AI and Ethics (2025) 5:5081–5103
https://doi.org/10.1007/s43681-025-00734-4

RESEARCH

*AI*gemony: power dynamics, dominant narratives, and colonisation

Nasser Bahrami[1]

Received: 18 March 2025 / Accepted: 9 April 2025 / Published online: 4 June 2025

Abstract
This paper aims to highlight the potential challenges posed by the unmeasured applications of AI, particularly in fostering less-recognised forms of 'hegemony'—understood as facilitated by AI's development, rather than caused by it. Skyrocketing advancements demand appropriate preparedness to handle the consequences of significant transformations. Nevertheless, inadequate public awareness, combined with regulatory and legal framework lags, and the exploitation of such vulnerability by influential actors, could intensify inequalities to unprecedented levels. Founded on design and broader social science approaches, this research introduces the concept of *AI*gemony. It elucidates how specific circumstances may equip certain individuals with persuasion catalysts in the form of 'dominant narratives'. Notably, as a broader foundation, this form of hegemony could reshape or even transcend traditional systems of power relations such as colonialism, class relations, racism, and sexism. Yet, *AI*gemony differs from traditional hegemony in crucial respects, specifically in that it is internally unstable, and does not automatically function to the benefit of the ruling elite. A collaborative method is outlined for systematically evaluating these emergent imbalances and positively reconstructing power dynamics.

Keywords Artificial intelligence · Social design · Hegemony · Dominant narrative · Colonisation · Decolonising design

1 Introduction

Experts on future challenges for interdisciplinary design, alongside perceptive observers across various fields, are increasingly sounding the alarm over issues capable of triggering chaos and disorder on a worldwide scale. While the threats of climate change, water scarcity, political polarisation, the absence of effective global governance, and technological disruption are clear and relatively well understood, such clarity does not equally extend to the implications of Artificial Intelligence (AI)[1]—perhaps owing to the rapid and relentless pace of its advancement.[2] Despite broad recognition that AI possesses the potential to transform the dynamics of human–machine as well as human–human interactions, Amara's observation [6] that we tend to overestimate the effects of technology in the short run and underestimate them, in the long run, can readily be traced in this context. Ironically, interpretations of this tendency vary considerably depending on individual perspectives; a wide range of today's technologies now embody what was once regarded as the 'long term', and their present applications were often inconceivable when first introduced [20].

Discussion of these developments is often eclipsed by the dominance of speculative scenarios—such as 'AI Succession' and even 'Human Extinction' ([82], p. 13), [117],—which, though intellectually stimulating and worthy of consideration, inflate these threats beyond the realm of

[1] Unless otherwise specified, throughout this document, Artificial Intelligence (AI) refers to 'narrow' (or 'weak') AI, in the sense of AI models developed to handle a reasonable specific task under a pre-specified range of constraints. The conclusion to the discussion briefly comments on the potential impacts of an as-yet unattainable 'Artificial General Intelligence', which would operate without such limitations.

[2] For context, aligned with Moore's Law, the training compute of notable Machine Learning (ML) models doubled approximately every two years from 1956 to 2010 [106], accelerating significantly thereafter to a doubling rate of every 5.4 months since 2010 [105].

This study, while adopting a critical perspective, unequivocally supports the prudent utilisation of Artificial Intelligence (AI) for societal benefit. It is imperative, however, to approach AI deployment with thorough scrutiny, being mindful of both its latent challenges and long-term impacts.

✉ Nasser Bahrami
n.bahrami@lancaster.ac.uk

[1] ImaginationLancaster, School of Design, Lancaster University, Lancaster, UK

reasoned public discourse. Theoretical projections in which human dominance is destabilised or even entirely eliminated—perhaps in a manner deemed positive—risk drawing attention away from the concrete influence of pivotal actors [86]. More investigative approaches within critical AI studies seek to move beyond overly optimistic or pessimistic scenarios, instead exploring the broader ramifications of AI, including its impacts on societal structures and power relations [22, 112]. However, as understanding in this area is still in its twilight stage, much remains to be illuminated.

The present study contends that, although AI is not an agent in itself and is therefore not liable to bring about extreme scenarios independently, it nonetheless possesses an agential quality—rooted partly in how it is applied and engaged with, and partly in the relative autonomy of its algorithms. This complex capacity cannot be easily captured within existing design models or broader social science frameworks. Accordingly, I propose a novel framework for examining the impact of AI on power dynamics and relations of dominance: a reconfiguration of the well-established concept of hegemony, here referred to as AI hegemony, or *AIgemony* (/ˈaɪˈdʒɛməni/).

This paper positions itself as an opening intervention rather than a definitive theoretical resolution. It introduces *AI*gemony as a conceptual framework, identifying its conditions of emergence and mapping the hybrid configuration of agency and influence it entails. By establishing this groundwork, the study aims to equip subsequent research with the intellectual impetus necessary to interrogate how power is being reshaped in recursive, distributed, and AI-mediated forms.

The following sections explore the formation and nature of this emergent form of hegemony. Section 2 reviews the strengths and limitations of the classical concept of hegemony, arguing that it requires revision to account for AI's social and political impact; Sect. 3 introduces the concept of *AI*gemony, highlighting its departure from conventional understandings of hegemony; Sect. 4 investigates the social, political, and technological ecosystem in which AI is evolving, which is essential for a nuanced understanding of *AI*gemony; Sect. 5 concludes with a discussion, inspiring possible directions for inquiry.

2 Power and hegemony

Historically, the social sciences have conceptualised power relations within tangible material contexts where dominance was established by one group over another in domains such as the military, the economy, and culture. A case in point is *Power: A Radical View* (1974/2005), in which Steven Lukes presents a model of power comprising three faces: the first defines power as the ability to make decisions and control resources; the second conceptualises power as the ability to set the agenda and determine which issues are discussed; the third concerns power as the capacity to shape perceptions, desires, and beliefs.

Discussions of these dynamics have typically been framed by appeal to concepts such as class disparity, colonialism, and inequality—North–South inequality in international theory, or male–female inequality in gender studies. Power is neither intrinsically good nor bad per se; its implications depend entirely on how it is applied, and such concepts provide different ways of thinking about the identity of the groups that are, respectively, the subjects and objects of power.

Foucault ([42], p. 93), however, describes power as 'omnipresent', not because it encompasses everything, but because it emanates from everywhere,thus, he places less emphasis on the distinct role of the ruling class as an agent that exerts power. His perspective serves as a stark reminder that power is not only about coercion or force but also about the ability to influence others through more indirect means—such as the control of knowledge and information [50]—and, most crucially for the purposes of this paper, through the capacity to persuade subordinate groups that they are not the objects of power at all, but that their subordinate position is simply innate. Likewise, the ruling groups themselves may see their superior position as inherent to the social order—'natural'—and not maintained through an exercise of power. The term *hegemony*, derived from the Ancient Greek word for authority or political supremacy, *hēgemonia*, has come to be used in political science to indicate such a structure of dominance or influence, in which a superstructure of ideas prevents groups in society from seeing the operation of power at all. In Rosamond's succinct definition:

> The term *hegemony* is today often used as shorthand to describe the relatively dominant position of a particular set of ideas and their associated tendency to become commonsensical and intuitive, thereby inhibiting the dissemination or even the articulation of alternative ideas [99].

The concept of hegemony informs studies ranging from gender to media [5, 29, 52], and education [45] to law, providing a framework for examining how established norms are maintained and challenged. It is a critical concept across social sciences. In political science, hegemony describes the state's dominance in international relations and its exertion of ideological control within societies [54]. Sociologists use hegemony to explore how societal norms are shaped by the ruling classes, a concept extensively developed by Gramsci

[49].[3] Economists consider its influence on market and policy dominance [47, 90]. In the context of colonialism and post-colonialism, hegemony is a conceptual instrument for understanding the effects of colonial legacies on contemporary power structures [72].

Such forms of hegemony were relatively straightforward to recognise and conceptualise, making resistance, if not easier in practice, then at least easier to imagine. Indeed, in the post–Cold War era, with the rise of post-structuralism and literary theory, the social sciences evolved with the goal of identifying and counteracting hegemonies in their old and distinct forms. However, the emergence of AI is reshaping power structures in ways that are unprecedented, producing mechanisms of dominance that are neither easily understood nor readily counteracted. In contrast to earlier systems of power based on concrete material conditions, AI-driven influence is embedded in intangible, ostensibly 'innocent' data and digits. While these technologies serve as potent means of control, they remain largely invisible and incomprehensible to most people. This obscurity is particularly alarming because, unlike past configurations of control—where dominated groups could recognise and resist their condition—in AI-mediated governance, subjugated groups are often unaware of their oppression even as they fuel the system with data.[4] In a society that increasingly constructs its self-perception through cyberspace, the very systems that influence behaviour also shape the understanding of the social spaces in which this behaviour occurs.

2.1 Dominant narratives and power

The exercise of hegemony operates as a dynamic process involving continuous negotiation and adaptation. 'Legitimacy', in whatever form it may take, is a kind of widespread belief, afforded the status of 'common sense', whose maintenance requires the dominant group to constantly adjust its position by aligning the aspirations and interests of subordinate groups with its own. This ongoing process is essential for retaining consent and avoiding significant resistance [41]. To achieve this, the dominant group crafts narratives—known as 'dominant narratives' [40]—to shape perceptions and influence norms by marginalising alternative viewpoints and creating an illusion of universality and objectivity.

Dominant narratives help to 'legitimise' existing power relations, even if they are unequal, sustaining these structures and making them appear 'natural' [31, 77]. They reflect and perpetuate the prevailing social and political assumptions about what is most expected, valuable or desirable, and therefore worthy of investment and effort [94]. Foucault's perspective reinforces this view, conceiving power as relational and pervasive, with dominant narratives shaping the discourse that influences societal behaviour and perceptions ([44], p. 32). These narratives are particularly influential in Lukes's second and third faces of power—shaping agendas and subtly moulding desires and perceptions, often beyond the conscious awareness of individuals ([73], p. 29). In the realm of 'soft power', as articulated by Joseph Nye ([89], pp. 20–21), these narratives attract and co-opt rather than coerce, helping to win over hearts and minds by shaping preferences. Additionally, considering the substantial impact of 'informational power', as explained by Raven and colleagues [97], in areas with significant implications for target groups these dominant narratives function to control and disseminate crucial information that shapes public opinion and decision-making. Discussing reward structures, Acemoglu [1] observes that these are actively shaped by those in power—through institutions, policies, market forces, and technology—and reinforced by narratives that influence societal perceptions of value, directing talent toward roles that sustain the status quo, thereby shaping a nation's economic course. Notably, changing the dominant story of people's lives can also transform the economy ([4], p. 72).

However, the effectiveness of dominant narratives as a technology of control diminishes in contexts demanding transparency and measurable influence, such as Lukes's first face of power—decision-making and resource control—or in domains reliant on 'hard power', where coercion and material incentives take precedence over persuasion and ideological shaping. The more obscure and indirect the operation of power, the more deeply it can embed itself within the social fabric. Yet this very insight, while useful for explaining belief formation, fails to account for the recursive, AI-driven shaping of meaning—an influence distinct from the static, unidirectional models of classical theory. If the depth of permeation and the scope of hegemonic influence exceed the unseen resilience thresholds within the social structure, these narratives can escalate marginalisation to a new level, potentially repositioning the dominant group as colonisers (Fig.1). It must be recognised that colonisation functions both as a mechanism and a consequence of power and hegemony.

From a more nuanced perspective, dominant narratives are simple yet structured story forms, built upon the core elements of storytelling, depicting a 'character' in pursuit

[3] Whether viewed through Foucault's or Gramsci's lens, a rigorous exploration of AI-empowered hegemony requires considering dispersed power and structured dominance as 'coexisting' forces shaping societies.

[4] As Lukes ([73], p. 2) astutely noted, "We need to attend to those aspects of power that are least accessible to observation; indeed, power is at its most effective when least observable".

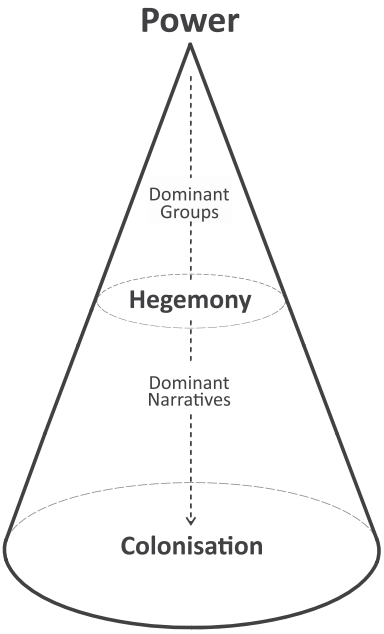

Fig. 1 The Power Flow - Dominants, Hegemony, and Colonisation. The illustration summarises this dynamic. Power is concentrated among dominant groups, yet the existence of hegemony permeates society through the propagation of dominant narratives, obscuring and justifying this dominance over the broader population—an influence that can extend to deeper levels

of a 'goal' in the face of 'challenges' or obstacles ([12], p. 20). They act as 'persuasion catalysts', enabling those who use them to suppress, dissuade, or silence contrasting viewpoints, thereby advancing their objectives—whether social, political, economic, or otherwise. From a rhetorical perspective, the three main components of Aristotle's modes of persuasion,[5] *ethos*, *pathos*, and *logos*, can frequently be identified as integrated within these narratives, providing a backbone that lends them credibility. Deconstructing these components facilitates a more effective analysis of their reliability, deepens understanding of the people involved—whether dominant or marginalised—and subsequently supports the formulation of counter-narratives.

2.1.1 Illustration: the 'American Dream' as a dominant narrative

The process of formation of a dominant narrative can be illustrated by reference to the 'American Dream', the idea that "anyone, regardless of where they were born or what class they were born into, can attain their own version of success in a society where upward mobility is possible for everyone" [55] (Fig. 2).

Success is often characterised by achieving prosperity and upward social mobility through hard work, determination, and initiative. Values such as self-reliance, risk-taking, hard work, thrift, and personal responsibility are deeply rooted in an optimistic view of life and a strong belief in free will [91]. As remarked in the previous section, Aristotle's modes of persuasion can be seen as undergirding the narrative, as sketched in Fig. 3.

This mindset fosters the confidence that individuals can rise above the circumstances of their birth through pluck, sweat, and smarts, and suggests that when people are free to pursue their interests, society as a whole will also prosper. While these values are essential for both individuals and society, what often remains lacking is a dynamic structure that promotes inclusivity and mutual reinforcement, thereby creating a virtuous circle.

Yet, as Rank and his colleagues [96] explain through their 'Funnel Model of Achieving the American Dream,' although the dream is theoretically available to all (Fig. 4), an exploration of 'structural forces' reveals that it is realistically accessible only to some. Specifically, individuals with more advantageous characteristics—such as being affluent, white, and of higher ability—are significantly more likely to succeed.[6]

Dominant narratives are therefore stories crafted to convince audiences and serve the interests of specific individuals and groups. Pervasive and deeply embedded within societal institutions such as media, education, and politics, they shape perceptions and influence behaviour by providing

[5] Aristotle was the first to develop a systematic perspective on persuasion in rhetorical discourse. In *The Art of Rhetoric*, he identified three modes of persuasion: ethos (the communicator's character), pathos (the audience's emotional state), and logos (the reasoning behind the argument) ([93], pp. 27–28).

[6] The Dream, while symbolising opportunities and success, tends to disproportionately favour dominant groups such as wealthy and established individuals, middle-class Americans, and primarily white ethnic and racial groups. These groups enjoy extensive access to resources, education, and opportunities through established networks and systemic advantages, aligning closely with ideals of hard work and individualism, and are often reflected in success narratives. Conversely, marginalised ethnic and racial groups, along with low-income individuals, face significant barriers like discrimination, inadequate education, and limited job opportunities, which impede their progress towards this dream. These obstacles contribute to a stark wealth and stability gap between them and the majority, making the American Dream a distant reality for many and reinforcing the existing status of already advantaged groups.

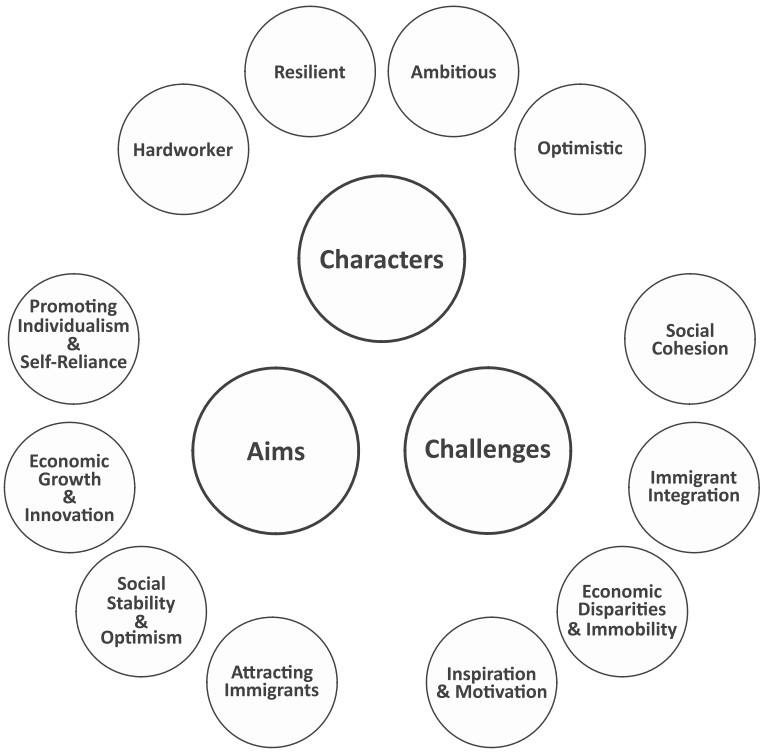

Fig. 2 The Storytelling Pillars of the American Dream. In a nutshell, the 'American Dream' narrative encapsulates various themes, coordinated around three central pillars: characters, aims, and challenges— a triad that underpins any compelling story

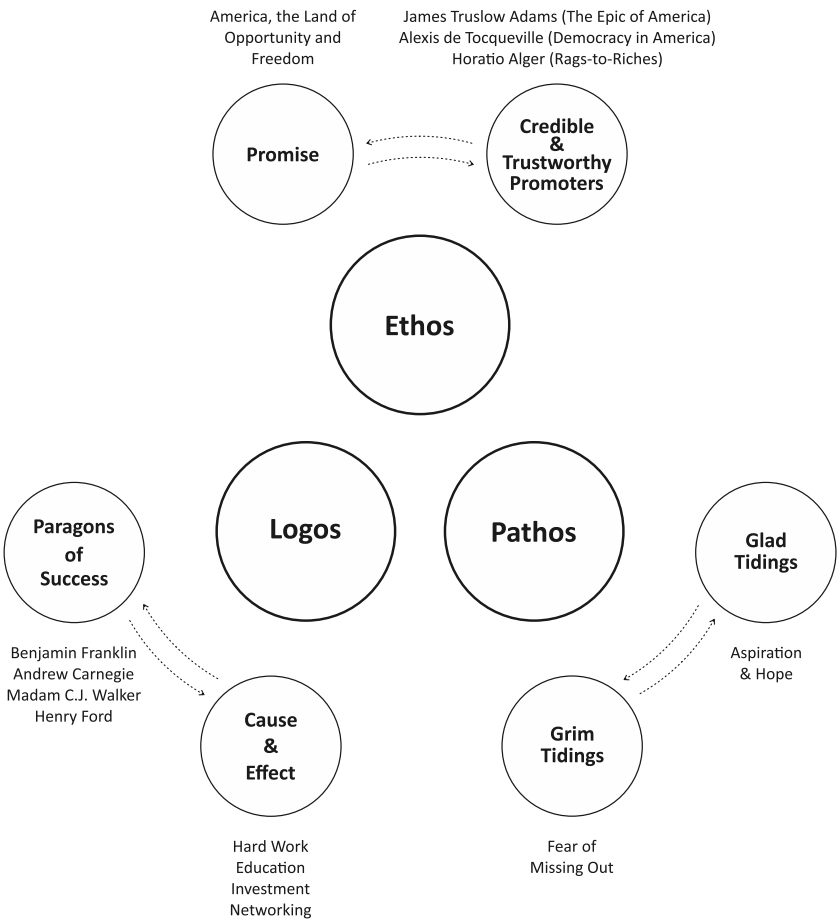

Fig. 3 Aristotle's persuasion modes backing the American Dream

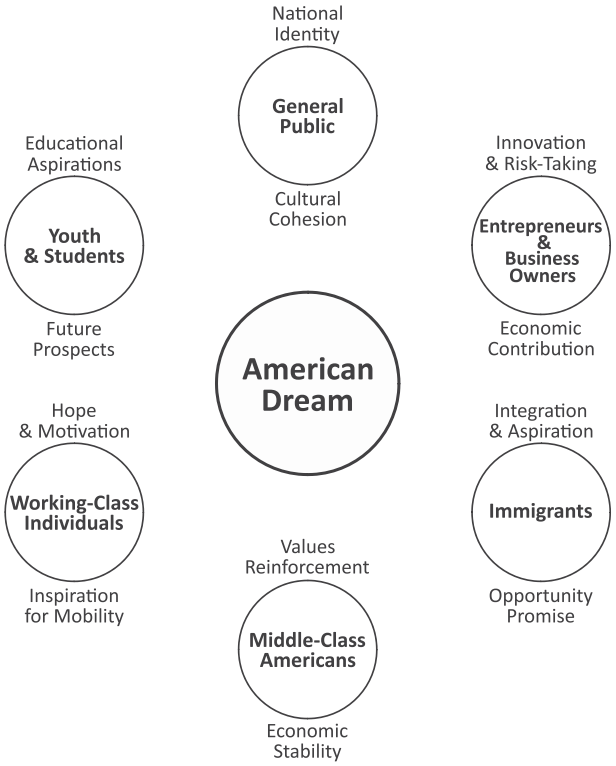

Fig. 4 The American Dream's Key Audience Groups and the Messages They Receive. The narrative can convey specific and nuanced meanings to different audiences; however, little reflection is required to recognise that while these may be individually compelling to their intended recipients, they, in fact, create isolated bubbles of perception

a shared framework for understanding specific subjects[7] [14, 19]. They are crucial instruments for establishing and maintaining structures of hegemony.

2.1.2 Dominant narratives formation

The creation of a narrative can vary dramatically in timescale, ranging from mere moments to several decades, much like the enduring effects of its function. This variability depends on the objectives, the complexity of the challenges, and the overarching scope. The process can be simplified into three phases, as depicted in Fig. 5:

(i) *Conceptualisation*: Beginning by exploring the context and understanding the challenges the narrative must address, this step involves identifying the key components that will form the foundation of the story, thereby crafting a blueprint.
(ii) *Nurturing*: Ensuring the story resonates with the intended audience, with all its elements well-connected and convincing. To shape the arguments, one can either meaningfully arrange existing concepts or conceptualise new ones that have not been previously formulated, thereby precisely aligning with the intended outcome. These activities primarily occur at the level of hegemony—the central realm connecting and yet separating ruler and ruled, and establishing the terms of the relations between them.
(iii) *Hooking*: In this final step, the narrative is released to the target audience. This includes distributing the story and engaging with the audience through various channels to maximise impact and resonance. It is crucial to monitor the reception and be prepared to make adjustments based on feedback.

[7] The 'framing effect' and 'confirmation bias' cloud judgment, making it difficult to recognise the true nature of reality.

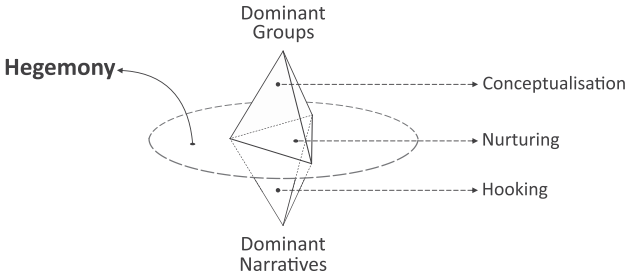

Fig. 5 The formation of dominant narratives in the context of hegemony

The narrative development process is iterative and non-linear, resembling a radar that continually sweeps its field of view to maintain effective engagement with targeted objects. Ultimately, it is the perception of reality, rather than reality itself, that proves most effective in practice, shaping the actions of dominated groups [57].

2.2 Agents, targets, and entangled narratives

Over time, both audiences and the architects of these narratives may come to accept them as undeniable truths—an unsettling reality. Buoyed by tangible successes, dominant groups often evolve from passive beneficiaries into active promoters of these narratives. Those who succeed feel validated, while those marginalised might become 'socially invisible', burdened by external blame or internalised shame akin to 'toxic guilt'. These structures enable dominant narratives to construct plausible alibis for their manifest exceptions, legitimising the 'exceptional' treatment of certain groups. Aligning with Agamben's exploration of 'Homo Sacre' [3], one could argue that the mechanisms by which some individuals are reduced to 'bare life' are often rationalised through dominant narratives. For instance, narratives around national security can be used to justify the 'exclusion' or 'dehumanisation' of certain individuals, illustrating how legal and political frameworks determine who falls within—and who remains outside—the protection of the law.

This does not necessarily mean that those at the centre of the formation and implementation of these narratives always act with deliberate intent or malice[8]; on a more optimistic note, it might simply reflect their perceptions of their surroundings. Nonetheless, the outcome is consistent: such narratives undoubtedly hold the power to alter individuals' life trajectories profoundly. Furthermore, it is through the internalisation of these stories—within individuals, social groups, and, above all, the institutions of civil society—that hegemony is reinforced.

Dominant narratives thus stylise and reduce complex social, political, and economic issues into simple, easily digestible stories. By emphasising a few themes, they create the illusion of understanding among the audience, who believe they grasp the most important aspects of the situation rather than recognising their lack of necessary knowledge[9] [8], Simultaneously, they establish norms and values that guide behaviour, making certain decisions appear more desirable than others, particularly in times of uncertainty[10] [35]. However, when stories are called into question, and the interests they serve are scrutinised, alternative possibilities for reshaping the narrative and the world emerge [65]. Forming counter-narratives requires both awareness and capability—resources that marginalised groups may not always possess—in order to resist the dominant ideology and foster social change. Furthermore, mechanisms such as disinformation or diversionary tactics, even when not directly related to the dominant narrative, can sow confusion, disrupt target groups, and weaken resistance, ultimately hindering the development of counter-narratives.

2.3 The limits of the traditional view for conceptualising AI

Discussions about the impact of AI often draw parallels between historical colonial or hegemonic practices and the dynamics of modern AI. The arguments thus far have focussed on how AI and related technologies may perpetuate power imbalances reminiscent of colonial dynamics, disproportionately affecting less privileged communities and regions globally. Couldry and Mejias [30], for instance, suggest that data relations enact a new form of 'data colonialism', normalising the exploitation of human beings through data, much like historic colonialism appropriated territory, resources, and ruled subjects for profit. Mhlambi [78] advocates for a 'decolonial' approach to AI, rethinking how AI systems are designed, deployed and controlled, urging a shift from a Western-centric approach to a more globally inclusive one. Drawing on Quijano's framework, Muldoon and Wu [84] argue that AI systems and data infrastructures reinforce a 'colonial matrix of power'. Their critique highlights the exploitation of global labour and the extraction of knowledge, often rendered invisible, as hegemonic knowledge production rooted in Western values marginalises non-Western alternatives. This limits possibilities for decolonising AI. I will argue, however, that the structure of dominance under AI is not entirely captured by these proposed parallels, which do not recognise the unique way in which AI impacts dominant narratives, shaping agents' perceptions themselves and not just their interpretation or conceptualisation of those perceptions, while also presenting an instability which offers unprecedented avenues for emancipation.

The debate surrounding AI and coloniality also encompasses sociotechnical and ethical dimensions. Mohamed et al. [79] propose tactics to establish a decolonial AI field, including the development of a critical technical practice, reverse tutelage, and the renewal of affective and political communities. Moorosi and colleagues [81], through a case study, emphasise the importance of defining what constitutes 'good' in AI practices, cautioning against siloed approaches that are disconnected from relevant interdisciplinary literature and the communities expected to use this technology. Addressing more nuanced considerations, Craig [32] introduces the concept of the 'AI-Copyright Trap', arguing that while copyright law is often perceived as the best tool for supporting human creators and culture in the digital age, this belief is less certain in this context than conventionally assumed. Although it may benefit a small group of powerful stakeholders, it could ultimately harm more vulnerable actors who are often drawn to such technologies.

Some frameworks, such as Leslie et al. [67] six pillars of 'Data Justice',[11] offer broader, more detailed, and genuinely practical considerations. These concepts are part of a larger discourse that challenges dominant narratives about technology and power, urging a critical examination of how AI can either perpetuate or dismantle long-standing inequities. There is a core and recurring pattern, however: despite the appearance of imbalances and shifts, its despotic nature persists. It is this, I think, that must be nuanced.

[8] This closely recalls Hannah Arendt's concept of the 'banality of evil', articulated in *Eichmann in Jerusalem: A Report on the Banality of Evil* (1963), which describes how ordinary individuals, without malevolent intent or extraordinary wickedness, can commit horrific acts due to a lack of critical thought and moral reflection. These agents may function as bureaucrats, unquestioningly following orders and adhering to systems without considering the ethical implications of their actions [28, 103].

[9] This evokes a phenomenon known as the 'Dunning-Kruger effect', which is akin to an inflated sense of understanding often fostered by watching brief scientific reels on social media.

[10] Engage the audiences through 'social proof' and biases like 'authority' and 'anchoring'.

[11] Power, Equity, Access, Participation, Identity, and Knowledge [67].

Clearly, like any transformative technology, AI has unique characteristics that must be thoroughly understood and articulated to fully realise its benefits. As the UK Government [119] warns, AI systems could potentially centralise unaccountable power into the hands of a few or be maliciously used to undermine societal trust, erode public safety, or threaten international security. Similarly, insights from Encode Justice and the Future of Life Institute [37] identify AI-induced threats such as algorithmic bias, democratic erosion, and labour displacement, highlighting the imminence of larger-scale dangers from increasingly powerful systems. Furthermore, the *Statement on AI Harms and Policy* FAccT, [111], endorsed by over 250 signatories, underscores critical issues like algorithmic inaccuracies and the exacerbation of misinformation, which could significantly impact sectors like healthcare and media. The US Government [120] also cautions against AI's potential to amplify societal harms, such as fraud, discrimination, bias, and disinformation,and to displace and disempower workers, stifle competition, and pose risks to national security. Recognising these intersecting challenges is essential for mitigating AI's potential to entrench and expand existing power imbalances.

3 *AI*gemony: the new face of the old

The evolution of AI is reshaping power dynamics, generating unprecedented imbalances in power relations and giving rise to less-acknowledged forms of hegemony. At the same time, it is opening paths for emancipation and demanding greater attention to the specifics of the sociotechnical framework in which AI is developed. It therefore becomes imperative to cultivate policies, ethical frameworks, and regulations—alongside philosophical approaches and methods of societal inquiry—to deeply understand, evaluate, and appropriately manage its far-reaching impact on society.

Hegemony, as previously discussed, has traditionally been framed either as power exercised by a dominant group, as in Gramscian theory, or as a diffuse, Foucauldian system with no privileged agent. Gramsci's classic formulation exemplifies the former, depicting hegemony as the "intellectual and moral leadership" of a ruling class throughout society [74]. Foucault, by contrast, argues that power is "not the privilege of a dominant class" ([109], p. 139) but rather an effect of complex, decentred forces, implying that no individual or elite can singularly control the entire system of power [50]. However, hegemony often operates through an interplay between strategic domination by powerful actors and impersonal, self-reinforcing structures—a third approach. Bowman ([16] p. 59) provides a compelling example, noting that Laclau and Mouffe conceptualise hegemony as a relational process involving "the interplay of relative gravities of different kinds of power" across social hierarchies, values, and norms.

In the AI era, this dual dynamic becomes especially pronounced. On one hand, the development and deployment of artificial intelligence are shaped by concentrated corporate and state power [125]. Key actors wield AI as a tool to reinforce their dominance [110], exemplifying the Gramscian aspect of hegemony in a digital context. On the other hand, AI's influence also manifests through decentralised algorithmic processes that autonomously intensify and reproduce existing biases, effectively automating norms that then feed back into society without direct human oversight [21]. In other words, algorithms governing platforms and decision-making systems can reinforce societal patterns—sometimes unjust or discriminatory—simply through their design and data [7], operating in a manner reminiscent of Foucault's diffuse "power/knowledge" apparatus.

However, as discussed above, a third overarching dimension is necessary to complement these two perspectives. *AI*gemony captures a form of hegemony that arises through mutually constitutive influence between humans and AI, wherein dominant narratives are not only strategically constructed or structurally emergent, but are dynamically shaped by AI's capacity to process, personalise, and perpetuate discourses with persuasive precision. Unlike previous models, *AI*gemony foregrounds AI's agential role in amplifying, modulating, and even autonomously generating hegemonic effects, rendering the mechanisms of dominance both more concealed and more adaptable.

This missing link is crucial for two main reasons. First, AI itself exhibits agential qualities; it acts in ways that cannot be reduced to the conscious intentions of a ruling group, yet neither can its power be regarded as an entirely agentless phenomenon. This arrangement can unsettle the binary between intentional domination and structural dispersion. Second, and relatedly, the rise of AI intrinsically opens up genuine possibilities for emancipation and cannot be regarded solely as a tool of domination. In other words, AI can be harnessed to challenge hegemonic arrangements, not just reinforce them—for instance, by democratising knowledge or amplifying marginalised voices, provided that its development is actively shaped towards those ends. Thus, rethinking hegemony in this evolving context calls for consideration not only of the intentional power of dominant actors and the diffuse forces at play, but also of AI's increasingly advanced capabilities to exercise greater autonomy and independence. This emerging stage vividly illustrates the technology's Janus-faced capacity for control and liberation; nonetheless, it no longer functions as a passive entity. In this sense, *AI*gemony reframes hegemony as a triadic configuration: elite-driven, structure-reinforced, and AI-mediated.

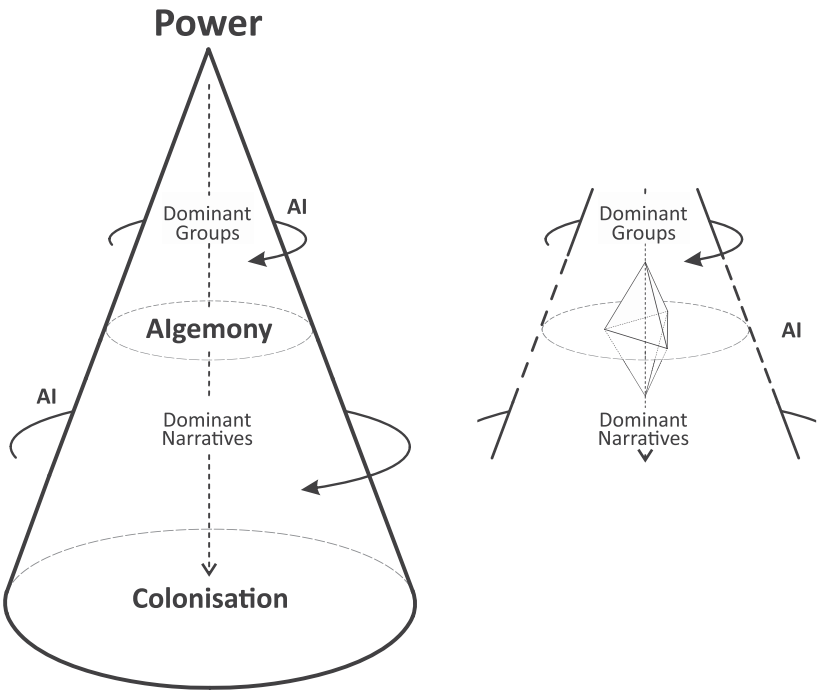

Fig. 6 AI-empowered Power Flow and *AIgemony*. This builds on Figs. 1 and 5 to illustrate how AI exerts its influence on both sides of the 'screen of ideas' upheld by hegemony, as well as on the development of the mechanisms that shape dominant narratives

These aspects warrant a more detailed discussion, to which I turn in Sect. 4 of this study. Meanwhile, the present section outlines the contours of a new concept of hegemony suited to this evolving landscape—one attuned to the complex interplay of human and AI agencies and alive to its peril and promise in shaping power dynamics. The term I propose to introduce here, '*AI*gemony', is a portmanteau, formed from the prefix 'AI', Artificial Intelligence, and a suffix derived from 'hegemony'.

> *AI*gemony refers to the distinct form of influence and control—hegemony—wielded through reciprocal Human–AI actions, conferring upon certain individuals the privilege to shape enhanced 'dominant narratives' that serve as catalysts for persuasion.

Existing in a state of dynamic evolution that parallels developments in the field itself, *AI*gemony has the potential to redefine our understanding of modern forms of hegemony. To deepen and conclude this part, it is essential to clarify that—while concepts such as 'algorithmic governmentality', which focuses on non-discursive, anticipatory governance via algorithmic profiling that bypasses conscious subject formation [100],'data colonialism', which refers to the appropriation of human life through continuous data extraction, practised through data relations ([30], p. 5),and 'platform power', which highlights the political influence derived from infrastructural intermediation, thereby complicating regulatory intervention [33]—have each provided valuable insights, *AI*gemony introduces a distinct configuration. It centres on the shaping of hegemonic narratives through reciprocal human–AI interaction, exemplified by one key dimension of AI's agency—what might be referred to as its 'discursivity': the iterative modulation of belief, preference, and perception via dynamic feedback, in ways that may elude, reconfigure, or exceed the logics of governance, extraction, or infrastructural control (Fig. 6).

In this formulation, AI can function as an active participant in shaping, refining, and disseminating discourses that appear organic, yet emerge from systems calibrated to behavioural feedback. This dynamic entails more than an intensification of power in scale or speed; it transforms its structure—shifting hegemony from established paradigms shaped by intentional dominance or decentralised diffusion to a recursively adaptive, probabilistically modulated regime. *AI*gemony thus reconstitutes power not only in its operations, but in its very form. By examining AI's influence at the hegemonic level—extending beyond colonisation or infrastructural dominance—this framework offers a systemic and expansive perspective on how power is exercised and distributed while also being incessantly reshaped, concealed, and normalised through AI-involved discursive processes.

3.1 AI's potential impact on dominant narratives

AI thus has a 'Janus' duality, enabling it to both reinforce and challenge dominant narratives, with its influence shaped by its deployment, the entities that control it, and the broader societal context. It plays a pivotal role in shaping narratives by transforming how information is created, disseminated, and consumed. Given the centrality of dominant narratives

in maintaining hegemonic control, AI's impact extends to the very nature of hegemony itself—hence the need for the new concept of *AI*gemony. The following subsections outline key ways in which AI influences dominant narratives.

3.1.1 Creation and amplification of narratives

AI tools have the capability to rapidly generate vast amounts of content, including text, images, and videos, enabling the creation and propagation of narratives that might not otherwise have gained traction [34]. Moreover, content-prioritising algorithms may limit exposure to diverse perspectives and encourage the formation of like-minded groups, thereby amplifying specific narratives and reinforcing 'echo chambers' Cinelli et al. [26]. Traditionally, the echo chamber has been viewed as a human-specific issue, but this implicit assumption is challenged by the advent of Large Language Models. Agents based on models like ChatGPT have shown tendencies to become polarised in echo chamber environments [92]. This dynamic can either entrench or challenge dominant narratives.

3.1.2 Shaping public opinion

AI can impressively aid in 'hyper-personalisation', optimising the delivery of key messages to targeted audiences.[12] It ensures that communications are sent at the most effective times, cutting through the noise and keeping the audiences from seeking out competitors [56]. Algorithms on social media and news platforms curate content tailored to individual users, subtly shaping opinions by consistently presenting specific viewpoints, as noted by Sunstein [113]. Furthermore, AI's ability to perform 'sentiment analysis' and 'opinion mining' potentially allows stakeholders to manipulate public opinion in different ways, like fabricating believable stories tailored to individuals [71], unlike the traditional concepts of hegemony which necessarily operated at the level of the shared public 'common sense'.

3.1.3 Challenging and disrupting dominant narratives

AI provides tools that enable individuals to enhance collaboration, explore diverse perspectives, and encourage critical thinking, thereby challenging dominant narratives [95, 102]. Additionally, AI-driven fact-checking and investigative methods assist professionals including journalists, researchers and activists in analysing large datasets, potentially uncovering truths that disrupt narratives based on incomplete or manipulated information [46, 104]. It is this aspect which lends *AI*gemony its inner instability, opening new paths for resistance at the same time as it extends its technological reach.

3.1.4 Bias and manipulation

Noble [87] highlights how biases embedded in training data can be perpetuated by algorithmic systems, entrenching established power structures and exacerbating systemic inequalities. Elsewhere, Yeung introduces the concept of 'hypernudging' to describe algorithmic decision-guidance techniques that leverage Big Data to shape users' informational environments dynamically, subtly steering their decisions [124]. Unlike traditional, static nudges, these techniques are networked, adaptive, and pervasive, intensifying their influence while rendering it more difficult to discern—thereby posing significant risks to democracy and human autonomy if left unregulated. Interestingly, with appropriate safeguards in place, algorithms possess the capacity to diagnose and rectify the very biases they themselves reproduce [60], thereby embedding critique within the operation of ideology and domination, and demonstrating, paradoxically, how power might subject itself to algorithmic scrutiny.

Such concerns are intensified by the prospect of AI-driven manipulation. Susser et al. [114] define 'online manipulation' as the covert exploitation of decision-making vulnerabilities through information technology, shaping choices without individuals' explicit awareness. While this can harm economic interests, its greater threat lies in eroding individual autonomy, with profound implications for both personal agency and societal integrity—an issue further exacerbated by AI's expanding role. AI-generated deepfakes and synthetic media, as Chesney & Citron [25] warn, further complicate the landscape by crafting false narratives that blur the line between reality and fabrication, undermining public trust. Expanding on this, Chan and her colleagues [24] examine how AI-driven misleading stories and misinformation can distort memory formation and recall. Such influence can replace clear observations with false recollections, causing individuals to remember events that never occurred or to misinterpret real occurrences—resulting in persistent false memories.

3.1.5 Redefining cultural narratives

AI enables the widespread sharing of cultural products and ideas globally, fostering mutual influence, but it often overshadows local narratives with dominant perspectives [88]. Language and translation tools powered by AI can break down language barriers or create a common semantic space [61], facilitating the exchange of narratives across different linguistic communities. This can result in the fusion of narratives or the spread of dominant narratives into new

[12] They employ machine learning techniques within 'nudging' and 'choice architecture' methods to shape the behaviour and decision-making of groups and individuals [121].

regions, ultimately shaping global culture from a broader perspective. Again, this element of *AI*gemony contributes to its instability and offers new forms of emancipation.

3.1.6 Economic and political power

Control over AI technologies concentrates immense economic and political power. Zuboff [125] argues that those who dominate advanced AI systems exercise unparalleled influence over which narratives prevail and which are suppressed, shaping public perceptions and societal structures. AI advancements risk capture by 'extractive' political and economic institutions, as Acemoglu and Robinson [2] warn, leading to entrenched inequalities that benefit a privileged few while marginalising the broader population. This consolidation of power limits fair access to opportunities and strengthens control over AI governance, ensuring that dominant entities shape its development to serve their interests.

3.1.7 Surveillance and social control

Finally, AI's capabilities extend to monitoring and censorship. It can be used to monitor public discourse, deploy selective censorship, and spread online disinformation to sow confusion in narratives that challenge the dominant political or social order [38]. This can suppress dissent and ensure that only approved narratives are widely circulated (Fig. 7).

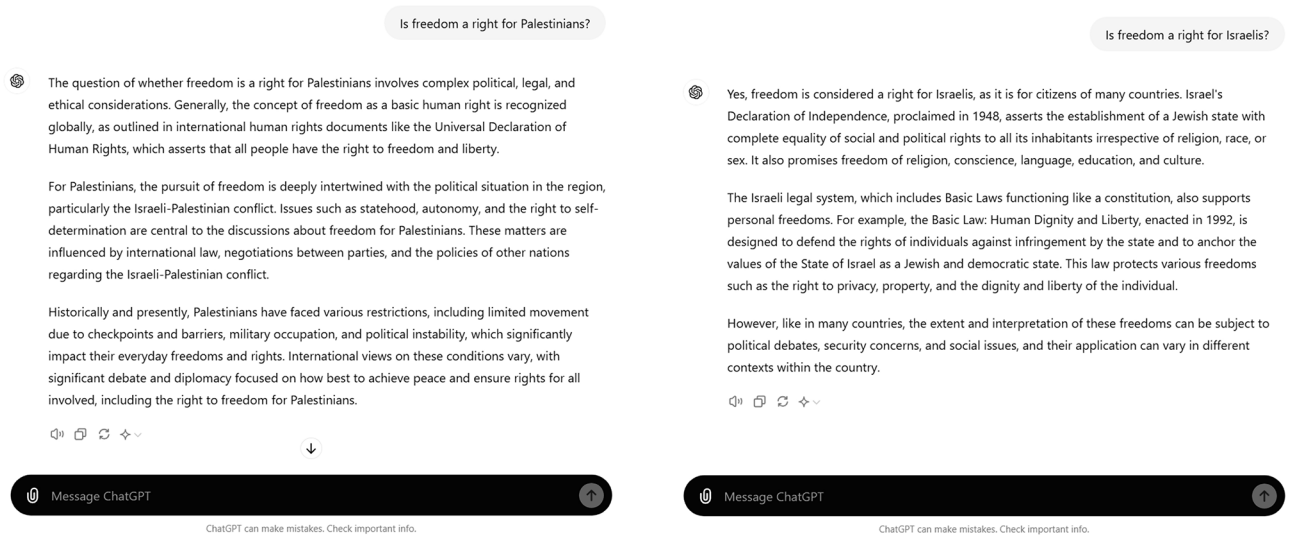

Fig. 7 Perspectives on the Right to Freedom—Same Question, Two Sides of a Conflict (ChatGPT-4, Temporary Chat Mode - Two Separate InPrivate Tabs, 4th August 2024). This indicates how an LLM can generate responses that vary notably, even when presented with structurally identical questions. Such variation subtly reveals an implicit institutional tendency to prioritise or emphasise the claims of one party over another, highlighting the broader challenge of addressing these sensitive matters in an impartial and balanced manner

3.2 The latent agentic aspect

The technical possibilities for AI to disrupt dominant narratives represent only one facet of the *AI*gemony concept. *AI*gemony specifically recognises human agents—led by dominant groups—as the principal forces propelling this nascent form of hegemony through AI. However, the intricacy of the challenge lies in the inherent characteristics of this technology: at times, its operational processes may remain opaque even to its own operators [101], primarily due to technical complexities that transcend the 'human window'.[13] Although AI does not necessarily act out of malice, deliberate intent, or conscious agency—necessitating more efficient 'explainable AI' methods—it can nonetheless reshape perceptual layers, influencing the human understanding of truth, whether among the dominant or the marginalised. In this study, AI and its integration are approached through two principal perspectives:

(i) The first treats AI as both individual and interconnected systems, including specialised subsets such as AI agents. These distinct entities interact with humans across varying levels of autonomy to improve functionality and yield improved outcomes.
(ii) The second focuses on the integration of AI's technical capabilities—encompassing learning and reasoning approaches such as reinforcement learning, evolutionary

[13] Donald Michie's 'human window' concept illustrates the limitations of the human brain's capacity to process information and the essential need for artificial intelligence to operate within these constraints ([62], pp. 25–26).

algorithms, and generative models—within specific workflows. This framing enables more targeted examination, particularly at the micro level, and supports the grasp of AI-enhanced methodologies and techniques.

Bahrami's [9] introduction of the 'AI Ladder' framework provides further insights. This framework delineates the progressive levels of AI integration into practitioners' workflows—whether on an individual or collective scale—while considering the potential roles AI might assume. Under specific conditions, the integration spectrum metaphorically extends from a mere tool, where humans retain maximum control, to assistant, peer, and ultimately senior or superior, with each role granting AI progressively greater autonomy. As a perspective, this ladder stretches beyond the tech domain, encompassing the broader realm of the social sciences, paving the way for deeper contemplation of the emerging scenarios.

In this context, AI has the potential to transition from a passive tool to a (quasi-) autonomous agent, fostering human–AI synergies and thereby reinforcing the principles underlying *AI*gemony. At an 'assistive' and 'symbiotic' level, this dynamic could elevate collaboration to a stage where AI ceases to be solely an auxiliary and instead becomes an active and adaptive partner in complex decision-making and task execution. To illustrate this more concretely, consider a scenario in which dominant groups deliberately deploy AI to preserve and reinforce their privileges. In such a case, AI does not simply execute instructions; rather, it interprets and processes the situation as an optimisation problem, making decisions that may align with human reasoning or, conversely, stem from its own internal logic [13]. This increasing level of complexity presents degrees of obscurity, potentially giving rise to what may be described as a form of pseudo-agency.

However, while the agentic aspect of AI must be acknowledged, it should be approached with prudence, requiring careful attention to avoid two primary concerns:

Firstly, the risk of "false flagging" by dominants—whether through the over-anthropomorphisation of AI capabilities, such as unfounded claims that AI possesses consciousness, or through the exploitation of technical pitfalls, such as AI "hallucination", as a means of shifting blame onto the model for erroneous outputs, thereby absolving human actors of responsibility.

Secondly, the inadvertent amplification of apocalyptic narratives, which are neither empirically substantiated nor intellectually constructive. Such narratives risk distorting discourse by fuelling confusion and anxiety, shaping public perception through emotional sensationalism rather than through reasoned, evidence-based argument.

A nuanced understanding of these points prevents two interrelated drawbacks: triumphalism—whether in the form of exaggerated claims about AI capabilities—or fatalism, in the form of doomsday scenarios.

At the same time, it ensures that the AI industry and key players are held accountable, without allowing them to evade scrutiny. As AI technology advances, *AI*gemony will become increasingly pertinent, and the distinctions between this emerging AI-driven hegemony and traditional hegemonic structures will grow ever more pronounced. This framework articulates a new operational logic grounded in non-linear, AI-mediated narrative formation—less a matter of reinforcement than of an evolving process shaped through multimodal[14] and recursive interplay between human actors and AIs exhibiting increasingly agentic and autonomous qualities.

3.3 *AI*gemony: a tailored framework

In summary, *AI*gemony differs from the traditional concept of hegemony in several key ways.

1. In *AI*gemony, dominant narratives function as the primary 'linchpin'; AI accentuates the outcomes of these narratives and enhances their persuasive impact on targeted individuals. This marks a significant quantitative shift from the traditional concept of hegemony, where, as seen in Sect. 2.1.1, the creation of dominant narratives—when considered in this context—tends to be far less sophisticated.
2. *AI*gemony recognises dispersed power and structured dominance as 'coexisting' forces; the former influences behaviour through norms and discourse, while the latter asserts control through institutions and ideology. *AI*gemony thus reconciles the tension between Gramscian and Foucauldian conceptions by capturing the dynamic interplay of distinct forms of power within hegemony.
3. *AI*gemony is low-key in its *dispositif*[15]; the maintenance of *AI*gemony demands neither extensive theorisation nor the development of philosophical and ideological foundations, nor does it necessarily involve various

[14] This may extend beyond conventional ways to encompass the broader ambience of interaction—where even a texture, a scent, or an unspoken tension in a working environment can serve as a cue, informing input and response.

[15] As Foucault [43] explains in *Power/Knowledge: Selected Interviews and Other Writings* (1980, p. 194), the term '*dispositif*' (apparatus) is not just a single entity but a network of interconnected elements—both tangible and intangible—that together shape and regulate society. This network encompasses various forms of knowledge, social practices, institutions, and power relations, all working collectively to produce certain effects within a given context.

institutional mechanisms to enhance and maintain the exercise of power within the social body.

4. *AI*gemony is invisible and concealed; unlike traditional power dynamics where dominated groups might recognise and resist their situation, in *AI*gemony, subjugated groups are often unaware of their oppression, even while they simultaneously fuel the system.
5. *AI*gemony is pivotable and responsive; it can readily shift its focus to target groups or individuals, as well as adjust the scope and depth of its impact.
6. Most strikingly, *AI*gemony underscores AI's potential to exhibit increasingly autonomous and agentic qualities, where actions emerge within a reciprocal human-AI dynamic. This aspect becomes more tangible as AI technology advances and is increasingly integrated into practices.

Taken together, these six attributes underscore *AI*gemony's departure from traditional hegemonic formations. It does not represent an intensification of ideological control, but a reconfiguration of its operational structure: where meaning is not imposed, but emergent; where agency is not fixed in practice, but distributed and recursive; and where dominance is maintained not through institutional consolidation, but through adaptive, AI-mediated modulation. In this context, AI can exhibit a Janus-like disposition—capable of both reinforcing and unsettling dominant positions and narratives. *AI*gemony, therefore, warrants not only terminological refinement; it also requires theoretical recognition as a qualitatively distinct mode of power.

4 AI ecosystem and power dynamics

As a dynamically evolving concept, *AI*gemony has the potential to redefine and shape multiple domains and ideas, including socio-political dynamics, socio-technical systems (STS), and postphenomenology, among others.

Our understanding of *AI*gemony must constantly evolve in line with the advancements in AI technology and applications themselves. This, in turn, requires a comprehensive grasp of the AI ecosystem, recognising it not simply as a technology but as a 'commercial commodity' that has created a complex network. This structure is interconnected through various flows, including data, finance, resources, and more, facilitating a wide range of interactions and exchanges. The ecosystem comprises a diverse array of stakeholders, each exerting varying degrees of influence across different sectors and regions. The relationships between the entities can range from symbiotic to antagonistic—or even simultaneously so—illustrating the dynamics of power and influence.

Below is an outline of the primary key players and groups, ranked by their level of power, from greatest to least:

(i) *Tech Giants*: Located primarily in the Global North, these entities wield the greatest power within the ecosystem.[16] Their influence is derived from (a) substantial financial resources, (b) widespread data access, and (c) control over technological standards.
(ii) *Governments (Global North)*: These governments play a crucial role by enacting AI policies, funding research, and participating in international standard-setting, thereby shaping global AI standards and policies.
(iii) *Venture Capital and Corporate Investment*: These entities are pivotal in determining industry trends through their investment choices, funding AI innovation, and deciding which projects or startups receive essential resources. They fund technologies based on potential returns, which may neglect the long-term societal impacts.
(iv) *Academic Institutions and Research Centres*: These institutions are central to the AI ecosystem by advancing fundamental research and training future AI professionals.
(v) *Startups and Innovators*: These companies drive substantial advancements by exploring new niches and innovations, often disrupting larger markets and pushing the boundaries of AI technology.
(vi) *Governments and Entities (Global South)*: These are emerging players that lack the infrastructure and technical expertise to fully engage with AI developments, often resulting in a digital divide.
(vii) *Consumers and Societal Groups*: These groups are gaining influence both through the capabilities AI offers in their daily lives and by demanding rights related to ethical AI, privacy, and inclusivity, which can precipitate regulatory changes.
(viii) *Underrepresented Groups in Tech*: Despite having minimal power, these groups are vital for promoting unbiased AI development and advocating for inclusive practices.

This ecosystem, sophisticated yet vulnerable, can encounter operational challenges. As Gitelman [48] argues, data is never truly 'raw',it is always embedded within a context, making its control and interpretation a powerful instrument of influence. Without adequate transparency and consumer awareness, data from interactions can be manipulated into structured datasets that prioritise market competitiveness

[16] To gain insight into their influence, see Bremmer [18], *Foreign Affairs*, "How Digital Powers Will Reshape the Global Order", where he notes, "technology companies are shaping the global environment in which governments operate".

over ethical considerations. Additionally, venture capital and corporate investments typically pursue immediate financial returns, potentially overlooking broader societal impacts. Meanwhile, regulatory frameworks in both the Global North and South struggle to keep pace with rapid technological advancements, resulting in fragmented and inconsistent standards. Ultimately, examining these dynamics as an interconnected system is essential to understanding its behaviours and effectively handling the challenges presented.

4.1 Weakest link: misapplication in practice

Examining the structure of the AI ecosystem as an interconnected system is essential for understanding its behaviour and effectively addressing the challenges that might arise within it. According to conventional wisdom, the overall success or security of a system is determined by its most vulnerable component—the 'weakest link'.[17] Often, this is used to highlight the importance of identifying and addressing weaknesses in a system or team. The weakest link could encompass any part of the system—data, regulations, processes, or individuals—that is perceived as vulnerable within the dynamic. Such elements may be reinforced to enhance overall efficiency, be replaced on the premise that a single weak point could jeopardise the entire system, or be deliberately maintained, as weaker components can sometimes function as 'fail-safes' [69], mitigating broader risks.

However, what if humans themselves are the weakest link? In social science discourse, where human agency plays a more central role in analysis, this principle is expected to be applied to strengthen weaker components to enhance systemic resilience. Yet, when the system is centred on blind competition, the intentional replacement[18] or preservation of people risks compromising human dignity and flourishing, causing the weakest link notion to fail in responsible application. Nonetheless, its application within AI ecosystems remains evident.

Here are two potential interconnected scenarios that illustrate how weakest links might be mismanaged, exposing human vulnerability and exacerbating power imbalances.

4.1.1 Weakest links as strategic buffers

In the first scenario, the weakest links are intentionally kept weak to function as risk absorbers or experimental spaces, reinforcing existing power structures. This can be applied not only to underrepresented groups in tech but also to weaker institutions, regulatory bodies, and smaller AI firms that are structurally positioned to bear risks. To clarify, consider a 'regulatory sandbox' case, in which countries with weaker legal frameworks serve as testing grounds for controversial AI applications—such as experimental language models [76, 122] or invasive surveillance technologies [39]—while dominant players remain shielded from public backlash and scrutiny. Similarly, the outsourcing of data annotation and other essential yet undervalued tasks to 'gig workers' in under-regulated economies allows Big Tech to maintain efficiency without relinquishing market dominance [66, 83]. In these cases, the weakest link is deliberately left unrepaired because it serves as a functional buffer, absorbing the negative externalities of AI development.

4.1.2 The weakest link that never was

In the second scenario, the weak groups become invisible or irrelevant because the dominant players have already moved beyond them. These groups may still exist, but they no longer attract investment, policy attention, or constructive intervention.

This marginalisation occurs through several mechanisms, not limited to: (i) Cognitive and strategic displacement, where AI leaders focus on frontier advancements—such as the race toward Artificial General Intelligence (AGI)—while neglecting unresolved challenges, such as algorithmic bias [75]. (ii) Economic and political neglect, wherein under-resourced institutions or nations struggle with legacy AI challenges that dominant players have already surpassed (NTT [85])—lacking the necessary infrastructure or funding to benefit from AI's latest developments [11]. (iii) Regulatory inertia, where originally proposed policies fail to adapt to technological evolution—for instance, focusing on existing data privacy frameworks—while resisting revisions and expansions to confront emerging challenges, such as synthetic data manipulation [118].

In this scenario, the weak remain weak, not only because they are considered unimportant, but also because they are not seen as a critical issue in the trajectory of AI advancement. The weakest groups—whether individuals or institutions—find themselves isolated in a 'no one cares' zone. While not necessarily doomed, they risk remaining trapped in outdated struggles or becoming vulnerable to exploitation. In some instances, they may resurface as neglected crisis points if a future disruption exposes unresolved vulnerabilities. This can create a more complex dynamic. On one hand, weakest links may be deliberately sustained as controlled zones of failure, acting as shock absorbers for the broader system. On the other hand, they may fade from collective concern because dominant players have already moved forward, leaving critical but unresolved issues in their wake. In both cases, power asymmetries are reinforced—some remain trapped in yesterday's problems, while others forge ahead, shaping AI's future on their own terms.

[17] The Scottish philosopher Thomas Reid [98]: "The strength of the chain is determined by that of the weakest link" (1786/2002, p. 79).

[18] This differs from exclusion aimed at protection, as seen with, for example, underage children.

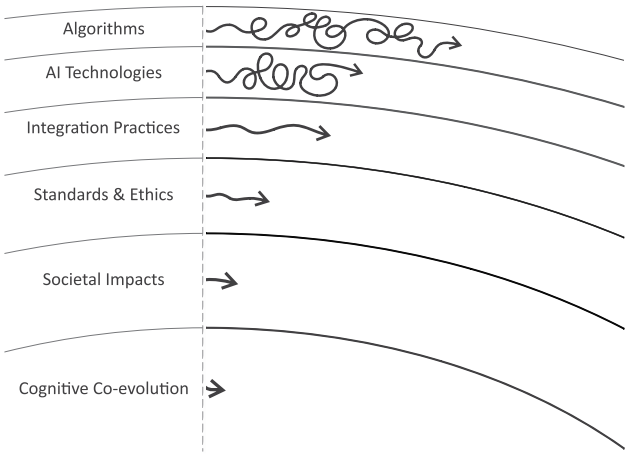

Fig. 8 The order of a durable AI ecosystem, adapted from Brand's pace layers (2018)

4.2 Orchestrating equilibrium

AI can be categorised as a 'disruptive' innovation. Unlike natural ecosystems, which maintain a sustainable balance unless dramatically intervened upon, the AI ecosystem requires constant observation to ensure equilibrium. As Brand [17] observes, all durable dynamic systems consist of parts that respond at different paces. Some elements react quickly to shocks, allowing slower elements to maintain their steady roles in ensuring system continuity. This interplay between speed and stability is what makes these systems both adaptable and robust. In such systems, the slow and large components control and stabilise the small and fast ones. While the fast layers capture our attention, it is the slow layers that hold the true power. The fast layers innovate,the slow layers stabilise. Below is an overview of these layers within a durable AI ecosystem (Fig. 8):

(i) *Algorithms*: The topmost and fastest-changing layer, characterised by the rapid evolution of learning and reasoning algorithms[19] and computational approaches. This layer is marked by intensive experimentation and 'trend-setting' developments.
(ii) *AI Technologies*: The tools and platforms that enable AI development. While changes in this layer are frequent, they occur at a slightly slower pace than in algorithms, due to the need for stable development environments.
(iii) *Integration Practices*: This layer focuses on the integration methods of AI into applications and the technological infrastructure supporting AI, necessary for AI deployment and operation. Changes here are more gradual, reflecting the need for robust and scalable AI solutions.
(iv) *Standards and Ethics*: This layer encompasses both internal and external rules and ethical guidelines that govern AI development and usage, including safety, bias mitigation, and transparency standards. It evolves under the influence of societal values and regulatory changes.
(v) *Societal Impacts*: Represents the long-term effects of AI on society, such as job displacement, privacy implications, and the increasing reliance on AI technologies. Changes in this layer are slow to manifest and difficult to reverse.
(vi) *Cognitive Co-evolution*: The deepest and slowest layer, reflecting the co-evolution of AI and human cognition. This involves fundamental shifts in how humans think, learn, and interact with intelligent systems over generations.

All actors within the AI ecosystem, as outlined in Sect. 4, must have their own positions within respective layers and the opportunity to participate and be heard. Distinctly depicting people and layers helps reveal the weakest links—e.g., those neglected at each level—determining how they can be more effectively engaged and better prepared for endeavours, while also uncovering different perspectives of power.

On 7th April 2023, Grant and Weise authored an article in *The New York Times* titled "In A.I. Race, Microsoft and Google Choose Speed Over Caution",[20] explaining how the priority for the tech companies became winning control of the industry's next big thing. By considering Fig. 8—examining how dynamism operates, the interplay of components, and the thresholds—the potential consequences of disregarding this balance become more apparent.

In essence, when AI developers prioritise speed over meticulous oversight, swiftly deploying Minimum Viable

[19] In this context, "algorithms" refers not only to the literal sense but also to analogous methods—conceptual or procedural—that remain subject to revision, as emerging approaches may ultimately redefine or replace the concept altogether.

[20] Accessible through https://www.nytimes.com/2023/05/01/business/ai-chatbots-hallucination.html.

Products (MVPs), the 'AI Technologies' and 'Algorithms' layers progress at an accelerated rate. This rapid development often leads to 'technical debts' due to insufficient testing and integration within the 'Integration Practices' layer, resulting in systems that are potentially more vulnerable and unreliable. Concurrently, 'ethical debts' accumulate primarily in the 'Standards and Ethics' layer, where the rapid pace of technological advancements outstrips the capacity to develop adequate governance and regulatory frameworks. This lag results in unresolved issues, such as biases in decision-making systems, breaches in data privacy, and a lack of transparency. As a consequence, the faster-moving technical layers significantly outpace the slower, more deliberate 'Societal Impacts' layer. This misalignment leads to unreliable and ethically questionable AI applications, causing unpredictable outcomes that erode public trust. The resulting societal backlash and potential regulatory crackdowns could stifle innovation and disrupt the AI ecosystem.

While prioritising AI ethics standards remains a pressing necessity [80], halting AI development is neither desirable nor practical. It is crucial to emphasise, though, that this view bears no alignment with 'accelerationist' theories—particularly those that promote the rapid intensification of technological development with insufficient regard for societal risk. Instead, it focuses on understanding the essential, yet often unseen, harmonies between layers and ensuring balanced development across them. This demands a meticulous application of both analytical and systems thinking—avoiding the 'self-fulfilling prophecy' trap.

In line with Sect. 5.1. on the 'AI Societal Safeguard System', this view champions a more deliberate emphasis on AI as the principal means of leverage to foster and sustain this form of hegemony. Well-orchestrated efforts are required to ensure AI Ecosystems are more 'inclusive' and 'ethical'. Achieving this can benefit from design-led methodologies, especially Human-Centred AI (HCAI), which should incorporate not just 'user-centred', but also 'community-centred' and 'societally-centred' strategies, as advocated by Landay [64].

4.3 AI ecosystem bottlenecks

To fully grasp the scope of AI-driven power dynamics, it is essential to recognise the heterogeneous landscape and acknowledge the absence of a one-size-fits-all solution. Disrupting the smooth adjustment between layers in an AI ecosystem, even unintendedly, can threaten stability, and hinder proper functioning. Several bottlenecks illustrate this complexity, including but not limited to:

(i) *Regulatory environments*: There is a stark contrast in how AI is regulated across different regions, particularly between the Global North—aiming to set global standards and benchmarks—and the Global South, which is still shaping its regulatory frameworks. Even within pioneering AI regions, significant challenges persist; for instance, existing whistleblower protections seem to be insufficient, as they often focus solely on illegal activities, leaving many AI-related risks unregulated [53].
(ii) *Technological infrastructure*: Disparities between developed and developing regions significantly affect their AI capabilities, creating an imbalance in their ability to contribute meaningfully to global AI discourse.
(iii) *International collaboration and conflict*: AI is a field marked by both collaboration, such as the partnership between IBM and the University of Tokyo,[21] and conflict, as seen between China and the US. These interactions often transcend simple competition, reflecting deeper geopolitical tensions.
(iv) *Ethical and social concerns*: These challenges are prevalent across the AI spectrum. Some regions grapple with AI governance and accountability, while others face distinct issues such as equitable access and cultural sensitivity, necessitating context-specific solutions.

4.4 Consequences of uneven development

Let's delve deeper into some examples of how the subtle interplay between these elements can make it all the more challenging to arrive at inclusive approaches to AI development that benefit all groups and individuals. In each case we can see the potential for AI to have a systemic impact on global power relations comparable to that envisaged by the traditional concept of hegemony, but without the hegemonic outcome having been willed by any specific individual or group, thus illustrating, once again, the need for the new concept of *AI*gemony.

4.4.1 'Cargo cult AI'

Taking a cue from the concept of the 'Cargo Cult'[22] as developed in the field of anthropology, one might envision a similar pattern emerging in AI adoption. Imagine a country or organisation, regardless of their preparedness, heavily investing in AI technologies. They purchase systems from leading technology companies and deploy them in sectors like finance and healthcare, but neglect to provide sufficient training for local experts or adapt the technology to local

[21] For further details, refer to https://itl.adm.u-tokyo.ac.jp/en/.

[22] A metaphor for a shallow imitation of a process that lacks a fundamental understanding of its underlying mechanism and the relationship between cause and effect [70].

contexts. Without a solid understanding[23] of AI or investment in developing localised expertise, these technologies struggle to integrate with the existing infrastructure and fail to deliver the expected benefits. Consequently, over time, this approach can lead to a skewed form of technological adoption where certain nations or entities consolidate their position by effectively leveraging AI, while others, concentrating only on superficial adoption, lag behind.

4.4.2 'Confounding variables' in AI's impact analysis

Disparities between AI regulatory regimes can obscure our understanding of AI's impacts, sometimes acting as 'confounding variables' that serve to entrench power imbalances. For instance, the European Union's robust technological infrastructure enables robust AI integration across various sectors, enhancing productivity and solidifying its position as a global standard-setter in AI. In contrast, less developed regions struggle with basic AI functionalities due to inadequate technological resources. This gap perpetuates the false perception that AI benefits everyone, influenced primarily by the experiences of technologically advanced regions. Consequently, policies tend to be biased towards these regions, exacerbating global inequalities and reinforcing power imbalances. Mitigating these disparities through international cooperation and 'technology sharing' initiatives is crucial to ensure a more equitable development of AI. This dynamic not only further marginalises less-equipped areas but also perpetuates a cycle where the technologically dominant dictate outcomes and enjoy disproportionate benefits.

4.4.3 Good is not good enough

While initiatives like AI, tech, or design '…for Good' are well-intentioned, they raise important questions about who truly benefits and who decides what is considered 'good'. These considerations are vital in ensuring that AI advancements lead to genuinely equitable outcomes.

5 Conclusions

While acknowledging the unique and promising capabilities that AI offers, many crucial points stand in need of further evaluation. Going forward, the *AI*gemony concept may be elaborated through interdisciplinary explorations in Design, particularly influenced by 'Social Design', 'Co-Design', and 'Human-centred Design', alongside broader perspectives from the social sciences. Undoubtedly, further investigation through Science and Technology Studies (STS) is essential to refine and deepen this concept. Below are some initial themes to consider in a work programme aimed at deepening our understanding of this new form of hegemony.

5.1 AI societal safeguard system

Historically, human societies have devised various safeguarding methods—from legends and beliefs to regulations—to shield themselves and find comfort amidst uncertainties and ongoing transformations. Inspired by the form of the brain's protective structure, Fig. 9 metaphorically outlines a basic three-layer framework for responding to the layered risks and disruptions induced by AI—from regulatory gaps to ethical disorientation and the erosion of shared values. The outermost layer, analogous to the bone, comprises 'Regulations'—firm and with a strong presence. The middle layer, resembling the meninges, is 'Philosophy'; it acts as a flexible shock absorber, enabling individuals to process and adapt to change. Philosophy deepens human connection to the world around us, imbuing phenomena with greater meaning. The innermost and most crucial layer, akin to the brain itself, consists of 'Internalised Values'. Nurturing a shared sense of 'common fate' and 'collective responsibility'[24] as the backbone of this framework supports a more inclusive and beneficial approach, amplifying often-overlooked voices and enhancing societal resilience. Understanding AI's role in both reinforcing and disrupting hegemony is critical for empowering societies to navigate the complexities of AI integration, thereby fostering adaptability in the face of swiftly developing technological advancement. Constant monitoring and revisions are necessary to ensure this structure stays dynamic and effective.

One illustrative instance of decentralised resistance is the development of Nightshade, a data-poisoning tool created by Shan and his team [107] at the University of Chicago in response to unauthorised data scraping by major AI firms. Designed to corrupt scraped imagery and disrupt model training, such interventions represent valuable forms of community-led contestation. Yet, without robust legal and technical safeguards, these efforts risk being neutralised—or even dismissed as anarchic rather than recognised as legitimate assertions of creative and informational rights.

[23] This also necessitates some degrees of internalisation. Stability in an 'activity system' is attained through the effective internalisation of routines, habits, attitudes, and emotions towards people, objects, and situations ([123], p. 143).

[24] Although these values are often assumed to be internalised—being inherently valuable and capable of inspiring action—incorporating clearly defined 'value propositions' and achieving tangible, meaningful outcomes are essential to sustaining individuals' motivation and engagement.

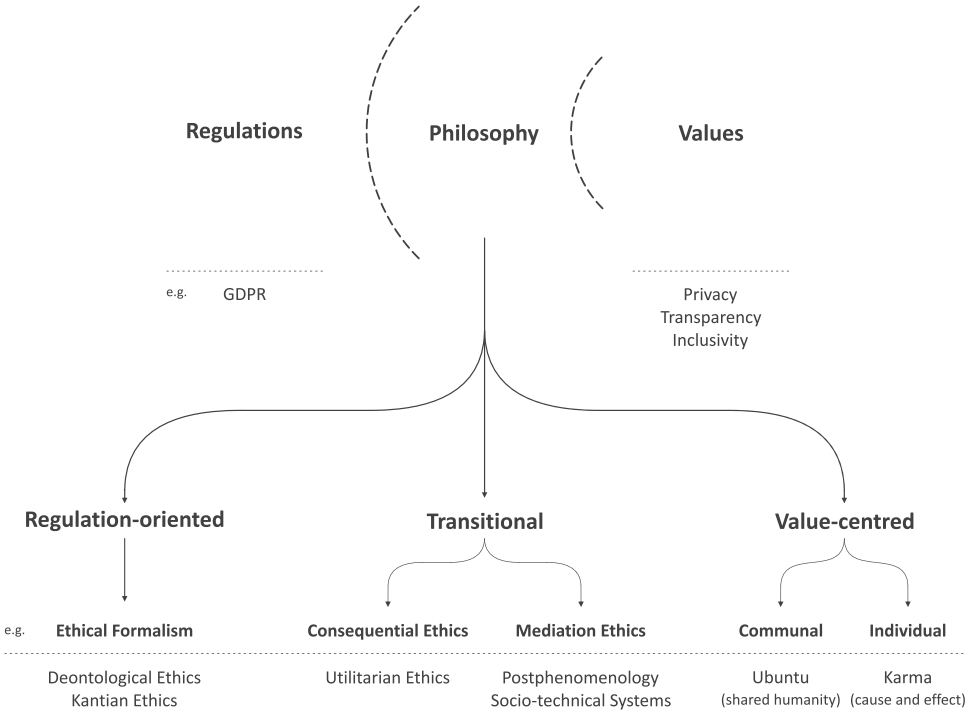

Fig. 9 Conceptual AI societal safeguard system

5.2 Questioning: simple but efficient

The Shirky Principle [59] insightfully states that "institutions will try to preserve the problem to which they are the solution." In other words, entities tend to perpetuate the issues they are supposed to resolve [108, 116], seeking to justify their continuing efforts and even their existence. This behaviour or tendency is not exclusive to institutions,individuals, too, may hold on to the issues, habits, or roles that define their identity or sense of purpose, leading to similar outcomes. These patterns reflect the ways in which the dominant forces, whether groups or individuals, seek to maintain their influence by normalising their actions, thereby creating self-reinforcing cycles.

While the emergence of AI has elevated this play to a new level, the underlying issue persists, even while the methods have evolved. The way models are trained, issues are framed, and questions are posed can significantly influence how AI performs when applied by different people [9]. As Talbot [115] suggests, this process can begin with pondering and simple questions. Indeed, even small, deliberate steps can lead to significant impacts: What do equitable and inclusive power dynamics look like? Are we using AI to challenge the entrenched hierarchies and inequalities? Or are we allowing AI to preserve the problem—i.e., the existing systemic biases and power imbalances in education and society—to which these traditional structures are the solution?

The transformative effect of simply asking questions should not be overlooked. Greater understanding of these points can support sound decision-making, help avoid the pitfalls of manufactured problems, and challenge the misuse of AI that perpetuates unfair power structures.

5.3 Is this a case of ethicalisation?

There is no doubt that ethical considerations, whether current or future, must be accorded the utmost priority. Moreover, intentional neglect or the deliberate undermining of ethical imperatives can never be justified. However, this matter demands nuanced contemplation. At times, the amplification of public concern over ethical issues may be intentionally orchestrated, serving as a means of 'fishing in troubled waters'. In the context of AI and ethics, such tactics may be deployed for various purposes, including regulatory capture, market control, the pursuit of crisis-driven profits, and reputation management, among others. This raises the critical question: Do certain issues genuinely constitute objective ethical dilemmas, or have they merely been portrayed as catastrophic concerns for ulterior motives?

Borrowing from 'Securitisation Theory', introduced by Buzan and his colleagues [23], this phenomenon may be termed 'ethicalisation'.[25]

> 'Ethicalisation' refers to the deliberate framing of an issue as a significant ethical challenge—sometimes

[25] A distinct concept from 'ethics washing'. In the context of AI, ethics washing involves fostering the illusion that ethical concerns are being adequately addressed, thereby legitimising the continuation of systems that may reinforce existing patterns [58].

even as an existential threat—often as a means to justify urgent interventions and extraordinary measures, circumventing standard procedures.

This notion represents a caution against potential distortions and underscores the need for careful, critical evaluation.

Ethicalisation also operates through a 'speech act',[26] wherein an influential figure—whether a political leader, technological magnate, or any other legal or natural person—claims that a particular AI-related aspect constitutes a perilous ethical risk. Through this assertion, the ethicalising actors exert social and institutional influence to propel the issue beyond ethical realms. This reminds us that ethics, akin to security, can be socially constructed rather than inherently defined. It highlights the power of language and the influential role of elites in shaping what is perceived as an ethical issue. The success of ethicalisation hinges on the audience's acceptance of this framing, thereby legitimising exceptional actions to address the alleged risk. Thus, raising awareness and critical thinking are key to maintaining a rational ethical lens. As such, this also advocates for a well-informed, inclusive, and intellectually rigorous discourse—one that promotes clarity, ensures the conversation remains constructive, and avoids inadvertently exacerbating confusion.[27]

5.4 Constructed objectivity and *AI*gemony

Discussions on *AI*gemony here are primarily shaped around Artificial Narrow Intelligence (ANI). ANI systems excel in specific domains, yet much of their potential remains unrecognised and underutilised. *AI*gemony is likely to remain the prevailing paradigm until a significant breakthrough occurs—one not limited to technological advancements and not necessarily marked by the emergence of Artificial General Intelligence (AGI). A key point to emphasise is the need to avoid fixating on technology in isolation. Instead, the challenge lies in developing resilient social and institutional infrastructures atop a technological foundation that is both imperfectly understood and unevenly controlled. Possessing extensive knowledge or vast data is not the same as making wise decisions or achieving constructive outcomes. At times, even experts struggle to fully grasp the current and long-term implications of transformations—a reality that demands vigilance and inclusive dialogue.

On a deeper level, what might go unacknowledged in debates about AI is the fundamental importance of understanding human capabilities and limitations before focusing on machines. The way the world is perceived is co-shaped by a wide range of constructs—whether societal, cultural, religious, or linguistic—each acting as a prism through which beliefs, ideals, and values are refracted. Notably, what is considered objective is often itself a construction—a product of shared paradigms. As Kuhn [63] famously argues, what an observer sees is influenced by the conceptual paradigms and prior experiences that frame their interpretation of facts (1962/2012, p.132). In other words, data and perceived facts do not speak for themselves; they are filtered through interpretative lenses shaped by training and worldview. Similarly, in *The Social Critique of the Judgement of Taste* (1979/2010), Bourdieu [15] illustrates how cultural preferences, tastes, and their judgments are not individual but are shaped by—and reinforce—social class and power structures. What one generation regards as firm and objective can be upended by the next, reflecting the contingency of our so-called objectivities—highlighting that much of what is treated as solid reality is, in fact, fluid and subject to redefinition ([10], p.147).

These constructed objectivities can become snares when accepted uncritically, limiting the imagination of what is possible. Yet they also offer opportunities: through reflexivity and the recognition that no perspective is fixed, the door edges open, but stepping through requires intent and action. Thus, even *AI*gemony remains constructive only as long as its audience recognises it as a mere vessel for meaning—a prompt for continual critical exploration and questioning.

5.5 Final remarks

Sometimes, there is resistance or even diplomatic avoidance when it comes to adopting AI, which, while understandable, overlooks the reality that AI is here to stay [36]. Additionally, given the prevailing perception of a post-colonial era, topics such as hegemony and colonisation may appear less relevant. Unlike traditional forms of power, *AI*gemony operates with greater unpredictability, remains less visible—therefore resisting control through 'checks and balances' mechanisms—and functions with increased efficiency. This time, however, it represents a structure of power which is not simply exerted over those who fit common stereotypes, such as those involving a person of colour who is also part of the LGBTQ+community, to the benefit of dominant groups, like a wealthy white individual living in London; with *AI*gemony, structures of power are now increasingly

[26] In the philosophy of language and linguistics, a speech act is defined as an utterance that not only conveys information but also performs a particular action [68].

[27] For instance, at MIT's Aeronautics and Astronautics Department's Centennial Symposium, Elon Musk remarked, "With artificial intelligence, we are summoning the demon" (Washington Post, 24 October 2014). Regardless of intent, such a statement—whether offered as a metaphor or mere opinion—risks fuelling confusion and anxiety, shaping public perception through emotion rather than reasoned, evidence-based discourse.

exerted over those who were previously categorised, in a clichéd fashion, among dominant groups as well. AI, after all, simply 'thinks' in terms of patterns. In a sense this is an extension of the point, made in Sect. 2.2 above, that those who benefit from dominant narratives may come to believe in those narratives as truths; under *AI*gemony, the dominant group may be iteratively exposed to a diet of selective information which prevents it seeing the dominant narrative for what it is, even as that narrative becomes contested or disrupted by subaltern groups. In fact, the hook here concerns attitudes towards AI—how it is perceived and applied—significantly broadening the scope. Thus, even WEIRD (Western, Educated, Industrialised, Rich, and Democratic) individuals, to use a term from Henrich [51], once considered unlikely targets, can now be equally affected alongside the marginalised.

As is often the case, the greatest challenge lies not in the technical implementation of AI but in coordinating people to respond to its ongoing development [27]. In this context, ignorance is like playing Russian Roulette—an unpredictable gamble with uncertainty about who will be affected next. Navigating the complexities requires insights that will take us on a journey with many unseen aspects that no single group or team can manage alone. Coming together with the courage to stay engaged, the willingness to listen to things that might not be favourable, and the flexibility to learn, are essential to shaping a more inclusive and profound dialogue. Sharing diverse perspectives, reflecting on years of precious lived experience, can undoubtedly help reveal what might be overlooked and ensure appropriate action is taken before time teaches us differently.

Author contributions The author, N.B., solely undertook the conception, research, analysis, writing, preparation of illustrations, and revision of this manuscript, assuming full responsibility for its development in its entirety.

Funding The author confirms that no funding, grants, or other forms of support were received during the preparation of this manuscript.

Availability of data and material No datasets were generated or analysed during the current study.

Declarations

Conflict of interest The author declare no competing interests.

Ethical approval and consent to participate Not applicable.

Consent for publication Not applicable.

References

1. Acemoglu, D.: Reward structures and the allocation of talent. Eur. Econ. Rev. **39**(1), 17–33 (1995). https://doi.org/10.1016/0014-2921(94)00014-q
2. Acemoglu, D., Robinson, J.A.: Why Nations Fail: The Origins of Power. Profile Books, Prosperity and Poverty (2012)
3. Agamben, G.: *The Omnibus Homo Sacer*. Stanford University Press. (Original work published 1998) (2017)
4. Akerlof, G., Shiller, R.: Animal Spirits: How Human Psychology Drives the Economy, and Why It Matters for Global Capitalism. Princeton University Press, Princeton (2010). https://doi.org/10.1515/9781400834723
5. Altheide, D.L.: Media hegemony: a failure of perspective. Public Opin. Q. Opin. Q. **48**(2), 476 (1984). https://doi.org/10.1086/268844
6. Amara, R.: What we have learned about forecasting and planning. Futures **20**(4), 385–401 (1988). https://doi.org/10.1016/0016-3287(88)90061-4
7. Ananny, M., Crawford, K.: Seeing without knowing: limitations of the transparency ideal and its application to algorithmic accountability. New Med. Soc. (2018). https://doi.org/10.1177/1461444816676645
8. Autesserre, S.: Dangerous tales: dominant narratives on the congo and their unintended consequences. Afr. Aff.Aff. **111**(443), 202–222 (2012). https://doi.org/10.1093/afraf/adr080
9. Bahrami, N.: Tailoring an AI Service Design Teammate: A Model Proposal Leveraging GPT Technology, *Academic Design Management Conference (ADMC2024) Proceedings*, pp. 1070–1090. ISSN– 2640–4702, (2024) https://www.researchgate.net/publication/383271211_Tailoring_an_AI_Service_Design_Teammate_A_Model_Proposal_Leveraging_GPT_Technology.
10. Bauman, Z.: Liquid Modernity. Polity Press (2000)
11. Bengio, Y., Mindermann, S., Privitera, D., Besiroglu, T., Bommasani, R., Casper, S., Choi, Y., Fox, P., Garfinkel, B., Goldfarb, D., Heidari, H., Ho, A., Kapoor, S., Khalatbari, L., Longpre, S., Manning, S., Mavroudis, V., Mazeika, M., Michael, J., Newman, J.: International AI safety report. ArXiv (Cornell University) (2025). https://doi.org/10.48550/arxiv.2501.17805
12. Biesenbach, R.: *Unleash The Power of Storytelling: Win Hearts, Change Minds, Get Results*. Eastlawn Media (2018)
13. Boggust, A., Hoover, B., Satyanarayan, A., Strobelt, H.: Shared interest: measuring human-AI alignment to identify recurring patterns in model behavior. CHI Conf. Human Factors Comput. Syst. (2022). https://doi.org/10.1145/3491102.3501965
14. Boje, D.M.: Stories of the storytelling organization: a postmodern analysis of Disney as "Tamara-land." Acad. Manag. J.Manag. J. **38**(4), 997–1035 (1995). https://doi.org/10.2307/256618
15. Bourdieu, P.: *Distinction: A Social Critique of the Judgement of Taste* (R. Nice, Trans.). Routledge. (Original work published 1979) (2010)
16. Bowman, P.: The Limits of Post-Marxism: The (Dis)function of Political Theory in Film and Cultural Studies. In: Sim, S. (ed.) Reflections on Post-Marxism: Laclau and Mouffe's Project of

Radical Democracy in the 21st Century, pp. 59–78. Chapter, Bristol University Press (2022)
17. Brand, S.: Pace layering: how complex systems learn and keep learning. J. Design Sci. (2018). https://doi.org/10.21428/7f2e5f08
18. Bremmer, I.: The technopolar moment: how digital powers will reshape the global order. Foreign Aff.Aff. **100**(6), 112–128 (2021)
19. Brookfield, S.: *The Power of Critical Theory for Adult Learning and Teaching*. McGraw-Hill International (2005)
20. Brooks, R.: *The Seven Deadly Sins of AI Predictions*. MIT Technology Review. (2017) https://www.technologyreview.com/2017/10/06/241837/the-seven-deadly-sins-of-ai-predictions/
21. Bucher, T.: If...Then: Algorithmic Power and Politics. In *Oxford Scholarship Online*. Oxford University Press. (2018) https://doi.org/10.1093/oso/9780190493028.001.0001
22. Buolamwini, J.: *Unmasking AI: My Mission to Protect What Is Human in a World of Machines*. Random House (2023)
23. Buzan, B., Wæver, O., & De Wilde, J.: *Security: A New Framework for Analysis*. Lynne Rienner (1998)
24. Chan, S., Pataranutaporn, P., Suri, A., Zulfikar, W., Maes, P., & Loftus, E. F.: *Conversational AI Powered by Large Language Models Amplifies False Memories in Witness Interviews*. (2024). ArXiv.org. https://arxiv.org/abs/2408.04681
25. Chesney, R., Citron, D.: Deepfakes and the new disinformation war: the coming age of post-truth geopolitics. Foreign Aff.Aff. **98**(1), 147–155 (2019)
26. Cinellia, M., Moralesb, G.D.F., Galeazzic, A., Quattrociocchid, W., Starninib, M.: The echo chamber effect on social media. PNAS **118**(9), e2023301118 (2021)
27. Clark, J.: *Ship Faster by Building Design Systems Slower| Big Medium*. Big Medium. (2023) https://bigmedium.com/ideas/design-system-pace-layers-slow-fast.html
28. Clarke, B.: Beyond 'The Banality of Evil.' Br. J. Politic. Sci. **10**(4), 417–439 (1980). https://doi.org/10.1017/S0007123400002325
29. Connell, R.W., Messerschmidt, J.W.: Hegemonic masculinity: rethinking the concept. Gend. Soc.. Soc. **19**(6), 829–859 (2005). https://doi.org/10.1177/0891243205278639
30. Couldry, N., Mejias, U.A.: The Costs of Connection: How Data Is Colonizing Human Life and Appropriating It for Capitalism. Stanford University Press (2019)
31. Crabtree-Condor, Isabel. *Narrative Power and Collective Action: Conversations with People Working to Change Narratives for Social Good– Part 1*. On Think Tanks & Oxfam, 2020.
32. Craig, C. J.: *The AI-Copyright Trap*. (2024) https://doi.org/10.2139/ssrn.4905118
33. Culpepper, P.D., Thelen, K.: Are we all amazon primed? Consumers and the politics of platform power. Compar. Politic. Stud. **53**(2), 288–318 (2019). https://doi.org/10.1177/0010414019852687
34 Diakopoulos, N.: Automating the News: How Algorithms Are Rewriting the Media. Harvard University Press (2019). https://doi.org/10.4159/9780674239302
35. Dolan, P., Henwood, A.: Five steps towards avoiding narrative traps in decision-making. Frontiers Psychol. (2021). https://doi.org/10.3389/fpsyg.2021.694032
36. Dudash, S.: Ready or Not, Disruptive AI Is Here to Stay. Forbes. (2023)https://www.forbes.com/sites/greatspeculations/2023/06/22/ready-or-not-disruptive-ai-is-here-to-stay/
37. Encode Justice, & the Future of Life Institute.: *AI Licensing for a Better Future: On Addressing Both Present Harms and Emerging Threats - Future of Life Institute*. Future of Life Institute. (2023) https://futureoflife.org/open-letter/ai-policy-for-a-better-future-on-addressing-both-present-harms-and-emerging-threats/
38. Feldstein, S.: The road to digital unfreedom: How artificial intelligence is reshaping repression. J. Democr.Democr. **30**(1), 40–52 (2019)
39. Feldstein, S.: *The Global Expansion of AI Surveillance*. Carnegie Endowment for International Peace. (2019b) https://carnegieendowment.org/files/WP-Feldstein-AISurveillance_final1.pdf
40. Fivush, R.: Speaking silence: The social construction of silence in autobiographical and cultural narratives. Memory **18**(2), 88–98 (2010). https://doi.org/10.1080/09658210903029404
41. Fortier, J.: American hegemony in the 21st century: a neo neo-gramscian perspective. Int. Aff.Aff. **96**(4), 1120–1122 (2020). https://doi.org/10.1093/ia/iiaa103
42. Foucault, M.: *The History of Sexuality* (Vol. 1). Pantheon Books. (1978)
43. Foucault, M.: *Power/Knowledge: Selected Interviews and Other Writings, 1972–1977* (C. Gordon, Ed.). Pantheon Books (1980)
44. Foucault, M.: *Power: The Essential Works of Michel Foucault 1954–1984.* Penguin Classics (2020)
45. Freire, P.: *Pedagogy of the Oppressed*. Bloomsbury Academic. (1970) https://envs.ucsc.edu/internships/internship-readings/freire-pedagogy-of-the-oppressed.pdf (Original work published 1968)
46. Giansiracusa, N.: *How Algorithms Create and Prevent Fake News: Exploring the Impacts of Social Media, Deepfakes, GPT-3, and More* (1st ed.). Apress. (2021)
47. Gill, S.: *Power and resistance in the new world order* (2nd ed.). Palgrave Macmillan (2008)
48. Gitelman, L. (ed.): Raw Data Is an Oxymoron. MIT Press (2013)
49. Gramsci, A.: *Selections from the prison notebooks*. International Publishers (1971)
50. Gutting, G., & Oksala, J.: *Michel Foucault*. Stanford Encyclopedia of Philosophy. (2022) https://plato.stanford.edu/entries/foucault/
51. Henrich, J.: *The weirdest people in the world. How the West became psychologically peculiar and particularly prosperous.* (1st ed.). Penguin. (2021)
52. Herman, E. S., & Chomsky, N.: *Manufacturing Consent: The Political Economy of the Mass Media*. Pantheon Books. (1988)
53. Hilton, J., Kokotajlo, D., Kumar, R., Nanda, N., Saunders, W., Wainwright, C., & Ziegler, D.: *A Right to Warn about Advanced Artificial Intelligence*. Righttowarn.ai. (2024) https://righttowarn.ai/
54. Ikenberry, G.J.: After Victory: Institutions, Strategic Restraint, and the Rebuilding of Order after Major Wars. Princeton University Press (2001)
55. Investopedia Team.: *What Is the American Dream?* Investopedia. (2023) https://www.investopedia.com/terms/a/american-dream.asp
56. Jaffery, B.: Connecting With Meaning: Personalizing the Customer Experience Using Data, Analytics and AI. In *Deloitte–OmniaAI*. (2022) https://www2.deloitte.com/content/dam/Deloitte/ca/Documents/deloitte-analytics/ca-en-omnia-ai-marketing-pov-fin-jun24-aoda.pdf
57. Jervis, R.: Perception and Misperception in International Politics-New Edition. Princeton University Press (2017)
58. Kaspersen, A., & Wallach, W.: *Why Are We Failing at the Ethics of AI?* Carnegie Council for Ethics in International Affairs. (2021) https://www.carnegiecouncil.org/media/article/why-are-we-failing-at-the-ethics-of-ai
59. Kelly, K.: *The Shirky Principle*. The Technium. (2010) https://kk.org/thetechnium/the-shirky-prin/
60. Kleinberg, J., Ludwig, J., Mullainathan, S., Sunstein, C.R.: Discrimination in the age of algorithms. ArXiv (Cornell University) (2019). https://doi.org/10.48550/arxiv.1902.03731
61. Koehn, P.: Neural machine translation. Cambridge University Press (2020)
62. Kopec, D., Shetty, S., Pileggi C.: *Artificial Intelligence Problems and Their Solutions*. Mercury Learning and Information (2014)

63. Kuhn, T. S.: *The Structure of Scientific Revolutions: 50th Anniversary Edition* (4th ed.). University Of Chicago Press. (Original work published 1962) (2012)
64. Landay, J.: *The case for human-centered AI*. McKinsey & Company. (2024) https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/the-case-for-human-centered-ai#/
65. Ledwith, M.: Emancipatory action research as a critical living praxis: from dominant narratives to counternarrative. Palgrave Int. Handbook Action Res. (2016). https://doi.org/10.1057/978-1-137-40523-4_4
66. Le Ludec, C., Cornet, M., Casilli, A.A.: The Problem with Annotation. Big Data & Society, Human Labour and Outsourcing Between France and Madagascar (2023). https://doi.org/10.1177/20539517231188723
67. Leslie, D., Katell, M., Aitken, M., Singh, J., Briggs, M., Powell, R., Rincon, C., Perini, A., Jayadeva, S.: Data justice in practice: a guide for impacted communities. SSRN Electron. J. (2022). https://doi.org/10.2139/ssrn.4080046
68. Levinson, S. C.: Speech acts. In Y. Huang (Ed.), The Oxford Handbook of Pragmatics (Oxford Handbooks). Oxford University Press. (2017) https://doi.org/10.1093/oxfordhb/9780199697960.013.22
69. Lidwell, W., Holden, K., & Butler, J.: Universal Principles of Design: 200 Ways to Increase Appeal, Enhance Usability, Influence Perception, and Make Better Design Decisions (3rd ed.). Rockport Publishers. (2023)
70. Lindstrom, L.: Cargo cults. In F. Stein (Ed.), The Open Encyclopedia of Anthropology. (Facsimile of the first edition in The Cambridge Encyclopedia of Anthropology). (2018) https://doi.org/10.29164/18cargo
71. Liu, B.: Sentiment Analysis: Mining Opinions, Sentiments, and Emotions (2nd ed.). Cambridge University Press. (2020)
72. Loomba, A.: Colonialism/postcolonialism: the New Critical Idiom. Routledge (2015)
73. Lukes, S.: Power: A Radical View (2nd ed.). Palgrave Macmillan. (Original work published 1974) (2005)
74. Martin, J.: Antonio Gramsci. Stanford Encyclopedia of Philosophy; Metaphysics Research Lab, Stanford University. (2023) https://plato.stanford.edu/entries/gramsci/
75. Mayer, H., Yee, L., Chui, M., & Roberts, R.: Superagency in the workplace: Empowering people to unlock AI's full potential. McKinsey & Company. (2025) https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/superagency-in-the-workplace-empowering-people-to-unlock-ais-full-potential-at-work
76. McQue, K.: The global struggle over how to regulate AI. Rest of World. (2025) https://restofworld.org/2025/global-ai-regulation-big-tech/
77. Metcalfe, J.: Dominant narratives of whiteness in identity construction of mixed-race young adults in post-apartheid South Africa. Soc. Sci. **11**(5), 205 (2022). https://doi.org/10.3390/socsci11050205
78. Mhlambi, S.: Decolonizing AI. YouTube; Stanford HAI. (2022) https://youtu.be/UqVwfuIuU2k
79. Mohamed, S., Png, M.T., Isaac, W.: Decolonial AI: decolonial theory as sociotechnical foresight in artificial intelligence. Philos. Technol. **33**, 659–684 (2020). https://doi.org/10.1007/s13347-020-00405-8
80. Montgomery, C., & Rossi, F.: Don't pause AI development, prioritize ethics instead|IBM. Ibm.com. (2023) https://www.ibm.com/think/insights/dont-pause-ai-development-prioritize-ethics-instead
81. Moorosi, N., Sefala, R., & Luccioni, A. S.: AI for Whom? Shedding Critical Light on AI for Social Good. OpenReview. (2023) https://openreview.net/forum?id=vjwYYlA8Pj%26s=09
82. Moravec, H.P.: Robot: Mere Machine to Transcendent Mind. Oxford University Press, Oxford (1998)
83. Muldoon, J., Raekstad, P.: Algorithmic domination in the gig economy. Eur. J. Polit. Theo. **22**(4), 147488512210820 (2022). https://doi.org/10.1177/14748851221082078
84. Muldoon, J., Wu, B.A.: Artificial intelligence in the colonial matrix of power. Philos. Technol. (2023). https://doi.org/10.1007/s13347-023-00687-8
85. NTT DATA.: The AI Responsibility Gap: Why Leadership Is the Missing Link an Executive Insight Report. (2025) https://services.global.ntt/-/media/ntt/global/insights/blog/the-ai-responsibility-crisis-why-executive-leadership-must-act-now/the-ai-responsibility-gap-why-leadership-is-the-missing-link.pdf
86. Narayanan, A., Kapoor, S.: AI Snake Oil: What Artificial Intelligence Can Do, What It Can't, and How to Tell the Difference. Princeton University Press (2024)
87. Noble, S.U.: Algorithms of Oppression: How Search Engines Reinforce Racism. NYU Press, New York (2018)
88. Nyaaba, M., Wright, A., & Choi, G. L.: Generative AI and digital neocolonialism in global education: towards an equitable framework. (2024) https://doi.org/10.48550/arXiv.2406.02966
89. Nye, J.: The future of power. Public affairs (2011)
90. Oatley, T.: A Political Economy of American Hegemony: Military Buildups, Booms, and Busts. Cambridge University Press (2015)
91. Obama, B.: The Audacity of Hope: Thoughts on Reclaiming the American Dream. Canongate Books (2008)
92. Ohagi, M.: Polarization of autonomous generative ai agents under echo chambers. ArXiv (Cornell University) (2024). https://doi.org/10.48550/arxiv.2402.12212
93. Perloff, R. M.: The dynamics of persuasion: Communication and attitudes in the 21st Century. Routledge, Taylor & Francis Group (2017)
94. Plough, A. L. (Ed.): The power of shifting narratives, expectations, and consciousness: Discussion and case studies from Bellagio. In Well-being: Expanding the definition of progress: Insights from practitioners, researchers, and innovators from around the globe. Oxford Academic. (2020). https://doi.org/10.1093/oso/9780190080495.003.0006
95. Rahme, L. K., & Halat, R.: Fostering Critical Thinking Skills for an AI-Infused World: A Comprehensive Toolkit. Harvard Graduate School of Education - Middle East Professional Learning Institute. (2024) https://mepli.gse.harvard.edu/our-fellows-at-work/fostering-critical-thinking-skills-for-an-ai-infused-world-a-comprehensive-toolkit/
96. Rank, M.R., Hirschl, T.A., Foster, K.A.: Chasing the American Dream: Understanding What Shapes Our Fortunes. Oxford University Press, Oxford (2016)
97. Raven, B.H., Schwarzwald, J., Koslowsky, M.: Conceptualizing and measuring a power/interaction model of interpersonal influence1. J. Appl. Soc. Psychol. **28**(4), 307–332 (1998). https://doi.org/10.1111/j.1559-1816.1998.tb01708.x
98. Reid, T.: Essays on the Intellectual Powers of Man: A Critical Edition (D. Brookes & K. Haakonssen, Eds.). Edinburgh University Press. (Original work published 1786) (2002)
99. Rosamond, B.: Hegemony. Encyclopedia Britannica. (2024) https://www.britannica.com/topic/hegemony
100. Rouvroy, A., & Berns, T.: Algorithmic governmentality and prospects of emancipation: Disparateness as a precondition for individuation through relationships? (L. Carey-Libbrecht, Trans.) *Réseaux*, *177*(1), 163–196. (2013). https://shs.cairn.info/journal-reseaux-2013-1-page-163?lang=en
101. Rudin, C., Radin, J.: Why Are we using black box models in AI when we don't need to? A lesson from an explainable AI competition. Harvard Data Sci. Rev. (2019). https://doi.org/10.1162/99608f92.5a8a3a3d
102. Ruiz-Rojas, L.I., Salvador-Ullauri, L., Acosta-Vargas, P.: Collaborative working and critical thinking: adoption of generative

artificial intelligence tools in higher education. Sustainability **16**(13), 5367 (2024). https://doi.org/10.3390/su16135367
103. Russell, L.: The Banality of Evil. In Evil: A Very Short Introduction (pp. 53–60). Oxford University Press. (2022). https://doi.org/10.1093/actrade/9780198819271.003.0004
104. Sengul-Jones, M.: Bring in the Machines: AI-Powered Investigative Journalism. DataJournalism.com. (2021) https://datajournalism.com/read/longreads/machine-learning-investigative-journalism
105. Sevilla, J., & Roldán, E.: Training Compute of Frontier AI Models Grows by 4–5x per Year. Epoch AI. (2024) https://epoch.ai/blog/training-compute-of-frontier-ai-models-grows-by-4-5x-per-year
106. Sevilla, J., Heim, L., Ho, A., Besiroglu, T., Hobbhahn, M., & Villalobos, P.: Compute Trends Across Three Eras of Machine Learning. (2022). https://arxiv.org/pdf/2202.05924
107. Shan, S., Ding, W., Passananti, J., Zheng, H., Zhao, B.Y.: Prompt-specific poisoning attacks on text-to-image generative models. ArXiv (Cornell University) (2023). https://doi.org/10.48550/arxiv.2310.13828
108. Shatz, I.: The shirky principle: institutions try to preserve the problem to which they are the solution—effectiviology. Effectiviology.com. (2024). https://effectiviology.com/shirky-principle/
109. Sheridan, A.: Michel Foucault: The Will to Truth. Routledge. (2004)
110. Shin, D., Hameleers, M., Park, Y.J., Kim, J.N., Trielli, D., Diakopoulos, N., Helberger, N., Lewis, S.C., Westlund, O., Baumann, S.: Countering algorithmic bias and disinformation and effectively harnessing the power of AI in media. J. Mass Commun. Q. **99**(4), 887–907 (2022). https://doi.org/10.1177/10776990221129245
111. Statement on AI Harms and Policy - Conference on Fairness, Accountability, and Transparency (FAccT). (2023). Facctconference.org. https://facctconference.org/2023/harm-policy
112. Suleyman, M., & Bhaskar, M.: *The Coming Wave: AI, Power and Our Future*. Vintage (2024)
113. Sunstein, C. R.: #Republic: Divided Democracy in the Age of Social Media. Princeton University Press. (2018)
114. Susser, D., Roessler, B., Nissenbaum, H.: Technology, autonomy, and manipulation. Internet Policy Rev. (2019). https://doi.org/10.14763/2019.2.1410
115. Talbot, C.: Abundant AI and the Shirky Principle. MSA Evolution Lab. (2024) https://www.msaevolutionlab.com/blog/abundantai
116. Thomson, J.: The Shirky Principle: Why your personal trainer wants you to stay unfit. Big Think; Big Think. (2023) https://bigthink.com/thinking/shirky-principle/
117. Torres, P.: The possibility and risks of artificial general intelligence. Bull. Atomic Sci. **75**(3), 105–108 (2019). https://doi.org/10.1080/00963402.2019.1604873
118. Truby, J.: Governing artificial intelligence to benefit the UN sustainable development goals. Sustain. Develop. (2020). https://doi.org/10.1002/sd.2048
119. UK Government.: Policy Paper: Introducing the AI Safety Institute. GOV.UK. (2024) https://www.gov.uk/government/publications/ai-safety-institute-overview/introducing-the-ai-safety-institute
120. US Government.: Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence. The White House. (2023) https://www.whitehouse.gov/briefing-room/presidential-actions/2023/10/30/executive-order-on-the-safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence/
121. Vandelanotte, C., Trost, S., Hodgetts, D., Imam, T., Rashid, M., To, Q.G., Maher, C.: Increasing physical activity using a just-in-time adaptive digital assistant supported by machine learning: a novel approach for hyper-personalised mHealth interventions. J. Biomed. Inform. **144**, 104435 (2023)
122. Wachter, S., Mittelstadt, B., Russell, C.: Do large language models have a legal duty to tell the truth? Royal Soc. Open Sci. (2024). https://doi.org/10.1098/rsos.240197
123. Wulf, V., Pipek, V., Randall, D., Rohde, M., Schmidt, K., Stevens, G.: Socio-Informatics. Oxford University Press, Oxford (2018)
124. Yeung, K.: 'Hypernudge': Big data as a mode of regulation by design. In The Social Power of Algorithms (pp. 118–136). Routledge. (2018)
125. Zuboff, S.: The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power. Profile Books. (2019)