INSTITUTE
FOR LAW & AI

AI is like…

A literature review of AI metaphors and why they matter for policy

AI Foundations Report 2 | Matthijs Maas | October 2023

law-ai.org

AI is like…

A literature review of AI metaphors and why they matter for policy

Institute for Law & AI – AI Foundations Report 2

October 2023 | Matthijs Maas[1]

Abstract

As AI systems have become increasingly capable and impactful, there has been significant public and policymaker debate over this technology's impacts—and the appropriate legal or regulatory responses. Within these debates many have deployed—and contested—a dazzling range of analogies, metaphors, and comparisons for AI systems, their impact, or their regulation.

This report reviews why and how metaphors and analogies matter to both the study and practice of AI governance, in order to contribute to more productive dialogue and more reflective policymaking. It first reviews five stages at which different foundational metaphors play a role in shaping the processes of technological innovation, the academic study of their impacts, the regulatory agenda, the terms of the policymaking process, and legislative and judicial responses to new technology. It then surveys a series of cases where the choice of analogy materially influenced the internet policy issues as well as (recent) AI law issues. The report then provides a non-exhaustive survey of 55 analogies that have been given for AI technology and some of their policy implications. Finally, it discusses the risks of utilizing unreflexive analogies in AI law and regulation.

By disentangling the role of metaphors, analogies, and frames in these debates, and the space of analogies for AI, this survey does not aim to argue against the use or role of analogies in AI regulation—but rather to facilitate more reflective and productive conversations on these timely challenges.

Cite as: Maas, Matthijs, 'AI is like… A literature review of AI metaphors and why they matter for policy.' *Institute for Law & AI*. AI Foundations Report 2. (October 2023). https://www.law-ai.org/ai-policy-metaphors

[1] Senior Research Fellow – Institute for Law & AI | Research Affiliate, Centre for the Study of Existential Risk, University of Cambridge. ORCID iD: 0000-0002-6170-9393. Email: matthijs.maas@law-ai.org. For comments and input on various sections, I thank Laurens Prins, Kevin Frazier, John Danaher, Esben Kran, Marco Almada, John-Clark Levin, Carlos Ignacio Gutierrez, Aishwarya Saxena, and Gavin Leech. I especially thank Suzanne Van Arsdale, José Jaime Villalobos, and Kevin Frazier for in-depth substantive, structural, and editorial insights that have greatly improved this document. I also thank Daisy Newbold-Harrop, Erin Cooper, and Wes Cowley for editorial assistance in preparing this report for publication. All views expressed, and especially any remaining errors, are my own.

Executive Summary

This report provides an overview, taxonomy, and preliminary analysis of the role of basic metaphors and analogies in AI governance.

Aim: The aim of this report is to contribute to improved analysis, debate, and policy for AI systems by providing greater clarity around the way that analogies and metaphors can affect technology governance generally, around how they may shape AI governance, and about how to improve the processes by which some analogies or metaphors for AI are considered, selected, deployed, and reviewed.

Summary: In sum, this report:

I. Draws on technology law scholarship to review five ways in which metaphors or analogies exert influence throughout the entire cycle of technology policymaking by shaping:
   1. patterns of technological innovation;
   2. the study of particular technologies' sociotechnical impacts or risks;
   3. which of those sociotechnical impacts make it onto the regulatory agenda;
   4. how those technologies are framed within the policymaking process in ways that highlight some issues and policy levers over others; and
   5. how these technologies are approached within legislative and judicial systems.

II. Illustrates these dynamics with brief case studies where foundational metaphors shaped policy for cyberspace, as well as for recent AI issues.

III. Provides an initial atlas of 55 analogies for AI, which have been used in expert, policymaker, and public debate to frame discussion of AI issues, and discusses their implications for regulation.

IV. Reflects on the risks of adopting unreflexive analogies and misspecified (legal) definitions.

Below, the reviewed analogies are summarized in Table 1.

Table 1: overview of surveyed analogies for AI (brief, without policy implications)

Theme	**Frame (varieties)**
Essence	Field of science
Terms focusing on *what AI is*	IT technology (just better algorithms, AI as a product)
	Information technology
	Robots (cyber-physical systems, autonomous platforms)
	Software (AI as a service)
	Black box
	Organism (artificial life)
	Brain

	Mind (digital minds, idiot savant)
	Alien (shoggoth)
	Supernatural entity (god-like AI, demon)
	Intelligence technology (markets, bureaucracies, democracies)
	Trick (hype)
Operation Terms focusing on *how AI works*	Autonomous system
	Complex adaptive system
	Evolutionary process
	Optimization process
	Generative system (generative AI)
	Technology base (foundation model)
	Agent
	Pattern-matcher (autocomplete on steroids, stochastic parrot)
	Hidden human labor (fauxtomation)
Relation Terms focusing on *how we relate to AI, as (possible) subject*	Tool (just technology)
	Animal
	Moral patient
	Moral agent
	Slave
	Legal entity (digital person, electronic person, algorithmic entity)
	Culturally revealing object (mirror to humanity, blurry JPEG of the web)
	Frontier (frontier model)
	Our creation (mind children)
	Next evolutionary stage or successor
Function	Companion (social robots, care robots, generative chatbots, cobot)

Terms focusing on *how AI is or can be used*	Advisor (coach, recommender, therapist)
	Malicious actor tool (AI hacker)
	Misinformation amplifier (computational propaganda, deepfakes, neural fake news)
	Vulnerable attack surface
	Judge
	Weapon (killer robot, weapon of mass destruction)
	Critical strategic asset (nuclear weapons)
	Labor enhancer (steroids, intelligence forklift)
	Labor substitute
	New economic paradigm (fourth industrial revolution)
	Generally enabling technology (the new electricity / fire / internal combustion engine)
	Tool of power concentration or control
	Tool for empowerment or resistance (emancipatory assistant)
	Global priority for shared good
Impact Terms focusing on the *unintended risks, benefits or side-effects of AI*	Source of unanticipated risks (algorithmic black swan)
	Environmental pollutant
	Societal pollutant (toxin)
	Usurper of human decision-making authority
	Generator of legal uncertainty
	Driver of societal value shifts
	Driver of structural incentive shifts
	Revolutionary technology
	Driver of global catastrophic or existential risk

Table of Contents

Introduction

Everyone loves a good analogy like they love a good internet meme—quick, relatable, shareable,[2] memorable, and good for communicating complex topics to one's family.

Background: As AI systems have become increasingly capable and have had increasingly public impacts, there has been significant public and policymaker debate over the technology. Given the breadth of the technology's application, many of these discussions have come to deploy—and contest—a dazzling range of analogies, metaphors, and comparisons for AI systems in order to understand, frame, or shape the technologies' impact and its regulation.[3] Yet the speed with which many often jump to invoke particular metaphors—or to contest the accuracy of others—leads to frequent confusion over these analogies, how they are used, and how they are best evaluated or compared.[4]

Rationale: Such debates are not just about wordplay—metaphors matter. Framings, metaphors, analogies, and (at the most specific end) definitions can strongly affect many key stages of the world's response to a new technology, from the initial developmental pathways for technology, to the shaping of policy agendas, to the efficacy of legal frameworks.[5] They have done so consistently in the past, and we have reason to believe they will especially do so for (advanced) AI. Indeed, recent academic, expert, public, and legal contests around AI often already strongly turn on "battles of analogies."[6]

Aim: Given this, there is a need for those speaking about AI to better understand (a) when they speak in analogies—that is, when the ways in which AI is described (inadvertently) import one or more foundational analogies; (b) what it does to utilize one or another metaphor for AI; (c) what different analogies could be used instead; (d) how the appropriateness of one or another metaphor is best evaluated; and (e) what, given this, might be the limits or risks of jumping at particular analogies.

This report aims to respond to these questions and contribute to improved analysis, debate, and policy by providing greater clarity around the role of metaphors in AI governance, the range of possible (alternate) metaphors, and good practices in constructing and using metaphors.

---

[2] Analogy suggested by ChatGPT (24/10/2023). No other material in this report has involved the use of generative AI models in its creation.

[3] For some recent discussions of common analogies, see for instance: Stern, Jacob. 'AI Is Like … Nuclear Weapons?' The Atlantic, 26 March 2023. https://www.theatlantic.com/technology/archive/2023/03/ai-gpt4-technology-analogy/673509/. (discussing analogies such as "gain-of-function research," "social media," "nuclear weapons," and "electricity" and Edward Teller's 1942 fears that the first nuclear test might lead to atmospheric ignition); Crawford, Jason. 'Four Lenses on AI Risks'. The Roots of Progress, 29 March 2023. https://rootsofprogress.org/four-lenses-on-ai-risks (discussing the implications of viewing AI through the lens of "software," as "complex system," as "an agent with unaligned interests," or as "a separate, advanced culture or species"); *Windows on Theory* (blog), 28 June 2023. https://windowsontheory.org/2023/06/28/metaphors-for-ai-and-why-i-dont-like-them/. (reviewing the implications, and limits, of the metaphors "stochastic parrots", "(Blurry) JPEG of the web", "The New McKinsey", and "Markets, Bureaucracies, Democracies", amongst others). For a fuller overview, see Part III.

[4] This is not solely the case for explicit analogies. Indeed, this report can be read alongside a separate report that focuses more specifically on the use and definitions of a wide range of terms and concepts for "advanced AI." See Maas, Matthijs, 'Concepts in Advanced AI Governance: a Literature Review of Key Terms and Definitions.' Institute for Law & AI. AI Foundations Report 3. (2023). https://www.law-ai.org/advanced-ai-gov-concepts

[5] See also the discussion throughout Part I, below.

[6] As coined by Rebecca Crootof in: Thomson-DeVeaux, Amelia. 'The Supreme Court Is Stubbornly Analog — By Design'. *FiveThirtyEight* (blog), 29 May 2018. https://fivethirtyeight.com/features/the-supreme-court-is-stubbornly-analog-by-design/. See also Downing, Kate. 'Battle of the AI Analogies'. *Law Offices of Kate Downing* (blog), 21 June 2023. https://katedowninglaw.com/2023/06/21/battle-of-the-ai-analogies/.

Caveats: The aim here is not to argue against the use of any analogies in AI policy debates—if that were even possible. Nor is it to prescribe (or dismiss) one or another metaphor for AI as "better" (or "worse") per se. The point is not that one particular comparison is the best and should be adopted by all, or that another is "obviously" flawed. Indeed, in some sense, a metaphor or analogy cannot be "wrong," only more tenuous and more or less suitable when considered from the perspective of some values or some (regulatory) purpose. As such, different metaphors may work best in different contexts. Given this, this report highlights the diversity of analogies in current use and provides context for more informed future discourse and policymaking.

Terminology: Strictly speaking, there is a difference between a metaphor—"an implied comparison between two things of unlike nature that yet have something in common"—and an analogy—"a non-identical or non-literal similarity comparison between two things, with a resulting predictive or explanatory effect."[7] However, while in legal contexts the two can be used in slightly different ways, cognitive science suggests that humans process information by metaphor and by analogy in similar ways.[8] As a result, within this report, "analogy" and "metaphor" will be used relatively interchangeably to refer to (1) communicated *framings of an (AI) issue* that describe that issue (2) through terms, similes, or metaphors which *rely on, invoke, or import references* to a different phenomenon, technology, or historical event, which (3) is (assumed to be) *comparable in one or more ways* (e.g., technical, architectural, political, or moral) (4) which are *relevant to evaluating or responding* to the (AI) issue at hand. Furthermore, the report will use the term "foundational metaphor" to discuss cases where a particular metaphor for the technology has become deeply established and embedded within larger policy programs, such that the nature of the metaphor as a metaphor may even become unclear.

Structure: Accordingly, this report now proceeds as follows. In Part I, it discusses why and how definitions matter to both the study and practice of AI governance. It reviews five ways in which analogies or definitions can shape technology policy generally. To illustrate this, Part II reviews a range of cases in which deeply ingrained foundational metaphors have shaped internet policy as well as legal responses to various AI uses. In Part III, this report provides an initial atlas of 55 different analogies that have been used for AI in recent years, along with some of their regulatory implications. Part IV briefly discusses the risks of using analogies in unreflexive ways.

[7] Carpenter, Jacob. 'Persuading with Precedent: Understanding and Improving Analogies in Legal Argument'. *Capital University Law Review* 44 (1 January 2016): 461. https://scholarship.law.marquette.edu/facpub/705 pg 464 (citing sources).
[8] ibid.

I. How metaphors shape technology governance

Given the range of disciplinary backgrounds in debates over AI, we should not be surprised that the technology is perceived and understood differently by many.

Nonetheless, it matters to get clarity, because terminological and analogical framing effects happen at all stages in the cycle from technological development to societal response. They can shape the initial development processes for technologies as well as the academic fields and programs that study their impacts.[9] Moreover, they can shape both the policymaking processes and the downstream judicial interpretation and application of legislative texts.

1. Metaphors shape innovation

Metaphors and analogies are strongly rooted in human psychology.[10] Even some nonhuman animals think analogically.[11] Indeed, human creativity has even been defined as "the capacity to see or interpret a problematic phenomenon as an unexpected or unusual instance of a prototypical pattern already in one's conceptual repertoire."[12]

Given this, metaphors and analogies can shape and constrain the ability of humans to collectively create new things.[13] In this way, technology metaphors can affect the initial human processes of invention and investment that drive the development of AI and other technologies in the first place. It has been suggested that foundational metaphors can influence the organization and direction of scientific fields—and even that all scientific frameworks could to some extent be viewed as metaphors.[14] For example, the fields of cell biology and biotechnology have for decades been shaped by the influential foundational metaphor that sees biological cells as "machines," which has led to sustained debates over the scientific use and limits of that analogy in shaping research programs.[15]

More practically, at the development and marketing stage, metaphors can shape how consumers and investors assess proposed startup ideas[16] and which innovation paths attract engineer, activist, and policymaking interest

[9] In this sense metaphors and analogies can be relevant to consider even if one is not interested in the question of (AI) policy or regulation per se—but if one's purpose for defining AI is instead technological (to build a given technology) or analytical (to study the sociotechnical impacts of that technology). For more on this, see Maas, Matthijs, 'Concepts in Advanced AI Governance: a Literature Review of Key Terms and Definitions.' *Institute for Law & AI*. AI Foundations Report #3. (2023). (Part I(1)).

[10] Lakoff, George, and Mark Johnson. *Metaphors We Live By*. Edited by With a new Afterword. Chicago, IL: University of Chicago Press, 2003. https://press.uchicago.edu/ucp/books/book/chicago/M/bo3637992.html.

[11] Holyoak, Keith J., and Paul Thagard. *Mental Leaps: Analogy in Creative Thought*. Cambridge, Mass.: Bradford Books, 1996.

[12] Churchland, Paul M. *The Engine of Reason, the Seat of the Soul: A Philosophical Journey Into the Brain*. MIT Press, 1995. Pg. 278.

[13] Brand, Charlotte Olivia, Alex Mesoudi, and Paul E. Smaldino. 'Analogy as a Catalyst for Cumulative Cultural Evolution'. PsyArXiv, 24 August 2020. https://doi.org/10.31234/osf.io/ynkqf.

[14] Honeck, Richard P., and Robert R. Hoffman, eds. *Cognition and Figurative Language*. Routledge, 1980. https://www.routledge.com/Cognition-and-Figurative-Language/Honeck-Hoffman/p/book/9781138361096.

[15] Bongard, Joshua, and Michael Levin. 'Living Things Are Not (20th Century) Machines: Updating Mechanism Metaphors in Light of the Modern Science of Machine Behavior'. *Frontiers in Ecology and Evolution* 9 (2021). https://doi.org/10.3389/fevo.2021.650726.

[16] Mollick, Ethan. 'Blinded by Analogies'. *One Useful Thing*, 23 February 2023. https://oneusefulthing.substack.com/p/blinded-by-analogies., citing: Gompers, Paul A., Will Gornall, Steven N. Kaplan, and Ilya A. Strebulaev. 'How Do Venture Capitalists Make Decisions?' *Journal of Financial Economics* 135, no. 1 (1 January 2020): 169–90. https://doi.org/10.1016/j.jfineco.2019.06.011.

and support. In some such cases, metaphors can support and spur on innovation; for instance, it has been argued that through the early 2000s, the coining of specific IT metaphors for electric vehicles—as a "computer on wheels"—played a significant role in sustaining engineer support for and investment in this technology, especially during an industry downturn in the wake of General Motors' sudden cancellation of its EV1 electric car.[17]

Conversely, metaphors can also hold back or inhibit certain pathways of innovation; for instance, in the Soviet Union in the early 1950s, the field of cybernetics (along with other fields such as genetics or linguistics) fell victim to anti-American campaigns, which characterized it as "an 'obscurantist', 'bourgeois pseudoscience'".[18] While this did not affect the early development of Soviet computer technology (which was highly prized by the state and the military), the resulting ideological rejection of the "man-machine" analogy by Marxist-Leninist philosophers led to an ultimately dominant view, in Soviet sciences, of computers as solely "tools to think with" rather than "thinking machines," holding back the consolidation of the field (such that even the label "AI" would not be recognized by the Soviet Academy of Sciences until 1987) and shifting research attention into projects that focused on the "situational management" of large complex systems rather than the pursuit of human-like thinking machines.[19] This stood in contrast to US research programs, such as DARPA's 1983–1993 Strategic Computing Initiative, an extensive, $1 billion program to achieve "machines that think."[20]

2. Metaphors inform the study of technologies' impacts

Particular definitions also shape and prime academic fields that study the impacts of these technologies (and which often may uncover or highlight particular developments as issues for regulation). Definitions affect which disciplines are drawn to work on a problem, what tools they bring to hand, and how different analyses and fields can build on one another. For instance, it has been argued that the analogy between software code and legal text has supported greater and more productive engagement by legal scholars and practitioners with such code at the level of its (social) meaning and effects (rather than narrowly on the level of the techniques used).[21] Given this, terminology can affect how AI governance is organized as a field of analysis and study, what methodologies are applied, and what risks or challenges are raised or brought up.

3. Metaphors set the regulatory agenda

More directly, particular definitions or frames for a technology can set and shape the policymaking agenda in various ways.

---

[17] See Lloyd, Jay. 'Computers on Wheels?' *Issues in Science and Technology* (blog), 6 February 2023. http://issues.org/computers-on-wheels-electric-vehicles-eisler/.

[18] Kirtchik, Olessia. 'The Soviet Scientific Programme on AI: If a Machine Cannot "Think", Can It "Control"?' *BJHS Themes*, 7 August 2023, 1–15. https://doi.org/10.1017/bjt.2023.4. Pg. 4.

[19] Ibid.

[20] See: Roland, Alex, and Philip Shiman. *Strategic Computing: DARPA and the Quest for Machine Intelligence, 1983-1993*. Edited by William Aspray. UK ed. edition. The MIT Press, 2002.

[21] See Grimmelmann, James. 'The Structure and Legal Interpretation of Computer Programs'. *Journal of Cross-Disciplinary Research in Computational Law* 1, no. 3 (11 May 2023). https://papers.ssrn.com/abstract=4445484; Shaeffer, John. 'Software as Text'. *Santa Clara High Technology Law Journal* 33, no. 3 (20 February 2017): 324. https://digitalcommons.law.scu.edu/chtlj/vol33/iss3/1. As discussed in: Almada, Marco. 'Metaphors and Analogies in Law & Tech'. Substack newsletter. *AI, Law, and Otter Things*, 1 June 2023. https://marcoalmada.substack.com/p/metaphors-and-analogies-in-law-and.

For instance, terms and frames can raise (or suppress) policy attention for an issue, affecting whether policymakers or the public care (enough) about a complex and often highly technical topic in the first place to take it up for debate or regulation. For instance, it has been argued that framings that focus on the viscerality of the injuries inflicted by a new weapon system have in the past boosted international campaigns to ban blinding lasers and antipersonnel mines, yet they ended up being less successful in spurring effective advocacy around "killer robots."[22]

Moreover, metaphors—and especially specific definitions—can shape (government) perceptions of the empirical situation or state of play around a given issue. For instance, the particular definition used for "AI" can directly affect which (industrial or academic) metrics are used to evaluate different states' or labs' relative achievements or competitiveness in developing the technology. In turn, that directly shapes downstream evaluations of which nation is "ahead" in AI.[23]

Finally, terms can frame the relevant legal actors and policy coalitions, enabling (or inhibiting) inclusion and agreement at the level of interest or advocacy groups that push for (or against) certain policy goals. For instance, the choice for particular terms or framings that meet with broad agreement or acceptance amongst many actors can make it easier for a diverse set of stakeholders to join together in pushing for regulatory actions. However, such agreement may be fostered by definitional clarity, when terms or frames are transparent and meet with wider acceptance, or because of definitional ambiguity, when a broad term (such as "ethical AI") allows for sufficient ambiguity that different actors can meet on an "incompletely theorized agreement"[24] to pursue a shared policy program on AI.

4. Metaphors frame the policymaking process

Terms can have a strong overall effect on policy issue-framing, foregrounding different problem portfolios as well as regulatory levers. For instance, early societal debates around nanotechnology were significantly influenced by analogies with asbestos and genetically modified organisms.[25]

Likewise, regulatory initiatives that frame AI systems as "products" imply that these fit easily within product safety frameworks—even if that may be a poor or insufficient model for AI governance, for instance because it

[22] Rosert, Elvira, and Frank Sauer. 'How (Not) to Stop the Killer Robots: A Comparative Analysis of Humanitarian Disarmament Campaign Strategies'. *Contemporary Security Policy* 42, no. 1 (30 May 2020): 4–29. https://doi.org/10.1080/13523260.2020.1771508. On the flexibility of public attitudes to "killer robots" based on their framing or context, see also: Rosendorf, Ondřej, Michal Smetana, and Marek Vranka. 'Algorithmic Aversion? Experimental Evidence on the Elasticity of Public Attitudes to "Killer Robots"'. *Security Studies* 0, no. 0 (2023): 1–31. https://doi.org/10.1080/09636412.2023.2250259.

[23] Murdick, Dewey, James Dunham, and Jennifer Melot. 'AI Definitions Affect Policymaking'. Center for Security and Emerging Technology, 2 June 2020. https://cset.georgetown.edu/research/ai-definitions-affect-policymaking/. (noting that "the competitive landscape varies significantly in sub-areas such as computer vision (where China leads), robotics (where China has made significant progress), and natural language processing (where the United States maintains its lead).", at 2).

[24] Sunstein, Cass R. 'Incompletely Theorized Agreements'. *Harvard Law Review* 108, no. 7 (1995): 1733–72. https://doi.org/10.2307/1341816.; Note that while this situation has drawbacks, it need not always be counterproductive. See Stix, Charlotte, and Matthijs M. Maas. 'Bridging the Gap: The Case for an "Incompletely Theorized Agreement" on AI Policy'. *AI and Ethics*, January 2021. https://doi.org/10.1007/s43681-020-00037-w.

[25] Schwarz-Plaschg, Claudia. 'The Power of Analogies for Imagining and Governing Emerging Technologies'. *NanoEthics* 12, no. 2 (1 August 2018): 139–53. https://doi.org/10.1007/s11569-018-0315-z.

is a model that fails to address any risks at the developmental stage[26] or because it fails to accurately focus on fuzzier impacts on fundamental rights if those cannot be easily classified as consumer harms.[27]

This is not to say that the policy-shaping influence of terms (or explicit metaphors) is absolute and irrevocable. For instance, in a different policy domain, a 2011 study found that using metaphors that described crime as a "beast" led study participants to recommend law-and-order responses, whereas describing it as a "virus" led them to put more emphasis on public-health-style policies. However, even under the latter framing, law-and-order policy responses still prevailed, simply commanding a smaller majority than they would otherwise.[28]

Nonetheless, metaphors do exert sway throughout the policymaking process. For instance, they can shape perceptions of the feasibility of regulation by certain routes. As an example, framings of digital technologies that emphasize certain traits of technologies—such as the "materiality" or "seeming immateriality," or the centralization or decentralization, of technologies like submarine cables, smart speakers, search engines, or the bitcoin protocol—can strongly affect perceptions of whether, or by what routes, it is most feasible to regulate that technology at the global level.[29]

Likewise, different analogies or historical comparisons for proposed international organizations for AI governance—ranging from the IAEA and IPCC to the WTO or CERN—often import tacit analogical comparisons (or rather constitute "reflected analogies") between AI and those organizations' subject matter or mandates in ways that shape the perceptions of policymakers and the public regarding which of AI's challenges require global governance, whether or which new organizations are needed, and whether the establishment of such organizations will be feasible.[30]

5. Metaphors and analogies shape the legislative and judicial response to tech

Finally, metaphors, broad analogies, and specific definitions can frame legal and judicial treatment of a technology in both the ex ante application of AI-focused regulations and the ex post subsequent judicial interpretation of either such AI-specific legislation or of general regulations in the context of cases involving AI.

[26] See for instance: Korzekwa, Rick. 'Product Safety Is a Poor Model for AI Governance'. AI Impacts, 1 February 2023. https://aiimpacts.org/product-safety-is-a-poor-model-for-ai-governance/.
[27] Almada, Marco. 'Metaphors and Analogies in Law & Tech'. Substack newsletter. *AI, Law, and Otter Things*, 1 June 2023. https://marcoalmada.substack.com/p/metaphors-and-analogies-in-law-and.
[28] Thibodeau, Paul H., and Lera Boroditsky. 'Metaphors We Think With: The Role of Metaphor in Reasoning'. *PLOS ONE* 6, no. 2 (23 February 2011): e16782. https://doi.org/10.1371/journal.pone.0016782. As discussed in *The Economist*. 'The Dangers of Misleading Metaphors'. 3 November 2018. https://www.economist.com/books-and-arts/2018/11/03/the-dangers-of-misleading-metaphors?etear=nl_special_5.
[29] Beaumier, Guillaume, Kevin Kalomeni, Malcolm Campbell‐Verduyn, Marc Lenglet, Serena Natile, Marielle Papin, Daivi Rodima‐Taylor, Arthur Silve, and Falin Zhang. 'Global Regulations for a Digital Economy: Between New and Old Challenges'. *Global Policy* 11, no. 4 (September 2020): 515–22. https://doi.org/10.1111/1758-5899.12823.
[30] Maas, Matthijs, and Villalobos, José Jaime. 'International AI Institutions: a literature review of models, examples, and proposals.' Institute for Law & AI, AI Foundations Report 1. (2023). https://www.law-ai.org/international-ai-institutions

Indeed, much of legal reasoning, especially in court systems, and especially in common law jurisdictions, is deeply analogical.[31] This is for various reasons.[32] For one, legal actors are also human, and strong features of human psychology can skew these actors towards the use of analogies that refer to known and trusted categories: as such, as Mandel has argued, "availability and representativeness heuristics lead people to view a new technology and new disputes through existing frames, and the status quo bias similarly makes people more comfortable with the current legal framework."[33] This is particularly the case because much of legal scholarship and work aims to be "problem-solving" rather than "problem-finding"[34] and to respond to new problems by appealing to pre-existent (ethical or legal) principles, norms, values, codes, or laws.[35] Moreover, from an administrative perspective, it is often easier and more cost-effective to extend existing laws by analogy.

Finally, and more fundamentally, the resort to analogy by legal actors can be a shortcut that aims to apply the law, and solve a problem, through an "incompletely theorized agreement" that does not require reopening contentious questions or debates over the first principles or ultimate purposes of the law,[36] or renegotiating hard-struck legislative agreements. This is especially the case at the level of international law, where either negotiating new treaties or explicitly amending multilateral treaties to incorporate a new technology within an existing framework can be wrought, drawn-out processes[37] such that many actors may prefer ultimately addressing new issues (such as cyberwar) within existing norms or principles by analogizing them to well-established and well-regulated behaviors.[38]

Given this, when confronted with situations of legal uncertainty—as often happens with a new technology[39]—legal actors may favor the use of analogies to stretch existing law or to interpret new cases as

[31] Sunstein, Cass R. 'Analogical Reasoning'. SSRN Scholarly Paper. Rochester, NY, 7 October 2021. https://doi.org/10.2139/ssrn.3938546. And see previously: Sunstein, Cass R. 'On Analogical Reasoning'. *Harvard Law Review* 106, no. 3 (January 1993): 741. https://doi.org/10.2307/1341662.

[32] For a discussion of these, see also Crootof, Rebecca, and B. J. Ard. 'Structuring Techlaw'. *Harvard Journal of Law & Technology* 34, no. 2 (2021): 347–417. https://jolt.law.harvard.edu/assets/articlePDFs/v34/1.-Crootof-Ard-Structuring-Techlaw.pdf pg. 395-396. See also Maas, Matthijs M. 'Artificial Intelligence Governance Under Change: Foundations, Facets, Frameworks'. University of Copenhagen, 2020. http://www.legalpriorities.org/documents/Maas-PhD-Dissertation.pdf, Pg. 214.

[33] Mandel, Gregory N. 'Legal Evolution in Response to Technological Change'. *The Oxford Handbook of Law, Regulation and Technology*, 20 July 2017, 225–45. https://doi.org/10.1093/oxfordhb/9780199680832.013.45. Pg. 238.

[34] Liu, Hin-Yan, and Matthijs M. Maas. '"Solving for X?" Towards a Problem-Finding Framework to Ground Long-Term Governance Strategies for Artificial Intelligence'. *Futures* 126 (1 February 2021): 22. https://doi.org/10.1016/j.futures.2020.102672.

[35] Hopster, Jeroen K. G., and Matthijs M. Maas. 'The Technology Triad: Disruptive AI, Regulatory Gaps and Value Change'. *AI and Ethics*, 28 June 2023. https://doi.org/10.1007/s43681-023-00305-5. Pg. 2 (referring to these as 'first-order problems' or challenges).

[36] Sunstein, Cass R. 'Analogical Reasoning'. SSRN Scholarly Paper. Rochester, NY, 7 October 2021. https://doi.org/10.2139/ssrn.3938546.

[37] That is not to say that such changes cannot happen through soft law pathways: see for instance Israel, Brian. 'Treaty Stasis'. *AJIL Unbound* 108 (ed 2014): 63–69. https://doi.org/10.1017/S2398772300001860. And Smith, Bryant Walker. 'New Technologies and Old Treaties'. *AJIL Unbound* 114 (ed 2020): 152–57. https://doi.org/10.1017/aju.2020.28.

[38] Eichensehr, Kristen E. 'Cyberwar & International Law Step Zero'. *Texas International Law Journal* 50, no. 2 (2015): 357–80. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2611198 ;See also Deeks, Ashley. 'The Geography of Cyber Conflict: Through a Glass Darkly'. *International Law Studies* 89, no. 1 (14 March 2013). https://papers.ssrn.com/abstract=2233560. ("[the US government has] an inherent institutional instinct . . . to anchor novel legal situations in existing bodies of law and practice, and to reason by analogy. . . . Particularly where the analogies are quite reasonable (as they are between kinetic and cyber activities), it often is easier to draw from existing rules than to craft new ones from whole cloth.").

[39] Liu, Hin-Yan, Matthijs Maas, John Danaher, Luisa Scarcella, Michaela Lexer, and Leonard Van Rompaey. 'Artificial Intelligence and Legal Disruption: A New Model for Analysis'. *Law, Innovation and Technology* 12, no. 2 (16 September 2020): 205–58. https://doi.org/10.1080/17579961.2020.1815402.; and for this argument in the international law context, see: Maas, Matthijs M. 'International Law Does Not Compute: Artificial Intelligence and The Development, Displacement or Destruction of the Global Legal Order'. *Melbourne Journal of International Law* 20, no. 1 (2019): 29–56. https://law.unimelb.edu.au/__data/assets/pdf_file/0005/3144308/Maas.pdf

falling within existing doctrine. That does not mean that courts need immediately settle or converge on one particular "right" analogy. Indeed, there are always multiple analogies possible, and these can have significantly different implications for how the law is interpreted and applied. That means that many legal cases involving technology will involve so-called "battles of analogies."[40] For example, in recent class action lawsuits that have accused generative AI providers such as Stable Diffusion and Midjourney of copyright infringement, plaintiffs have argued that these generative AI models are "essentially sophisticated collage tools, with the output representing nothing more than a mash-up of the training data, which is itself stored in the models as compressed copies."[41] Some have countered that this analogy suffers some technical inaccuracies, since current generative AI models do not store compressed copies of the training data, such that a better analogy would be that of an "art inspector" that takes every measurement possible—implying that model training either is not governed by copyright law or constitutes fair use.[42]

Finally, even if specific legislative texts move to adopt clear, specific statutory definitions for AI—in a way that avoids (explicit) comparison or analogy with other technologies or behavior—this may not entirely avoid framing effects. Most obviously, legislative definitions for key terms such as "AI" obviously affect the material scope of regulations and policies that use and define such terms.[43] Indeed, the effects of particular definitions have impacts on regulation not only ex ante but also ex post: in many jurisdictions, legal terms are interpreted and applied by courts based on their widely shared "ordinary meaning."[44] This means, for instance, that regulations that refer to terms such as "advanced AI," "frontier AI," or "transformative AI"[45] might not

[40] As discussed by Rebecca Crootof in: Thomson-DeVeaux, Amelia. 'The Supreme Court Is Stubbornly Analog — By Design'. *FiveThirtyEight* (blog), 29 May 2018. https://fivethirtyeight.com/features/the-supreme-court-is-stubbornly-analog-by-design/.

[41] Downing, Kate. 'Battle of the AI Analogies'. *Law Offices of Kate Downing* (blog), 21 June 2023. https://katedowninglaw.com/2023/06/21/battle-of-the-ai-analogies/.

[42] Lindberg, Van. 'Building and Using Generative Models Under US Copyright Law'. *Rutgers Business Law Review* 18, no. 2 (30 May 2023). https://papers.ssrn.com/abstract=4464001.

[43] Schuett, Jonas. 'Defining the Scope of AI Regulations'. *Law, Innovation and Technology* 15, no. 1 (3 March 2023): 1–23. https://doi.org/10.1080/17579961.2023.2184135. Note that in legislative texts, the inclusion of particular statutory definitions can play various roles: (1) communicative roles (clarifying legislative intent) and (2) performative roles (investing groups or individuals with rights or obligations). See Price, Jeanne. 'Wagging, Not Barking: Statutory Definitions'. *Cleveland State Law Review* 60, no. 60 (2013): 999–1055. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2288824 . Moreover, within legislation there are different types of definitions that play distinct roles, such as (1) delimiting definitions that establish the limits or boundaries on an otherwise ordinary meaning of a term, (2) extending definitions that broaden a term's meaning to expressly include elements or components that might not normally be included in the ordinary meaning of a term, (3) narrowing definitions that aim to set limits or expressly exclude particular understandings, and (4) mixed definitions that use several of these approaches to clarify components. See Government of Canada, Department of Justice. 'Legistics - Definitions', 2 December 1999. https://www.justice.gc.ca/eng/rp-pr/csj-sjc/legis-redact/legistics/p1p5.html. I thank Suzanne Van Arsdale and Kevin Frazier for highlighting this taxonomy.

[44] Martínez, Eric, and Christoph Winter. 'Ordinary Meaning of Existential Risk. LPP Working Paper No. 7-2022, 15 December 2022. https://doi.org/10.2139/ssrn.4304670. Others have suggested that courts will interpret definitions in ways that align with the median public opinion. See Dorf, Michael C. 'Majoritarian Difficulty and Theories of Constitutional Decision Making'. *Journal of Constitutional Law* 13, no. 2 (2010): 283–304. Note that in certain circumstances, a court may refer to a technical meaning of a term to resolve ambiguity. See for instance Sullivan, Ruth. 'Technical Meaning and Meanings Fixed by Law'. In *Statutory Interpretation*, 73–95. Irwin Law, 2016. However, even if a technical definition is invoked, this may not always be an easy resolution if there are many competing or overlapping technical definitions for the same term. In some cases where the meaning of a word is ambiguous, then in some contexts such as the US, courts may apply a series of additional substantive canons of interpretation. See Baude, William, and Ryan D. Doerfler. 'The (Not So) Plain Meaning Rule'. *The University of Chicago Law Review* 84, no. 2 (Spring 2017): 539–66. https://lawreview.uchicago.edu/print-archive/not-so-plain-meaning-rule. I thank Suzanne Van Arsdale for this suggestion and both Suzanne and Kevin Frazier for work in this space.

[45] For definitions of these terms, see also: Maas, Matthijs, 'Concepts in Advanced AI Governance: a Literature Review of Key Terms and Definitions.' *Institute for Law & AI*. AI Foundations Report #3. (2023).

necessarily be interpreted or applied in ways that are in line with how the term is understood within expert communities.[46]

All of this underscores the importance of our choice of terms and frames—whether broad and indirect metaphors or concrete and specific legislative definitions—when grappling with the impacts of this technology on society.

II. Foundational metaphors in technology law: Cases

Of course, these dynamics are not new and have been studied in depth in fields such as cyberlaw, law and technology, and technology law.[47] For instance, we can see many of these framing dynamics within societal (and regulator) responses to other cornerstone digital technologies.

1. Metaphors in internet policy: Three cases

For instance, for the complex sociotechnical system[48] commonly called *the internet,* foundational metaphors have strongly shaped regulatory debates, at times as much as sober assessments of the nuanced technical details of the artifacts involved have.[49] As noted by Rebecca Crootof:

> "A 'World Wide Web' suggests an organically created common structure of linked individual nodes, which is presumably beyond regulation. The 'Information Superhighway' emphasizes the import of speed and commerce and implies a nationally funded infrastructure subject to federal regulation. Meanwhile, 'cyberspace' could be understood as a completely new and separate frontier, or it could be viewed as yet one more kind of jurisdiction subject to property rules and State control."[50]

For example, different terms (and the foundational metaphors they entail) have come to shape internet policy in various ways and domains. Take for instance the following cases:

---

[46] See also the discussion of the importance of these terms in Maas, Matthijs, 'Concepts in Advanced AI Governance: a Literature Review of Key Terms and Definitions.' *Institute for Law & AI*. AI Foundations Report #3. (2023). https://www.law-ai.org/advanced-ai-gov-concepts (Section I.(2)).

[47] See Ard, BJ, and Rebecca Crootof. 'The Case for "Technology Law"'. *Nebraska Governance & Technology Center* (blog), 16 December 2020. https://ngtc.unl.edu/blog/case-for-technology-law. Note that this is a broad and expansive set of fields; for an overview and selection of classic work, see also Crootof, Rebecca, and B. J. Ard. 'Structuring Techlaw'. *Harvard Journal of Law & Technology* 34, no. 2 (2021): 347–417. https://jolt.law.harvard.edu/assets/articlePDFs/v34/1.-Crootof-Ard-Structuring-Techlaw.pdf, ftn 10.

[48] Dafoe, Allan. 'On Technological Determinism: A Typology, Scope Conditions, and a Mechanism'. *Science, Technology, & Human Values* 40, no. 6 (1 November 2015): 1047–76. https://doi.org/10.1177/0162243915579283. Pg. 1052 ("technology can refer to vast sociotechnical systems, such as the Internet, as well as specific artifacts, standards, routines, and beliefs that make up these systems, such as computers, the Internet protocol, e-mail routines, and beliefs about the reliability of online information.").

[49] See for instance Cohen, I. Glenn, and Jonathan H. Blavin. 'Gore, Gibson, and Goldsmith: The Evolution of Internet Metaphors in Law and Commentary'. *Harvard Journal of Law and Technology* 16, no. 1 (2002): 265. https://doi.org/10.2139/ssrn.479742 (reviewing the implications of metaphors that frame the internet as an information superhighway, cyberspace, or as "real" space, respectively).

[50] Crootof, Rebecca. 'Regulating New Weapons Technology'. In *The Impact of Emerging Technologies on the Law of Armed Conflict*, edited by Eric Talbot Jensen and Ronald T.P. Alcala, 1–25. Oxford University Press, 2019. https://oxford.universitypressscholarship.com/view/10.1093/oso/9780190915322.001.0001/oso-9780190915322-chapter-1. Pg. 17.

Institutional effects of framing cyberwar policy within cyber-"space": For over a decade, the US military framed the internet and related systems as a "cyberspace"—that is, just another "domain" of conflict along with land, sea, air, and space—leading to strong consequences institutionally (expanding the military's role in cybersecurity and supporting the creation of US Cyber Command) as well as for how international law has subsequently been applied to cyber operations.[51]

Issue-framing effects of regulating data as "oil," "sunlight," "public utility," or "labor": Different metaphors for "data" have drastically different political and regulatory implications.[52] The *oil* metaphor emphasizes data as a valuable traded commodity that is owned by whoever "extracts" it and that, as a key resource in the modern economy, can be a source of geopolitical contestation between states. However, the oil metaphor implies that the history of data prior to its collection is not relevant and so sidesteps questions of any "misappropriation or exploitation that might arise from data use and processing."[53] Moreover, even within an regulatory approach that emphasizes geopolitical competition over AI, one can still critique the "oil" metaphor as misleading, for instance because of the ways in which it skews debates over how to assess "data competitiveness" in military AI.[54] By contrast, the *sunlight* metaphor emphasizes data as a ubiquitous public resource that ought to be widely pooled and shared for social good, de-emphasizing individual data privacy claims; the *public utility* metaphor sees data as an "infrastructure" that requires public investment and new institutions, such as data trusts or personal data stores, to guarantee "data stewardship"; and the *labor* frame asserts the ownership rights of the individuals generating data against what are perceived as extractive or exploitative practices of "surveillance capitalism."[55]

Judicial effects of treating search engines as "newspaper editorials" in censorship cases: In the mid-2000s, US court rulings involving censorship on search engines tended to analyze them by analogy to older technologies such as the newspaper editorial.[56]

As these examples suggest, different terms and their metaphors matter. They serve as intuition pumps for key audiences (public, policy) that otherwise may have significant disinterest in, lack of expertise in, inferential distance to, or limited bandwidth for new technologies. Moreover, as seen in social media platforms and online content aggregators' resistance to being described as "media companies" rather than "technology companies,"[57] even seemingly innocuous terms can carry significant legal and policy implications—in doing so, such terms

---

[51] Branch, Jordan. 'What's in a Name? Metaphors and Cybersecurity'. *International Organization* 75 (2021): 39–70. https://doi.org/10.1017/S002081832000051X.

[52] See Maas, Matthijs M. 'Artificial Intelligence Governance Under Change: Foundations, Facets, Frameworks'. University of Copenhagen, 2020. http://www.legalpriorities.org/documents/Maas-PhD-Dissertation.pdf. Pg. 215–16 (citing sources).

[53] Scholz, Lauren. 'Big Data Is Not Big Oil: The Role of Analogy in the Law of New Technologies'. *Tennessee Law Review* 85, no. 2020 (2019): 863–93. https://doi.org/10.2139/ssrn.3252543. Pg. 865.

[54] See Chahal, Husanjot, Ryan Fedasiuk, and Carrick Flynn. 'Messier than Oil: Assessing Data Advantage in Military AI'. Center for Security and Emerging Technology, July 2020. https://cset.georgetown.edu/research/messier-than-oil-assessing-data-advantage-in-military-ai/.

[55] Maas, 'Artificial Intelligence Governance Under Change: Foundations, Facets, Frameworks'. pg. 215–216.

[56] Whitney, Heather. 'Search Engines, Social Media, and the Editorial Analogy'. Knight First Amendment Institute, 27 February 2018. http://knightcolumbia.org/content/search-engines-social-media-and-editorial-analogy.; see also Lakier, Genevieve. 'The Problem Isn't the Use of Analogies but the Analogies Courts Use'. *Knight First Amendment Institute at Columbia University* (blog), 26 February 2018. https://knightcolumbia.org/content/problem-isnt-use-analogies-analogies-courts-use.

[57] Napoli, Philip, and Robyn Caplan. 'Why Media Companies Insist They're Not Media Companies, Why They're Wrong, and Why It Matters'. *First Monday* 22, no. 5 (2 May 2017). https://doi.org/10.5210/fm.v22i5.7051. (discussing a range of cases that illustrate how tech and social media companies have been aversive to classification as "media companies," since such classification "has historically meant more intensive government oversight, in the form of affirmative obligations to serve the public interest and more stringent regulation in areas such as concentration of ownership.").

can serve as a legal "sorter," determining whether a technology (or the company developing and marketing it) is considered as falling into one or another regulatory category.[58]

2. Metaphors in AI law: Three cases

Given the role of metaphors and definitions to strongly shape the direction and efficacy of technology law, we should expect them to likewise play a strong role in affecting the framing and approach of AI regulation in the future, for better or worse. Indeed, in a range of domains, they have already done so:

Autonomous weapons systems under international law: International lawyers often aim to subsume new technologies under (more or less persuasive) analogies to existing technologies or entities that are already regulated.[59] As such, different analogies have been drawn between autonomous weapons systems to *weapons*, *combatants*, *child soldiers*, or *animal combatants*—all of which lead to very different consequences for their legality under international humanitarian law.[60]

Release norms for AI models with potential for misuse: In debates over the potential misuse risks from emerging AI systems, efforts to attempt to restrict or slow publication of new systems with potential for misuse have found themselves challenged by framings that pitch the field of AI as being intrinsically an *open science* (where new findings should be shared whatever the risks) versus those that emphasize analogies to *cybersecurity* (where dissemination can help defenders protect against exploits). Critically, however, both of these analogies may misstate or underappreciate the dynamics that affect the offense-defense balance of new AI capabilities: while in information security the disclosure of software vulnerabilities has traditionally favored defense, this cannot be assumed for AI research, where (among others) it can be much more costly or intractable to "patch" the social vulnerabilities exploited by AI capabilities.[61]

Liability for inaccurate or unlawful speech produced by AI chatbots, large language models, and other generative AI: In the US, Section 230 of the 1996 Communications Decency Act protects online service providers from liability for user-generated content that they host and has accordingly been considered a cornerstone to the business model of major online platforms and social media companies.[62] For instance, in Spring 2023, the US Supreme Court took up two lawsuits—*Gonzales v. Google* and *Twitter v. Taamneh*—which could have shaped Section 230 protections for algorithmic recommendations.[63] While the

[58] I thank Kevin Frazier for this point.
[59] Crootof, Rebecca. 'Regulating New Weapons Technology'. In *The Impact of Emerging Technologies on the Law of Armed Conflict*, edited by Eric Talbot Jensen and Ronald T.P. Alcala, 1–25. Oxford University Press, 2019. https://oxford.universitypressscholarship.com/view/10.1093/oso/9780190915322.001.0001/oso-9780190915322-chapter-1 . See generally; Eichensehr, Kristen E. 'Cyberwar & International Law Step Zero'. *Texas International Law Journal* 50, no. 2 (2015): 357–80. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2611198
[60] Crootof, Rebecca. 'Autonomous Weapon Systems and the Limits of Analogy'. *Harvard National Security Journal* 9 (2018): 51–83. https://doi.org/10.2139/ssrn.2820727.; Horowitz, Michael C. 'Why Words Matter: The Real World Consequences of Defining Autonomous Weapons Systems'. *Temp. Int'l & Comp* 30 (2016): 14. https://sites.temple.edu/ticlj/files/2017/02/30.1.Horowitz-TICLJ.pdf
[61] Shevlane, Toby, and Allan Dafoe. 'The Offense-Defense Balance of Scientific Knowledge: Does Publishing AI Research Reduce Misuse?' In *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*, 173–79. AIES '20. New York, NY, USA: Association for Computing Machinery, 2020. https://doi.org/10.1145/3375627.3375815.
[62] Kosseff, Jeff. 'A User's Guide to Section 230, and a Legislator's Guide to Amending It (or Not)'. *Berkeley Technology Law Journal* 37, no. 2 (2022). https://papers.ssrn.com/abstract=3905347.
[63] Robertson, Adi. 'The Supreme Court Battle for Section 230 Has Begun'. The Verge, 1 December 2022. https://www.theverge.com/2022/12/1/23487958/supreme-court-gonzalez-google-twitter-taamneh-section-230.

Court's rulings on these cases avoided addressing the issue,[64] similar court cases (or legislation) could have strong implications for whether digital platforms or social media companies will be held liable for unlawful speech produced by large language model-based AI chatbots.[65] If such AI chatbots are analogized to existing *search engines*, they might be able to rely on a measure of protection from Section 230, greatly facilitating their deployment, even if they link to inaccurate information. Conversely, if these chatbot systems are considered so novel and creative that their output goes beyond the functions of a search engine, they might instead be considered as "information content providers" within the remit of the law—or simply held to be beyond the law's remit (and protection) entirely.[66] This would mean that technology companies would be held legally responsible for their AI's outputs. If that were the case, this reading of the law would significantly restrict the profitability of many AI chatbots, given the tendency of the underlying LLMs to "hallucinate" facts.[67]

All this again highlights that different definitions or terms for AI will frame how policymakers and courts understand the technology. This creates a challenge for policy, which must address the transformative impact and potential risks of AI as they are (and as they may soon be), and not only as they can be easily analogized to other technologies and fields. What does that mean in the context of developing AI policy in the future?

III. An atlas of AI analogies

Development of policy must contend with the lack of settled definitions for the term "AI," with the varied concepts and ideas projected onto it, and with the pace at which new terms —from "foundation models" to "generative AI"—are often coined and adopted.[68]

Indeed, this breadth of analogies that are coined around AI should not be surprising, given that even just the term "artificial intelligence" has a number of aspects that support conceptual fluidity (or alternately, confusion). This is for various reasons.[69] In the first place, the term invokes a term—"intelligence"—which is in widespread and everyday use, and which for many people has strong (evaluative or normative) connotations. It is essentially a suitcase word that packages together many competing meanings,[70] even while it hides deep and

[64] Willard, Lauren Willard, Madeline Salinas, Beth Brinkmann, Yaron Dori, Megan, Madeline Salinas, Beth Brinkmann, Yaron Dori, and Megan Crowley. 'The U.S. Supreme Court Punts on Section 230 in Gonzalez v. Google LLC'. Global Policy Watch, 19 May 2023. https://www.globalpolicywatch.com/2023/05/the-u-s-supreme-court-punts-on-section-230-in-gonzalez-v-google-llc/.

[65] Robertson, Adi. 'The Supreme Court Could Be about to Decide the Legal Fate of AI Search'. The Verge, 16 February 2023. https://www.theverge.com/2023/2/16/23591290/supreme-court-section-230-gonzalez-google-bard-bing-ai-search-algorithms.

[66] Lima, Cristiano. 'AI Chatbots Won't Enjoy Tech's Legal Shield, Section 230 Authors Say'. *Washington Post*, 17 March 2023. https://www.washingtonpost.com/politics/2023/03/17/ai-chatbots-wont-enjoy-techs-legal-shield-section-230-authors-say/.

[67] Perault, Matt. 'Section 230 Won't Protect ChatGPT'. Lawfare, 23 February 2023. https://www.lawfareblog.com/section-230-wont-protect-chatgpt.; Hutton, Christopher. 'AI Chatbots Aren't Protected by Section 230, Gorsuch Says'. *Washington Examiner*, 21 February 2023, sec. SCOTUS & US District Court News. https://www.washingtonexaminer.com/policy/courts/gorsuch-chatgpt-section-230.

[68] Toner, Helen. 'What Are Generative AI, Large Language Models, and Foundation Models?' *Center for Security and Emerging Technology* (blog), 12 May 2023. https://cset.georgetown.edu/article/what-are-generative-ai-large-language-models-and-foundation-models/.

[69] The following builds on: Maas, Matthijs M. 'Artificial Intelligence Governance Under Change: Foundations, Facets, Frameworks'. University of Copenhagen, 2020. http://www.legalpriorities.org/documents/Maas-PhD-Dissertation.pdf pg. 34-35.

[70] See Minsky, Marvin. Consciousness is a Big Suitcase. Interview by John Brockman, 26 February 1998. https://www.edge.org/conversation/marvin_minsky-consciousness-is-a-big-suitcase. (on terms like "consciousness," "learning," or "memory").

perhaps even intractable scientific and philosophical disagreement[71] and significant historical and political baggage.[72]

Secondly, and in contrast to, say, "blockchain ledgers," AI technology comes with a baggage of decades of depictions in popular culture—and indeed centuries of preceding stories about intelligent machines[73]—resulting in a whole genre of tropes or narratives that can color public perceptions and policymaker debates.

Thirdly, AI is an evocative general-purpose technology that sees use in a wide variety of domains and accordingly has provoked commentary from virtually every disciplinary angle, including neuroscience, philosophy, psychology, law, politics, and ethics. As a result of this, a persistent challenge in work on AI governance—and indeed, in the broader public debates around AI—has been that different people use the word "AI" to refer to widely different artifacts, practices, or systems, or operate on the basis of definitions or understandings which package together a range of implicit assumptions.[74]

Thus, it is no surprise that AI has been subjected to a diverse range of analogies and frames. To understand potential implications of AI analogies, we can draw a taxonomy of common framings of AI (see Table 2), whereby we can distinguish between analogies that focus on:

1. the *essence* or nature of AI (what AI "is"),
2. AI's *operation* (how AI works),
3. our *relation* to AI (how we relate to AI as subject),
4. AI's societal *function* (how AI systems are or can be used),
5. AI's *impact* (the unintended risks, benefits, and other side-effects of AI).

[71] In this way, 'intelligence' might be considered an 'essentially contested concept'. For the classic account of this concept, see Gallie, W. B. 'Essentially Contested Concepts'. *Proceedings of the Aristotelian Society* 56 (1955): 167–98.
[72] Cave, Stephen. 'The Problem with Intelligence: Its Value-Laden History and the Future of AI'. In *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*, 29–35. New York NY USA: ACM, 2020. https://doi.org/10.1145/3375627.3375813.
[73] See Cave, Stephen, Kanta Dihal, and Sarah Dillon, eds. *AI Narratives: A History of Imaginative Thinking about Intelligent Machines*. New York: Oxford University Press, 2020.
[74] As one theorist has quipped; "by far the greatest danger of Artificial Intelligence is that people conclude too early that they understand it": Yudkowsky, Eliezer. 'Artificial Intelligence as a Positive and Negative Factor in Global Risk.' In *Global Catastrophic Risks*, by Eliezer Yudkowsky, 308–45. New York: Oxford University Press, 2008. https://oxford.universitypressscholarship.com/view/10.1093/oso/9780198570509.001.0001/isbn-9780198570509-book-part-21. Pg. 308.

Table 2: Atlas of AI analogies, with framings and selected policy implications.

Theme	**Frame** (examples)	**Emphasizes to policy actors (e.g.)**
Essence Terms focusing on *what AI is*	Field of science[75]	Ensuring scientific best practices; improving methodologies, data sharing, and benchmark performance reporting methodologies to avoid replicability problems;[76] ensuring scientific freedom and openness rather than control and secrecy.[77]
	IT technology (just better algorithms, AI as a product[78])	Business-as-usual; industrial applications; conventional IT sector regulation. Product acquisition & procurement processes; product safety regulations.
	Information technology[79]	Economic implications of increasing returns to scale and income distribution vs. distribution of consumer welfare; facilitation of communication and coordination; effects on power balances.
	Robots (cyber-physical systems,[80] autonomous platforms)	Physicality; embodiment; robotics; risks of physical harm;[81] liability; anthropomorphism; embedment in public spaces.

[75] See broadly: Simon, Herbert A. 'Artificial Intelligence: An Empirical Science'. *Artificial Intelligence* 77, no. 1 (1 August 1995): 95–127. https://doi.org/10.1016/0004-3702(95)00039-H. For a more specific counter-argument of why (some) types of AI research might not be best analogized with the usual processes of science, see: Shevlane, Toby, and Allan Dafoe. 'The Offense-Defense Balance of Scientific Knowledge: Does Publishing AI Research Reduce Misuse?' In *AIES '20: Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*, 2020. http://arxiv.org/abs/2001.00463. (suggesting that the "offense-defense balance" of AI research can be similar to biological research, hardware vulnerabilities, or nuclear engineering research, instead).

[76] Hutson, Matthew. 'Artificial Intelligence Faces Reproducibility Crisis'. *Science* 359, no. 6377 (16 February 2018): 725–26. https://doi.org/10.1126/science.359.6377.725. Burnell, Ryan, Wout Schellaert, John Burden, Tomer D. Ullman, Fernando Martinez-Plumed, Joshua B. Tenenbaum, Danaja Rutar, et al. 'Rethink Reporting of Evaluation Results in AI'. *Science* 380, no. 6641 (14 April 2023): 136–38. https://doi.org/10.1126/science.adf6369.

[77] For a historical example of this dynamic at work in a different technology, see: Wellerstein, Alex. *Restricted Data: The History of Nuclear Secrecy in the United States*. Chicago, IL: University of Chicago Press, 2021. https://press.uchicago.edu/ucp/books/book/chicago/R/bo15220099.html. (arguing that Cold War efforts to control the spread of nuclear weapons information and the newly discovered scientific facts that made such weapons possible spurred significant debates over the effects or appropriateness of such policies on American science).

[78] European Commission. 'Proposal for a Regulation of the European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts'. European Commission, 21 April 2021. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52021PC0206. Article 6(1)(a-b). But for critiques, see: Almada, Marco, and Nicolas Petit. 'The EU AI Act: Between Product Safety and Fundamental Rights'. SSRN Scholarly Paper. Rochester, NY, 20 December 2022. https://doi.org/10.2139/ssrn.4308072. Korzekwa, Rick. 'Product Safety Is a Poor Model for AI Governance'. AI Impacts, 1 February 2023. https://aiimpacts.org/product-safety-is-a-poor-model-for-ai-governance/.

[79] Dafoe, Allan. 'AI Governance: Overview and Theoretical Lenses'. In *The Oxford Handbook of AI Governance*, edited by Justin B. Bullock, Yu-Che Chen, Johannes Himmelreich, Valerie M. Hudson, Anton Korinek, Matthew M. Young, and Baobao Zhang, 0. Oxford University Press, 2023. https://doi.org/10.1093/oxfordhb/9780197579329.013.2.

[80] Calo, Ryan. 'Artificial Intelligence Policy: A Primer and Roadmap'. *UC Davis Law Review* 51, no. 2 (2017): 37. https://lawreview.law.ucdavis.edu/issues/51/2/Symposium/51-2_Calo.pdf

[81] See for instance Cave, Stephen, Kate Coughlan, and Kanta Dihal. '"Scary Robots": Examining Public Responses to AI'. In *Proceedings of AAAI / ACM Conference on Artificial Intelligence, Ethics and Society 2019*, 8, 2019. https://dl.acm.org/doi/abs/10.1145/3306618.3314232.

Software (AI as a service)	Virtuality; digitality; cloud intelligence; open-source nature of development process; likelihood of software bugs.[82]
Black box[83]	Opacity; limits to explainability of a system; risks of loss of human control and understanding; problematic lack of accountability. But also potentially de-emphasizes human decisions and their value judgments behind an algorithmic system, and presents the technology as monolithic, incomprehensible, and unalterable.[84]
Organism (artificial life)	Ecological "messiness"; ethology of causes of "machine behavior" (development, evolution, mechanism, function).[85]
Brains	Applicability of terms and concepts from neuroscience; potential anthropomorphization of AI functionalities along human traits.[86]
Mind (digital minds,[87] idiot savant[88])	Philosophical implications; consciousness, sentience, psychology.
Alien (shoggoth[89])	Inhumanity, incomprehensibility, deception in interactions

[82] Crawford, Jason. 'Four Lenses on AI Risks'.
[83] Pasquale, Frank. *The Black Box Society: The Secret Algorithms That Control Money and Information*. Reprint edition. Cambridge, Massachusetts London, England: Harvard University Press, 2016.
[84] Sommerer, Lucia. 'From Black Box to Algorithmic Veil: Why the Image of the Black Box Is Harmful to the Regulation of AI'. *Better Images of AI Blog* (blog), 1 February 2022. https://blog.betterimagesofai.org/from-black-box-to-algorithmic-veil-why-the-image-of-the-black-box-is-harmful-to-the-regulation-of-ai/.
[85] Rahwan, Iyad, Manuel Cebrian, Nick Obradovich, Josh Bongard, Jean-François Bonnefon, Cynthia Breazeal, Jacob W. Crandall, et al. 'Machine Behaviour'. *Nature* 568, no. 7753 (April 2019): 477. https://doi.org/10.1038/s41586-019-1138-y.
[86] For a critique, see Salles, Arleen, Kathinka Evers, and Michele Farisco. 'Anthropomorphism in AI'. *AJOB Neuroscience* 11, no. 2 (2 April 2020): 88–95. https://doi.org/10.1080/21507740.2020.1740350. Pg. 92-93.
[87] Shulman, Carl, and Nick Bostrom. 'Sharing the World with Digital Minds'. In *Rethinking Moral Status*, edited by Steve Clarke, Hazem Zohny, and Julian Savulescu. Oxford University Press, 2021. https://academic.oup.com/book/41245/chapter/350760172.
[88] Rohit.Krishnan. 'AI Is an Idiot Savant'. Strange Loop Canon, 10 April 2022. https://www.strangeloopcanon.com/p/ai-is-an-idiot-savant.
[89] The Economist. 'Artificial Intelligence Is a Familiar-Looking Monster, Say Henry Farrell and Cosma Shalizi'. *The Economist*, 21 June 2023. https://www.economist.com/by-invitation/2023/06/21/artificial-intelligence-is-a-familiar-looking-monster-say-henry-farrell-and-cosma-shalizi.; Farrell, Henry. 'Shoggoths amongst Us'. Substack newsletter. *Programmable Mutter* (blog), 30 June 2023. https://programmablemutter.substack.com/p/shoggoths-amongst-us.

	Supernatural entity (god-like AI,[90] demon[91])	Force beyond human understanding or control.
	Intelligence technology[92] (markets, bureaucracies, democracies[93])	Questions of bias, principal-agent alignment and control.
	Trick (hype)	Potential of AI exaggerated; questions of unexpected or fundamental barriers to progress, friction in deployment; "hype" as smokescreen or distraction from social issues.
Operation Terms focusing on *how AI works*	Autonomous system	Different levels of autonomy; human-machine interactions; (potential) independence from "meaningful human control"; accountability & responsibility gaps.
	Complex adaptive system	Unpredictability; emergent effects; edge case fragility; critical thresholds; "normal accidents".[94]
	Evolutionary process	Novelty, unpredictability, or creativity of outcomes;[95] "perverse" solutions and reward hacking.
	Optimization process[96]	Inapplicability of anthropomorphic intuitions about behavior.[97] Risks of the system optimizing for the

[90] Hogarth, Ian. 'We Must Slow down the Race to God-like AI'. *Financial Times*, 13 April 2023. https://www.ft.com/content/03895dc4-a3b7-481e-95cc-336a524f2ac2.

[91] McFarland, Matt. 'Elon Musk: "With Artificial Intelligence We Are Summoning the Demon."' *Washington Post*, 5 December 2021. https://www.washingtonpost.com/news/innovations/wp/2014/10/24/elon-musk-with-artificial-intelligence-we-are-summoning-the-demon/. ; Salmon, Paul M, Brandon King, Gemma J M Read, Jason Thompson, Tony Carden, Neville A Stanton, and Scott McLean. 'Summoning the Demon? Identifying Risks in a Future Artificial General Intelligence System'. *Human Factors*, 2023. https://research.usc.edu.au/esploro/outputs/conferencePaper/Summoning-the-demon-Identifying-risks-in/99721898702621.

[92] Dafoe, Allan. 'AI Governance: Overview and Theoretical Lenses'. In *The Oxford Handbook of AI Governance*, edited by Justin Bullock, Yu-Che Chen, Johannes Himmelreich, Valerie M. Hudson, Anton Korinek, Matthew Young, and Baobao Zhang, 0. Oxford University Press, 2023. https://doi.org/10.1093/oxfordhb/9780197579329.013.2

[93] The Economist. 'Artificial Intelligence Is a Familiar-Looking Monster, Say Henry Farrell and Cosma Shalizi'. 2023.

[94] Maas, Matthijs M. 'Regulating for "Normal AI Accidents": Operational Lessons for the Responsible Governance of Artificial Intelligence Deployment'. In *Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society*, 223–28. AIES '18. New York, NY, USA: Association for Computing Machinery, 2018. https://doi.org/10.1145/3278721.3278766.

[95] Lehman, Joel, Jeff Clune, Dusan Misevic, Christoph Adami, Lee Altenberg, Julie Beaulieu, Peter J. Bentley, et al. 'The Surprising Creativity of Digital Evolution: A Collection of Anecdotes from the Evolutionary Computation and Artificial Life Research Communities'. *Artificial Life* 26, no. 2 (21 November 2019). http://arxiv.org/abs/1803.03453.

[96] Yudkowsky, Eliezer. 'Artificial Intelligence as a Positive and Negative Factor in Global Risk.' In *Global Catastrophic Risks*, by Eliezer Yudkowsky, 308–45. New York: Oxford University Press, 2008. https://oxford.universitypressscholarship.com/view/10.1093/oso/9780198570509.001.0001/isbn-9780198570509-book-part-21. Pg. 5 ("The term 'Artificial Intelligence' refers to a vastly greater space of possibilities than does the term 'Homo sapiens.' When we talk about 'AIs' we are really talking about minds-in-general, or optimization processes in general").

[97] ibid.

	wrong targets or metrics;[98] Goodhart's Law;[99] risks from "reward hacking".
Generative system (generative AI)	Potential "creativity" but also unpredictability of system; resulting "credit-blame asymmetry" where users are held responsible for misuses, but can claim less credit for good uses, shifting workplace norms.[100]
Technology base (foundation model)	Adaptability of system to different purposes; potential for downstream reuse and specialization, including for unanticipated or unintended uses; risk that any errors or issues at the foundation-level seep into later or more specialized (fine-tuned) models;[101] questions of developer liability.
Agent[102]	Responsiveness to incentives and goals; incomplete-contracting and principal-agent problems;[103] surprising, emergent, and harmful multi-agent interactions[104] systemic, delayed societal harms and diffusion of power away from humans.[105]
Pattern-matcher (autocomplete on steroids,[106] stochastic parrot[107])	Problems of bias; mimicry of intelligence; absence of "true understanding"; fundamental limits.

[98] Thomas, Rachel, and David Uminsky. 'The Problem with Metrics Is a Fundamental Problem for AI'. *ArXiv:2002.08512 [Cs]*, 19 February 2020. http://arxiv.org/abs/2002.08512.

[99] Manheim, David, and Scott Garrabrant. 'Categorizing Variants of Goodhart's Law', 24 February 2019. http://arxiv.org/abs/1803.04585.

[100] Porsdam Mann, Sebastian, Brian D. Earp, Sven Nyholm, John Danaher, Nikolaj Møller, Hilary Bowman-Smart, Joshua Hatherley, et al. 'Generative AI Entails a Credit–Blame Asymmetry'. *Nature Machine Intelligence*, 4 May 2023, 1–4. https://doi.org/10.1038/s42256-023-00653-1.

[101] Jones, Elliot. 'Explainer: What Is a Foundation Model?' Ada Lovelace Institute, 17 July 2023. https://www.adalovelaceinstitute.org/resource/foundation-models-explainer/.

[102] Chan, Alan, Rebecca Salganik, Alva Markelius, Chris Pang, Nitarshan Rajkumar, Dmitrii Krasheninnikov, Lauro Langosco, et al. 'Harms from Increasingly Agentic Algorithmic Systems'. arXiv, 20 February 2023. https://doi.org/10.48550/arXiv.2302.10329.; for another overview of types of intelligent agents, see also Kilian, Kyle A., Christopher J. Ventura, and Mark M. Bailey. 'Examining the Differential Risk from High-Level Artificial Intelligence and the Question of Control'. *Futures* 151 (1 August 2023): 103182. https://doi.org/10.1016/j.futures.2023.103182. (pg. 3).

[103] Hadfield-Menell, Dylan, and Gillian Hadfield. 'Incomplete Contracting and AI Alignment'. In *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society*, 2019. http://arxiv.org/abs/1804.04268. For an accessible discussion of how this could emerge in modern deep learning models, see Cotra, Ajeya. 'Why AI Alignment Could Be Hard with Modern Deep Learning'. Cold Takes, 21 September 2021. https://www.cold-takes.com/why-ai-alignment-could-be-hard-with-modern-deep-learning/.

[104] Clifton, Jesse. 'Cooperation, Conflict, and Transformative Artificial Intelligence - A Research Agenda'. Center on Long-Term Risk, March 2020. https://longtermrisk.org/files/Cooperation-Conflict-and-Transformative-Artificial-Intelligence-A-Research-Agenda.pdf.

[105] Chan, Alan, et al. 'Harms from Increasingly Agentic Algorithmic Systems'. arXiv, 20 February 2023., pg. 11-14.

[106] Mollick, Ethan. 'Blinded by Analogies'. *One Useful Thing*, 23 February 2023. https://oneusefulthing.substack.com/p/blinded-by-analogies.

[107] Bender, Emily M., Timnit Gebru, Angelina McMillan-Major, and Shmargaret Shmitchell. 'On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?'; In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 610–23. FAccT '21. New York, NY, USA: Association for Computing Machinery, 2021. https://doi.org/10.1145/3442188.3445922.

	Hidden human labor (fauxtomation[108])	Potential of AI exaggerated; "hype" as a smokescreen or distraction from extractive underlying practices of human labor in AI development.
Relation Terms focusing on *how we relate to AI, as (possible) subject*	Tool (just technology, intelligent system[109])	Lack of any special relation towards AI, as AI is not a subject; questions of reliability and engineering.
	Animal[110]	Entities capable of some autonomous action, yet lacking full competence or ability of humans. Accordingly may be potentially deserving of empathy and/or (some) rights[111] or protections against abusive treatment, either on their own terms[112] or in light of how abusive treatment might desensitize and affect social behavior amongst humans;[113] questions of legal liability and assignment of responsibility to robots,[114] especially when used in warfare.[115]
	Moral patient[116]	Potential moral (welfare) claims by AI, conditional on certain properties or behavior.
	Moral agent	Machine ethics; ability to encode morality or moral rules.
	Slave[117]	AI systems or robots as fully owned, controlled, and directed by humans; not to be humanized or granted standing.

[108] Taylor, Astra. 'The Automation Charade'. *Logic Magazine*, 1 August 2018. https://logicmag.io/failure/the-automation-charade/.

[109] Molina, Martin. 'What Is an Intelligent System?' arXiv, 18 December 2022. https://doi.org/10.48550/arXiv.2009.09083. Pg. 1 ("computer-based tools are being created to automate tasks that require mental effort. [...] This evolution has generated a type of tool that we call intelligent system.").

[110] Darling, Kate. 'Robots Are Animals, Not Humans'. *Wired UK*, 14 April 2021. https://www.wired.co.uk/article/robots-animals-kate-darling.; though for a critical perspective, see: Johnson, Deborah G., and Mario Verdicchio. 'Why Robots Should Not Be Treated like Animals'. *Ethics and Information Technology* 20, no. 4 (1 December 2018): 291–301. https://doi.org/10.1007/s10676-018-9481-5.

[111] Turner, Jacob. *Robot Rules: Regulating Artificial Intelligence*. New York, NY: Springer Berlin Heidelberg, 2018. Pg. 137–143.

[112] Danaher, John. 'Welcoming Robots into the Moral Circle: A Defence of Ethical Behaviourism'. *Science and Engineering Ethics*, 20 June 2019. https://doi.org/10.1007/s11948-019-00119-x.

[113] Darling, Kate. 'Extending Legal Rights to Social Robots: The Effects of Anthropomorphism, Empathy, and Violent Behavior Towards Robotic Objects'. In *We Robot Conference 2012, University of Miami*, 2012. https://doi.org/10.2139/ssrn.2044797.

[114] Kelley, Richard, Enrique Schaerer, Micaela Gomez, and Monica Nicolescu. 'Liability in Robotics: An International Perspective on Robots as Animals'. *Advanced Robotics* 24, no. 13 (1 January 2010). https://papers.ssrn.com/abstract=2271471.

[115] Crootof, Rebecca. 'Autonomous Weapon Systems and the Limits of Analogy'. *Harvard National Security Journal* 9 (2018): 51–83. https://doi.org/10.2139/ssrn.2820727.

[116] For a critical argument, see: Moosavi, Parisa. 'Will Intelligent Machines Become Moral Patients?' *Philosophy and Phenomenological Research* n/a, no. n/a (9 September 2023). https://doi.org/10.1111/phpr.13019.

[117] See: Bryson, Joanna J. 'Robots Should Be Slaves'. In *Close Engagements with Artificial Companions: Key Social, Psychological, Ethical and Design Issue*, edited by Yorrick Wilks, 63–74, 2010. http://www.cs.bath.ac.uk/~jjb/ftp/Bryson-Slaves-Book09.html. Pg.1. ("My thesis is that robots should be built, marketed and considered legally as slaves, not companion peers."). Though for one critique, see Estrada, Daniel. 'Human Supremacy as Posthuman Risk'. *Computer Ethics - Philosophical Enquiry (CEPE) Proceedings* 2019, no. 1 (29 May 2019). https://doi.org/10.25884/6q27-6t77.

Legal entity (digital person, electronic person,[118] algorithmic entity[119])	Potential of assigning (partial) legal personhood to AI for pragmatic reasons (e.g., economic, liability, or risks of avoiding "moral harm"), without necessarily implying deep moral claims or standing.
Culturally revealing object (mirror to humanity,[120] blurry JPEG of the web[121])	Generally, implications of how AI is featured in fictional depictions and media culture.[122] Directly, AI's biases and flaws as a reflection of human or societal biases, flaws, or power relations. May also imply that any algorithmic bias derives from society rather than the technology per se.[123]
Frontier (frontier model[124])	Novelty in terms of both capabilities (increased capability and generality) and/or in form (e.g., scale, design, or architectures) compared to other AI systems; as a result, new risks because of new opportunities for harm, and less well-established understanding by the research community. Broadly, implies danger and uncertainty but also opportunity; may imply operating within a wild, unregulated space, with little organized oversight.
Our creation (mind children[125])	"Parental" or procreative duties of beneficence; humanity as good or bad "example."

[118] Nevejans, Nathalie. 'European Civil Law Rules in Robotics'. Study for the JURI Committee. Legal Affairs. European Parliament: Directorate-General for Internal Policies: POLICY DEPARTMENT C: CITIZENS' RIGHTS AND CONSTITUTIONAL AFFAIRS, 2016. http://www.europarl.europa.eu/RegData/etudes/STUD/2016/571379/IPOL_STU(2016)571379_EN.pdf.

[119] LoPucki, Lynn M. 'Algorithmic Entities'. *Washington University Law Review* 95, no. 4 (2018): 887–953. https://openscholarship.wustl.edu/law_lawreview/vol95/iss4/7/.

[120] But for a critique, see: Vallor, Shannon. 'The AI Mirror: Reclaiming Our Humanity in an Age of Machine Thinking'. In *Proceedings of the 2022 AAAI/ACM Conference on AI, Ethics, and Society*, 6. AIES '22. New York, NY, USA: Association for Computing Machinery, 2022. https://doi.org/10.1145/3514094.3539567.

[121] Chiang, Ted. 'ChatGPT Is a Blurry JPEG of the Web'. *The New Yorker*, 9 February 2023. https://www.newyorker.com/tech/annals-of-technology/chatgpt-is-a-blurry-jpeg-of-the-web.

[122] Cave, Stephen, Kanta Dihal, Eleanor Drage, and Kerry McInerney. 'Who Makes AI? Gender and Portrayals of AI Scientists in Popular Film, 1920–2020'. *Public Understanding of Science* 32, no. 6 (1 August 2023): 745–60. https://doi.org/10.1177/09636625231153985.

[123] Press, Gil. 'AI Is A Mirror, Not A Master, Says Tim O'Reilly'. Forbes, 28 November 2022. https://www.forbes.com/sites/gilpress/2022/11/28/ai-is-a-mirror-not-a-master-says-tim-oreilly/.

[124] See amongst others: UK Government. 'AI Safety Summit: Introduction'. GOV.UK, 25 September 2023. https://www.gov.uk/government/publications/ai-safety-summit-introduction/ai-safety-summit-introduction-html.; Google. 'A New Partnership to Promote Responsible AI'. Google, 26 July 2023. https://blog.google/outreach-initiatives/public-policy/google-microsoft-openai-anthropic-frontier-model-forum/.; Anderljung, Markus, Joslyn Barnhart, Anton Korinek, Jade Leung, Cullen O'Keefe, Jess Whittlestone, Shahar Avin, et al. 'Frontier AI Regulation: Managing Emerging Risks to Public Safety'. arXiv, 11 July 2023. https://doi.org/10.48550/arXiv.2307.03718. Pg. 6; Shevlane, Toby, Sebastian Farquhar, Ben Garfinkel, Mary Phuong, Jess Whittlestone, Jade Leung, Daniel Kokotajlo, et al. 'Model Evaluation for Extreme Risks'. arXiv, 24 May 2023. https://doi.org/10.48550/arXiv.2305.15324. Pg. 3. For a discussion of the regulatory implications of this term, see also Maas, Matthijs, 'Concepts in Advanced AI Governance: a Literature Review of Key Terms and Definitions.' *Institute for Law & AI*. AI Foundations Report #3. (2023). https://www.law-ai.org/advanced-ai-gov-concepts

[125] Moravec, H. *Mind Children: The Future of Robot and Human Intelligence*. New Ed edition. Cambridge: Harvard University Press, 1990.

	Next evolutionary stage or successor	Macro-historical implications; transhumanist or posthumanist ethics & obligations.
Function Terms focusing on *How AI is-, or can be used*	Companion (social robots, care robots, generative chatbots, cobot[126])	Human-machine interactions; questions of privacy, human over-trust, deception, and human dignity.
	Advis**or** (coach, recommender, therapist)	Questions of predictive profiling, "algorithmic outsourcing" and autonomy, accuracy, privacy, impact on our judgment and morals.[127] Questions of patient-doctor confidentiality, as well as "AI loyalty" debates over fiduciary duties that can ensure AI advisors act in their users' interests.[128]
	Malicious actor tool (AI hacker[129])	Possible misuse by criminals or terrorist actors. Scaling up of attacks as well as enabling entirely new attacks or crimes.[130]
	Misinformation amplifier (computational propaganda,[131] deepfakes, neural fake news[132])	Scaling up of online mis- and disinformation; effect on "epistemic security";[133] broader effects on democracy, electoral integrity.[134]

[126] Knudsen, Mikkel, and Jari Kaivo-Oja. 'Collaborative Robots: Frontiers of Current Literature'. *Journal of Intelligent Systems: Theory and Applications* 3, no. 2 (22 September 2020): 13–20. https://doi.org/10.38016/jista.682479.

[127] Aguirre, Anthony, Peter B. Reiner, Harry Surden, and Gaia Dempsey. 'AI Loyalty by Design: A Framework for the Governance of AI'. In *The Oxford Handbook of AI Governance*, edited by Justin B. Bullock, Yu-Che Chen, Johannes Himmelreich, Valerie M. Hudson, Anton Korinek, Matthew M. Young, and Baobao Zhang, 0. Oxford University Press, 2022. https://doi.org/10.1093/oxfordhb/9780197579329.013.70.

[128] Köbis, Nils, Jean-François Bonnefon, and Iyad Rahwan. 'Bad Machines Corrupt Good Morals'. *Nature Human Behaviour*, 3 June 2021, 1–7. https://doi.org/10.1038/s41562-021-01128-2.; for empirical work into this question, see also Krügel, Sebastian, Andreas Ostermaier, and Matthias Uhl. 'ChatGPT's Inconsistent Moral Advice Influences Users' Judgment'. *Scientific Reports* 13, no. 1 (6 April 2023): 4569. https://doi.org/10.1038/s41598-023-31341-0.

[129] Schneier, Bruce. 'The Coming AI Hackers'. Council for the Responsible Use of AI, Belfer Center for Science and International Affairs, 2021. https://www.schneier.com/wp-content/uploads/2021/04/The-Coming-AI-Hackers.pdf.

[130] Brundage, Miles, Shahar Avin, Jack Clark, Helen Toner, Peter Eckersley, Ben Garfinkel, Allan Dafoe, et al. 'The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation'. *ArXiv:1802.07228 [Cs]*, 20 February 2018. http://arxiv.org/abs/1802.07228.; Barrett, Clark, Brad Boyd, Ellie Burzstein, Nicholas Carlini, Brad Chen, Jihye Choi, Amrita Roy Chowdhury, et al. 'Identifying and Mitigating the Security Risks of Generative AI'. arXiv, 28 August 2023. https://doi.org/10.48550/arXiv.2308.14840.

[131] Woolley, S., and P. Howard. 'Political Communication, Computational Propaganda, and Autonomous Agents'. *International Journal Of Communication* 10, no. 9 (2016). https://demtech.oii.ox.ac.uk/research/posts/ijoc-political-communication-computational-propaganda-and-autonomous-agents-introduction/.

[132] Zellers, Rowan, Ari Holtzman, Hannah Rashkin, Yonatan Bisk, Ali Farhadi, Franziska Roesner, and Yejin Choi. 'Defending Against Neural Fake News'. *ArXiv:1905.12616 [Cs]*, 29 May 2019. http://arxiv.org/abs/1905.12616.

[133] Seger, Elizabeth, Shahar Avin, Gavin Pearson, Mark Briers, Seán Ó hÉigeartaigh, and Helena Bacon. 'Tackling Threats to Informed Decisionmaking in Democratic Societies: Promoting Epistemic Security in a Technologically-Advanced World'. The Alan Turing Institute, October 2020. https://www.turing.ac.uk/research/publications/tackling-threats-informed-decision-making-democratic-societies.

[134] But for a counterargument, see also: Simon, Felix M., Sacha Altay, and Hugo Mercier. 'Misinformation Reloaded? Fears about the Impact of Generative AI on Misinformation Are Overblown'. *Harvard Kennedy School Misinformation Review*, 18 October 2023. https://doi.org/10.37016/mr-2020-127.

Vulnerable attack surface[135]	Susceptibility to adversarial input, spoofing, or hacking.
Judge[136]	Questions of due process and rule of law; questions of bias and potential self-corrupting feedback loops based on data corruption.[137]
Weapon (killer robot,[138] weapon of mass destruction[139])	In military contexts, questions of human dignity,[140] compliance with laws of war, tactical effects, strategic effects, geopolitical impacts, and proliferation rates. In civilian contexts, questions of proliferation, traceability, and risk of terror attacks.
Critical strategic asset (nuclear weapons)[141]	Geopolitical impacts; state development races; global proliferation.
Labor enhancer (steroids,[142] intelligence forklift[143])	Complementarity with existing human labor and jobs; force multiplier on existing skills or jobs; possible unfair advantages & pressure on meritocratic systems.[144]

[135] Miessler, Daniel. 'The AI Attack Surface Map v1.0'. Unsupervised Learning, May 2023. https://danielmiessler.com/p/the-ai-attack-surface-map-v1-0. See also: Hayward, Keith J, and Matthijs M Maas. 'Artificial Intelligence and Crime: A Primer for Criminologists'. *Crime, Media, Culture* 17, no. 2 (30 June 2020): 209–33. https://doi.org/10.1177/1741659020917434.

[136] D'Amato, Anthony. 'Can/Should Computers Replace Judges?' *Northwestern University School of Law Scholarly Commons*, 1977, 20. https://scholarlycommons.law.northwestern.edu/cgi/viewcontent.cgi?article=1128&context=facultyworkingpapers; Chesterman, Simon, Lyria Bennett Moses, and Ugo Pagallo. 'All Rise for the Honourable Robot Judge? Using Artificial Intelligence to Regulate AI: A Debate'. *Technology and Regulation* 2023 (3 October 2023): 45–57. https://doi.org/10.26116/techreg.2023.005.

[137] Kamyshev, Pasha. 'Machine Learning In The Judicial System Is Mostly Just Hype'. *Palladium Magazine* (blog), 30 March 2019. https://palladiummag.com/2019/03/29/machine-learning-in-the-judicial-system-is-mostly-just-hype/.

[138] Sparrow, Robert. 'Killer Robots'. *Journal of Applied Philosophy* 24, no. 1 (2007): 62–77. https://doi.org/10.1111/j.1468-5930.2007.00346.x.; Crootof, Rebecca. 'The Killer Robots Are Here: Legal and Policy Implications'. *CARDOZO LAW REVIEW* 36 (January 2015): 80.

[139] Kallenborn, Zachary. 'Are Drone Swarms Weapons of Mass Destruction?' The Counterproliferation Papers, Future Warfare Series. U.S. Air Force Center for Strategic Deterrence Studies, Air University, 6 May 2020. https://media.defense.gov/2020/Jun/29/2002331131/-1/-1/0/60DRONESWARMS-MONOGRAPH.PDF. Bahçecik, Şerif Onur. 'Civil Society Responds to the AWS: Growing Activist Networks and Shifting Frames'. *Global Policy* 0, no. 0 (2019). https://doi.org/10.1111/1758-5899.12671.

[140] Rosert, Elvira, and Frank Sauer. 'How (Not) to Stop the Killer Robots: A Comparative Analysis of Humanitarian Disarmament Campaign Strategies'. *Contemporary Security Policy* 42, no. 1 (30 May 2020): 4–29. https://doi.org/10.1080/13523260.2020.1771508.

[141] Stern, Jacob. 'AI Is Like … Nuclear Weapons?' (2023); though for critiques of the comparison, see also Kaushik, Divyansh. 'Panic about Overhyped AI Risk Could Lead to the Wrong Kind of Regulation'. Vox, 3 July 2023. https://www.vox.com/future-perfect/2023/7/3/23779794/artificial-intelligence-regulation-ai-risk-congress-sam-altman-chatgpt-openai.; Matthews, Dylan. 'AI Is Supposedly the New Nuclear Weapons — but How Similar Are They, Really?' Vox, 29 June 2023. https://www.vox.com/future-perfect/2023/6/29/23762219/ai-artificial-intelligence-new-nuclear-weapons-future.

[142] Attiah, Karen. 'For Writers, AI Is like a Performance-Enhancing Steroid'. *Washington Post*, 13 January 2023. https://www.washingtonpost.com/opinions/2023/01/13/ai-writers-performance-enhancing-steroid/.

[143] Barak, Boaz. 'GPT as an "Intelligence Forklift."' *Windows On Theory* (blog), 19 May 2023. https://windowsontheory.org/2023/05/19/gpt-as-an-intelligence-forklift/.

[144] Porsdam Mann, Sebastian, Brian D. Earp, Sven Nyholm, John Danaher, Nikolaj Møller, Hilary Bowman-Smart, Joshua Hatherley, et al. 'Generative AI Entails a Credit–Blame Asymmetry'. *Nature Machine Intelligence*, 4 May 2023, 1–4. https://doi.org/10.1038/s42256-023-00653-1.

	Labor substitute	Erosive to or threatening of human labor; questions of retraining, compensation, and/or economic disruption.
	New economic paradigm (fourth industrial revolution)	Changes in industrial base; effects on political economy.
	Generally enabling technology (the new electricity / fire / internal combustion engine[145])	Widespread usability; increasing returns to scale; ubiquity; application across sectors; industrial impacts; distributional implications; changing the value of capital vs. labor; impacting inequality.[146]
	Tool of power concentration or control[147]	Potential for widespread social control through surveillance, predictive profiling, perception control.
	Tool for empowerment or resistance (emancipatory assistant[148])	Potential for supporting emancipation and/or civil disobedience.[149]
	Global priority for shared good	Global public good; opportunity; benefit & access sharing.
Impact Terms focusing on the *unintended risks,*	Source of unanticipated risks (algorithmic black swan[150])	Prospects of diffuse societal-level harms or catastrophic tail-risk events, unlikely to be addressed by market forces; accordingly highlights paradigms of "algorithmic preparedness"[151] and risk regulation more broadly.[152]
	Environmental pollutant	Environmental impacts of AI supply chain;[153] significant energy costs of AI training.

[145] Goode, Lauren. 'Google CEO Sundar Pichai Says AI Is More Profound than Electricity or Fire'. The Verge, 19 January 2018. https://www.theverge.com/2018/1/19/16911354/google-ceo-sundar-pichai-ai-artificial-intelligence-fire-electricity-jobs-cancer.

[146] Dafoe, Allan. 'AI Governance: Overview and Theoretical Lenses'. In *The Oxford Handbook of AI Governance*, edited by Justin Bullock, Yu-Che Chen, Johannes Himmelreich, Valerie M. Hudson, Anton Korinek, Matthew Young, and Baobao Zhang. Oxford University Press, 2022. https://doi.org/10.1093/oxfordhb/9780197579329.013.2 Pg. 7.

[147] Lazar, Seth. 'Power and AI: Nature and Justification'. In *The Oxford Handbook of AI Governance*, by Seth Lazar, edited by Justin Bullock, Yu-Che Chen, Johannes Himmelreich, Valerie M. Hudson, Anton Korinek, Matthew Young, and Baobao Zhang. Oxford University Press, 2022. https://doi.org/10.1093/oxfordhb/9780197579329.013.12.; Liu, Hin-Yan. 'The Power Structure of Artificial Intelligence'. *Law, Innovation and Technology* 10, no. 2 (3 July 2018): 197–229. https://doi.org/10.1080/17579961.2018.1527480.; Chapman, David. *Better without AI*, 2023. https://betterwithout.ai/.

[148] Kane, Gerald C., Amber Young, Ann Majchrzak, and Sam Ransbotham. 'Avoiding an Oppressive Future of Machine Learning: A Design Theory for Emancipatory Assistants'. *MIS Quarterly*, December 2020. https://misq.umn.edu/avoiding-an-oppressive-future-of-machine-learning-a-design-theory-for-emancipatory-assistants.html.

[149] Savaget, Paulo, Tulio Chiarini, and Steve Evans. 'Empowering Political Participation through Artificial Intelligence'. *Science and Public Policy* 46, no. 3 (1 June 2019): 369–80. https://doi.org/10.1093/scipol/scy064.; Lampell, Zach, and Lily Liu. 'How Can AI Amplify Civic Freedoms?' OpenGlobalRights, 18 December 2018. https://www.openglobalrights.org/how-can-AI-amplify-civic-freedoms/.

[150] Kolt, Noam. 'Algorithmic Black Swans'. *Washington University Law Review* 101 (25 February 2023). https://papers.ssrn.com/abstract=4370566.

[151] Ibid.

[152] For a study of the risk-regulation approach, see: Kaminski, Margot E. 'Regulating the Risks of AI'. *Boston University Law Review* 103 (19 August 2022). https://doi.org/10.2139/ssrn.4195066.

[153] Crawford, Kate. *Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence*. New Haven: Yale University Press, 2021.

benefits or side-effects of AI	Societal pollutant (toxin[154])	Erosive effects of AI on quality and reliability of the online information landscape.
	Usurper of human decision-making authority	Gradual surrender of human autonomy and choice and/or control over the future.
	Generator of legal uncertainty	Driver of legal disruption to existing laws;[155] driving new legal developments.
	Driver of societal value shifts	Driver of disruption to and shifts in public values;[156] value erosion.
	Driver of structural incentive shifts	Driver of changes in our incentive landscape; lock-in effects; coordination problems.
	Revolutionary technology[157]	Macro-historical effects; potential impact on par with agricultural or industrial revolution.
	Driver of global catastrophic or existential risk	Potential catastrophic risks from misaligned advanced AI systems or from nearer-term "prepotent" systems;[158] questions of ensuring value-alignment; questions of whether to pause or halt progress towards advanced AI.[159]

Different terms for AI can therefore invoke different frames of reference or analogies. Use of analogies—by policymakers, researchers, or the public—may be hard to avoid, and they can often serve as fertile intuition pumps.

[154] Chapman, David. *Better without AI*, 2023. https://betterwithout.ai/AI-is-public-relations ("demand that they remove their creepy 'neural' systems, and stop trying to read and manipulate your mind. Like microplastics in your water, these are the insidious, hidden toxins in your phone").

[155] Liu, Hin-Yan, Matthijs Maas, John Danaher, Luisa Scarcella, Michaela Lexer, and Leonard Van Rompaey. 'Artificial Intelligence and Legal Disruption: A New Model for Analysis'. *Law, Innovation and Technology* 12, no. 2 (16 September 2020): 205–58. https://doi.org/10.1080/17579961.2020.1815402. See also: Crootof, Rebecca, and B. J. Ard. 'Structuring Techlaw'. *Harvard Journal of Law & Technology* 34, no. 2 (2021): 347–417. https://jolt.law.harvard.edu/assets/articlePDFs/v34/1.-Crootof-Ard-Structuring-Techlaw.pdf

[156] Hopster, Jeroen K. G., and Matthijs M. Maas. 'The Technology Triad: Disruptive AI, Regulatory Gaps and Value Change'. *AI and Ethics*, 28 June 2023. https://doi.org/10.1007/s43681-023-00305-5.

[157] Garfinkel, Ben. 'The Impact of Artificial Intelligence: A Historical Perspective'. In *The Oxford Handbook of AI Governance*, edited by Justin B. Bullock, Yu-Che Chen, Johannes Himmelreich, Valerie M. Hudson, Anton Korinek, Matthew M. Young, and Baobao Zhang, 0. Oxford University Press, 2023. https://doi.org/10.1093/oxfordhb/9780197579329.013.5.

[158] Critch, Andrew, and David Krueger. 'AI Research Considerations for Human Existential Safety (ARCHES)', 29 May 2020. https://arxiv.org/abs/2006.04948.

[159] Future of Life Institute. 'Pause Giant AI Experiments: An Open Letter'. *Future of Life Institute* (blog), 30 March 2023. https://futureoflife.org/open-letter/pause-giant-ai-experiments/.; Grace, Katja. 'Let's Think about Slowing down AI'. AI Impacts, 22 December 2022. https://aiimpacts.org/lets-think-about-slowing-down-ai/.; Maas, Matthijs. 'Paths Untaken: The History, Epistemology and Strategy of Technological Restraint, and Lessons for AI'. *Verfassungsblog* (blog), 9 August 2022. https://verfassungsblog.de/paths-untaken/.

IV. The risks of unreflexive analogies

However, while metaphors can be productive (and potentially irreducible) in technology law, they also come with many risks. Given that analogies are shorthands or heuristics that compress or highlight salient features, challenges can creep in the more removed they are from the specifics of the technology in question.

Indeed, as Crootof and Ard have noted, "[a]n analogy that accomplishes an immediate aim may gloss over critical distinctions in the architecture, social use, or second-order consequences of a particular technology, establishing an understanding with dangerous and long-lasting implications."[160]

Specifically:

1. The selection and foregrounding of a certain metaphor hides that there are always multiple analogies possible for any new technology, and each of these advances different "regulatory narratives."

2. Analogies can be misleading by failing to capture a key trait of the technology or by alleging certain characteristics that do not actually exist.

3. Analogies limit our ability to understand the technology—in terms of its possibilities and limits—on its own terms.[161]

The challenge is that unreflexive drawing of analogies in a legal context can lead to ineffective or even dangerous laws,[162] especially once inappropriate analogies become entrenched.[163]

However, even if one tries to avoid explicit analogies between AI and other technologies, apparently "neutral" definitions of AI that seek to focus solely on the technology's "features" can and still do frame policymaking in ways that may not be neutral. For instance, Krafft and colleagues found that whereas definitions of AI that emphasize "technical functionality" are more widespread among AI researchers, definitions that emphasize "human-like performance" are more prevalent among policymakers, which they suggest might prime policymaking towards future threats.[164]

As such, it is not just loose analogies or comparisons that can affect policy, but also (seemingly) specific technical or legislative terms. The framing effects of such terms do not only occur at the level of broad policy debates but can also have strong legal implications. In particular, they can create challenges for law when narrowly specified regulatory definitions are suboptimal.[165]

---

[160] Crootof, Rebecca, and B. J. Ard. 'Structuring Techlaw'. *Harvard Journal of Law & Technology* 34, no. 2 (2021): 347–417. https://jolt.law.harvard.edu/assets/articlePDFs/v34/1.-Crootof-Ard-Structuring-Techlaw.pdf pg. 396.

[161] Maas, Matthijs M. 'Artificial Intelligence Governance Under Change: Foundations, Facets, Frameworks'. University of Copenhagen, 2020. http://www.legalpriorities.org/documents/Maas-PhD-Dissertation.pdf, Pg. 214.; drawing on: Crootof, Rebecca. 'Regulating New Weapons Technology'. In *The Impact of Emerging Technologies on the Law of Armed Conflict*, edited by Eric Talbot Jensen and Ronald T.P. Alcala, 1–25. Oxford University Press, 2019. https://oxford.universitypressscholarship.com/view/10.1093/oso/9780190915322.001.0001/oso-9780190915322-chapter-1 .

[162] Crootof, Rebecca, and B. J. Ard. 'Structuring Techlaw'. (2021). pg. 396-398.

[163] Ibid. pg. 398.

[164] Krafft, P. M., Meg Young, Michael Katell, Karen Huang, and Ghislain Bugingo. 'Defining AI in Policy versus Practice'. In *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*, 72–78. New York NY USA: ACM, 2020. https://doi.org/10.1145/3375627.3375835.

[165] This is discussed in further detail in: Maas, Matthijs, 'Concepts in Advanced AI Governance: a Literature Review of Key Terms and Definitions.' *Institute for Law & AI*. AI Foundations Report #3. (October 2023). https://www.law-ai.org/advanced-ai-gov-concepts

This creates twin challenges. On the one hand, picking suitable concepts or categories can be difficult at an early stage of a technology's development and deployment, when its impacts and limits are not always fully understood.[166] At the same time, the costs of picking and locking in the wrong terms or framings within legislative texts can be significant.

Specifically, beyond the opportunity costs of establishing better concepts or terms, unreflexively establishing legal definitions for key terms can create the risk of later, downstream "governance misspecification."[167] Such misspecification can occur when regulation is originally targeted at a particular artifact or (technological) practice through a particular material scope and definition for those objects. The implicit assumption here is that the term in question is a meaningful proxy for the underlying societal or legal goals to be regulated. While that may be appropriate in many cases, there is a risk that the law becomes less efficient, ineffective, or even counterproductive if either initial misapprehension of the technology or subsequent technological developments lead to that proxy term coming apart from the legislative goals.[168] Such misspecification can be seen in various cases of technology governance and regulation, including 1990s US export control thresholds for "high-performance computers" that treated the technology as far too static;[169] the Outer Space Treaty's inability to anticipate later Soviet Fractional Orbital Bombardment System (FOBS) capabilities, which were able to position nuclear weapons in space without, strictly, putting them "in orbit";[170] or initial early-2010s regulatory responses to drones or self-driving cars, which ended up operating on under- and overinclusive definitions of these technologies.[171]

Given this, the aim should not be to find the "correct" metaphor for AI systems. Rather, a good policy is to consider when and how different frames can be more useful for specific purposes, or for particular actors and/or (regulatory) agencies. Rather than aiming to come up with better analogies directly, this focuses regulatory debates on developing better processes for analogizing and for evaluating these analogies. For instance, such processes can depart from broad questions, such as:

1. What are the foundational metaphors used in this discussion of AI? What features do they focus on? Do these matter in the way they are presented?

---

[166] This can be understood as a more specific case or instance of the general 'Collingridge Dilemma'. See Collingridge, David. *The Social Control of Technology*. New York: Palgrave Macmillan, 1981. See also: Maas, Matthijs M. 'Innovation-Proof Governance for Military AI? How I Learned to Stop Worrying and Love the Bot'. *Journal of International Humanitarian Legal Studies* 10, no. 1 (2019): 129–57. https://doi.org/10.1163/18781527-01001006. Pg. 132-135. Nonetheless, there are also arguments in favor of the general feasibility of forward-looking, "anticipatory" regulation, even at an early stage. See Guston, David H. 'Understanding "Anticipatory Governance"'. *Social Studies of Science* 44, no. 2 (April 2014): 218–42. https://doi.org/10.1177/0306312713508669. Armstrong, Harry, and Jen Rae. 'A Working Model for Anticipatory Regulation'. Nesta, 2017. https://media.nesta.org.uk/documents/working_model_for_anticipatory_regulation_0_TpDHt7z.pdf.

[167] I thank Christoph Winter for introducing the term and concept.

[168] In a legal context, this echoes HLA Hart's classic "no vehicles in the park" dilemma–the situation where a certain rule (say, at a city park) was originally formulated to ban certain objects (e.g., motor vehicles) from a park, but where it was phrased without awareness of other objects (e.g., bicycles, roller skates, electric wheelchairs, or drones) that might fall under this terminology, creating later uncertainty over whether it would—or why it should—apply to these new objects. See Hart, H. L. A. 'Positivism and the Separation of Law and Morals'. *Harvard Law Review* 71, no. 4 (February 1958): 593. https://doi.org/10.2307/1338225. Pg. 607. See also Schlag, Pierre. 'No Vehicles in the Park'. *Seattle University Law Review* 23 (1999): 381–89. https://digitalcommons.law.seattleu.edu/cgi/viewcontent.cgi?article=1623&context=sulr .

[169] Picker, Colin B. 'A View from 40,000 Feet: International Law and the Invisible Hand of Technology'. *Cardozo Law Review* 23 (2001): 151–219. Pg. 212.

[170] Maas, 'Artificial Intelligence Governance Under Change: Foundations, Facets, Frameworks'. pg. 197-205.

[171] See Calo, Ryan. 'The Case for a Federal Robotics Commission'. Brookings Institute Center for Technology Innovation, 1 September 2014. https://papers.ssrn.com/abstract=2529151. Pg. 6, 8 (discussing a 2011 incident where Nevada passed accidentally overinclusive self-driving car regulations, which had to be repealed after it turned out that they inadvertently imposed stringent obligations on existing vehicles with partially-autonomous features, as well as cases where US laws against drone surveillance ended up focusing far too much on flying drones, rather than other mobile robots).

2. What other metaphors could have been chosen for these same features or aspects of AI?
3. What aspects or features of AI do these metaphors foreground? Do they capture these features well?
4. What features are occluded? What are the consequences of these being occluded?
5. What are the regulatory implications of these different metaphors? In terms of the coalitions they enable or inhibit, the issue and solution portfolios they highlight, or of how they position the technology within (or out of) the jurisdiction of existing institutions?

Improving these ways in which we analogize AI clearly needs significantly more work. However, it is critical that we do so to improve how we draw on frames and metaphors for AI and to ensure that—whether we are trying to understand AI itself, appreciate its impacts, or govern them effectively—our metaphors aid rather than lead us astray.

Conclusion

As AI systems have received significant attention, many have invoked a range of diverse analogies and metaphors. This has created an urgent need for us to better understand (a) when we speak of AI in ways that (inadvertently) import one or more analogies, (b) what it does to utilize one or another metaphor for AI, (c) what different analogies could be used instead for the same issue, (d) how the appropriateness of one or another metaphor is best evaluated, and (e) what, given this, might be the limits or risks of jumping at particular analogies.

This report has aimed to contribute to answers to these questions and enable improved analysis, debate, and policymaking for AI by providing greater theoretical and empirical backing to how metaphors and analogies matter for policy. It has reviewed 5 pathways by which metaphors shape and affect policy and reviewed 55 analogies used to describe AI systems. This is not meant as an exhaustive overview but as the basis for future work.

The aim here has not been to argue against the use of metaphors but for a more informed and reflexive and careful use of these metaphors. Those who engage in debate within and beyond the field should at least have greater clarity about the ways that these concepts are used and understood, and what are the (regulatory) implications of different framings.

The hope is that this report can contribute foundations for a more deliberate and reflexive choice over what comparisons, analogies, or metaphors we use in talking about AI—and for the ways we communicate and craft policy for these urgent questions.

Also in this series

→ Maas, Matthijs, and Villalobos, José Jaime. 'International AI institutions: A literature review of models, examples, and proposals.' *Institute for Law & AI*, **AI Foundations Report 1**. (September 2023). https://www.law-ai.org/international-ai-institutions

→ Maas, Matthijs, 'Concepts in advanced AI governance: A literature review of key terms and definitions.' *Institute for Law & AI*. **AI Foundations Report 3**. (October 2023). https://www.law-ai.org/advanced-ai-gov-concepts

→ Maas, Matthijs, "Advanced AI governance: A literature review." *Institute for Law & AI*, **AI Foundations Report 4**. (November 2023). https://law-ai.org/advanced-ai-gov-litrev

INSTITUTE
FOR LAW & AI

law-ai.org