SciencesPo

TECH & GLOBAL AFFAIRS
INNOVATION HUB

FIVE AIMS:

Lessons from Internet Governance for Artificial Intelligence Management Strategies

Wendy HALL

University of Southampton, Regius Professor of Computer Science, Associate Vice President and Director of the Web Science Institute

Kieron O’HARA

University of Southampton, Emeritus Fellow in Electronics and Computer Science

Pierre NORO

SciencesPo Tech & Global Affairs Innovation Hub, Advisor

Dame Wendy Hall, Dame Wendy Hall, DBE, FRS, FREng, is Regius Professor of Computer Science, Associate Vice President (International Engagement) and Director of the Web Science Institute at the University of Southampton. She was co-Chair of the UK government's AI Review in 2017 and a member of the AI Council. In 2023 she was appointed to the United Nations high-level advisory body on artificial intelligence. Her latest book, Four Internets, co-written with Kieron O'Hara, was published by OUP in 2021.

Kieron O'Hara is emeritus fellow in Electronics and Computer Science at the University of Southampton. His latest book, *Blockchain Democracy: Ideology and the Crisis of Social Trust*, will be published by Edward Elgar in Summer 2025**.**

Pierre Noro is adjunct faculty member at SciencesPo Paris and Université Paris-Cité, advisor to the PSIA Tech & Global Affairs Hub, and part of the editorial team of the AIAAIC. His works are mostly dedicated to decentralized governance, digital ethics and the social and environmental impact of digital technologies.

The authors are grateful for the initial contributions and thoughtful feedback from **Constance de Leusse**, Senior Advisor of the Tech & Global Affairs Innovation Hub, whose support greatly enhanced the elaboration of this policy brief.

This document is part of the Policy Brief series published by the Paris School of International Affairs (PSIA) Technology and Global Affairs Innovation Hub.

The Hub's core mission is to accelerate collaborative technology and international governance to address global challenges. Its activities are specifically focused on technology and democracy, defence and security, sustainability and prosperity.

Learn more about the Hub on our website: www.sciencespo.fr/psia-innovation-hub/

Published in January 2025.

Suggested citation: Hall, Wendy, O'Hara, Kieron & Noro, Pierre, *Five AIMS: Lessons From Internet Governance For Artificial Intelligence Management Strategies*, policy brief commissioned by the Paris School of International Affairs Tech & Global Affairs Innovation Hub, January 2025.

INTRODUCTION

Discussions of artificial intelligence (AI) ethics and governance have a long history,[1] appearing and reappearing in public debate, sometimes prompted by works of art, such as Karel Čapek's play *R.U.R. (Rossum's Universal Robots)* or Stanley Kubrick's film *2001: A Space Odyssey*, sometimes by academic works that achieve public visibility, such as Alan Turing's 'Computing Machinery and Intelligence' or Douglas Hofstadter's *Gödel, Escher, Bach: an Eternal Golden Braid*, sometimes by events, such as the defeat of Go champion Lee Sedol in 2016 by the AlphaGo program, and sometimes by products, such as MIT's "social robot" Kismet, or the ChatGPT chatbot. 2024 saw five Nobel Prizes awarded for contributions to -or applications of- machine learning. In recent years, the technologies of deep learning, followed by generative AI, have made the regulation and governance of AI appear imperative, bringing AI ethics into the mainstream of political thinking about technology.

The acceleration of development, funding and adoption of AI technologies during this time has foregrounded the imperative for companies and governments to rapidly establish governance frameworks to support responsible innovation, protect individual rights, and promote societal well-being. In the absence of effective governance, many are concerned that the unchecked proliferation of AI could exacerbate existing inequalities, undermine democratic values and institutions, and pose significant risks to global security and stability.

But from a less negative perspective, AI promises enormous advances in all sorts of areas where finding patterns in data is vital, from medicine to climate change to civil administration to defence to science, especially as research productivity declines and its complexity increases.[2] On this view, ensuring the safety of AI is necessary less for the reduction of risk, than for earning the social licence that legitimises its operation – a general societal acceptance that the technology will remain within the bounds of societal expectations.

Yet at the moment, AI governance is somewhat undertheorised, and *ad hoc* measures and institutions are being proposed and being assembled into a *de facto* governance regime without much consideration for coherence, practicality or enforceability. In this paper, we discuss one particular framework for Internet governance, and consider its plausibility in application to AI governance in particular, with an eye to it being a tool to address this question of coherence. We extrapolate its principles for the AI domain, and suggest a series of research questions arising from the differences we identify between Internet and AI governance.

FOUR INTERNETS

In a series of publications, O'Hara and Hall developed a geopolitical and ideological theory of Internet governance,[3] arguing that a few specific normative ideas about the Internet had become particularly influential. Their influence is realised through, first, an **engineering interpretation of their tenets** reflected in the Internet technology stack; and secondly, **the backing of one or more geopolitical entities** (or large technology companies) with the power to modify the sociotechnical context. There are many ideas about what the Internet should be for and how it should be governed, but in the analysis only five were sufficiently supported by engineering and political power to be considered significant in Internet terms.

The Internet is a network of networks, and four of these ideas could be seen as describing a network of networks of similar types, i.e. an Internet of their own: an **Open Internet**, a **Bourgeois Internet**, a **Paternal Internet**, and a **Commercial Internet**. These four together, through their deep interconnections, make up the Internet as a whole. The project was therefore entitled *Four Internets* (no definite article, as it is possible for further models to emerge). Additionally, there is a subversive **Spoiler model**, parasitic on the other Internets. It therefore did not describe an Internet, although it required an Internet to subvert. These models can be thought of as ideal types in a Weberian sense.[4]

Each Internet was associated with a geopolitical exemplar, but it is important to emphasise that all geopolitical actors **are supportive of more than one Internet type**. The same government might pursue the open vision via its support for its transformative startup sector, the bourgeois vision via its commitments to human rights, the paternal vision via imperatives of national security and the wellbeing of its population, the commercial vision via its business policies and the spoiler model via its cyber-defence and cyberattack strategy.

The Four Internets (4Is) can be described as follows:

- **The Open Internet** – A **libertarian** vision based on the free and efficient flow of information, ideally uncensored, with open and universal standards, interoperability and permissionless access, discouraging ‘walled gardens’ where incommensurable rules are enforced. It perceives itself as serving the public good through supporting collaborative innovation.
- **The Bourgeois Internet** – This vision is focused on **securing human rights and civility**, especially around privacy, private and intellectual property, free expression, rights to information, and rights to Internet access. Innovation is legitimate when it does not breach these. It means to serve the public good partly through its support for rights, and partly through support for innovation that does not breach rights.[5]
- **The Paternal Internet** – This is a top-down concern for the security of society, demanding of the Internet that it **directly serves public aims**, such as economic growth and social stability, and that it is pro-social (e.g. not disseminating pornography or political extremism, or not hosting sites that promote self-harm or suicide).
- **The Commercial Internet** – This vision conceives the Internet both as a **type of property and an infrastructure for new types of property**. Innovation should therefore be monetisable, and the creation of profit signals that innovation has added social value. Walled gardens are acceptable if they provide services to meet tangible demand. The role of governance is to preserve property rights.

In addition:

- **The Spoiler Model** – This is based on the hacker ethic,[6] celebrating the expertise of gifted programmers to create elegant, innovative code that subverts traditional narratives, and creates new and anarchic realities through the power inherent in software. It serves the public good by promoting **liberty and resistance, and challenging existing power structures**.

ARTIFICIAL INTELLIGENCE MANAGEMENT STRATEGIES AND INTERNET GOVERNANCE: CONVERGENCES

The 4Is framework was created to describe the development and practice of Internet governance, via soft and hard regulation, codes of practice, technical protocols etc., in its geopolitical and ideological context. The question to be addressed in this paper is whether it affords any advantages in thinking about AI governance or management strategies. At this stage, we take this as a descriptive question, rather than the normative question of whether the 4Is framework *ought* to be used.[7] We will call the analogues of the 4Is ideal types **Artificial Intelligence Management Strategies** (AIMS).

It is important to note that the genealogies of the Internet and AI are very different. The Internet was developed in the context of a libertarian, open ideology, manifesting through open standards and interoperability, and connectivity is fundamental for it to flourish. The other ideal governance types evolved in response to collective action problems that had appeared when the Open Internet scaled up. Hence they were necessarily reactive in nature to an existing (and demonstrably successful) set of open standards. Openness thus has a certain priority with respect to the Internet compared to its competing governance types. Not only was it first on the scene, but, as it informed the design of its technical and institutional infrastructures, a lack of openness may cause fragmentation or critical inefficiency by raising barriers to joining the Internet and growing the network.

In contrast, nothing about AI makes openness essential for its operation. Furthermore, the various standards which the community is debating do not yet exist. The core of the debate about AI governance is what standards ought there to be in the first place, whereas by the time Internet governance became a topic of general political interest, standards had already been defined and institutions crafted. The Internet demanded a conservative, incremental approach, while the field of AI is still open to bold ideas.

The most obvious analogy between the Internet models and the current AI landscape is that both are at the nexus of digital modernity, digital technologies and society.[8] The Internet Models and the AI ecosystem deal with similar stakeholders, primarily states, international organisations such as the EU, and large technological companies, but also special-purpose institutions, smaller private actors (including startups), open source communities and civil society organisations. It is safe to assume similar stakeholders would logically reuse some of the Internet models for AI governance, either because of the successes of Internet governance, because of institutional inertia, or alternatively because of the need to source 'off-the-shelf' management strategies owing to the rapid development and adoption of the technology. Furthermore, each model contains an account of the/a social good which would influence their views about both the potential and the risks of AI.

The interactions between stakeholders in regard to AI also inherit from their previous relationships in respect to the Internet. For instance, the asymmetry of technical knowledge between technology companies and public actors remains structural and contributes to their strategies. AI and Internet technologies also share similar geographical concentrations of power: the 4Is framework emphasises the close relationships between particular areas and governance styles – the Open Internet emerged from Silicon Valley, the Commercial Internet was pioneered in Washington, the Bourgeois Internet has gained traction via the 'Brussels effect',[9] while China has always engaged with the Internet with paternal, even authoritarian, concern about outcomes. Russia, with its radical ideology of the information space,[10] has been prominent in refining Spoiler Models.

Nevertheless, there are new centres of AI power emerging, especially including India, Qatar, the United Arab Emirates and Saudi Arabia. States are unlikely to contain the expertise to develop AI models independently, while the provision of sufficient computing power to support the creation of LLMs and other AI models is likely to require coordination between states and private entities. The private entities need not be based in the relevant states (witness Microsoft's recent $1.5bn investment in the UAE's G42[11]). Nevertheless, any private sector entity trying to develop such models would need to be large and well-funded,[12] and hence would be likely to be visible to regulators.

Current discourses around AI governance are reminiscent of those that structured the emergence of the different Internet Models in some ways, with some possibly significant differences. Both tended to downplay the materiality of digital technologies, to emphasise instead ideological and economic issues (e.g. existential risk or the impact of AI on human work). Secondly, while the infrastructure enabling the Internet was initially largely provided by public actors, empowering states to implement control strategies, private stakeholders invest in their own material capacity for AI development, including data centres, special purpose chip design and manufacturing facilities, and even low carbon emission energy sources for processing.[13] On the other hand, the crucial early investments of the United States in Internet infrastructure and funding were accompanied by a relatively 'hands off' governance strategy, resulting in the devolution of a large amount of Internet infrastructure to private or non-state hands (while ensuring most were incorporated in the US). Thirdly, as a permissionless system, exercising total control over the whole Internet is at least expensive, if not impossible within a particular jurisdiction, whereas the barriers to entry into Generative AI exclude all but those with the deepest pockets. The latter therefore make a smaller (if powerful) target.

FIVE AIMS

The five ideal types of the 4Is governance framework each have an intuitive and *prima facie* plausible translation into an Artificial Intelligence Management Strategy, although with indeterminacies and blurred boundaries between them. Even though the AIMS (like the 4Is ideal types) have a narrative sensemaking element that will resist precise formulation, further research is needed to establish the boundaries and contrast between them with as much precision as the theory will allow. As with the 4Is, governance models cover various actors, including governments, supranational bodies such as the EU, companies, interest groups, as well as specialised institutions set up for particular purposes (in the case of AI, this may include safety institutes and other clearinghouses for testing, benchmarking and disseminating best practice).

The Five AIMS can be characterised according to **competing ideas of the public good** and legitimate ways of achieving it. The AIMS are also conditioned by factors including societies' and governments' legacy attitudes to industrial strategy, scientific research, intervention in markets, social control, geopolitical advantage and vulnerability, and previous public policy stances toward technology. There is an ongoing debate in the background about whether regulation tends mostly to limit innovation, or alternatively is needed to guide it to positive outcomes (and, of course, whether an outcome is perceived to be positive will depend on whatever idea of the public good is currently sanctioned).

We suggest the translation might look like the following:

- **Open AIMS:** Fostering openness and transparency, common ownership and collaboration, interoperability.
- **Bourgeois AIMS:** Fostering rights and civility with procedural rules and codes.
- **Paternal AIMS:** Mandating outcomes and confining uses.
- **Commercial AIMS:** Allowing market solutions to resource allocation problems.
- **Hacker AIMS:** Libertarian, anti-authoritarian, decentralised approach valorising software skills, resisting censorship, and empowering individuals and communities to make and reshape the information space.

Note that, whereas in the 4Is framework the spoiler model did not create an Internet so much as free ride on the others, the Hacker AIMS *will* produce actual AI systems that function. Hence, in the Five AIMS framework, the analogue of the spoiler model is a first-order object.

While these AIMS are all driven by different ideas of the public good, of course it would be quite possible for AI systems of many kinds, or for many purposes, to be built within each governance regime. For instance, an AI system that generated abstracts for scientific papers would probably be consistent with each of the AIMS, although some of them might, for instance, emphasise the protection of intellectual property or the enforcement of data protection rights more strongly than others. On the other hand, a deepfake algorithm to create non-consensual sexual content from images of existing people would probably be excluded by the Bourgeois and Paternal AIMS, and possibly from the Open and Commercial AIMS as well.

Examples of each of the AIMS can be found in recent proposals for regulation or governance. For instance, the movement for open source generative AI models would be an example of Open AIMS (see for instance Mark Zuckerberg's statements, although his interpretation and implementation of "open source" AI is debated, as discussed below).[14] There are many different approaches to the preservation of rights in the style of Bourgeois AIMS, including parts of the EU's AI Act.[15] However, the focus on risk and safety, also present in the AI Act, as well as President Biden's executive order of 2023[16] and Chinese attempts to regulate generative AI, shows influence of the Paternal AIMS.[17]

Concerns that premature regulation could hinder innovation follow from the Commercial AIMS, as are the promotion of regulatory means to protect business models, for example, by raising barriers to entry to the industry.[18] At the time of writing, some analysts anticipate that the incoming Trump administration in the US is likely to relax agency regulation on AI, repeal President Biden's executive order, and reduce antitrust enforcement, all from the Commercial AIMS playbook.[19] President Trump may, however, also follow some Paternal AIMS, with export controls to restrict Chinese access to cutting-edge technology.

Meanwhile, those approaches aligned with the Hacker AIMS (i) see disruption as the real opportunity for AI, as opposed to viewing it conservatively as a breach of contextual integrity to be managed, (ii) are concerned that only a small number of tech giants and governments will have effective control and would prefer a more widespread developer base, or (iii) actively revel in its subversive potential.[20] Some are enthused by the possibility of independent, decentralised communities resisting censorship,[21] while at the governmental scale, the use of AI in defence and warfare for decision support may invoke speeds and scales that would seriously challenge the protections of current international law.[22]

QUESTIONS FOR RESEARCH AND CLARIFICATION

The above is a sketch of how the 4Is framework might apply to the governance of AI. The AIMS, like the ideal types of 4Is, are narratives at a high level of abstraction. As narratives, the AIMS do not aspire to crystal clarity, but rather are intended as sensemaking aids; the characterisations above, while they may already help categorize and interpret discourses, stances, and proposals, are doubtless too broad to determine actual policies. However, they are also, as they stand, open to a number of competing interpretations within the AI context; they remain anchored to the Internet governance context and need to be fine-tuned to the relevant issues in AI.

The purpose of this briefing is not to resolve such issues, but to set out some of the research challenges required to cement the Five AIMS as a suitable framework for understanding the governance of AI. There are no doubt many of these, but for the purposes of this paper we can illustrate them by concentrating on four: the significance of orthogonal goods within the Five AIMS; the understanding of openness within the generative AI paradigm; safety; and nationalism.

Orthogonal Goods

The AIMS focus on identifying the primary beneficiaries of the AI, their autonomy in determining how they are affected by AI, and setting out standards or arbiters for measuring its positive effects. This focus on the agents and patients of the applications of AI supplies structural desiderata, but is consistent with divergent attitudes to the actual goods/evils it produces.

Examples of such orthogonal goods might include:

- **Distribution.** E.g. should the benefits of AI disseminate across the world, including the global South, and conversely does risk management include the global South?[23]
- **Explainability.** E.g. what attitude should the proponents of these AIMS have toward the black-box concerns of observers, and what solutions might be proposed?[24]
- **AI’s carbon footprint and other environmental parameters.** E.g. should the development of AI models be curbed or slowed in order to reduce model builders’ power consumption?[25] Or is AI an essential tool for predicting and designing policy interventions for addressing climate change?[26]

On all these issues, will it be possible to create the requisite international frameworks in an increasingly transactional world where rule-based order is falling out of fashion? These are additional issues to those that the AIMS try to address, and divergent stances might be taken within each AIMS.

On the other hand, the nature of the AIMS might tend to favour certain approaches to some of these problems to others. For instance, with the explainability of AI, the Bourgeois AIMS might focus on the rights of those affected by AI to understand (and thus assess responsibility in) the processes by which decisions were made about them. The Open AIMS might be similarly concerned with the transparency of the algorithm, though more generally focused on the rights of all to audit (and perhaps copy) it. The Paternal AIMS might favour explanations for governance bodies, administrators and policymakers, rather than those affected. Within the Commercial AIMS, the concern might be that insisting on explainability could prevent valuable but opaque calculations from being implemented, or alternatively that legal clarity would let firms deploy AI-driven processes knowing precisely how to avoid liability, especially with regulations limiting developers' responsibility. The Hacker AIMS may be relatively uninterested in explanation, preferring instead to exploit the potential of the technology without qualms about detail. Indeed, from the Hacker AIMS perspective, a requirement for explanation may be seen as a means to render programmers open to intervention and control by government. None of these preferences is necessarily dictated by the AIMS, although we can see there are affinities between certain AIMS and certain policies.

Similarly with environmental impact, it may be that Open AIMS suggests transparency without dictating a particular response. Bourgeois AIMS and Paternal AIMS might mandate policies to reduce carbon emissions. Commercial AIMS may favour mechanisms such as tradeable carbon credits, or alternatively may support the scaling up of AI model development to integrate energy provision into the process (as, for example, the deal in 2024 between Microsoft and Constellation Energy to restart a retired nuclear facility at Three Mile Island).[27] Such deals as this may raise the barriers to entry to competitors in the production of AI models and applications, but whether that is a concern of the Commercial AIMS will depend on which commercial interests are held paramount – e.g. should the commercial imperative be total profit however distributed, or maximised competition?

Openness

The notion of openness is reasonably well understood in many computational and commercial domains, but tends to be treated in isolation within these domains, without drawing connections to a general organising principle.[28] In particular, while open source development is an important species of openness in computing, exactly how it translates into foundational deep learning model development is disputed, because training AI models is a very different process from software development.

To meet the condition of traditional open source software - in simple terms, that the source code is free to access, modify, and redistribute for everyone - AI developers would have to give full access to not only the code, but also, among other things, the weights associated with its nodes, as well as the data upon which it is trained.[29] Current approaches, such as restrictively releasing certain aspects of a model (often its weights), or imposing access controls or APIs, may be more practical from the point of view of building the industry, collecting innovation rent, and protecting competition advantage, but it is hardly frictionless. Nonetheless, touting a model as open source can be a marketing or legal argument (for instance, the EU AI Act provides exemptions for open source AI systems), and many AI developers have been criticised for trading openness for 'just-open-enoughness' and engaging in 'open-washing'.[30] While the Open Source Initiative (OSI) recently offered a definition of open source AI as systems that can be used, studied, modified, and shared,[31] there is an ongoing struggle over the meaning of 'openness' at the core of the Open AIMS.

Another point of contention is that, in the context of the Internet, 'open' often means de- or unregulated, but this does not seem to square with the ideas of those who promote openness in AI. States have often advanced regulatory frameworks while encouraging open source AI ecosystems, while many companies developing closed models lobby for light-touch regulation or self-regulation. In the case of open source systems, even on the OSI definition, there are ambiguities: open source models could be promoted as more auditable and thus *more* compatible with regulation, since compliance can be more transparently assessed and enforced, or, conversely, *less* prone to governance, as open licences make models available for anyone to reuse and adapt, making it easier to circumvent technical safeguards and hijack a model for malicious uses, decreasing the control of initial developers over their systems and thus diluting their legal responsibility regarding harmful downstream impacts.

Furthermore, there are also clear overlaps between Open AIMS and Hacker AIMS, in that the latter also supports certain aspects of the openness agenda, including transparency, the ability to share, decentralisation, and so on. Quite how to make the distinction between Open AIMS and Hacker AIMS, or even if there is a principled distinction between them at all, will require further investigation.

Safety

AI safety is strongly associated with Paternal AIMS, even when the notion of safety is underspecified. For China, safety involves social stability, a view whose Confucian roots are integrated with the imperative of preserving the Chinese Communist Party's position at the apex of the state hierarchy.[32] In the UK, the Online Harms Act focuses on harm prevention to individuals. Hence there are variations of scale within safety as conceived.

Furthermore, there may also be variations in the temporal dimension as well. Some concerns are reactions to long-term impacts that may have existential import, such as the proliferation of misinformation undermining the functioning of democracy. Others may be driven by immediate risks or impacts, such as new types of AI-enabled cybercrime like deepfake phishing.[33]

Nationalism

Paternal AIMS also covers nationalistic plans to reshore aspects of the AI industry, such as training, semiconductor production and infrastructure.[34] In India, the state created a digital infrastructure via the ID system Aadhaar and the UPI payments interface, and mandated the holding of data on Indian citizens in India, now bolstering its project of e-government chatbots that work with Devanagari and other scripts. Control of the AI stack is also a strategy to support India's export of digital and AI technology to developing countries, and so a component of its strategy for leadership of the Global South. The EU is giving incentives to small startups to use its burgeoning fleet of supercomputers for AI development.[35] Gulf States, including UAE, Qatar and Saudi Arabia also encourage their domestic AI champions; even Bhutan is making a play to be a hub for AI and fintech.[36] Paternal AIMS encourages states and civil society to structure the AI ecosystem in congenial ways, from semiconductors to super-apps, as well as supporting education, training and partnerships between industry and academia, thereby spanning the entire AI value chain.

The US has previous experience of taking a regulatory journey starting with commissioning and guiding research on emerging technologies (often through military support), holding back from regulation until the potential of those is better known, and then moving to a more active regulatory stance, often negotiated with industry leaders, or even based on self-regulation. China and India both have fostered strong domestic tech companies, but in China they are subject to arbitrary swings of policy from the Communist Party, while in India it is unlikely to be coincidence that its tech giants work closely with government and align with its *Hindutva* principles of national self-reliance and championing elite companies.[37] On this pattern, more open approaches can morph into paternalism along the way.

Finally, in those countries such as China, India and the United States which promote nationalist industrial policy by encouraging national champions, inevitably the Paternal AIMS will be influenced by Commercial AIMS and *vice versa*. In such economies, the distinction between the two may be hard to make out.

CONCLUSION

To conclude, the 4Is framework may carry over in interesting ways into the context of AI governance. This paper has explored some of the possibilities of understanding 4Is in terms of Five AIMS (Artificial Intelligence Management Strategies), and argued that there is a *prima facie* case for the value of such a translation. However, the Five AIMS are at a high level of abstraction, and greater precision will require the investigation of a number of open issues.

[1] Borenstein, Jason, et al. "AI Ethics: A Long History and a Recent Burst of Attention." *Computer*, vol. 54, no. 1, 2021, pp. 96-102.

[2] Azhar, Azeem. "Why Humanity Needs AI." *Exponential View*, 7 Sept. 2024.

[3] See O'Hara, Kieron, and Wendy Hall. "Four Internets: The Geopolitics of Digital Governance." *CIGI Papers*, no. 206, Centre for International Governance Innovation, 2018, and the full book O'Hara, Kieron, and Wendy Hall. *Four Internets: Data, Geopolitics, and the Governance of Cyberspace*. Oxford UP, 2021.

[4] An ideal type is a methodological tool popularized by German sociologist Max Weber (1864 - 1920). In simple terms, it is a concept capturing common characteristics of different cases of a phenomenon, not to describe it perfectly but to facilitate categorization and comparison.

[5] The label 'bourgeois' was chosen, as opposed to, say, 'liberal' or 'civil', to emphasise the historical contingency of the growth of civil society alongside the emergence and flourishing of the bourgeoisie in late medieval/early modern Europe, and to distance the account from universalist narratives of liberalism and human rights.

[6] Himanen, Pekka. *The Hacker Ethic and the Spirit of the Information Age*. Vintage, 2001.

[7] See O'Hara and Hall's upcoming paper on the topic: O'Hara, Kieron, and Hall, Wendy. "Five AIMS – Artificial Intelligence Management Strategies." Atlantic Council GeoTech *Center AI Connect II* blog, 2025.

[8] O'Hara, Kieron. "Digital Modernity." *Foundations and Trends in Web Science*, vol. 9, nos. 1-2, 2022, pp. 1-254.

[9] Bradford, Anu. *The Brussels Effect: How the European Union Rules the World*. Oxford UP, 2020.

[10] Snyder, Timothy. *The Road to Unfreedom: Russia, Europe, America*. Bodley Head, 2018.

[11] Cornish, Chloe, and Hammond, George "Microsoft to Invest $1.5bn in Abu Dhabi AI Group G42", *Financial Times*, 16 Apr. 2024

[12] Although the need to fund and/or access enormous computing infrastructures might be overestimated, if the quality and training method of Chinese chatbot DeepSeek is verified and proves replicable (see Milmo, Hawkins and Kollewe, "Global tech shares fall as China AI chatbot DeepSeek spooks investors", *The Guardian*, 27 January, 2025).

[13] For example, see on data centers and energy supply, see the work of the International Energy Agency's What the data centre and AI boom could mean for the energy sector (2024), or Alex Lawson's "Google to buy nuclear power for AI datacentres in 'world first' deal" (*The Guardian*, 2024), on chip manufacturing, Michael Acton and Tim Bradshaw "Amazon steps up effort to build AI chips that can rival Nvidia" (*Financial Times*, 2024).

[14] Zuckerberg, Mark. "Open Source AI Is the Path Forward." *Meta Newsroom*, 23 July 2024.

[15] See for instance: Gregory, Sam. "Fortify the Truth: How to Defend Human Rights in an Age of Deepfakes and Generative AI." *Journal of Human Rights Practice*, vol. 15, no. 3, 2023, pp. 702-14, Geiger, Christophe. "Elaborating a Human Rights-Friendly Copyright Framework for Generative AI." *IIC*, vol. 55, 2024, pp. 1129-65, Kusche, Isabel. "Possible Harms of Artificial Intelligence and the EU AI Act: Fundamental Rights and Risk." *Journal of Risk Research*, 2024, or Novelli, Claudio, et al. "Generative AI in EU Law: Liability, Privacy, Intellectual Property, and Cybersecurity." *arXiv*, Jan. 2024.

[16] United States, Executive Office of the President [Joseph Biden]. Executive Order 14110: on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence. 20 October 2023

[17] On this, see Sheehan, Matt. "China's AI Regulations and How They Get Made." *Horizons*, vol. 24, 2023, pp. 108-25, Novelli, Claudio, et al. "Taking AI Risks Seriously: A New Assessment Model for the AI Act." *AI and Society*, vol. 39, 2024, pp. 2493-97, or Pham, Bao-Chau, and Sarah R. Davies. "What Problems Is the AI Act Solving? Technological Solutionism, Fundamental Rights, and Trustworthiness in European AI Policy." *Critical Policy Studies*, 2024.

[18] To that regard, Andreessen, Marc. "Why AI Will Save the World." *Andreessen Horowitz*, 6 June 2023, Wheeler, Tom. "The Three Challenges of AI Regulation." *Brookings Commentary*, 15 June 2023, and Gikay, Asress Adimi. "Risks, Innovation, and Adaptability in the UK's Incrementalism versus the European Union's

Comprehensive Artificial Intelligence Regulation." *International Journal of Law and Information Technology*, vol. 32, no. 1, 2024, offer completing points of view.

[19] Villasenor, John, and Joshua Turner. "AI Policy Directions in the New Trump Administration." *Brookings Commentary*, 14 Nov. 2024,.

[20] Lu, Christina, et al. "Subverting Machines, Fluctuating Identities: Re-Learning Human Categorization." *FAccT '22: Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency*, ACM, 2022, pp. 1005-15, Manu, Alexander. *Transcending Imagination: Artificial Intelligence and the Future of Creativity*. Chapman and Hall/CRC, 2024, or Wong, Wilson Kia Onn. "The Sudden Disruptive Rise of Generative Artificial Intelligence?" *Journal of Open Innovation*, vol. 10, no. 2, 2024.

[21] St John, Rex. "Understanding Shoggoth Network." *Medium*, 22 Nov. 2023, then more convincingly Hagemann, Johannes, et al. "INTELLECT–1: Launching the First Decentralized Training of a 10B Parameter Model." *Prime Intellect*, 11 Oct. 2024.

[22] "How AI Is Changing Warfare." *The Economist*, 20 June 2024,

[23] Okolo, Chinasa T., et al. "Making AI Explainable in the Global South: A Systematic Review." *COMPASS '22: Proceedings of the 5th ACM SIGCAS/SIGCHI Conference on Computing and Sustainable Societies*, ACM, 2022, pp. 439-52, Png, Marie-Therese. "At the Tensions of South and North: Critical Roles of Global South Stakeholders in AI Governance." *FAccT '22: Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency*, ACM, 2022, pp. 1434-45.

[24] O'Hara, Kieron. "Explainable AI and the Philosophy and Practice of Explanation." *Computer Law and Security Review*, vol. 39, 2020, and previously cited Okolo, 2022.

[25] Faiz, Ahmad, et al. "LLMCarbon: Modeling the End-to-End Carbon Footprint of Large Language Models." *arXiv*, Sept. 2023.

[26] Chen, Lin, et al. "Artificial Intelligence-Based Solutions for Climate Change: A Review." *Environmental Chemistry Letters*, vol. 21, 2023, pp. 2525-57.

[27] Shaw, Alfie. "Microsoft and Constellation Sign PPA for Three Mile Island Restart." *Power Technology*, 23 Sept. 2024.

[28] Splitter, Violetta, et al. "Openness as Organizing Principle: Introduction to the Special Issue." *Organization Studies*, vol. 44, no. 1, 2023, pp. 7-27.

[29] Liesenfeld, Andreas, and Mark Dingemanse. "Rethinking Open Source Generative AI: Open-Washing and the EU AI Act." *FAccT '24: Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency*, ACM, 2024, pp. 1774-87.

[30] Nolan, Michael. "Llama and ChatGPT Are Not Open-Source: Few Ostensibly Open-Source LLMs Live Up to the Openness Claim." *IEEE Spectrum*, 27 July 2023.

[31] See the Open Source Initiative's definition of open source in AI.

[32] See Kieron O'Hara & Wendy Hall (2021). Four Internets: Data, Geopolitics, and the Governance of Cyberspace, New York: Oxford University Press, pages 127-133, or Kirk, H.R., Lee, K. & Micallef, C. The Nuances of Confucianism in Technology Policy: an Inquiry into the Interaction Between Cultural and Political Systems in Chinese Digital Ethics. *Int J Polit Cult Soc* 35, 129–152 (2022)

[33] Banzaca, Jennifer. "Your Zoom Call with Your Executive Team? They Were AI Bots." *Private Funds CFO*, 3 June 2024.

[34] Saheb, Tahereh, and Tayebeh Saheb. "Topical Review of Artificial Intelligence National Policies: A Mixed Method Analysis." *Technology in Society*, vol. 74, 2023.

[35] Jülich & Seattle "Europe wants startups to do AI with supercomputers", *The Economist*, March 2024

[36] "Bhutan prays it can be India's Hong Kong", *The Economist,* October 2024

[37] On the Chinese strategy, see de La Bruyère & Picarsic, "A 'techlash' with Chinese characteristics", *TechCrunch*, November 2021 and on India, see Priya Chacko, "Disciplining India: paternalism, neo-liberalism and Hindutva civilizationalism", *International Affairs*, 99(2), 551-565, 2023.