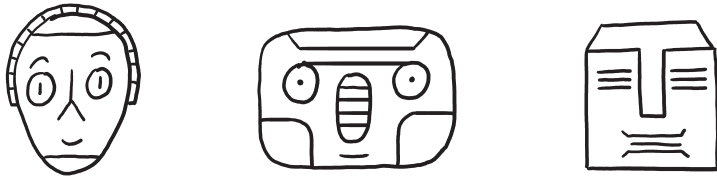
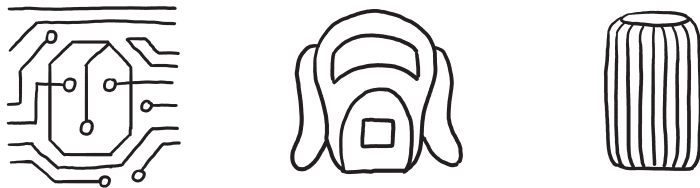


making sense of the unknown



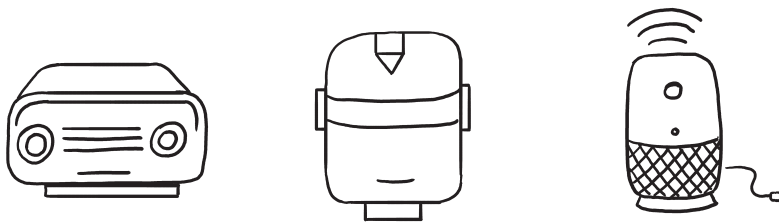
From *Ex Machina* to *Black Mirror*,



metaphors color our thinking about AI.



How can we use it to our benefit?



We all know what artificial intelligence (AI) looks like, right? Like HAL 9000, in *2001: A Space Odyssey*—a disembodied machine that turns on its “master.” Less fatal but more eerie AI is Samantha in the movie *Her*. She’s an empathetic, sensitive and sultry-voiced girlfriend without a body—until she surprises with thousands of other boyfriends. Or perhaps AI blends the two, as an unholy love child of Hal and Samantha brought to “life” as the humanoid robot Ava in *Ex Machina*. Ava kills her creator to flee toward an uncertain freedom.

These images are a big departure from their benevolent precursors of more than half a century ago. In 1967, as a poet in residence at Caltech, Richard Brautigan imagined wandering through a techno-utopia, “a cybernetic forest / filled with pines and electronics / where deer stroll peacefully / past computers / as if they were flowers / with spinning blossoms.” In this post-naturalistic world, humans are “watched over / by machines of loving grace.” Brautigan’s poem painted a metaphorically expressed anticipatory mythology—a gleefully optimistic vision of the impact that the artificially intelligent products California’s emerging computer industry would make on the world.

But Brautigan’s poem captured only a small subset of the range of metaphors that over time have emerged to make sense of the radical promise—or is it a threat?—of artificial intelligence. Many other metaphors would later arrive not just from the birthplace of the computer industry. They jostled and competed to make sense of the profound possibilities that AI promised.

Today, those in the AI industry and the journalists covering it often cite cultural narratives, as do policy-makers grappling with how to regulate, restrict, or otherwise guide the industry. The tales range from ongoing invocations of Isaac Asimov's *Three Laws of Robotics* from his short story collection *I Robot* (about machine ethics) to the Netflix series *Black Mirror*, which is now shorthand for our lives in a datafied dystopia.

Outside Silicon Valley and Hollywood, writers, artists and policy-makers use different metaphors to describe what AI does and means. How will this vivid imagery shape the ways that human moving parts in AI orient themselves toward this emerging set of technologies?

More like Munich or more like Korea?

When we encounter a novel situation that defies established concepts, to make sense of the unknown we tend to search for analogies to familiar past situations. In other words, metaphors “tame” the new. They open it up to the imagination.

Famously, the debate about what to do about Vietnam in 1965 in Lyndon Johnson's presidential administration was ultimately a dispute between those who described the situation as “more like Munich”—thus demanding escalation rather than peace-making—or “more like Korea”—a quagmire to be avoided.¹ In fact, the truth lay in between, not as a blend of previous episodes. Both metaphors were misleading in different ways, and yet they were used extensively in debates and decision-making.

Indeed, as Richard Neustadt and Ernest May argued in their seminal *Thinking in Time*,² which offers a critical view of how policy-makers can best make use of history as a guide for both analysis and action, when faced with a novel challenge human nature inevitably leads us to analogies. We must thus both be aware of our implicit analogical frames *and* be explicit about such thinking by naming directly how the current situation *is* both and, perhaps more important, how it *is not* like the historical analogy one references.

metaphors “tame” the new
they open it up to the imagination

We don't only turn to metaphors when confronting the new. In fact, metaphors are vital thinking tools. As German philosopher Hans Blumenberg argued, because there can never be any direct, unmediated access to reality, metaphors are the irreducible lenses through which thought happens. Metaphors shape our knowledge of the world, doing what Blumenberg calls *thought work*, simplifying and thus making accessible complex concepts that without these comparisons we would fail to grasp.³

AI technologies challenge the notion of the human itself

When it comes to AI, metaphors abound because technology makes possible radically new forms of sense-making, perception, feeling and cognition co-produced through people's interactions with technical affordances and infrastructures.

What makes AI so new and unsettling? Unlike many post-industrial-revolution information and communication technologies, AI technologies challenge the notion of the human itself. Philosopher Tobias Rees writes that for people who believe that the human is distinct from and superior to the natural world, to animals and to Earth itself, and that they are more than *mere* machines, to suddenly face machines that appear lively, seductive, competent and clever could be annihilation. AI is like a cracked mirror, reflecting long-held notions of ourselves as humans *and* as humans living with, through and in, complicated conjunction with machines. That colliding image is long overdue for an update.

What are our AI metaphors?

Using metaphors helps nonexperts understand how we build, interact with and regulate technology. “Information just wants to be free,” “Data is the new oil” and computer security described as an infection or as a transgression by diseased or “foreign” bodies are some enduring metaphors.⁴

Metaphors often subsume and disguise the creation of technologies. For example, researchers Cornelius Puschmann and Jean Burgess found that the metaphors used for describing big data tend to obscure the material and political conditions of the production and ownership of that data.⁵ Through the use of a highly specific set of terms, the role of data as a valued commodity is effectively inscribed (e.g., “the new oil”)—most often by suggesting physicality, immutability, context independence and intrinsic worth.

Sometimes AI metaphors are explicit: We need “a Food and Drug Administration [FDA] for AI” or “a Motion Picture Association of America [MPAA] for AI.” Some people suggest AI represents a “new kind of market,” whereas others liken it to “a human rights challenge.” As for AI’s regulatory approaches, suggestions include “a peer-reviewed process” for algorithms, “an Institutional Review Board for data uses” or even “a Supreme Court” for algorithms.

Just as the lead-up to the Vietnam war differed from Munich or Korea, so the emergent reality of AI is more complex than any one metaphor suggests. Although we cannot help but think in metaphors about the technology, we need to be aware of how we are using those analogies to describe AI and the ways they emphasize certain features, obscure others and may distract us from what is truly novel about AI. Thus, for example, the metaphor of “an FDA for AI” suggests a government-defined regulatory approach that entails a systematic, defined process for inspecting and approving AI applications. By contrast, “an MPAA for AI” implies self-regulation.

What both of these metaphors hide is that AI is neither a distinct industry nor even a specific set of products. It is a general-purpose technology that has become embedded in myriad products and that is steadily penetrating and rapidly transforming every corner of the economy. A clearer and more explicit assessment of the metaphors used for governing AI should therefore reveal the limits of those metaphors for analysts to be more self-aware of the assumptions—even biases that these metaphors bring along.

AI and culture

Visions of AI are invariably linked to culture. In her collage *Faith and Trust* (2019), for example, artist Şerife Wong examines the dominant metaphors emerging through search results for the term *AI*. By collecting images to identify connections—and gaps—in the words and images we use to describe AI, Wong finds:

There is a preponderance of images that present AI as God, a divine creative force or as a spark—a connection between the human and the divine. Often there is a spark of light—electrical and like sunlight but also evocative of a Frankenstein-like creation. Those images of generative power all draw from Michelangelo’s *Creation of Adam* in that technology is a savior of mankind or can give us power over nature.

Wong points to images of robot and human shaking hands; sometimes their outstretched hands are on bodies wearing suits. In the images, AI is our partner, almost human, something to work with and toward. But that “AI is good” narrative creates a duality because it also implies that

there is a preponderance of images
that present AI as, variously, god,
a creative force or divine or as a spark—
a connection between the human and the divine

“AI is (also) evil.” The collage’s primary colors, Wong notes, are mostly cool tones of blues, greens and purples that convey the feeling of AI as cold. The tones suggest that we perceive AI to be nonhuman—not natural, warm or friendly. And herein lie contradictions: AI is creative, God-like, almost human but also unnatural, cold and not human. The real meaning of AI will emerge from these contradictions.

To echo Rees, AI unsettles our notion of ourselves as human and, by extension, our relationship to things other than human. That said, the range of AI images and AI metaphors is not geographically uniform. In fact, different metaphors about AI prevail in different countries. Americans discuss AI with metaphors that are materially different from those the Chinese or the Indians or the Europeans use. In China, AI is seen as less of a threat, less of a Frankensteinian monster or an engine of oppression the way it is often viewed in the West. In China, AI is more of a symbol of how the country as co-leader in AI development alongside the United States is regaining its rightful place on the world stage.

In Japan, AI-powered robots aren’t considered either vaguely or explicitly sinister. They are more often viewed as friendly, as souped-up versions of the famous Tamagotchi hand-held digital pets.⁸ Such personification of machines is not new. Buddhists administered funeral rites for the “dead” bionic pet AIBO in 2006, when Sony discontinued its animatronic dog after seven years of popularity. Since AIBO’s 2018 relaunch, AIBO owners of all ages gather at cafés marked with its logo. Japan is also the home of a range of other companion robots like LOVOT and PARO. These robots, modeled on a baby harp seal, are often used in medical care and mental health settings, with older people, and with children who have autism.

Against that backdrop, Japanese robotics scientists building *kansei* robots (loosely translated as *affective computing in robotics*) embed into their designs the concept of relationality between the robot and the human. They seek to imbue their robots with *kokoro*—Japanese for the *integration of emotion, intelligence and intention*. It is also the origin of intelligence

and emotion. That unique framing of robots—which is distinct from Western perceptions of robots as not quite human—has helped Japanese roboticists create a national context and a potential market for robotics in Japan and other East Asian countries.

A particular set of cultural norms shapes the contours and limits of human experience and feeling in these metaphors. Robotic traits might include amusement, play and curiosity of a deeply intimate kind with nonhumans. Such traits and attachments could be considered similar to those we ascribe to and form about our own pets. But dogs are not connected to the cloud and don't read our social media feeds. (Cats, however, might!) However, scholar Kate Devlin finds that when it comes to sex robots and fembots, we revert to traditional notions of gender, bodies and sexuality. Gone are the uncertainties related to nonhuman others. On the other end of the spectrum, some Japanese men warm to more conservative traditional holographic "girlfriend"/digital assistants like Azuma Hikari. The assistants embody subservient female attributes through their actions, such as turning on lights before their owners return to empty apartments, ordering their owners' favorite dinners and welcoming orders from their owners.

knowing the range of
global metaphors used
across the world to
make sense of AI will
also be valuable for
technologists developing
AI applications

limiting our engineers'
imagining of what these
technologies can do and
mean to their own norms
and habits will limit
AI's possibilities

Why consider metaphors?

Lessons from the global AI metaphors landscape stretch well beyond cultural insights to material implications for policy-makers—especially those negotiating global pacts to regulate AI. Because the metaphors bring assumptions about what AI can or may do today and in the future, they shape the debate about how national governments should regulate AI. These distinct beliefs about the role policy-makers in national governance of AI in turn inform—albeit often in inexplicit ways—the ambitions and boundaries that different governments consider for AI transnational regulations and treaty obligations. Understanding different AI metaphors is thus a precondition for understanding and agreeing on global AI regulations.

Knowing the range of global metaphors used across the world to make sense of AI will also be valuable for technologists developing AI applications. Limiting our engineers' imagining of what these technologies can do and mean to their own norms and habits will limit AI's possibilities. Cataloging and assessing metaphors used to describe and imagine AI's potential and prospects will help transfer color from these metaphors to creations, thus expanding these technologies' unique and unprecedented possibilities for their more human or multidimensional qualities.

More ambitiously, because AI calls into question our long-standing understanding of the human, assessment of the metaphoric foundations of the discourse around AI will help us imagine our own humanity in radically new ways. These metaphors comfort us in the face of the what we cannot yet conceptually grasp. As we grow our awareness of the unique affordances of AI, we may eventually develop new, more adequate concepts about ourselves. On that journey, AI metaphors help us navigate the unknown.

-
1. YF Khong, *Analogies at War: Korea, Munich, Dien Bien Phu, and the Vietnam Decisions of 1965*. Princeton, NJ: Princeton University Press, 1992.
 2. RE Neustadt, ER May. *Thinking in Time: The Uses of History for Decision Makers*. New York: Simon & Schuster, 2011 [1986].
 3. H Blumenberg, *Paradigmen zu einer Metaphorologie*. Berlin: Suhrkamp, 1999 [1960].
 4. S Helmreich, Flexible infections: computer viruses, human bodies, nation-states, evolutionary capitalism. *Sci Tech Hum Values* 2000;25L4:472-491.
 5. Cornelius Puschmann and Jean Burgess, "Big Data, Big Questions: Metaphors of Big Data," *International Journal of Communication* 8 (2014): 20. In reference to Perry Rotella, "Is Data the New Oil?" *Forbes*, 2 April 2012.
 6. S Šabanović, "Inventing Japan's "robotics culture": the repeated assembly of science, technology, and culture in social robotics." *Soc Stud Sci* 2014;44:342-367.