884.2015.1051307
- Kakalic, Y. & Hancock, J. (2023, August 31). Studying how Precarity, Technology, and AI intersect. The Inclusive Digital Economy Network+. https://includeplus.org/ studying-how-precarity-technology-and-ai-intersect/
- Kitchin, R. (2014). The Data Revolution: Big Data, Open Data, Data Infrastructures & Their Consequences. Sage Publications. https://doi.org/10.4135/9781473909472
- Knight, W. (2024, April 25). Meta's Open Source Llama 3 Is Already Nipping at OpenAI's Heels. Wired. https://www.wired.com/story/metas-open-source-llama-3-nipping-at-openaisheels/
- Korn, J. U. (2021). Connecting Race to Ethics Related to Technology: A Call for Critical Tech Ethics. *Journal of Social Computing*, 2(4), 357–364. https://doi.org/10.23919/ JSC.2021.0026
- Leaver, T., & Srdarov, S. (2023). ChatGPT Isn't Magic: The Hype and Hypocrisy of Generative Artificial Intelligence (AI) Rhetoric. *M/C Journal*, *26*(5). https://doi.org/10.5204/ mcj.3004

- Liesenfeld, A., Lopez, A., & Dingemanse, M. (2023). Opening up ChatGPT: Tracking openness, transparency, and accountability in instruction-tuned text generators. *arXiv.Org.* doi:10.1145/3571884.3604316
- Liesenfeld, A. & Dingemanse, M. (2024). Rethinking open source generative AI: open washing and the EU AI Act. In *The 2024 ACM Conference on Fairness, Accountability, and Transparency (FAccT '24)*. Association for Computing Machinery, New York, NY, USA, 1774–1787. doi: 10.1145/3630106.3659005
- Locke, T. (2004). Critical discourse analysis: Critical discourse analysis as research methodology. Bloomsbury Publishing Plc.
- Marshall, A., Higgins, L., Crysel, K., Dencik, J. & Fisher, L. (2023, October 11). *Ecosystems and open innovation: Co-create or stagnate*. IBM Institute for Business Value. https://www.ibm.com/thought-leadership/institute-business-value/en-us/report/ecosystems-open-innovation
- Maximilien, M., Topol, B. & Sepi, J. (2023, February 16). *IBM's approach to open technology*. IBM Developer. https://developer.ibm.com/articles/cl-open-architecture-update/
- Meijer, A. (2009). Understanding modern transparency. *International Review of Administrative Sciences*, 75(2), 255-269. https://doi.org/10.1177/0020852309104175

Meijer, A. J., Curtin, D., & Hillebrandt, M. (2012). Open government: Connecting vision and voice. *International Review of Administrative Sciences*, 78(1), 10–29. https://doi.org/ 10.1177/0020852311429533 Meijer, A. (2013). The History of Transparency: Analyzing the Long-term Socio-Political Construction of Transparency in the Netherlands. Paper prepared for the International Conference on Transparency Research, Paris, October 24th–26th 2013.

Merriam-Webster. (n.d.). Openness. https://www.merriam-webster.com/dictionary/openness

- Meta. (2023a, July 18). *Meta and Microsoft Introduce the Next Generation of Llama*. Meta Newsroom. https://about.fb.com/news/2023/07/llama-2/
- Meta. (2023b, July 18). *Statement of Support for Meta's Open Approach to Today's AI*. Meta Newsroom. https://about.fb.com/news/2023/07/llama-2-statement-of-support/
- Meta. (2023c, November 30). *Responsible AI: FAIR progress and learnings across socially responsible AI research*. Meta Blog. https://ai.meta.com/blog/fair-progress-and-learnings-across-socially-responsible-ai-research/
- Meta. (2023d, December 4). AI Alliance Launches as an International Community of Leading Technology Developers, Researchers, and Adopters Collaborating Together to Advance Open, Safe, Responsible AI. Meta Blog. https://ai.meta.com/blog/ai-alliance/
- Meta. (2024a, April 18). *Introducing Meta Llama 3: The most capable openly available LLM to date*. Meta Blog. https://ai.meta.com/blog/meta-llama-3/
- Meta. (2024b, June 18). *Sharing new research, models, and datasets from Meta FAIR*. Meta Blog. https://ai.meta.com/blog/meta-fair-research-new-releases/
- Meta. (n.d.-a). *Discover the possibilities with Meta Llama*. Meta Llama. Retrieved July 19, 2024, from https://llama.meta.com/

- Meta. (n.d.-b). *Making safety tools accessible to everyone: Enabling developers, advancing safety, and building an open ecosystem*. Meta Llama. Retrieved July 19, 2024, from https://llama.meta.com/trust-and-safety
- Meta. (n.d.-c). Request for Expressions of Interest (RFEI):Open Innovation AI Research Community. Meta Llama. Retrieved July 19, 2024, from https://llama.meta.com/openinnovation-ai-research-community/
- Mullaney, T. S., Peters, B., Hicks, M., & Philip, K. (2021). *Your Computer Is on Fire* (1st ed.). MIT Press. https://doi.org/10.7551/mitpress/10993.001.0001

Munn, L. (2022). Automation is a myth. Stanford University Press.

- Munn, L. (2023). The five tests: designing and evaluating AI according to indigenous Māori principles. AI & Society (2023). https://doi.org/10.1007/s00146-023-01636-x
- Noble, S. U. (2018). *Algorithms of oppression: how search engines reinforce racism*. New York University Press. https://doi.org/10.18574/9781479833641
- OpenAI. (2023a, April 5). *Our approach to AI safety*. https://openai.com/index/our-approach-toai-safety/
- OpenAI. (2023b, July 21). *Moving AI governance forward*. https://openai.com/index/moving-ai-governance-forward/
- OpenAI. (2023c, October 26). OpenAI's Approach to Frontier Risk: An Update for the UK AI Safety Summit. https://openai.com/global-affairs/our-approach-to-frontier-risk/
- OpenAI. (2024a, May 7). *Our approach to data and AI*. https://openai.com/index/approach-todata-and-ai/

- OpenAI. (2024b, May 21). OpenAI safety update: Sharing our practices as part of the AI Seoul Summit. https://openai.com/index/openai-safety-update/
- OpenAI. (2024c, May 28). *OpenAI Board Forms Safety and Security Committee*. https:// openai.com/index/openai-board-forms-safety-and-security-committee/
- OpenAI. (n.d.-a). About. Retrieved July 20, 2024, from https://openai.com/about/
- OpenAI. (n.d.-b). *Developing beneficial AGI safely and responsibly*. Retrieved July 20, 2024, from https://openai.com/safety/
- OpenAI. (n.d.-c). OpenAI Charter. Retrieved July 20, 2024, from https://openai.com/charter
- Oxford University Press. (n.d.). *Openness*. Oxford Learner's Dictionaries. https:// www.oxfordlearnersdictionaries.com/definition/english/openness
- Pineau, J. (2023, November 30). Celebrating 10 years of FAIR: A decade of advancing the stateof-the-art through open research. Meta Blog. https://ai.meta.com/blog/fair-10-yearanniversary-open-science-meta/
- Puschmann, C. & Burgess, J. (2014). Big data, big questions: Metaphors of big data. International Journal of Communication, 8(20). https://ijoc.org/index.php/ijoc/ article/view/2169
- Ricaurte, P. (2022a, March 4). *Artificial intelligence and the feminist decolonial imagination*. BotPopuli. https://botpopuli.net/artificial-intelligence-and-the-feminist-decolonialimagination/
- Ricaurte, P. (2022b). Ethics for the majority world: AI and the question of violence at scale. *Media, Culture & Society*, *44*(4), 726–745. https://doi.org/ 10.1177/01634437221099612

- Robins-Early, N. (2024, March 9). *The feud between Elon Musk and Sam Altman explained*. The Guardian. https://www.theguardian.com/technology/2024/mar/09/why-is-elon-musk-suing-sam-altman-openai#
- Robinson, S. C. (2020). Trust, transparency, and openness: How inclusion of cultural values shapes Nordic national public policy strategies for artificial intelligence (AI).
 Technology in Society, 63, 101421-. https://doi.org/10.1016/j.techsoc.2020.101421
- Roderick, I. (2016). Critical Discourse Studies and Technology: A Multimodal Approach to Analysing Technoculture. Bloomsbury.
- Schlagwein, D., Conboy, K., Feller, J., Leimeister, J. M., & Morgan, L. (2017). "Openness" with and without Information Technology: A Framework and a Brief History. *Journal of Information Technology*, 32(4), 297-305. https://doi.org/10.1057/s41265-017-0049-3
- Schudson, M. (2015). The Rise of the Right to Know: Politics and the Culture of Transparency, 1945-1975. Harvard University Press. https://doi.org/10.4159/9780674915787
- Stark, L. & Hoffmann, A.L. (2019). Data Is the New What? Popular Metaphors & Professional Ethics in Emerging Data Culture. *Journal of Cultural Analytics 4*(1). https://doi.org/ 10.22148/16.036.
- Taylor, C. (2004). Modern Social Imaginaries. Duke University Press.
- The AI Alliance. (n.d.). Our Mission. AI Alliance. https://thealliance.ai/
- Truscello, M. (2005). The Rhetorical Ecology of the Technical Effect. *Technical Communication Quarterly: TCQ.*, *14*(3), 345–351. https://doi.org/10.1207/s15427625tcq1403_13
- Weller, M. (2014). *The Battle For Open: How openness won and why it doesn't feel like a victory*. Ubiquity Press.

Wyatt, S. (2004). Danger! Metaphors at work in economics, geophysiology, and the Internet. *Science, Technology, & Human Values, 29*(2), 242–261.

doi:10.1177/0162243903261947

- Wyatt, S. (2021). Metaphors in critical Internet and digital media studies. *New Media & Society,* 23(2), 406-416. https://doi.org/10.1177/1461444820929324
- Yusuf, K. (2023, May 9). *Introducing watsonx: The future of AI for business*. IBM Blog. https://www.ibm.com/blog/introducing-watsonx-the-future-of-ai-for-business/

Appendix

Table 1

Documents of Analysis

Meta	IBM	OpenAI
 Meta and Microsoft Introduce the Next Generation of Llama (Meta, 2023a). Statement of Support for Meta's Open Approach to Today's AI (Meta, 2023b). Responsible AI: FAIR progress and learnings across socially responsible AI research. (Meta, 2023c). AI Alliance Launches as an International Community of Leading Technology Developers, Researchers, and Adopters Collaborating Together to Advance Open, Safe, Responsible AI (Meta, 2023d). Introducing Meta Llama 3: The most capable openly available LLM to date (Meta, 2024a). Sharing new research, models, and datasets from Meta FAIR (Meta, 2024b). Discover the possibilities with Meta Llama (Meta, n.da). Making safety tools accessible to everyone: Enabling developers, advancing safety, and building an open ecosystem (Meta, n.db). Request for Expressions of Interest (RFEI):Open Innovation AI Research Community (Meta, n.dc). Celebrating 10 years of FAIR: A decade of advancing the state-of-the-art through open research (Pineau, 2023). 	 AI Alliance Launches as an International Community of Leading Technology Developers, Researchers, and Adopters Collaborating Together to Advance Open, Safe, Responsible AI (IBM, 2023). A strong history and commitment to open source (IBM, n.da). Innovating in the open (IBM, n.db). Meet Watsonx (IBM, n.dc). Open Projects (IBM, n.dd). Sharing our expertise in open source (IBM, n.de). Take your skills to the next level with generative AI (IBM, n.df). Ecosystems and open innovation: Co- create or stagnate (Marshall et al., 2023). IBM's approach to open technology (Maximilien et al., 2023). Introducing watsonx: The future of AI for business (Yusuf, 2023). 	 Planning for AGI and beyond (Altman, 2023). Our approach to AI safety (OpenAI, 2023a). Moving AI governance forward (OpenAI, 2023b). OpenAI's Approach to Frontier Risk: An Update for the UK AI Safety Summit. (OpenAI, 2023c). Our approach to data and AI (OpenAI, 2024a). OpenAI safety update: Sharing our practices as part of the AI Seoul Summit (OpenAI, 2024b). OpenAI Board Forms Safety and Security Committee (OpenAI, 2024c). About (OpenAI, n.da). Developing beneficial AGI safely and responsibly (OpenAI, n.db). OpenAI Charter (OpenAI, n.dc).

Note. Titles are listed in order of appearance in the reference list.