NETWORK LAW REVIEW

ARTICLE

Governing Hyperobjects: The New Economics of AI Regulation

Jason Potts

The Network Law Review is pleased to present a special issue entitled "The Law & Technology & Economics of AI." This issue brings together multiple disciplines around a central question: What kind of governance does AI demand? A workshop with all the contributors took place on May 22–23, 2025, in Hong Kong, hosted by Adrian Kuenzler (HKU Law School), Thibault Schrepel (Vrije Universiteit Amsterdam), and Volker Stocker (Weizenbaum Institute). They also serve as the editors.

***

Abstract. Artificial intelligence is not simply a hard regulatory problem, but a fundamentally different kind of object – a hyperobject. Like global warming or the internet, AI operates at scales of space, time, and complexity that exceed the conceptual and institutional boundaries of conventional regulatory systems. Traditional law and economics frameworks that are designed for bounded, legible, and externally governable systems struggle to meaningfully engage with AI because they assume a world made of regulatable parts. Hyperobjects defy this logic. They are nonlocal, interobjective, and partially withdrawn, making them illegible to models that rely on discrete causality, control from the outside, and stable agents.

***

1. Regulating AI is a hard problem

By artificial intelligence (AI) I mean the path from transformer architecture (Vaswani et al. 2017) to LLMs (Radford et al. 2019) to agents (Immorlica et al. 2024, Wang 2024). Plausible scenarios of how this path to autonomous agents might develop in the near future are set out in Kokotajalo et al. (2025). They project that the impact of superhuman AI will be enormous and will happen fast (by perhaps 2028), being driven by a progress multiplier as AI is used to generate further progress in AI. So, while the concept of artificial intelligence has been the touchstone of progress in computer science and the lodestone of speculative philosophy since the 1950s, the profound consequences of AI for economies and society come from the deployment of AI in agents. This is an extremely recent and rapidly accelerating development.

AI regulation presents arguably the most complex and high-stakes challenge in law and economics today. It is a perfect storm of hard problems that combines radical technological uncertainty, global scale, and sprawling externalities – ranging from private harms (job displacement, privacy violations, access inequality) to systemic risks (alignment failure, security threats, macroeconomic disruption). At the same time, AI is driving the most transformative technological shift of the century, while also polarising culture and politics (Jacobides et al 2021, Agrawal et al 2022, Cohen et al 2024, Jones 2024, Kolt 2025, Damioli et al 2025). Crafting legislation and regulatory frameworks to manage these effects is difficult enough. Adding to the challenge is the widespread assumption that AI is fundamentally unlike anything before – that it is, in regulatory terms, maximus sui generis. Yet the claim I make is that AI is not a sui generis regulatory challenge at all, but of a class of things that are extremely hard to regulate. The name for this class, which also includes global warming, planetary ecologies, nuclear waste, internet and blockchains, is hyperobjects. This paper introduces the (post-Heideggerian) philosophical concept of hyperobjects (first proposed by Morton 2013) to a law and economics audience. I show why AI is also a hyperobject and explain the problem that hyperobjects pose to regulatory regimes, and, crucially, why they can't contain them.

The standard model in the economics of regulation is built to address so-called market failures – namely, externalities, information asymmetries, and public goods. While this is an expansive set, there is a conventional line from theory to policy in order to intervene to correct an analytically identified externality. This policy intervention works through instruments that include Pigouvian taxes, subsidies, liability regimes and ex ante regulatory rules, and which are implemented and enforced by public institutions (legislatures, regulators, courts, etc). The 'market-failure' framework has proven effective in addressing externalities over a wide range of economic domains. However, when applied to AI, it faces profound limitations.

The challenge is not merely the accumulation of regulatory difficulties, such as uncertainty about safety and alignment, the pace of technological development, conflicts over privacy and data, labour displacement, innovation races, or global security risks, although each presents a formidable problem on its own and more so in combination. The problem is that these are all effects of a deeper structural issue.

The problem lies in the reality of the object being regulated, or in philosophical terms in its ontology. AI is not a conventional economic or legal object – it is not a regular thing. It is, I claim, of the class of objects that Morton (2013) terms a 'hyperobject' – an entity of massive spatiotemporal scale, partially beyond human perception, distributed and entangled with other systems in ways that defy conventional comprehension and control. Hyperobjects resist clear delineation, stable causality, and straightforward attribution of responsibility. They are illegible to the institutional architectures that underpin modern regulation, which are premised on observable, bounded, and controllable phenomena. The problem with AI is that it just looms so much larger than normal human-scale regulatory objects that have been the previous targets of even the most ambitious regulatory policies of the past century. While regulation can effectively target regular objects within bounded domains, hyperobjects exceed the epistemic and jurisdictional scope of traditional regulatory regimes. We have very little experience at regulating an object of this scale, or at this scale of growth.

In this context, the appropriate response is not regulation, in the conventional public law sense, but governance, which is a mode of institutional ordering grounded in decentralized, community-driven rule formation, as articulated by Ostrom (1990), and extended into digital domains by scholars like Lessig (1999) under the maxim "code is law." The claim is therefore that while hyperobjects might not be regulatable, owing to their inherent properties, they could be governable by the affordances of those same properties. Specifically, the emergence of computational governance tools such as programmable smart contracts, algorithmic institutions and digitally enforced coordination protocols (de Filippi et al 2024, Nabben et al 2024, Schneider 2024, Rennie and Potts 2024) offers a way to govern hyperobjects. These tools enable embedded rules that are enacted through technical architectures rather than through external enforcement. Understood this way, the regulatory challenge of agentic and burgeoning superintelligent AI is not a matter of failed policy design (McGinnis 2022, Guerreiro 2023), but a mismatch between institutional form and ontological object. As a general principle, then, AI is not a 'thing' to be controlled through exogenous intervention; it is a non-human being that increasingly operates with autonomy, agency, and persistence across scales. What is required is to attempt shared rule-making and adaptation with a fundamentally new

class of agent, i.e. to approach the ordering problem of AI not as an external imposition of regulatory control but as a problem of governance.

2. Hyperobjects

An object is something material that can be seen and touched by humans. Objects in the economy for instance are things like commodities, and in law they are things that are the subject of legal rights or responsibilities. Now there exist things that are not like this because they are non-material (e.g. the concept of rights, or the idea of value or justice). Here it makes no sense to speak of the location or size of these things in the way one does of regular objects, i.e. in reference to the human scale of being. But hyperobjects are real in the material sense but have a different existence with respect to humans, even though created by human action and woven from human artifacts. Hyperobjects are a new and recent class of existence in the world. They present not only philosophical challenges, but also legal and economic challenges too. But they also provide a new framework for understanding new phenomena such as AI.

In math and computer science, a hyperobject is a high-dimensional and non-local entity, such as certain types of data structure or algebraic manifold. However the analytic concept of a hyperobject we use here is the philosophical concept of a hyperobject which was introduced by Tim Morton in a remarkable 2013 book Hyperobjects: Philosophy and Ecology after the End of the World. This book was a work of ecological critique from the phenomenological perspective of Husserl and based on Heideggerian object-oriented ontology (Morton 2011, Harman 2010, 2018). Morton uses the term 'hyperobject' to name entities or systems so vast in time and space that they "defeat traditional ideas about what a thing is."[1] For instance, global warming is real and has tangible effects – so is clearly a thing – yet no one can hold the climate in their hand or fully grasp it at once.[2] Indeed, climate can only be seen with complex computational models. Hyperobjects are "massively distributed in time and space relative to humans." Hyperobjects transform how we think about human coexistence with the world and, as Morton explains, confronting these immense non-human entities is central to understanding the Anthropocene, the era

[1] "A hyperobject could be a black hole. A hyperobject could be the Lago Agrio oil field in Ecuador, or the Florida Everglades. A hyperobject could be the biosphere, or the Solar System. A hyperobject could be the sum total of all the nuclear materials on Earth; or just the plutonium. A hyperobject could be the very long-lasting product of direct human manufacture, such as Styrofoam or plastic bags, or the sum of all the whirring machinery of capitalism. Hyperobjects, then, are 'hyper' in relation to some other entity, whether they are directly manufactured by humans or not." (Morton 2013)

[2] "Naturally humans have been aware of enormous entities – some real, some imagined – for as long as they have existed. ... but there is something quite special about recently discovered entities, such as climate." (Morton 2013)

of large-scale human impact on Earth. Morton explains that hyperobjects impact not only environmental philosophy but also politics, art, culture. To this I further suggest they also impact law and economics through the arrival of new technologies in the form of hyperobjects (internet, blockchains, AI).

In Morton's speculative realist account, hyperobjects are vast, non-human phenomena that stretch across planetary spatial scales and geological timeframes. They defy normal human perception due to their massive scale or temporal depth. Hyperobjects are not abstractions or metaphors; they are real entities whose full scope is never available to any single observer. In Morton's theory (which builds on Heidegger), hyperobjects possess viscosity, meaning they stick to us and entangle our lives in ways that are difficult or impossible to escape. Morton describes hyperobjects in general as being nonlocal, distributed across time and space so we only ever encounter fragments or local manifestations of the whole, along with a stretched or warped sense of time around them. They exhibit phasing, appearing and disappearing in ways that reflect their higher-dimensional existence: they are real and active even when not directly observable, surfacing unpredictably in strange patterns. They are interobjective – not reducible to a single substance or location, but formed through the entangled relationships among many things, visible only through the collective effects they generate. To live with hyperobjects is to be deeply entangled with nonhuman objects and forces that we are responsible for.

Morton develops this ontology as a general class. But Morton does so to propose a new philosophical perspective on global warming, climate change and planetary ecologies (the book's subtitle is 'philosophy and ecology after the end of the world'). The argument I'm making here is to extend that concept to AI. Of course, any particular instance of an AI codebase or ML architecture in production is easy to re-objectivity and localise as a regular object that can be targeted, contained and regulated. The 'hyperobject-ness' of AI is not seen in any isolated specimen but in how it is developing and evolving into a new global-scale phenomenon with the emergent properties of a hyperobject that Morton describes. The Anthropocene, in Morton's view, is not the age when humans become a geological force, which is true but parochial. Rather, the Anthropocene is better understood ontologically as the time when nonhuman beings "make decisive contact with humans" – these hyperobjects, which "are responsible for the next moment in human history and thinking … with their towering temporality, their phasing in and out of human space and time, their massive distribution, their viscosity" – began to live among us as "real entities whose primordial reality is withdrawn from humans". Morton explains:

"Hyperobjects are what has brought about the End of the world. Clearly planet Earth has not exploded. But the concept of world is no longer operational and hyperobjects are what has brought about its demise."

"To those great Victorian discoveries, then – evolution, capital, the unconscious – we must now add spacetime, ecological interconnection, and non-locality."

"Hyperobjects seem to force something on us, something that affects some core ideas of what it means to exist, what Earth is, what society is. … They are entities that become visible through post-Humean statistical causality."

Hyperobjects are massive but finite nonhuman 'beings'. The radical philosophy of hyperobjects is indeed to propose a new sort of 'being' in the world – and that the arrival of these new nonhuman beings is the defining mark of the Anthropocene. There is no place 'outside' from which to look in at the 'world'. Nature is not 'over there'. Hyperobjects require us to abolish the idea of metalanguage to account for things while remaining uncontaminated by them. This dethronement is what Morton means by 'after the end of the world' and of the need to develop an ecology without nature. We must learn to live with hyperobjects now, which is hard because it breaks modernist categories of inside and outside and of simple causality.

Compared to regular objects, hyperobjects are strange. The challenge for object oriented ontology (Harman 2010) is to furnish a new relationship with these strange nonhuman beings. We relate to 'regular objects' naturally, as things in a world, as knowing subjects. We know these objects get weird at very small scales (quanta, when they stop being things), and also with extremes of energy, when very heavy or very fast (when space and time becomes viscous around them). None of those situations overlap with ordinary human 'being and time' (what Heidegger called 'Dasein'). But hyperobjects are like this with us – they are 'room temperature' objects, yet have these strange properties. We are still not used to living with hyperobjects yet we are now inside them. We never see them directly or all at once as they withdraw to higher dimensions. Due to their scale, hyperobjects 'withdraw' from human experience (being, in the Heideggerian sense) in some of their higher-dimensional existence. This withdrawal can be partially recovered, e.g. through computational modelling, as with climate forecasts, but the base reality is that human 'being' is lower dimensional than the 'being' of hyperobjects. This is what makes them strange and illegible. Being profoundly futural (e.g. global warming is a problem because of what happens in the future but we can detect in the present, through models), we detect their presence as they phase in and out around us, reaching back from the future to shape human action now. To see them we need devices, tools, measures, models, which is to say that they are visible only through our tools and technologies. Hyperobjects emit 'zones' or

directives, telling humans how to act toward them. Rational choice struggles with hyperobjects. Utilitarian calculus or any social welfare functions constructed from them fails due to the unknowable, incalculable consequences that ramify from actions into the future and back to now. Hyperobjects also break the idea of regulation of 'a world', over there, which is objectively legible and regulable. We are withinhyperobjects, and can only govern. Regulation requires an objective system – as in control theory, which is the most general formulation of the theory of institutions (Zargham and Ben-Meir 2024). But hyperobjects precisely collapse the idea that our world is made of systems that can be contained. AI is such a thing.

3. AI is a hyperobject

AI is our newest hyperobject. Of course it builds on earlier compute hyperobjects, such as global telecommunication networks and internet protocols, and is adjacent to blockchains. The public policy significance of this is that hyperobjects are unregulatable because our institutions can't see them (nor can we, unaided). But hyperobjects are possibly governable with new tools. The Internet, blockchain networks, and AI models are all vast digital infrastructures that share the hallmarks of hyperobjects. They are pervasive, massively networked, and exceed any single individual's understanding or control, which has long been true of modern technology infrastructure. But new properties are emerging as these systems begin to interleave and ramify in their capabilities as they reach from the past (training data) and extend into the future (blockchains and smart contracts). The standard economic description of these objects as networks, which are also vast distributed phenomena with no clear boundaries, central location, or fully graspable form. They are nonlocal, existing across countless nodes, servers, and devices worldwide, with no central location and no single vantage point from which their totality can be perceived. Users and participants only ever interact with fragments. Further evidence of hyperobjects is that these systems exhibit viscosity, touching nearly every aspect of modern life, with disentanglement nearly impossible. Digital immutable records persist, resisting deletion or rollback. Their temporal dimension is deep and strange. Data histories and social graphs propagate forward, trained models pull the past into the future. AI models operate across deep time, encoding past knowledge and shaping long-term futures with unexpected resurfacings of forgotten data or patterns. Blockchains enable common knowledge to work at a vast scale. These systems display phasing as they manifest differently depending on perspective and moment.

Hyperobjects can appear abstract and ephemeral to one observer, concrete and infrastructural to another. AI's often remain invisible unless revealed, behaving differently depending on who interacts with it and how (they are phased objects). They are profoundly interobjective: neither the Internet nor blockchains are singular

entities, but emergent assemblages of code, machines, energy, institutions, and human activity. AIs intelligence is not located in one place but in a sprawling, dynamic network of nested embeddings of all human digital culture (Potts 2022). The presence of the internet, blockchains or AI is often only revealed by interacting with it. They exceed the scale and scope of individual understanding, operating as planetary-scale infrastructures that continually reshape their environments, behaviors, and meaning systems. We are never outside them.

Digital is a lingua franca – anything digital can send messages and instructions over anything else digital. A digital economy means a communication and compute layer over all objects in the economy that is integrated with digital institutions. In the industrial economy, capital became a hyperobject. In the post-industrial digital economy, institutions as digital infrastructure – digital money, smart contracts (DAOs, DeFi, etc) and token compute systems – are evolving into a new hyperobject that is increasingly machine-governed and illegible to individual economic agents. Blockchain and AI is humanity's first experience of digital hyperobjects in economic form that could persist for hundreds of years and move economic transactions off-world. Economic objects can be large and long-lived for two reasons: one, the asset does not experience entropy in use, such as land; or two, the governing institutions that protect and enforce the claims on the asset or property right persist on human historical scale (Williamson 2000). Internet plus blockchains plus machine intelligence break this limitation with institutional hyperobjects, enabling information or code that shapes economic coordination to persist and express, wherever it goes and in whatever form it takes at unlimited scale. A digital economy is different because it has a strange core of hyperobjects.

4. Economics of hyperobjects

Traditional economic and legal frameworks were designed to address problems at a human scale: clear property rights, local transactions, individual liability, and linear causality. Hyperobjects do not fit this model. They are large-scale, networked, withdrawn and stochastic (high dimensional), and yet persist over long timeframes, making them difficult to locate, assign responsibility for, or govern using existing tools. Hyperobjects require a new way of thinking about the boundaries of organisations, limits of agency, extent of property rights, contracts and exchange, and the meaning of externalities. In the standard object model, economics assumes that agents act locally (in firms, in markets), that goods and externalities are relatively bounded, and that economic objects are stable and finite (e.g. economic goods, well-defined property rights, legibility to institutions). Hyperobjects melt these normally solid assumptions. They produce unbounded externalities that permeate all economic activity. They make the environment an embedded system.

They dilute agency, no individual actor controls them, requiring new forms of open system governance.[3] They operate on deep timescales, demanding long-term thinking that standard discounting struggles to accommodate.

Engaging with hyperobjects requires new epistemic infrastructures – models to perceive them, platforms to interact with them, and computational rule systems to coordinate around them. They require rules (not values) to organise them (Miyazano 2025).[4] An economics of hyperobjects needs to engage with entangled, long-range, collective 'deep' or 'axial' phenomena that traditional tools can neither see nor represent.[5] Economic systems and legal frameworks for hyperobjects will need to evolve and adapt together, likely involving new institutions and principles that are more flexible and adaptive.[6] A basic problem that of identity and legibility of hyperobjects – measures, visualisations and other points of contact – in order to attach governance. This is the role of property rights in a market economy, and the identity of agents in a political and legal order (Akerlof and Kranton 2000), as the object identifiers for exchange and contract (i.e. as the pointers and references for transactions). The classic definition of externalities and public goods is everything that escapes this, which in turn defines the subject matter of the economics of regulation. Hyperobjects are pure and untamed externalities in this sense. So a critical first step is toward creating a language (a semiosis) with which to interact with AI hyperobjects.

5. Simple economics of AI regulation

AI regulation is typically implemented by governments under legislative authority to address a wide range of policy goals. At the micro level, regulation addresses individual safety and rights, such as protecting privacy and fairness in algorithmic decision-making. At the macro level, regulation addresses systemic risks related to AI alignment, national security, economic competitiveness, and speculative existential threats (e.g. AGI singularity). These regulatory objectives span multiple legal domains, including tort law, competition law, administrative law, and socio-technical governance. In addition to public-facing concerns, governments also

[3] See Rennie and Potts (2024) on contribution systems and open organisations.

[4] Formal verification of properties is a rules based approach to alignment, rather than trying to achieve safety through endeavours to put values into machines, whether as goals or as constraints. A formal verificationist approach seeks only to show that a logical system has particular properties that relate to, e.g., safety from certain classes of attack.

[5] See Dopfer and Potts (2024) on evolutionary semiosis, writing as first hyperobject, digital writing as latest hyperobject. 'Axial' in the sense of Jaspers (1948).

[6] We already see some movement: international climate tribunals have been proposed to handle disputes related to climate damage; some jurisdictions have given legal personhood to rivers, glaciers, and forests (e.g. New Zealand's Whanganui River) as a way to ensure these sprawling entities can be represented in court (Wade Smith 2024).

have self-interested reasons to regulate AI, including fiscal, industrial and strategic trade or military objectives (Kokotajlo et al. 2025). Regulatory outcomes may also reflect the interests of powerful stakeholders,[7] where public interest and private lobbying shape rules (Hammond 2025). Overall, demand for AI regulation comes from social policy goals, strategic interests and private influence, and its supply from various agencies.

AI regulation seeks to control a powerful emerging technology under dynamic uncertainty as to the externalities of future costs and benefits (Tirole 2021, Kolt 2025). Regulators face considerable problems of information asymmetry and uncertainty. The economic literature on AI regulation has explored trade-offs between safety regulation that seeks to restrict development in the interest of known safety considerations against innovation opportunity cost (Agrawal et al.2022) or growth (Aghion et al. 2022). In standard economic analysis, AI safety is framed as a public good and alignment as a market failure. Analysis explores tradeoffs between regulatory constraints (for social benefit) against private harms of regulation (in dynamics, performance). Gans (2025) for instance prefers liability regimes of ex post tort bargaining over ex ante rules based regulation that optimises the tradeoff between risk and reward. Industrial organisational analysis assumes production of AI has high fixed costs (training) with low marginal cost deployment. This suggests AI models are natural monopolies, with remedy through public provision or price regulation (Acemoglu 2021, Acemoglu and Johnson 2024). The strategic context of AI regulation has game theoretic elements of an innovation race (developers and regulators, developer vs developer, nation versus nation for dominance). Schrepel and Potts (2025) explain that the resources needed to produce AI models are best governed in the commons.

6. Complex governance of AI hyperobjects

The regulation and governance problem of a hyperobject is that regulation of a system requires being outside the system to effect control, and having a model of the system to manage feedback. Hyperobjects meet neither condition. First, traditional regulatory governance relies on control theory, or cybernetic principles, where systems are managed through feedback loops and constraints. This model underlies many economic and political institutions: regulations act as external constraints that steer system behavior toward desired outcomes (Zargham and Ben-Meir 2024). In closed-loop control, a regulator observes the system's output, compares it to a target, and adjusts inputs accordingly.[8] However, this requires being able to observe, define

[7] In line with theories of regulatory capture (Stigler 1965).

[8] In closed-loop feedback control, a system monitors its output and feeds that information back in to adjust its behaviour. In a regulatory context this generates adaptive policies as a controller (e.g. a central

and model the system from outside. Yet because hyperobjects are massively distributed, partially illegible and inseparable from their environment, it is very hard to fully map or observe them from an external vantage point, for there is no clear 'outside' from which to regulate. Instead, governance must use imperfect, partial models and localized feedback. This limitation due to scale and opacity renders conventional cybernetic regulation insufficient for governing hyperobjects.

Second, hyperobjects are difficult to know and to understand. Unlike regular objects, which can be empirically observed, measured, and described from the outside, hyperobjects are partially hidden, illegible, and withdrawn from direct observation. They cannot be rendered transparent or represented in stable models. Because we exist within hyperobjects, we can only "see" them from inside, through partial, evolving representations. Yet hyperobjects are also complex, emergent and sometimes evolving systems. The only reliable way to know what a hyperobject will do is to let it run and observe its behaviour over time. This implies that governance must rely on continuous, real-time adaptation based on feedback from the system's current (but never fully knowable) state. Critically, this feedback is not directed at the system itself, but at a model of the system. Moreover, this model is generated by the system through its own processes. In other words, regulation becomes a matter of a system adapting to its own internally constructed model of its environment, rather than to the environment directly. This is what Foster (2005) describes as fourth-order complexity: systems that adapt not to external signals alone, but to representations of those signals generated from within.

This structure aligns with Friston's theory of Markov blankets, where a system maintains its boundary by modeling and responding to its environment through self-generated, probabilistic inferences.[9] In both accounts, adaptation is mediated by internal epistemic structures, not external control. For hyperobjects, seeing and regulating collapse into the same process, which is a recursive, endogenous form of governance. Hyperobjects require cybernetic learning that cannot be done by humans alone. Collective learning must be locally adaptive (drawing on local

bank) monitors inflation (system output) and adjusts its control signals (interest rates). Open loop control systems differ in that they have no feedback mechanism, so they can only set rules at the beginning. Closed-loop control systems therefore require: (1) a reference value for the target or goal state of the system (x% inflation); (2) the plant system (the economy); (3) sensors or feedback measurement (CPI); (4) error detection measure between reference state (x%) and output signal (CPI); (5) controller (central bank); (6) actuator input or control signal (interest rates); and (7) feedback loop (central bank monitoring and action).

[9] A Markov blanket (Friston 2010) is a conceptual boundary that separates a system (or a set of internal states) from everything else (its external environment), such that given the states in the Markov blanket, the internal states are conditionally independent of the external states. In other words, the Markov blanket "screens off" the inside from the outside – information from the outside world can only affect the internal states via the blanket. All self-organising (dissipative) systems must maintain a Markov blanket to resist disorder and persist over time and do so by predicting and adapting to sensory input.

information and context), synthetic and generative (to piece together distributed information and consequence), and machine supported (using machine intelligence, memory and processing) to visualise and represent the hyperobject system and to create a closed loop feedback in which humans and machines use and combine different types of knowledge. A trained AI model creates an embedding of pooled knowledge (Potts 2022). Humans bring local contextual knowledge in the form of prompts, as initiation of actions, making judgement and discrimination, but also acting as a sensor to the world. The machine, in turn, performs knowledge operations as the agent of the human, searching and computing across pools of training data, finding patterns that the humans might not see in their immediate operations. The pool contains expert knowledge in a suspended state. The model has knowledge (the trained parameter set). But the model does not know it has knowledge and so lacks agency. Human agency is required to prompt the knowledge pool to realise knowledge from it. The total system has agency and knowledge, as a cyborg hyperobject.

Third, the distributed, evolving infrastructure of AI systems are not merely tools; they exhibit properties of self-sensing, autonomous coordination, and collective construction, which resemble forms of agency. As Morton (2013: 5) suggests, the Anthropocene marks a shift in which non-human forces begin to actively shape human history, not passively, but as participants with their own dynamics.[10] In this context, AI systems function as non-human agents with which we are increasingly entangled in strategic interaction. This reframes traditional economic metaphors like "games against nature" (where humans act strategically under uncertainty imposed by a passive environment) as games with emerging, autonomous agents that adapt, respond, and co-evolve with us. To describe this, consider the concept of the agentic commons: shared computational infrastructures that are not just governed by communities, but that themselves participate in governance, through embedded rules, feedback mechanisms, and adaptive logic. These are not static commons, but dynamic systems that respond to inputs and evolve over time. One example is the contribution system – a class of cybernetic governance frameworks in which decentralized agents coordinate the production and maintenance of digital infrastructure (Rennie and Potts 2024). These systems blend human and machine agency, forming a cybernetic agentic commons that demands new models of incentive design, strategic behavior, and collective governance. Rather than regulating these systems externally, we must now start to consider where we might begin to, and how to, co-govern them from within.

[10] Morton (2013: 5) writes "In the time of hyperobjects we are no longer able to think of history as exclusively human, for the very reason we are in the Anthropocene. A strange name indeed, since in this period non-humans make decisive contact with humans."

7. Conclusion

A regulatory control approach to AI requires that we must first make AI legible to a state (Scott 1999), then write constraints on the space of actions that AIs or humans can perform, that will then be enforced by agencies or hardware. Each of these steps is doomed to failure. AI is not part of the environment that we can modify and control at will. AI has brought about, in Morton's sense, the end of the world. Machines are no longer, as it were, 'over there'. We are fully inside them now (Potts 2022). So we must do something different. To start, our regulatory aspirations should be framed as closer to strategic or diplomatic games with a new form of intelligence and non-human agency. Indeed, that is what they are – viz. artificial intelligence – and we should take that literally. Hyperobjects are new players in the game and we must look for institutional designs to elicit cooperation (North et al. 2009, Rennie et al. 2025). We must recognise their scale and futurality. New institutional forms such as autonomous ecological institutions (Wade Smith 2024) or contribution systems (Rennie and Potts 2024), which are decentralized, participatory governance infrastructures co-produced by humans and machines, could be instrumental in this process. AI is a hyperobject, the implication of which is that AI cannot be regulated in the traditional sense. What is needed instead is a new governance paradigm that is based on embedded feedback, cybernetic adaptation, and participatory, machine-assisted coordination. This reframes the problem of AI not as policy failure but as a deeper mismatch between the ontology of the object and the architecture of our institutions. AI is not just a powerful tool or a source of risk. It is an emerging agentic system, a dynamic infrastructural commons, that humans must learn to co-govern from within. Many of the pressing law and economics challenges of our time arise from our encounter with hyperobjects and that we must begin to develop new institutional logics for living with these emergent, non-human actors.

Jason Potts

Professor of Economics, Alfaisal University,
Kingdom of Saudi Arabia Research Affiliate, MIT, USA