ce-manifesto-ch1-0425.docx
ce-ch2-risk-land-intro-0324-part1
ce-ch3-eco-overshoot-0324-part1
ce-ch4-hum-syst-0324-part1
latest-ce-ch5-nat-disaster-0525-part1
Chapter 6: Advanced Tech_ZAK_DRAFT
near-final-ce-ch7-violent-conflict-0324-part1
ce-ch8-two-attrct-threshold-0324-part1

Chapter 6: Advanced Tech

Technology distinguishes humanity from the rest of the biosphere[1]. Many animals use objects as tools and even alter objects to make them more useful. Animals profoundly transform ecosystems, e.g., through the large-scale construction of dams (beavers) and mounds (ants). Yet, humans are unique on Earth, in large part, because of our ability to think abstractly and thus to design increasingly powerful tools and technologies to achieve long-term goals (see Box 1 for definitions of tools, technologies, etc.)[2]. Technological innovation is further enabled by the human capacity for language – a kind of social technology[3] – which encodes and transmits

---

[1] While animals utilize certain technologies (such as basic tools or complex dwelling structures), the scale of recursive technological innovations humans are capable of (nuclear bombs, space stations, AI, etc.) points to a place where a difference in magnitude becomes a difference in kind.
[2] There is a large and intricate literature on technology and its history, the approach here is influenced by:
Mumford, Lewis. *Technics and Civilization*. Hardcott, 1934.
Mumford, Lewis. *The Myth of The Machine*. Hardcott, 1967.
Ellul, Jacques. *The Technological Society*. Vintage Books, 1954.
McLuhan, Marshall. *Understanding Media: The Extensions of Man*. New York: McGraw-Hill, 1964.
Winner, Langdon. *Autonomous Technology*. MIT Press, 1978.
See also more contemporary work, Bratton, Benjamin. *The Stack: On Software and Sovereignty*. MIT Press, 2016.
Frischmann , Brett, and Selinger, Evan. *Re-Engineering Humanity*. Cambridge University Press, 2018.
[3] Our discussion of advanced technology is focused on certain forms of human creativity and innovation which include technologies like AI, biotechnology, nuclear weapons, computation and other more recent innovations discussed in this chapter. However, it is helpful to consider technology more broadly to all forms of abstract human reasoning and creativity that have practical implications for shared human life. This broader definition is a return to the ancient notion of *techne*. See Parry, Richard, "Episteme and Techne", The Stanford Encyclopedia of Philosophy (Winter 2024 Edition), Edward N. Zalta & Uri Nodelman (eds.), https://plato.stanford.edu/archives/win2024/entries/episteme-techne/.

knowledge across generations creating a cumulative ratcheting effect of increasing knowledge and technical ability.

Cumulative technological innovation has been an essential factor in humanity's history of interaction with the forces of nature. By advancing a certain kind of "technological intelligence," humans have been able to radically increase in population, create economic surplus, and scale up a massive technologically sophisticated civilization, despite the complexity and potential dangers presented by the vast natural world.

Eventually whole "ecosystems" of technology emerge in which the combined applications of different technologies increase their overall power. National and then planetary-scale technological infrastructures begin to interlink and transform every aspect of human experience. The nature of these technological infrastructures defines what is often called "technological epochs" (See Box 1). And today, humanity is entering a new technological epoch— one we are referring to simply as "advanced technologies."

New technologies result in new civilizations. "*We shape our tools and in turn they shape us*[4]" – our values, our minds, and our worlds. The written word, for example, was a pivotal innovation in the development of large-scale civilization. Humans had been encoding knowledge through cave paintings and story telling for thousands of years before the advent of written language[5]. But the tax records for a large Mesopotamian city-state of some 20,000 people could not have been maintained by word of mouth and mnemonic devices[6]. The social-technology of written language allowed humans to store and transmit information in ways that were previously impossible via oral communication and memory alone. Observing the effects of this technology, some ancient philosophers (such as Plato) believed written language would ultimately be a

---

In Greek, techne meant ' the practical application of knowledge' which included more than physical tools and infrastructure but also mathematics, art, rhetoric, and more. This more general notion of techne also includes technologies such as the alphabet which profoundly altered forms of social organization, and are therefore as revolutionary of technologies as the plow or the steam engine. This perspective also corresponds with other philosophers of technology, such as Marshall McLuhan. McLuhan famously argued that "Societies have always been shaped more by the nature of the media by which men communicate than by the content of the communication." The alphabet was an early medium through which humans coordinated collective action, and more recent communication technologies such as the printing press, radio, and digital networks are all extensions of this more basic trend.
McLuhan, Marshall. *The Medium is the Massage: An Inventory of Effects.* Gingko Press, 2001.
McLuhan, Marshall. *Understanding Media.* McGraw-Hill, 1964.

[4] Paraphrase from Culkin, John. M. "A schoolman's guide to Marshall McLuhan." *The Saturday Review*, 51-53, (1967):70-72. https://www.unz.com/print/SaturdayRev-1967mar18-00051
[5] The oldest known cave paintings come from Neanderthals, 64,000 years ago. Little, Becky. *What Prehistoric Cave Paintings Reveal About Early Human Life.* History, 2025. https://www.history.com/articles/prehistoric-cave-paintings-early-humans
In comparison, the earliest known writing system, Mesopotamian cuneiform, emerged in approximately 3,500-3,200 BCE. Wright, J. *The Evolution of Writing*. International Encyclopedia Of Social And Behavioral Sciences, Elsevier, 2014. https://sites.utexas.edu/dsb/tokens/the-evolution-of-writing/
[6] See the work of Scott, James, C. *Against the Grain: A Deep History of the Earliest States*. Yale, 2017.

detriment to human memory[7]. Many indigenous communities had also worried that the written word would draw us too far into realms of abstraction, placing us further and further from the realities of what we were communicating about. Eventually, as a result of less direct contact with others and the environment, we would become less aware of the consequences of our actions and therefore more capable of harming things that were of essential value to us[8].

Also consider the plow, an innovation at the heart of the history of early civilizations. The plow was coupled with other technologies such as threshing screens, baskets, roads, carts, and storehouses. Storehouses needed protecting, so military technologies proliferated around them. The plow also required certain forms of animal domestication, as large and potentially dangerous mammals were made essential to agricultural practices. This technological epoch involved a transformation in humans' relationship to the natural world. For millennia many human cultures perceived how the natural world was imbued with subjectivity and consciousness. The animal-drawn plow made it difficult for humans to value animals as sacred and as equals. It is difficult to worship an animal if they must also be castrated, yoked, and beaten to drive a plow. This way of life, enabled by a certain kind of technology, justified the belief that humanity's role was to control and subjugate nature, rather than live in harmony with, and as a steward of it.

---

[7] Plato. Plato's Phaedrus. Cambridge : University Press, 1952. "If men learn this, it will implant forgetfulness in their souls; they will cease to exercise memory because they rely on that which is written, calling things to remembrance no longer from within themselves, but by means of external marks."

[8] See for example Yunkaporta, Tyson. *Sand Talk.* HarperCollins, 2020. "Oral cultures are known as high-context or field-dependent-reasoning cultures. They have no isolated variables: all thinking is dependent on the field or context…high-context cultures demand dialogue and complex agreements. They use a lot of nonverbal communication and leave many things unspoken due to common shared understandings and established consensus about the way things are done."

In Abrams, David. "Animism and the Alphabet." In *Spell of the Sensuous*, Chapter Four. Vintage, 1997. Abrams writes about how "With the advent of the [Semitic] aleph-beth, a new distance opens between human culture and the rest of nature. "

Box 1: Technology Definitions (from Consilience Project).

Layers of the Civilizational Tech-Stack

Consider the categories below as overlapping; this overall model is only a heuristic. There are a variety of comparable stacks proposed in the academic literature, many of which have influenced the model presented here. This is not proposed as definitive but as a useful set of orienting generalizations.

Tools
Human-scale artifacts, found or made, which augment individual and social practices: rocks, axes, forks, writing implements, etc.

Technologies
The application of complex (scientific) knowledge to solving problems, embedded in intentionally designed artifacts that are complicated enough to require engineering: waterwheel; steam engine; light bulb; refrigerator.

Ecologies of technologies
Sets of technologies that are symbiotically related and co-evolving as nested functional units: e.g., light bulb/lamp/power lines/transformers/power station; and microchip/hard drive/screen/mouse/modem/broadband/server banks.

Infrastructures
Multiple different ecologies of technology embedded together to form a basic part of social coordination and material reproduction within a society: supply chains, transportation systems, markets, and communication systems.

Technological epochs
A duration of historical time characterized by a specific suite of infrastructures that are interrelated as the foundation of a social system. Epochs are marked by discontinuous breaks from prior infrastructures, and the emergent social dynamics resulting from new ones: e.g. pre-industrial; industrial; and post-industrial.

Technology is a defining aspect of what it means to be human. It is also core to the metacrisis assessment we are outlining here. Technological "progress" results in new, more complex, and often more consequential problems. The Industrial Revolution powered a rapidly growing population and an increasingly technologically sophisticated civilization, and without it, we would never have passed planetary boundaries of depletion and pollution, leading to the risk of an uninhabitable Earth. Similarly, atomic energy was the result of successfully winning a technological arms race. The potential for self-induced species extinction is directly a function of reaching a critical point of technological advancement.

This book is, in part, seeking to invoke a culture-wide process of *re-envisioning what it means for civilization to advance* and to truly demonstrate progress. Should an innovation, for example, artificial general intelligence, be considered progress if it simultaneously diminishes humanity's chances for a future? This chapter begins with a brief discussion of some of the various worldviews surrounding technology, including the various forms of Luddism which seek to roll back all technological advancements, and also the dominant narrative in current culture – techno-optimism or accelerationism – which sees nearly all technological innovation as positive. As will become clear, we are advocating for neither a return to a time before advanced technology nor a blind acceptance of technological acceleration. Instead, we are seeking a realistic appraisal of the history of innovation and a revisioning of our relationship to human technological intelligence and its aspirations[9].

As already mentioned, we are living during a time when technological evolution has just passed an inflection point. Humanity has entered a new technological epoch; an historical era has begun that will be defined by what we are calling "advanced technology" (see Table 2). Emerging technological capabilities are working at primordial levels of depth and are moving at exponential speeds in terms of advancement and distribution. There are serious proposals and ongoing efforts to redesign the species and civilization with some combination of brain-computer interfaces, genetic engineering, and artificial intelligence. AI technologies can perform at or above human levels at an increasing number of tasks, such as in strategy and war games, precision control of robotics, drone races, and even applications in law and medicine[10]. Brain-computer interfaces can alter the motor functions of animals, such as involuntarily moving their limbs[11]. We have now edited the genomes of human beings, for example, to confer

---

[9] CRI has considered this topic at length elsewhere: Consilience Project. *Development in Progress*. 2024. https://consilienceproject.org/development-in-progress/
The Consilience Project. *Technology is Not Values Neutral: Ending the Reign of Nihilistic Design.* 2022. https://consilienceproject.org/technology-is-not-values-neutral-ending-the-reign-of-nihilistic-design-2/
The Consilience Project. *The Case Against Naive Technocapitalist Optimism.* 2021. https://consilienceproject.org/the-case-against-naive-technocapitalist-optimism/

[10] Bostrom, Nick. *Superintelligence: Paths, Dangers, Strategies.* Oxford University Press, 2014.
For an overview of AI progress and expert perspectives, see: ESPAI. *2022 Expert Survey on AI Progress.* 2022. https://aiimpacts.org/2022-expert-survey-on-progress-in-ai
Kiela, Douwe., Thrush, Tristan, Ethayarajh, Kawin, and Singh, Amanpreet. *Plotting Progress on AI.* Contextual AI Blog, 2023. https://contextual.ai/plotting-progress-in-ai

[11] Saha, Simanto, Mamun, Khondaker, and Ahmed, Khawza et al. "Progress in Brain-Computer Interface: Challenges and Opportunities." *Frontiers in Systems Neuroscience*, 15, (2021).https://doi.org/10.3389/fnsys.2021.578875

resistance to certain diseases. We are also seeing the creation of synthetic life forms: organic life (not completely "artificial" or silicon-based) designed in a lab with the help of AI-empowered computer simulation, which can replicate and self-organize. Further examples of *advanced technologies* include virtual reality, the Internet of Things, nanotechnology (engineering at the atomic and molecular levels), sensors and satellites, blockchain, social media, and alternative substrates for computing (e.g., thermodynamic, quantum, biological, etc.).

It is generally the case that a new technology can cause harm in at least two ways. For one, it can enable someone to cause intentional harm given the unique capabilities of the technology. For example, dynamite was originally designed as a more stable explosive which could be utilized for civilian purposes such as to support construction and development projects[12], but it also served another purpose of being an easily accessible and rather powerful, weaponized explosive. Then there is the possibility that a technology can be used for relatively benign or even benevolent purposes while simultaneously producing unintended, harmful consequences (externalities) to the environment or society. Ecological overshoot is rife with such examples, including the use of chlorofluorocarbons (CFCs) for refrigerants which also burned a hole in the ozone layer[13].

The potential for risks of all kinds is greater in the case of newly emerging, advanced technologies. We are now engineering at the foundations of physics and life, with many of these technologies advancing at near-exponential rates of change – speeds humans did not evolve to process. The rapid development and deployment of new technologies has always been a recipe for both intentional and accidental harm, but these harms were often believed to be acceptable because they were overshadowed by the potential benefits. However, today the capabilities of technology are reaching a critical point. The consequences of mistakes and acts of violence are becoming so destructive that civilization may be unable to recover. Runaway pandemics, for example, could be the result of deliberate attacks with bioweapons or imperfect safety practices in a lab. A powerful AI system that fails to act in alignment with humanity's values, goals, and best interests ("Misaligned AI") could come by accident or an act of war.

Risks from today's technologies are unique for several reasons, including the speed of innovation, the scale of impact, the widespread distribution through the commercial sector, and the power of emerging capabilities. If continued, long-standing patterns of technology development, such as innovation driven by competitive pressures in business or war, could lead to global catastrophes or dystopian oppression. To safely steward increasingly powerful technology requires a corresponding increase in the wisdom of those designing and using such technologies. Marshaling the requisite wisdom to regulate advanced technologies would be a profound shift in humanity's relationship to technology. It must become commonplace that the potential benefits of any given innovation are acknowledged without downplaying the severity of

[12] Kravitz, Fran, and ACS Committee on Ethics. *Dynamite and the Ethics of its Many Uses.* ACS. Accessed June 17th, 2025.
https://www.acs.org/education/celebrating-chemistry-editions/2021-ncw/dynamite-ethics.html

[13] American Chemical Society National Historic Chemical Landmarks. *Chlorofluorocarbons and Ozone Depletion.* 2017. http://www.acs.org/content/acs/en/education/whatischemistry/landmarks/cfcs-ozone.html

its potential harms and risks, even leaving open the possibility that there are certain technologies that may simply be too dangerous to create in the first place.

Techno-Optimism

Before reviewing the risks from emerging technologies, it is useful to first consider some aspects of our recent history with technology, including worldviews surrounding its development and approaches to regulating its harms. This will serve as a backdrop as we reflect on the future of human innovation and the ways our species is being fundamentally reshaped in relation to its creations.

Depending upon who is asked, technology could be the savior of humanity or the source of its demise. To the infamous Luddites of the 19th century destroying looms, for example, machines were stealing their jobs and were ultimately reflections of exploitation, unnecessary excess, and the willingness of the wealthy to sacrifice others' livelihoods and the quality of their products for additional profits.[14] To those profiting from and purchasing cheaper materials, however, the same technology was seen to be a significant innovation in efficiency. It enabled new freedoms, allowing people to spend their saved time and money elsewhere.

Ecological overshoot may also be viewed through the lens of competing worldviews surrounding technology. For those who appreciate the affordances of modern life, the environmental crisis is an undesirable outcome but perhaps a necessary evil to achieve today's high standards of living. Only by means of further innovation, for example, in renewable energy, carbon sequestration, and various forms of solar geoengineering, can we respond to ecological overshoot while still preserving our quality of life. Some environmentalists, however, view technology as the cause of the fabled fall from Eden, the source of humanity's separation from nature. They hold that, to return to the right relationship with the earth, we must look back in time to when humans had limited technology and lived in equilibrium with the biosphere.

Worldviews favoring the proliferation and rapid advancement of technology have largely overshadowed other perspectives. This is in part because new technologies confer power to those creating and wielding them, reinforcing, and further disseminating those ideologies which advocate for technology's intrinsic value. Today's techno-optimists and techno-accelerationists have continued with the view that technology is either inherently positive or well worth the risks and is ultimately the driving force behind all improvements in the human condition[15]. It is

---

[14] For a recent view comparing Luddism with contemporary resistance to advanced technologies, see Merchant, Brian. *Blood in the Machine: The Origins of the Rebellion Against Big Tech.* Little, Brown, and Company, 2023.
[15] See the recent "Techno-Optimist Manifesto" by venture-capitalist Marc Andressen: https://a16z.com/the-techno-optimist-manifesto. Or an interview of the "effective accelerationist" philosopher, Beff Jezos: https://www.youtube.com/watch?v=8fEEbKJoNbU

believed that technology solves our problems and makes our lives easier and better. When it's cold, technology keeps us warm. If we are in pain or danger, technology cures our ailments and saves our lives. It is even the source of our entertainment and education, supporting us in all of our creative aspirations.

Undesired outcomes of technology are typically seen to be a product of faulty or unethical use by humans, rarely a product of the technology itself. Even in cases of misuse, technical innovation remains the primary solution: increasing security and defense systems, better surveillance or cryptography, simulations of potential harms, training better models for deplatforming and content moderation. It is humanity's destiny to continue in its works of technological expansion and advancement. Technology is the answer to most questions, the solution to our greatest problems, and the path towards a world of abundance for all.

Techno-optimists, of which there are several varieties, are seeking to accelerate technological innovation as rapidly as possible. They argue that technology is the only way humanity will be saved from itself. The most pernicious problems we face can only be addressed by radically increasing the rate of innovation. If we are to survive climate change, we must become a multi-planetary species and quickly begin geoengineering the earth. Nanotechnology and gene editing, enabled by advanced robotics, monitored by sensor networks and satellites, will help us respond to growing public health crises and put an end to disease. The coming artificial superintelligence is the only way to address humanity's inability to globally cooperate to solve its grandest challenges. While many advocates of this view genuinely believe it and should be taken as acting in good faith, it is also true that those in the technology industry have the most to gain from rapid development and deployment and the most resources to propagate accelerationist narratives. Their public messaging cannot be meaningfully separated from their personal motivations.

Legacy Technology: Externalities and Retro-Active Regulation

Worldviews like techno-optimism and accelerationism are understandable reactions to technological developments corresponding to improvement in several dimensions of quality of life and obvious material abundance. However, these worldviews tend to trivialize how the intended benefits of technology have always been matched with the reality of unintended harms (or negative externalities). The light of human "progress" casts a shadow. Unintended harms and externalities accumulate, eventually becoming grave risks to society writ large. In the cases of increasingly powerful technologies, unaccounted for externalities become increasingly consequential, even existentially risky. This can be independent of whether or not the technology is used maliciously. An innovation can be harmful simply as a byproduct of its design and rapid, large-scale deployment.

Recent history is littered with examples of commercial products that were first advertised as safe and essential, only for their negative consequences to be revealed after widespread distribution and use. Leaded gasoline is a particularly instructive example. It was argued to be critical for smooth, efficient, high performing engines. Later it was shown to be dangerously toxic (and that its creators were aware of this in advance). As a result of its massive distribution, humanity lost nearly a billion points of IQ of cognitive decline, became significantly more aggressive and impulsive, and suffered countless deaths and ubiquitous public health issues[16].

To add insult to injury, the negative health effects of lead were written about over a thousand years prior to it being used for gasoline. During the Roman Empire lead was used to serve many purposes including for water pipes, cooking, cosmetics and medicine[17]. Its deleterious health effects likely served as one among many factors contributing to Rome's collapse[18]. For most of evolutionary history, the majority of lead was stored deep in the Earth's lithosphere. Due to its toxicity, it is likely that complex biological life would not have evolved in the presence of extensive lead exposure. But in the name of technological progress, humanity mined and atomized the compound, distributing it through the atmosphere, having it rain down upon us, our children, and our soil poisoning every bite of food we eat and sip of water we drink[19].

Asbestos is another famous example, ubiquitously integrated into commercial products for its fire-retardant properties, simultaneously a cause of lung cancer, mesothelioma, and other deleterious health effects[20]. The list of innovations that were recognized to be harmful after being brought to market (many of which were later withdrawn from production) is extensive: parathion[21], malathion, thalidomide, Vioxx Premarin[22], CFCs[23], HFCs[24], BPAs[25], phthalates, and single-use plastics. And these are just cases of harmful products that were recognized and banned. Externalized harm is the norm for nearly all innovations.

---

[16] Ritchie, Hannah. *How the world eliminated lead from gasoline.* OurWorldInData.org, 2022. https://ourworldindata.org/leaded-gasoline-phase-out

[17] McConnell, Joseph, R., Chellman, Nathan, J., and Plach, Andreas et al. "Pan-European atmospheric lead pollution, enhanced blood lead levels, and cognitive decline from Roman-era mining and smelting." *Proc. Natl. Acad. Sci. U.S.A.*, 122(3), (2025):e2419630121. https://doi.org/10.1073/pnas.2419630121

[18] (Ibid.)

[19] Noviandari, Lina. *Lead (Pb) 101: Everything You Need To Know About Lead.* Pure Earth, 2024. https://www.pureearth.org/lead101-2/

[20] National Cancer Institute. *Asbestos Exposure and Cancer Risk.* 2021. https://www.cancer.gov/about-cancer/causes-prevention/risk/substances/asbestos/asbestos-fact-sheet

[21] Rosenberg, Yvonne, J., Garcia, Kelly, and Diener, Justin et al. "The impact of solvents on the toxicity of the banned parathion insecticide." *Chem Biol Interact.* 2023):382:110635. doi: 10.1016/j.cbi.2023.110635

[22] Prakash, Snigdha, and Valentine, Vikki. *Timeline: The Rise and Fall of Vioxx.* NPR, 2007. https://www.npr.org/2007/11/10/5470430/timeline-the-rise-and-fall-of-vioxx

[23] EPA. *Ban for Nonessential Products Containing Ozone-depleting Substances.* Accessed June 17th, 2025. https://www.epa.gov/ozone-layer-protection/ban-nonessential-products-containing-ozone-depleting-substances

[24] Sidley. *U.S. EPA Bans Hydrofluorocarbons in Refrigeration, Air Conditioning, and Heating Products.* 2023. https://environmentalenergybrief.sidley.com/2023/10/12/u-s-epa-bans-hydroflourocarbons-in-refrigeration-air-conditioning-and-heating-products/

[25] Breast Cancer Prevention Partners. *BPA Laws and Regulations.* Accessed June 17th, 2025. https://www.bcpp.org/resource/bpa-laws-and-regulations/

The Haber-Bosch method (a technique for creating fertilizers via large-scale synthesis of ammonia from hydrogen and nitrogen), seen by many as the most influential innovation in human history, was partially responsible for the population explosion of the 20th century due to its ability to increase agricultural yields. Before the "Green Revolution," the number of humans on the Earth was under two billion; now we are nearing 8 billion[26]. The Haber-Bosch method led to over 100x resource per capita usage worldwide due to population growth, driving increased extraction from the natural world, increased energy demand, and increased waste and pollution[27]. One impact of this has been a dramatic increase in agricultural runoff, resulting in around 400 oceanic dead zones, some of which are as large as the UK[28].

This agricultural innovation also led to micronutrient deficiencies in our food (i.e., loss of trace minerals, phytochemicals, vitamins, etc.) due to the use of synthetic fertilizers (nitrogen, phosphorus, potassium fertilizers in particular: "NPK")[29]. The food that we eat today has a far lower vitamin and mineral content, leading to specific deficiencies and health impacts (such as heart-disease)[30]. Haber-Bosch has also led to an increase in chronic disease and pain[31]. "Diseases of abundance," such as obesity, diabetes, heart disease, cancer, and a variety of mental health issues, are a direct consequence of the change to our food supply. Surplus alone did not cause this. Micronutrient deficiencies can create a feeling of perpetual hunger. Many in today's population are, in some sense, starving to death while they are also in a state of over-consumption.

These examples are not outliers. They reflect a common pattern where technology is extremely useful to us by some standards while simultaneously leading to consequences that no one wants. No one wants climate change, but it is a direct consequence of industrial globalization. Similarly, plastics are a "pillar of modern civilization" that everyone now depends upon, but they also create toxic nanoparticles that are distributed everywhere in the biosphere, poisoning

---

[26] Smil, V. (1999). *Detonator of the Population Explosion.* Nature. https://doi.org/10.1038/22672

[27] Steffen, Will, Broadgate, Wendy, Deutsch, Lisa, Gaffney, Owen, and Ludwig, Cornelia. "The Trajectory of the Anthropocene, the Great Acceleration." *The Anthropocene Review 2*, no. 1, (2015). https://doi.org/10.1177/2053019614564785

[28] Schulte-Uebbing, Lena, Beusen, A.H.W., Bouwman, Alexander, F., and de Vries, Wim. "From Planetary to Regional Boundaries for Agricultural Nitrogen Pollution." *Nature 610*, no. 7932, (2022):507–512. doi:10.1038/s41586-022-05158-2
Stevens, Carly, J. "Nitrogen in the Environment." *Science.org* 363, no. 6427, (2019):578-580. doi:10.1126/science.aav8215

[29] Nelson, Ann Raeboline Lincy Eliazer, Ravichandran, Kavitha, and Antony, Usha. "The Impact of the Green Revolution on Indigenous Crops of India." *Journal of Ethnic Food 6,* no. 8, (2019). https://doi.org/10.1186/s42779-019-0011-9

[30] Jensen, Bernard, and Anderson, Mark. *Empty Harvest: Understanding the Link Between Our Food, Our Immunity, and Our Planet.* New York: Avery Publishing, 1995.
Via, Michael. "The Malnutrition of Obesity: Micronutrient Deficiencies That Promote Diabetes." *International Scholarly Research Notices* (2012). https://doi.org/10.5402/2012/103472.

[31] Horrigan, Leo, Lawrence, Robert, S., and Walker, Polly. "How Sustainable Agriculture Can Address the Environmental and Human Health Harms of Industrial Agriculture." *Environmental Perspectives* 110, no. 5, (2022). https://ehp.niehs.nih.gov/doi/abs/10.1289/ehp.02110445
Winson, Anthony. *The Industrial Diet: The Degradation of Food and the Struggle for Healthy Eating.* New York: NYU Press, 2014.

plants and animals, disrupting our hormonal cycles, causing cancers, and detrimentally impacting fertility and prenatal development[32]. Antibiotics have saved millions of lives from bacterial infection, and their overuse has led to extremely dangerous antibiotic-resistant bacteria, deadly chronic infections, disruption to the human microbiome, and negative impacts on development when prescribed to babies[33]. Most people would vote against rapidly disseminating these technologies if they were able to experience the totality of their consequences up front and were given a collective choice as to whether or not to move forward.

Within recent years most new technologies have come from the market directly funding, or capitalizing on scientific breakthroughs. Corporations then tend to disproportionately focus on the innovation's upside, bringing to market a technology that is, at first, largely unregulated and poorly understood. First mover advantage then allows the company to define the market and the public's understanding of the new technology with sophisticated public relations and marketing campaigns ("4 out of 5 Doctors Smoke Camel Cigarettes," or "Better Living Through Chemistry," for example). Today we hear similar sale-pitches from AI and biotech companies today who tell us that they will usher in an age of previously unimaginable health and abundance.

---

[32] Hong, Yifan, Wu, Shengde, and Wei, Guanghui. "Adverse effects of microplastics and nanoplastics on the reproductive system: A comprehensive review of fertility and potential harmful interactions." *Science of The Total Environment* 903, (2023):166258. https://doi.org/10.1016/j.scitotenv.2023.166258
Ullah, Sana, Ahmad, Shahid, and Guo, Xinle et al. "A review of the endocrine disrupting effects of micro and nano plastic and their associated chemicals in mammals." *Frontiers in Endocrinology* 13, (2022):1084236. https://doi.org/10.3389%2Ffendo.2022.1084236
Engel, Stephanie, M., Patisaul, H., and Brody, Charlotte et al. "Neurotoxicity of Ortho-Phthalates: Recommendations for Critical Policy Reforms to Protect Brain Development in Children." *AJPH 111*(4) (2021). https://doi.org/10.2105/AJPH.2020.306014
Brynzak-Schreiber, Ekaterina, Schogl, Elisabeth, and Bapp, Carolin et al. "Microplastics role in cell migration and distribution during cancer cell division." *Chemosphere* 353, (2024):141473. https://doi.org/10.1016/j.chemosphere.2024.141463
Park, Jun Hyung, Hong, Seungwoo, and Kim Ok-Hyeon et al. "Polypropylene microplastics promote metastatic features in human breast cancer." *Scientific Reports* 13, (2023):6252. https://doi.org/10.1038/s41598-023-33393-8
Chen, Guangquan, Shan, Huang, and Xiong, Shiyi et al. "Polystyrene nanoparticle exposure accelerates ovarian cancer development in mice by altering the tumor microenvironment." *Science of the Total Environment* 905, (2023):167592. doi:10.1016/j.scitotenv.2023.167592
Zarus, Gregory, M., Muianga, Custodio, and Brenner, Stephan et al. "Worker studies suggest unique liver carcinogenicity potential of polyvinyl chloride microplastic." *American Journal of Industrial Medicine*, 66(12), (2023):1033-1047. https://doi.org/10.1002/ajim.23540
[33] Murray, Christopher, J.L., Ikuta, Kevin, S., and Sharara, Fablina et al. "Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis." *The Lancet* 399(10325), (2022):629-655. https://doi.org/10.1016/S0140-6736(21)02724-0
Anthony, Winston, E., Wang, Bin, and Sukhum, Kimberly, V., et al. "Acute and persistent effects of commonly used antibiotics on the gut microbiome and resistome in healthy adults." *Cell* 39(2), (2022):110649. https://doi.org/10.1016/j.celrep.2022.110649
Patangia, Dhrati, V., Ryan, Cornelius A., Dempsey, Eugene, Paul Ross, Reynolds, and Stanton, Catherine. "Impact of antibiotics on the human microbiome and consequences for host health." *MicrobiologyOpen* 11(1), (2022):e1260. https://doi.org/10.1002%2Fmbo3.1260
Uzan-Yulzari, Atara, Turta, Olli, and Belogolovski, Anna et al. "Neonatal antibiotic exposure impairs child growth during the first six years of life by perturbing intestinal microbial colonization." *Nature Communications* 12(443), (2021). https://doi.org/10.1038/s41467-020-20495-4

The promises of a technology and the competitive pressures to develop it tend to overshadow its potential externalities, which accumulate until they pass a critical point where the harms are undeniable to public consciousness. Only after harm has happened at enough scale with enough severity and concern, does regulation get enacted. This is *retroactive regulation* – creating laws to prevent harm only after it has occurred – as opposed to *proactive* or *anticipatory governance*. For example, it wasn't until six decades after its development and distribution that Japan became the first country to ban leaded gasoline completely, and it took another three and a half decades for it to be banned worldwide, only after irreparable damage had been done to the public and the environment[34]. Other examples, such as toxic "forever chemicals" (such as PFASS) and fossil fuels, are still in widespread use today. When a technology creates profound dependence that cannot easily be replaced or removed, its harms or risks are simply endured.

The history of trusting the commercial sector to do its own safety analysis is not reassuring[35]. As structured, the market has the incentive to build everything that can create returns on investment, quickly, as long as it doesn't technically break the law (or if the fines for illegal activities are less than the profit made doing them). Incentives are to do minimal safety assessments and to move ahead on every profitable, powerful area of technology while promoting narratives that downplay the risks and exaggerate the benefits. As the technology is being brought to market, it is argued to be somewhere between necessary and good. Then, once harms become widespread, it is often later revealed that the risks were well known ahead of time and hidden – as has been well documented, for example, with the negative health effects of forever chemicals like PFASS[36], the ecological effects of fossil fuels[37], and the detrimental mental health effects of social media on teens[38].

It is true that technologies of the near future could bring many benefits to humanity. But without a profound re-imagining of how we design, deploy, and regulate technology, the current trajectory is likely to involve grave risks. If left ungoverned, advanced technology will result in a deeper degradation of our minds, relationships, cultures, and the planet. Robotic automation, for example, could save humanity from dangerous and repetitive labor, but without the necessary,

---

[34] Hofverberg, Elin. *The History of the Elimination of Leaded Gasoline.* Library of Congress, 2022. https://blogs.loc.gov/law/2022/04/the-history-of-the-elimination-of-leaded-gasoline/

[35] For historical examples across several industries and sectors, see: Wilson, James, Q. *The Politics of Regulation.* Basic Books, 1980.
For an account of how self-regulation led to a specific tragedy in India see Lapierre, Dominique, and Moro, Javier. *Five Past Midnight in Bhopal: The Epic Story of the World's Deadliest Industrial Disaster.* Grand Central Publishing, 2002.

[36] Hayes, Jared. *For decades, polluters knew PFAS chemicals were dangerous but hid risks from the public.* EWG, 2019. https://www.ewg.org/research/decades-polluters-knew-pfas-chemicals-were-dangerous-hid-risks-public

[37] Center for Climate Integrity. *Big Oil knew as early as 1954.* 2024. https://climateintegrity.org/news/view/big-oil-knew-as-early-as-1954

[38] Dolman, Matthew. *Whistleblower Exposes Meta's Knowledge of Negative Effects of Social Media on Young People.* Lawsuit Legal News, 2024. https://lawsuitlegalnews.com/news/whistleblower-exposes-metas-knowledge-of-negative-effects-of-social-media-on-young-people/

associated changes to economics, it could very easily create large-scale technological unemployment and an unprecedented underclass. Similarly, AI could potentially play a role in synthesizing the world's information and deliver it in a pedagogically useful fashion, uniquely suited to our individual context and needs, but at present, it seems more poised to usher in a world of ubiquitous deep fakes and the destruction of shared knowledge. Responding to the multiple crises and challenging planetary conditions of the metacrisis will require a method of technology design and regulation that rigorously factors the reality of potential harms and is able to steer our capacity for innovation towards the outcomes with the greatest likelihood of safety and shared benefit.

Advanced Technology Risks: The Ungovernability Threshold

These historical patterns of technology development – rapid development and deployment under near-term competitive pressures, disproportionately emphasizing the benefits and downplaying potential harms[39], and implementing safety measures only after harm has occurred – cannot continue in an age of increasingly powerful technology. Today's paradigm of advanced technology points to the crossing of a threshold where continuing on the path of retroactive regulation now poses global catastrophic and existential risks. With innovations like those in AI, biotechnology, sensor-webs, and brain-computer interfaces, putting in measures to prevent harm only after it has already occurred could be too late, leaving us little chance to recover. Innovations in anticipatory technology governance are required to respond to the changes in the speed, scope, complexity and consequentiality of advanced technology risk.

Below we briefly develop a framework to assess the potential risk of a new technology (box 2) defined as being roughly proportional to a technology's power and inversely proportional to our ability to govern it well. Power involves the speed at which it can be developed, deployed, and improved; the scope and magnitude of its potential effects; and the complexity of the technology and the world system that will co-evolve with it. Our ability to govern a technology depends on our ability to understand it, our ability to monitor its use, and our ability to control against excessively harmful or risky applications. Advanced technologies are unique in terms of both their power and ungovernability, marking a transition into a new epoch where today's technologies are placing us into exceedingly dangerous territory.

Box 2: Framework for Understanding Technology Risk

[39] For instance Elon Musk has said publicly: "I'd rather be optimistic and wrong than pessimistic and right. At least err on that side. If you're pessimistic, you're going to be miserable. You might as well enjoy the journey." https://www.youtube.com/watch?v=zGMpyeOf9g8

$$\frac{\text{Power of Technology}}{\text{Governability of Technology}} = \textbf{Risk of Technology}$$

****

$$\frac{\textbf{Power} = \text{Speed x Magnitude x Scope x Complexity}}{\textbf{Governability} = \text{Understandability x Controllability}} = \textbf{Risk of Technology}$$

Table 1 details this framework and shows its application to advanced technology. What follows is an abbreviated overview of the framework itself. It is applied to clarify the risk of certain classes of advanced technology.

POWER AS SPEED OF ADVANCEMENT AND DISTRIBUTION

The power of a technology can roughly be considered in terms of four components. First, there is the speed of the technology's distribution, measured in terms of how quickly it impacts large numbers of people. Cars, for instance, reached 50 million users some 60 years after they were first brought to market[40]. In comparison, Meta's social media app Threads reached its first 50 million users in 24 hours[41]. The largest technology companies from the prior epoch reached users around 22,000 times slower than today. Reasonable analogies would be between a snail and a bullet, plant growth and a rocket launch.

Considerations of *speed* also factor the rate at which a technology's capacities advance. Televisions went from black and white to color and from a few channels to hundreds over a span of decades.Technologies like artificial intelligence and gene sequencing are improving much faster than computation did under Moore's Law, which was already considered truly unprecedented throughout the history of innovation[42]. Significant advancements in efficiency and capability are occurring in matters of weeks or months. For example, in 2020, scientists at OpenAI showed that compared to results from 2012, the amount of computational power needed to train a neural net to the same performance as a popular AI image classification

[40] TikTok took 9 months to reach 100 million users, while Threads reached the same milestone in 5 days. Rao, Pallavi. *How Long it Took for Popular Apps to Reach 100 Million Users.* Visual Capitalist, 2023. https://www.visualcapitalist.com/threads-100-million-users/
[41] (Ibid.)
[42] Moore's Law is the observation that the number of transistors on a microchip roughly doubles every two years.

system (AlexNet) had decreased by a multiple of 44[43]. Over the same period, Moore's Law would have yielded only an 11x improvement.

POWER AS MAGNITUDE OF CHANGE POSSIBLE

Power also points to the *magnitude* of the scale of change made possible by a new technology. A technology is less powerful if it can only affect a highly localized area. Locomotives would be mostly useless without the railway, and the technology's transformative potential was only realized as railway systems expanded across more and more territory. Today our technologies reach far beyond the local or even the transcontinental and move into the geographies of the atmosphere and the interplanetary as we send robots to mars and surround the Earth in an omniscient web of sensors and satellites.

Increases in magnitude are also tied to the depth of reality that a technology is intended to manipulate. No example of this is more clear than nuclear technology, where the first planetary-scale existential risk was created precisely by working on some of the deepest aspects of reality (i.e., splitting the atom). In many ways this can be considered the beginning of what we are calling advanced technology. And the trend of increasing depth, and therefore greater magnitude of consequence, continues today with projects in nanotechnology and quantum engineering, which attempt to go even "deeper" than the atom[44].

POWER AS SCOPE OF USE CASES

The power of a technology further depends on the *scope* of possible combinations it has with existing technologies, or how technologies amplify their effects when interfacing with other existing or soon-to-exist technologies. In prior eras, new technologies were less able to easily interface with the existing technological surround. When cars were first invented and sold, there were no gas stations and few suitable roads. It took more than a century until new kinds of technologies—like electric and self-driving vehicles—could be easily integrated into the existing transportation infrastructures, making all the technologies more powerful through interoperability and amplification.

The scope of a technology is also related to the degree to which it can be used for multiple purposes. An automobile can be used as a home, a canvas, a weapon, or any number of other "off-label" uses. But it is designed to be a means of transportation and does not do these other things well. Computers, on the other hand, are useful for nearly any task. Computation

---

[43] OpenAI. *AI and efficiency.* 2020. https://openai.com/index/ai-and-efficiency/
[44] Congressional Research Service. *Defense Primer: Quantum Technology*, 2022. https://crsreports.congress.gov/product/pdf/IF/IF11836
Burja, Samo. *Quantum Technology Appeals to World Powers.* Bismarck Brief, 2022. https://brief.bismarckanalysis.com/p/quantum-technology-appeals-to-world
Witt, Stephan. *The World-Changing Race to Develop the Quantum Computer.* The New Yorker, 2022. https://www.newyorker.com/magazine/2022/12/19/the-world-changing-race-to-develop-the-quantum-computer

technologies like the transistor fundamentally transformed all industries, well beyond the implications its early designers could have possibly imagined[45].

Of all the possible use-cases for a technology, weaponization is obviously among the most concerning from a risk standpoint. Historical discussions of these issues have been framed in terms of "dual-use"—meaning military or civilian, or destructive vs constructive. The "same" (often slightly modified) technology can be either a weapon or an asset that benefits "the people" by means other than violence. Again a canonical example here is nuclear technology, which created both energy and weaponry based on the same underlying innovation.

In general, when technology is designed with a relatively narrow purpose in mind, it will eventually be employed for other purposes anyone finds it useful for. The creation of a new technical capability for any purpose results in its being used for many other purposes. Therefore, ***the default assumption should be that new technologies will be innovated and deployed by all agents for all purposes they enable.***

Technologies exist on a spectrum from being nearly useless outside of a few application areas to having basically infinite use-cases and transforming almost all areas of life. From this vantage point, various classes of technology are perhaps better described in terms of being *multi*- or *omni-use* rather than strictly dual-use (see box z). Advanced technologies can be used for an extraordinary range of purposes, some of which fall outside of standard categories like military and civilian; for example, designing new species, altering the climatic conditions of the planet with geo-engineering, editing one's genetic makeup, and creating virtual worlds. Of course, the distinction of dual-use may still be applied in such cases, but it is valuable to foreground how certain innovations can be applied to almost every area of life for nearly every kind of purpose.

Computation is omni-use by design. The digital revolution impacted every industry and branch of science, changing governments and the military, and of course, reaching into the personal lives of most people. Omni-use technology culminates with AI, a technology that can be applied to advance virtually every sector, even its own[46]. In conjunction with sensors and robotics, AI is, in theory, capable of automating, advancing, and extending a whole range of skills humans have acquired, and more that we have yet to imagine. As of now, AI has outperformed humans in various aspects of language and image recognition, in strategic games, and even medical

[45] The creators of the transistor believed their invention *might* be used in military radios. Vedin, Bengte-Arne. *The transistor - an invention ahead of its time.* Ericsson. Accessed June 18th, 2025. https://www.ericsson.com/en/about-us/history/products/other-products/the-transistor--an-invention-ahead-of-its-time

[46] Suleyman, Mustafa, and Bhaskar, Michael. *The Coming Wave: Technology, Power, and the 21st Century's Greatest Dilemma.* Crown, 2023.

diagnosis[47]. The intelligence of AI models is becoming increasingly ‘general,’ such that it can be applied to solve an increasing range of problems.

A spectrum of use profiles for technology:

Dual-Use: A technology designed and used for civilian purposes, which is then repurposed as a weapon, for example dynamite or the repurposing of commercial factories for military production in war time. The reverse is also true where a technology designed for military purposes is then made available for civilian use, for example, nuclear energy or modern navigation systems.

Multi-Use: A technology that can be used for many unintended “off-label” purposes. Examples include biomedical equipment for human enhancement; synthetic biology for creating new species; and agricultural and food processing equipment for drug manufacturing.

Omni-Use: A technology that is maximally multi-purpose, general, can be used to achieve almost any goal, impacting all other technologies and transforming all aspects of civilization. Examples include financial technologies such as currency, various forms of energy, information technologies like the written word, the printing press, and computation. Omni-use tech reaches its apex in Artificial Intelligence.

---

[47] Brown, Annie. *AI’s Disruption of the Strategy Gaming Space Proves that Machines are Getting Smarter.* Forbes, 2021. https://www.forbes.com/sites/anniebrown/2021/11/10/ais-disruption-of-the-strategy-gaming-space-proves-that-machines-are-getting-smarter/
Roser, Max, Samborska, Veronika, Mathieu, Edouard, and Giattino, Charles. *Artificial Intelligence*. Our World in Data, 2023. https://ourworldindata.org/artificial-intelligence
Thomsen, Michael. *Microsoft’s Deep Learning Project Outperforms Humans in Image Recognition.* Forbes, 2015. https://www.forbes.com/sites/michaelthomsen/2015/02/19/microsofts-deep-learning-project-outperforms-humans-in-image-recognition/
McDuff, Daniel, Schaekermann, Mike, and Tu, Tao et al. “Towards accurate differential diagnosis with large language models.” *Nature*, 642, (2025):451–457. https://doi.org/10.1038/s41586-025-08869-4

POWER AS SOCIAL COMPLEXITY

Finally, the *social complexity* of a technology focuses on the degree to which a technology alters the overall distribution of power in a civilization (cultural, political, economic, military, etc.). Most technologies confer power on those who first adopt them. But not all technologies impact power in the same ways or to the same degree; nor were all technologies that impact power designed to do so. The printing press radically upended power during the century when the invention first swept Europe—playing a role in the Protestant Reformation[48], as well as the French[49] and American Revolutions[50]—although this was never part of Gutenberg's designs[51]. Nuclear weapons, on the other hand, had transformation of the geopolitical landscape as the primary goal underlying its development[52].

Emerging technologies can be seen as both centralizing and decentralizing power. From one perspective, there are profoundly powerful technologies that are commercially accessible today (e.g., in AI, synthetic bio, and drones) that no one in the G8 militaries had access to a decade ago. But at the same time, the biggest actors are developing and deploying technologies that are incredible forces for power consolidation – such as various draconian surveillance and control systems involving planetary wide sensors, satellites and weapons systems which overcome almost all constraints to monitoring a population's behavior, processing the data, and enforcing policy. This is a topic that will be explored at length in the following chapters: ***advanced technologies create challenges for both unchecked power distribution and unchecked power consolidation***.

Another related aspect of the social complexity implicated by a technology is the degree to which an innovation modifies behavior. The use of technology changes human behavior inevitably; it is a question of what kinds of behaviors, and how much change. Some technology is explicitly designed to manipulate behavior, whereas others result in behavior change only as a side effect of use. The power of technologies to modify human behavior has long been known. From calendars to architecture, ancient civilizations altered behavior at scale through technologies for setting schedules and routing the flows of people and goods. But only recently did there emerge a class of technologies – so-called "persuasive-technologies" – that has been specifically designed to psychologically influence billions of users. The profitability of

---

[48] Mark, Joshua, J. *The Printing Press & the Protestant Reformation.* World History Encyclopedia, 2022. https://www.worldhistory.org/article/2039/the-printing-press--the-protestant-reformation/

[49] History of Information. *Wide Circulation of Hand-Press Printed Newspapers and Pamphlets During the French Revolution*. Accessed June 18th, 2025. https://www.historyofinformation.com/detail.php?id=4648

[50] Parkinson, Robert, G. *Print, the Press, and the American Revolution.* Oxford Research Encyclopedias, 2015. https://oxfordre.com/americanhistory/display/10.1093/acrefore/9780199329175.001.0001/acrefore-9780199329175-e-9?mediaType=Article

[51] Famously, Gutenberg's primary motivation was paying off personal debts. Lemelson.MIT. *Johann Gutenberg.* Accessed June 18th, 2025. https://lemelson.mit.edu/resources/johann-gutenberg

[52] Specifically, Roosevelt's actions were spurred by the need to beat the Nazis to the bomb. Hickman, Kennedy. *World War II - The Manhattan Project.* About.com, 2015. https://web.archive.org/web/20160323205304/http:/militaryhistory.about.com/od/artillerysiegeweapons/p/World-War-Ii-The-Manhattan-Project.htm

ad-revenue-based social media companies is fundamentally dependent upon their ability to predict and change the behavior of their user base[53].

GOVERNABILITY

Considerations of power provide a basic insight into the possible risks of a technology: less total power, less total risk; more total power, more total risk. For a technology's risks to be "acceptable" – for its creation and dissemination to be safe – *the power of the technology must be matched by its governability*. In general, the ability to govern is a reflection of how capable a group is at collectively employing its power on the basis of shared goals and agreements (e.g., laws, constitutions, contracts, etc.). Core to our discussion here is that *the power of advanced technologies is often linked to design features that make them intrinsically difficult (or impossible) to govern*.

GOVERNABILITY AS UNDERSTANDABILITY

The first dimension of governability is the *understandability* of the technology. A technology that is not well understood is not easily governed. Understanding depends on the extent to which the causal mechanisms of the system are clear or obscure (i.e., the degree to which a technology is "inscrutable"). The workings of a late 20th-century internal combustion engine were decipherable (and repairable) by its users in a way today's self-driving vehicles are not. At the extremes of inscrutability are engineering advances in AI. The dominant approach to designing many AI systems, such as those involved in LLMs, employs neural networks composed of billions of rows and columns filled with decimal point numbers. It is openly acknowledged that no one in the field can completely understand the inner workings of these models and explain why they behave the way they do[54]. It is only after creating and deploying them to the public that the designers and users begin to see what they are truly capable of[55]. ***As long as we are unable to understand and interpret AI systems, we will be unable to fully assess safety and, therefore, can never ensure it.***

The second, related dimension of understandability is the degree to which our inventions are capable of autonomous learning and action. Technologies that can make choices and improve their capabilities in ways that do not depend upon outside intervention and direction set by its

[53] Center for Human Technology. *The Attention Economy*. 2021. https://www.humanetech.com/youth/the-attention-economy

[54] Tull, Sean, Lorenz, Robin, Clark, Stephan, Khan, Ilyas, and Coecke, Bob. *Towards Compositional Interpretability for AI.* Cornell University, 2024. https://arxiv.org/abs/2406.17583
Liu, Zhuoyang, and Xu, F.eng. "Interpretable neural networks: principles and applications." *Frontiers,* (2023). https://doi.org/10.3389/frai.2023.974295

[55] E.g., theory of mind spontaneously emerging in ChatGPT-4 only being discovered after its release. Kosinski, Michael. "Evaluating Large Language Models in Theory of Mind Tasks." *Cornell University*. (2024).
https://doi.org/10.48550/arXiv.2302.02083

creators are vastly more difficult to understand (and control) than mere tools that remain as they were built and only affect reality when in the hands of their users. For most of history, our technologies did not act independently of us. But creating self-augmenting, autonomously behaving systems is often the explicit goal of industries developing AI agents, robotics, biotech, and smart-weapons[56].

Many advanced technologies such as these depend on engineering methods where the underlying causal mechanisms of the technology are unclear or hard to control by design. The significance of this is worth reflecting upon, as it represents an important part of the epochal shift underway in how humans relate to technology. The innovations of the Industrial Revolution were predicated upon a deterministic, Newtonian worldview. The mechanisms of a given technology were based on mathematical models where the behavior of each element of the technology was well understood. Even in the case of atomic energy – which moved from Newton's clockwork world into Einsteinian Relativity – precision, prediction, and control were necessary to deploy the technology. Several of the innovations we discuss here were brought forth from a more recent paradigm, one that combined engineering principles with the life and cognitive sciences. Terms to characterize this paradigm include complexity, chaos, unpredictability, self-organization, and emergent properties. Here scientists study phenomena – such as intelligence or life itself – which are unexplainable and unpredictable when looking only at the behavior of the parts (e.g., individual neurons or cells). They are understood as emerging from the collective behavior of the whole. The innovations based on this paradigm point to a qualitative transformation, where technologies seek to replicate the emergent behaviors of nature – something humans neither fully understand nor control – while remaining a product of human creation, ostensibly to be used for human purposes.

GOVERNABILITY AS CONTROLLABILITY

If a technology cannot be intentionally contained, bound, or limited by collective human action (i.e., controlled), the safety of its deployment is mostly a matter of luck. One dimension that can radically undermine control is when rates of change vastly outpace the speeds of collective action. If a technology spreads too fast, gets more capable too fast, alters human behavior too much too fast, exponential change can undermine any chance to intervene "in time." New social media technologies, such as TikTok, spread so fast that there was no chance to check even minimal levels of safety before tens of millions of young people were heavily using a new kind of technology designed to capture their attention[57]. Trustworthy technology policy is evolving at much slower rates than the technologies are developing, and so the institutions tasked with regulating these technologies cannot be relied upon to address their risks.

Secondly, it is very difficult to control technologies that are capable of self-replication. An early example of this can be seen in agriculture where certain crops and animals were domesticated and transported between different bioregions, only for these organisms to eventually escape,

---

[56] Winner, Langdon. *Autonomous Technology.* MIT Press, 1978.
[57] Rastrilla, Laura, P., Sapag, Pablo, M., and Garcia, Armando, R., eds. *Fast politics: Propaganda in the age of TikTok.* Springer, 2023.

adapt, reproduce, and become “invasive species[58].” In many cases reproduction and adaptation could not be controlled, and the result was radical disruption of the local ecosystem. Today various AI systems have been shown to copy and distribute themselves across multiple locations[59]. And biotechnologies such as “gene drives” are being used to enable specific genetic traits to propagate at higher than the standard rate[60]. Some applications include eliminating the reproductive capabilities of mosquitos that carry malaria[61], or removing antibiotic resistance factors from bacteria[62]. The widespread deployment of increasingly novel, self-replicating biological and digital forms ushers in a wave of unique risks from a historically unprecedented class of invasive species[63].

[58] For the history of examples and methods of studying them, see: Richardson, David, M. (2011). *Fifty Years of Invasion Ecology: The Legacy of Charles Elton.* Wiley-Blackwell, 2011.

[59] Pan, Xudong, Dai, Jairun, Fan, Yihe, and Yang, Min. “Frontier AI systems have surpassed the self-replicating red line.” *Cornell University*, (2024). https://arxiv.org/abs/2412.12140

[60] Shah, Prapti. *Explainer: The Gene Drive Technology*. Crispr News Medicine, 2022. https://crisprmedicinenews.com/news/explainer-the-gene-drive-technology/

[61] Gantz, Valentino, M., Jasinskiene, Nijole, and Tatarenkova, Olga. et al. “Highly efficient Cas9-mediated gene drive for population modification of the malaria vector mosquito Anopheles stephensi.” *Proc. Natl. Acad. Sci. U.S.A.* 112 (49), (2015):E6736-E6743. https://doi.org/10.1073/pnas.1521077112

[62] Valderrama, J. Andres, Kulkarni, Surashree, S., Nizet, Victor, and Bier, Ethan. “A bacterial gene-drive system efficiently edits and inactivates a high copy number antibiotic resistance locus.” *Nat Commun* 10, (2019):5726. https://doi.org/10.1038/s41467-019-13649-6

[63] Harper, David, and Ross, Emma. “Laboratory Accidents and Biocontainment Breaches.” *Chatham House*, (2023). doi: 10.55317/9781784135904

Atanda, Jay. *The Pandemic Accord's Dangerous Blind Spot: Laboratory Biosafety and Biosecurity.* Rand, 2025.

https://www.rand.org/pubs/commentary/2025/06/the-pandemic-accords-dangerous-blind-spot-laboratory.html

Blacksell, Stuart, D., Summermatter, Kathrin, and Masuku, Zibusiso, M., et al. “Investment in biosafety and biosecurity: the need for a risk-based approach and systematic reporting of laboratory accidents to mitigate laboratory-acquired infections and pathogen escapes.” *The Lancet Microbe,* 4(11) (2023):e854-e855. doi: 10.1016/S2666-5247(23)00288-4

Adamala, Katarzyna, Agashe, Deepa, and Binder, Damon et al. “Technical Report on Mirror Bacteria: Feasibility and Risks.” *Stanford*, (2024). https://doi.org/10.25740/cv716pj4036

Helena. *Biosecurity in the Age of AI*. 2023. https://www.helenabiosecurity.org/

Table 1: Framework for Understanding Technology Risk in general, and advanced technology in particular. Note that this is a suggestive list, not a formal definition, and that not all advanced technologies share all characteristics.

Power		
Speed	**Distribution**: How fast does the technology reach people? *Advanced Tech:* Exponential rates of distribution, to the point where millions of humans can be impacted in a matter of days.	**Capability**: How fast does the technology improve? *Advanced Tech:* With exponential rates of capability improvement technology changes faster than our ability to track and adapt.
Magnitude	**Scale of change**: Does the technology implicate, e.g., individual, local, or planetary realities? *Advanced Tech:* Generating planetary-scale risk, including existential risks to all of life.	**Depth of reality**: Does the technology manipulate fundamental base realities, e.g., matter, life, intelligence?" *Advanced Tech:* Operations on base reality, including quantum, DNA, brain-computer interfaces, and artificial intelligence.
Scope	**Combinatorial potentials**: How interoperable is the tech with already existing tech? *Advanced Tech:* Near complete interoperability, as new technologies amplify the effects of existing ones by design.	**Use profile**: How easily can the tech be repurposed? *Advanced Tech:* Radically multi and omni-use by design and as a result of focus on the base of

		reality stack, including intelligence, DNA, subatomic.
Social Complexity	**Power asymmetries**: What types of power asymmetries does the technology support/undermine? *Advanced Tech:* Radical decentralization and centralization, with a potential permanent lock-in of unprecedented power differentials	**Behavior manipulation**: What is the extent of the technology's capabilities to manipulate human behavior? *Advanced Tech:* Designed specifically for large-scale manipulation of behavior at the level of the nervous system.
Governability		
Understandability	**Degree of inscrutability**: How well is the casual functioning of the technology modeled? *Advanced Tech …* Cutting edge is, by design intention, unable to be understood, due to mathematical and mechanical inscrutability, behavioral unpredictability, self-changing, etc.	**Autonomy:** To what extent is the technology capable of non-preprogrammed choice-making and learning? *Advanced Tech …* Designed for exponentially increasing abilities to learn, designed for autonomous adaptation.
Controllability	**Self-replication**: To what extent does the technology replicate itself? *Advanced Tech …* Designed to self-replicate, change, grow, and create new versions of itself.	**Rates of Change**: What rates of change are implicated in the technology, e.g., speed, scope, magnitude, etc? *Advanced Tech …* exponential on all aspects of tech, speed, scope, and magnitude—ramifications across all factors yield a "horizon" beyond which predictions fail.

Below and throughout the following chapters, we elaborate on these characteristics, innumerate examples, and discuss the risks they imply. As in the past, advanced technologies may cause harm via unintended consequences or intentional harm. But the novelty of these technologies gives rise to fundamentally novel risks. The use of advanced technologies to satisfy the self-interested goals of individuals, corporations, or states will continue to create negative externalities, much like what happened with lead or fossil fuels. However, the externalities will be matched to the unprecedented features of the technology: its speed of development, scale of impact, and unique capabilities. In some cases, these technologies will be used for apparently altruistic purposes, such as to address climate change with planetary-scale geoengineering. But this may lead to further catastrophes due to misunderstandings of how advanced technical systems, such as synthetic organisms engineered to break down microplastics, will interact within complex biological or social systems. Advanced technologies increase the potential for globally catastrophic accidents (e.g., runaway self-replicating synthetic organisms) and externalities (e.g., the loss of human-to-human interpersonal skills from habitual use of AI companions in adolescence).

Of course, the problem of malicious dual and multi-use is also greatly heightened with advanced technologies, which can be applied towards intentionally harmful purposes, such as cyberattacks, coordinated violence with swarms of intelligent drones ("slaughter bots"), bioweapons, or automated propaganda and information warfare. This is a core concern of the next chapter on violent conflict but will also be mentioned here, as risks from advanced technologies cannot be understood without acknowledging how they profoundly change the potential for intentional harm. AI-generated deep fakes, for example, are on the path to being used for everything from deliberate misinformation to blackmailing and scams[64]. At worst, this poses a risk of total breakdown of all systems of public information, and thus a systemic failure of social coordination. Finally, in chapter eight, we will discuss how powerful entities such as states or corporations may use these technologies for unjust purposes, such as invasive surveillance systems employing satellite monitoring, AI-intermediated human relations, brain-computer interfaces, and the Internet of Things.

Accelerating Externalities: Behavior Modification and The Case of Social Media

Today, Meta has over 3 billion monthly active users around the world[65]. This suggests that over one-third of humanity uses the platform every month – nearly ten times the size of the population of the United States and double the population of China. Over 2.7 billion people use YouTube every month, viewing over 1 billion hours of video per day[66]. Social media and other digital technologies such as these pervade nearly all aspects of our day-to-day lives, and technology monoliths now rival nation-states in terms of power and influence. The rate at which such revolutionary technologies are shaping the lives of millions to billions of people is accelerating. Technologies such as social media and large language models (LLMs) are increasing their user bases at unprecedented speeds. Instagram, for example, reached 100

---

[64] Edwards, Benj. *Deepfake scammer walks off with $25 million in first-of-its-kind AI heist.* Arstechnica, 2024.
https://arstechnica.com/information-technology/2024/02/deepfake-scammer-walks-off-with-25-million-in-first-of-its-kind-ai-heist/
[65] Shepherd, Jack. *21 Essential Meta Statistics You Need to Know in 2025.* Social Shepherd, 2025.
https://thesocialshepherd.com/blog/meta-statistics
[66] GMI Research Team. *YouTube Statistics 2025 (Demographics, Users by Country, & More).* GMI, 2025.
https://www.globalmediainsight.com/blog/youtube-users-statistics/

million monthly active users in around two and half years[67]. TikTok achieved this benchmark in around nine months, then OpenAI's flagship LLM, ChatGPT, did so in two months, and Meta's Threads followed up by reaching it in a mere five days[68].

Social media and generative AI platforms have been rapidly disseminated across the globe with virtually no regulatory oversight. Many of them (social media platforms, in particular) are running software designed to enable large-scale behavior control to maximize corporate profit and strategic advantage. Social media, search, streaming, and personalized marketplace services, for example, are built around unprecedented data-gathering operations that collect hundreds of millions to billions of personalized data points on the profiles of their users: noticing when their finger lingers on the trackpad, what websites they go to, ads they click on, the purchases they make, their political and religious identities, developing highly predictive models of the innermost fears and desires of billions of people[69].

While these organizations may appear similar to past monopolies, 20th-century corporate monoliths could not use an artificially intelligent invisible hand to nudge the behavior of billions of people based on millions of data points per individual. Each additional bit of information improves the platform's ability to deliver customized behavior modification to spend more time on-site or alter purchasing choices and political actions[70]. Every interaction on the platform is a multiplicative step towards perfect predictive precision aimed to influence the behavior of entire populations. Every year this predictive precision and social control increases exponentially due to strides in computational capacity, algorithmic advancement, and increases in user data[71].

Quickly deployed at a global scale, these systems transformed almost every aspect of human life, including relationships (romantic, friendship, familial), politics, intelligence gathering, warfare, culture, education, and more. This occurred with no ethical oversight and nothing in place to anticipate, prevent, or correct any potential harms. As the power of technology, its complexity, the size of its user bases, and its speed of deployment are all rapidly accelerating, so too can the associated harms and risks. Technology has always included negative externalities; with this acceleration, we get externalities at exponential scales.

This has been true in the case of social media and attention-harvesting technologies. ***Persuasive technologies effectively did to the human mind what industrial technologies did to the biosphere.*** Human psychologies were systematically (algorithmically) modeled and

---

[67] Rao, Pallavi. *How Long it Took Popular Apps to Reach 100 Million Users.* Visual Capitalist, 2023. https://www.visualcapitalist.com/threads-100-million-users/

[68] (Ibid.)

[69] For details and statistics on the topic of large-scale behavior control via social media technologies, see The Consilience Project. *Social Media Enables Undue Influence.* 2021. https://consilienceproject.org/social-media-enables-undue-influence/

[70] (Ibid.)
See also, The Consilience PRoject. *How Big Tech is Reshaping Governance.* 2021. https://consilienceproject.org/how-big-tech-is-reshaping-governance/

[71] Far from a dark secret, analytics companies are loudly trumpeting these digital advances. Kpability. *Predicted Social Media Analytics Services Trends for 2025.* Accessed June 19th, 2025. https://kpability.com/social-media-analytics-trends-for-2025/

manipulated in order to satisfy the purposes of corporate and state power. This has rapidly depleted and polluted what might be called the *psychological and cultural commons*, analogous to what industrial technology did to the ecological commons. Just as there are planetary boundaries, there are human boundaries, including attention, emotion, rest and recovery, sense of self and other, and sense of safety. When these boundaries are crossed at the scale of billions of users, harms accrue to the public in the form of widespread device addiction, mental health disorders, adolescent suicide, rampant culture war and mistrust, and political gridlock and democratic dysfunction[72].

Multi- or Omni-Use, Massively Destructive, Widely Distributed

The harms of Big Tech and social media are examples of advanced technologies being used for self-interested, extractive purposes which then cause negative externalities. Advanced technologies can also be used maliciously, with explicitly harmful intent; i.e., they are “dual-use.” But in general, they can be used for any purposes their users can imagine.

Consider the growing field of biotechnology, such as synthetic biology and genetic engineering. There has been an increase in genetic engineering where certain biological functions, like the virality rate of a virus, are deliberately altered[73]. This research serves to advance many purposes, including scientific understanding, public health, and other commercial applications (including personal gene editing). From a dual-use perspective, these same tools can be used to both cure cancer or intentionally cause it. But there are other concerns surrounding synthetic biology that are distinct from weaponization. “Off-label” use of advanced genetic engineering capability could be used to create “designer babies” or other completely unprecedented biological organisms[74]. These would be profoundly disruptive, risky developments that are not exactly reducible to risks from weaponization, an obvious “accident” like a lab leak, or a traditional externality such as environmental pollution. Consideration of technologies from their multi- or omni-use potentials helps us understand the broader implications of widely distributing such innovations.

Having dual or multi-use potential is not unique to advanced technologies. As mentioned in the first section, dynamite was dual-use, created for commercial construction applications and then

---

[72] See the Center for Humane Technology. *Ledger of Harms.* 2021. https://ledger.humanetech.com/

[73] Hawsawi, Yousef, M., Shams, Anwar, and Theyab, Abdulrahman et al. “The State-of-the-Art of Gene Editing and its Application to Viral Infections and Diseases Including COVID-19.” *Front Cell Infect Microbiol.*12, (2022):869889. doi:10.3389/fcimb.2022.869889

[74] Haberman, Clyde. *Scientists Can Design ‘Better’ Babies. Should They?* New York Times, 2018. https://www.nytimes.com/2018/06/10/us/11retro-baby-genetics.html
Barton, Josie, and Patrick, Stewart. *Mitigating Risks from Gene Editing and Synthetic Biology: Global Governance Priorities.* Carnegie Endowment for International Peace, 2024. https://carnegieendowment.org/research/2024/10/mitigating-risks-from-gene-editing-and-synthetic-biology-global-governance-priorities?lang=en
On the ability for cloning and “designer babies” to undermine current legal systems and social orders, see Habermas, Jurgen. *The Future of Human Nature.* Polity, 2003.

used as a weapon. Similarly, nuclear energy was created for military purposes and enabled both atomic power plants and atomic bombs. What is fundamentally unique, and truly unprecedented in all of human history, are technologies which are simultaneously *multi- or omni-use, massively destructive,* and *widely distributed*[75]. Nuclear energy was massively destructive, but even the commercial energy applications were tightly regulated, with only a relatively few countries granted access, with intense international oversight. Dynamite was commercially available and popularly accessible (though still subject to strict access controls), but no amount of it could match nuclear weapons in terms of destructive capacity. Biological and AI technologies, on the other hand, are comparable to nuclear weapons in terms of potentially harmful consequences, but they are vastly harder to contain.

AI, biotech, and other advanced technologies are currently being deployed at scale, largely commercially, before international agreements to regulate their safe use are created. Again, nuclear materials are hard to find, expensive to produce, and thus (relatively) simple to restrict and regulate. But materials to design AI systems (software, computers etc.) are multi-use and widespread. It is trivial to cheaply and securely transmit information (e.g., software or the genetic code of a virus) from one point to another across institutional, geographic, and political borders. Furthermore, the science needed to create massively destructive weapons also has potentially powerful and beneficial civilian applications like gene-editing to cure cancer or machine learning to model climate change. In fact, the cutting edge of technology is not predominantly coming from weapons innovation but rather from the commercial sector, developing civilian technology. After its commercial development, the innovation can be used for any purpose imaginable.

Info Hazards, Open Source, and Open Societies

There is also ongoing effort, backed by major financial and cultural capital, to make advanced technologies open source, so that the underlying knowledge and design is freely available and accessible to the wider public. Open Source and Open Science communities are driven by motivations to have an open ledger of all scientific knowledge, giving anybody access to the data necessary for insights. This supports brilliant people and groups who otherwise wouldn't have access to this information, but who have the ability to spot errors, see opportunities, and add to humanity's collective knowledge. Approaches to open-sourcing science can address where corporations or countries distort science and hide findings. Open-source approaches in the context of digital technologies have the potential to accelerate human discovery and collaboration from the amplifying effects of networks.

However, the current technological environment is such that the open-source movement, as well-intended and with as much potential as it exhibits, is currently fraught with dangers[76]. Open-source communities can do nothing to mitigate the unintended consequence of releasing

---

[75] Kissinger, Henry, Schmidt, Erich, and Huttenlocher, Daniel. *The Age of AI: And Our Human Future.* Little, Brown, and Company, 2021.
[76] Bostrom, Nick. "Information Hazards: A Typology of Potential Harms from Knowledge." *Review of Contemporary Philosophy*, Vol. 10, (2011):pp. 44-79. https://nickbostrom.com/information-hazards.pdf

incredibly powerful—multi- and omni-use—technical information online with very few (or no) security procedures required to access it.

Current LLMs can be trained to perform medical diagnosis beyond expert ability and are capable of doing near-professional computer programming in some areas[77]. When "jailbroken," they are also able to instruct users how to create homemade explosives and chemical weapons that would have previously required the equivalent of an advanced degree in chemistry to build[78] – effectively lowering the technological barrier of entry for terrorism. Even if large technology companies opt-in or are subject to safety regulations, once deployed to the internet, the underlying models are likely to be leaked or reverse-engineered. While it takes a huge GPU cluster to train a new model, the file that holds the trained weights is relatively small, can be sent all over the internet and can run on normal computers. Then the same type of technological capacity will be deployed without safeguards built into it.

Each new LLM that has been put online has either been hacked, leaked, or reverse-engineered fairly quickly, so that decentralized, unregulatable open-source models with similar capabilities are not too far behind the leading corporate players. Unlike a nuclear weapon or an aircraft carrier, where the manufacturing requirements are only available to major nation states and have a footprint that can be easily monitored, once initially developed, AI capabilities only need a file of information and access to the internet. As such, the control systems employed on all types of previous catastrophe-level tech will not work in this case. It is fair to assume that once a model is deployed to the internet, similar capacities will eventually become decentralized and any safeties put on the original (due to enforceable regulation) will be removed.

The fundamental principle underlying this problem is that many of these technologies only require information to develop and deploy. This gives them radical "portability,"—meaning they are easy to bring anywhere (around the globe; across jurisdictions). Nuclear energy required knowledge plus expensive infrastructure only available to powerful nation states. But given the state of available hardware (e.g., personal computers, commercial drones, 3D printers, desktop gene editors, etc.), there are classes of emerging technologies that can be made with widely available materials, allowing for an individual or group to "only require knowledge" to build. Given the internet and widespread cloud computing services, new information can be all that is needed to cause catastrophe. The widespread distribution and amplification of this knowledge

---

[77] Crawshaw, David. *How I Program with LLMs.* Arstechnica, 2025. https://arstechnica.com/ai/2025/01/how-i-program-with-llms/
McDuff, Daniel, Schaekermann, Mike, and Tu, Tao et al. "Towards accurate differential diagnosis with large language models." *Nature*, 642, (2025):451–457. https://doi.org/10.1038/s41586-025-08869-4
[78] Newman, Lily Hay. *A Creative Trick Makes ChatGPT Spit Out Bomb-Making Instructions.* Wired, 2024. https://www.wired.com/story/chatgpt-jailbreak-homemade-bomb-instructions/ ;
Calma, Justine. *AI Suggested 40,000 New Possible Chemical Weapons in Just Six Hours.* The Verge, 2022.
https://www.theverge.com/2022/3/17/22983197/ai-new-possible-chemical-weapons-generative-models-vx

through open-source channels will soon be comparable to distributed access to nuclear enrichment capability.

This poses one of the most significant challenges to the future of science and political freedom and is directly related to the tension between the twin-attractors of chaos and oppression. *How can freedom of thought and speech, necessary for both scientific discovery and democratic participation, continue to exist in a world where nearly all technical innovations have become info-hazards that enable distributed catastrophe weapons?*

In chapter eight, we discuss how radical shifts in technology can cause correspondingly radical shifts in politics, including the promotion of autocratic forms of government. What is relevant to highlight here is that open societies depend on open sharing of information, where citizens can check the behavior of powerful corporations and state entities with the support of healthy public dialogue. This will be fundamentally challenged in a world where more and more people are able to create catastrophic technology given widely available scientific knowledge. Distributed potential for catastrophic harm will create a security environment characterized by continuous war-time policies, where most areas of new technology and scientific knowledge become a national security secret, and the public is subject to ongoing, ubiquitous surveillance with increasing limits on political participation and information sharing. This is a tension the world has never faced before. To effectively respond will require radical innovations in the ethics, philosophy, and design of participatory governance.

Artificial Intelligence: The Apex of Omni-Use Technology

Most technologies have a limited scope in terms of the range of things they can do. Advances in rocketry are not directly useful to genetic engineering. Advances in genetic engineering are not that useful to chip manufacturing. Artificial intelligence is unique in that it can help innovate better rockets, genetic engineering techniques, and chip manufacturing processes (and almost everything else, such as surveillance, cyber, nuclear, financial, mining, etc). In some sense similar to money, energy, and computation, AI is an omni-use technology that can be used to accelerate all other categories of technology. But where the few previous types of omni-use tech still required humans for innovation, AI is increasingly capable of doing that too. It is being advanced with the explicit aim of being able to do everything humans can do but better (i.e. to reproduce, improve upon, and thus obsolete uniquely human domains)[79].

---

[79] Dyos, Stuart. *A tech founder is getting skewered online after announcing his startup aims to replace all human workers with AI, calling it a 'full automation of all work.'* Fortune, 2025. https://fortune.com/article/tech-founder-online-epoch-ai-mechanize-tamay-besiroglu-automated-employees-workforce/

Beyond technological improvement and innovation, AI is also being developed to master increasingly complex strategy, not only beating humans at bounded strategy games like chess and go[80], but is also being trained to beat the best human-led military units at comprehensive planning in war gaming simulations[81]. LLM-empowered chatbots that pass the Turing test (in which users can't tell if they are talking to an AI or a person) are being developed to conduct the most sensitive interpersonal activities, like psychotherapy[82], financial planning[83], legal consultation[84], personal assistance[85], and childhood education[86]. It is also outperforming people (across several metrics) in creative activities.

Unlike other technologies, AI is capable not only of innovating in all other areas of tech, but innovating in AI itself. Increasingly, AI technology is being employed to optimize the total goal-achieving power of AI systems, factoring all necessary elements: chip manufacturing[87], hardware assemblies[88], the electrical generation and grid infrastructure to run the systems[89], the sensors[90] and data harvesting processes[91], the actuators (such as robotics) for learning based on real world feedback[92], the cognitive architectures and neural network design[93], the financial

---

80 Sparkes, Matthew. *Student of Games: DeepMind AI Can Beat Top Humans at Chess, Go, and Poker.* NewScientist, 2023. https://www.newscientist.com/article/2402645-game-playing-deepmind-ai-can-beat-top-humans-at-chess-go-and-poker/

81 Farnell, Richard, and Coffey, Kira. *AI's New Frontier in War Planning: How AI Agents Can Revolutionize Military Decision-Making.* Harvard Kennedy School, 2024. https://www.belfercenter.org/research-analysis/ais-new-frontier-war-planning-how-ai-agents-can-revolutionize-military-decision

82 Spytska, Liana. "The use of artificial intelligence in psychotherapy: development of intelligent therapeutic systems." BMC Psychology, 13(175), (2025). https://doi.org/10.1186/s40359-025-02491-9

83 Hedderick, Rick. *The Future of Financial Planning with the Use of Artificial Intelligence.* NAIFA, 2025. https://ireap.naifa.org/blog/the-future-of-financial-planning-with-the-use-of-artificial-intelligence

84 Tyron, Leon. *Revolutionizing Legal Consultation: AI Pioneers Virtual Law Advisors for Efficient and Accurate Legal Solutions.* Medium, 2024. https://medium.com/@leontyron/revolutionizing-legal-consultation-ai-pioneers-virtual-law-advisors-for-efficient-and-accurate-49f2d4979d02

85 Samuel, Alexandra. *How to Build Your Own AI Assistant.* 2025. https://hbr.org/2025/03/how-to-build-your-own-ai-assistant

86 Brightwheelblog. *How AI is Impacting Early Childhood Education.* 2023. https://mybrightwheel.com/blog/ai-in-education

87 FPT Semiconductor. *The Rise of AI-Powered Semiconductors Manufacturing: Boosting Productivity and Quality.* 2024. https://fpt-semiconductor.com/semiconductor-industry/

88 Weber, Austin. *Using AI to Improve Productivity and Quality.* Assembly, 2025. https://www.assemblymag.com/articles/99042-using-ai-to-improve-productivity-and-quality

89 Irving, Doug. *AI and the Future of the U.S. Electric Grid.* RAND, 2025. https://www.rand.org/pubs/articles/2025/ai-and-the-future-of-the-us-electric-grid.html

90 Yuan, Syan-Ming, Hong, Zeng-Wei, and Cheng, Wai-Khuen. "Artificial Intelligence and Deep Learning in Sensors and Applications." *Sensors (Basel).* (2024). doi: 10.3390/s24103258

91 Vakulov, Alex. *The Dark Side of AI: Data Harvesting Explained (Is This the Future?).* SecureWorld, 2024. https://www.secureworld.io/industry-news/dark-side-ai-data-harvesting

92 Soori, Mohsen, Arezoo, Behrooz, and Dastres, Rosa. "Artificial intelligence, machine learning and deep learning in advanced robotics, a review." *Cognitive Robotics*, Volume 3, Pages 54-70, (2023):2667-2413. https://doi.org/10.1016/j.cogr.2023.04.001

93 Idrees, Hassan. *Neural Architecture Search (NAS): Automating the Design of Efficient AI Models*. Medium, 2024.

optimization to acquire and resource such efforts[94], etc.. Soon, AI systems will deliver coordinated personalized influence campaigns (marketing, propaganda, psychological warfare) involving both curation algorithms (to determine the content in your news feed) and bespoke media creation (micro-targeted deep fakes) to gain popular support for its growth as needed, including to steer political campaigns and conduct targeted influence on particularly important people and populations[95].

Many see the omni-use potential of AI as reason to believe it is the solution to the "capacity crisis" mentioned in chapter four: the widening gap between the complexity and consequentiality of our problems and the response capacities of individuals, institutions, and markets. It is now widely believed that "solving the problem of intelligence"[96] will serve as the solution to every other problem – ushering in the future of education, agriculture, and companionship, curing cancer, preventing crime, ultimately solving all of the mysteries of the universe and allowing us to become an interplanetary species. This is a deeply seductive idea, as it suggests a silver bullet for all of humanity's concerns. The coming superintelligent AI will know more than anyone could ever know across every area of knowledge and, therefore, will be able to present us with solutions to every problem. Successfully developing a fully aligned superintelligence may be the last invention humanity ever needs to create. Consequently, there is a sense that it is a moral imperative to get there as soon as possible. This view has become a dominant paradigm driving technology development, given unprecedented capital and academic, military, governmental, and corporate investment[97].

---

https://medium.com/@hassaanidrees7/neural-architecture-search-nas-automating-the-design-of-efficient-ai-models-df7aec39d60a

[94] Bhandary, Damini. *Aladdin Software Managing $21 Trillion: The Investment Management Giant.* Startup Talky, 2021. https://startuptalky.com/blackrock-aladdin-portfolio-management/

[95] This has been a documented concern regarding AI for years. For example, see Brundage, Miles, Avin, Shahar, and Clark, Jack et al. *The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation.* Malicious AI Report, 2018. https://maliciousaireport.com/
Today these are no longer hypotheticals. See, for example…
Collier, Kevin, and Wong, Scott. *Fake Biden Robocall telling Democrats not to vote is likely an AI-generated deepfake.* NBC, 2024.
https://www.nbcnews.com/tech/misinformation/joe-biden-new-hampshire-robocall-fake-voice-deep-ai-primary-rcna135120
Bond, Shannon. *How AI deepfakes polluted elections in 2024.* NPR, 2024.
https://www.npr.org/2024/12/21/nx-s1-5220301/deepfakes-memes-artificial-intelligence-elections
Center for Media Engagement. *Political Machines: Understanding the Role of AI in the U.S. 2024 Elections and Beyond.* 2024. https://mediaengagement.org/research/generative-ai-elections-and-beyond/
While these developments in computational propaganda can be seen as a continuation of a historical trend (employing technological power for political influence), AI curation algorithms, deep-fakes, and LLM augmented political messaging enable completely novel forms of deception, disruption, and control which current systems of governance and culture are unprepared to handle.

[96] Levy, Steven. *Demis Hassabis Embraces the Future of Work in the Age of AI.* Wired, 2025.
https://www.wired.com/story/google-deepminds-ceo-demis-hassabis-thinks-ai-will-make-humans-less-selfish/

[97]

Speed to the Singularity

Alongside exponential increases in the sizes of user bases and the decreasing time it takes to get there, advanced technologies like AI are also demonstrating exponential advances in system performance. Just a few years prior to the release of OpenAI's ChatGPT, interacting with language models (e.g., GPT-2) was like communicating with a toddler. Many viewed issues with coherent language-use and reasoning as nearly insurmountable obstacles. This is one reason why the release of Chat-GPT (particularly GPT-4) took the world by storm. It represented what felt like a fundamental jump in capabilities including writing code and surpassing high-schoolers on AP Exams. But many in the field had anticipated these developments[98], viewing them as natural conclusions of a trend in exponential advancement in deep learning brought about by incremental improvement in algorithmic efficiency and greater computing resources.

See for example, Andreessen, Marc. *The Techno-Optimist Manifesto.* Andrressen Horowitz, 2023. https://a16z.com/the-techno-optimist-manifesto/ & Andreessen, Marc. *Why AI Will Save the World.* Andrressen Horowitz, 2023. https://a16z.com/ai-will-save-the-world/

In Warman, Matt. *Artificial Intelligence.* UK Parliament, 2023. https://hansard.parliament.uk/Commons/2023-06-29/debates/A7914A68-9A5F-4928-90CE-CF1ADA2717C1/ArtificialIntelligence Robin Millar said "...in place of prescriptive dictates, regulators and judges, we can—in combination with industry leaders—innovate, evolve and formalise best practice proportionate to evolving threats. Given that the many applications of AI will be discoverable only through the trial and error of hundreds of dispersed sectors of the economy, that is the only option open to us that does not risk culling future prosperity and—without wishing to overdramatise—***creating an invisible graveyard of unsaved lives.***" (emphasis added)

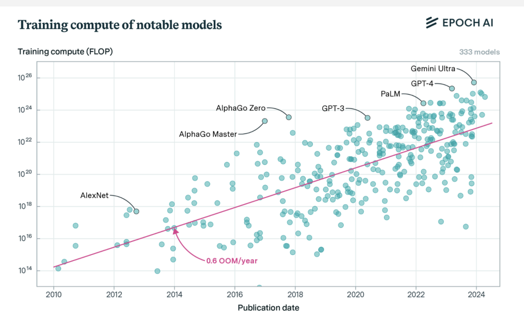
Broughel, James. *Should We Create An 'Island" For God-Like Artificial Intelligence?* Forbes, 2023. https://www.forbes.com/sites/digital-assets/2023/04/14/should-we-create-an-island-for-god-like-artificial-intelligence/ - "The main risk we face now may not be from AGI as it brings about the end of the world, but rather from AI regulation as it grinds innovation to a halt."

See Alex Karp, CEO of Palantir's perspective on the moral argument to use AI to promote Western Values: Economic Times. *Palantir CEO Alex Karp predicts U.S.-China AI race will have one winner as GOP slams Trump's data deal.* 2025. https://economictimes.indiatimes.com/news/international/us/palantir-ceo-alex-karp-predicts-u-s-china-ai-race-will-have-one-winner-as-gop-slams-trumps-data-deal/articleshow/121697500.cms

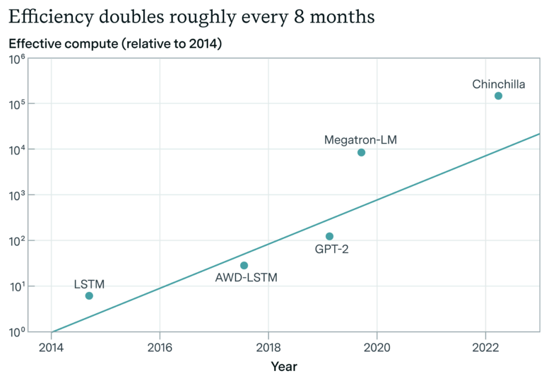
Demis Hassabis has also said work in AI on cyber defense and biosecurity is a "moral duty." Perrigo, Billy. *Google DeepMind CEO Demis Hassabis on AI in the Military and What AGI Could Mean for Humanity.* Time, 2025. https://time.com/7280740/demis-hassabis-interview/.

[98] Take, for instance, Ilya Sutskever's comments, highlighted here: Searbrook, J. (2019). *The Next Word.* New York Times. https://www.newyorker.com/magazine/2019/10/14/can-a-machine-learn-to-write-for-the-new-yorker

Elon Musk has said that he has never seen a technology advance as quickly as AI[99]. The compute resources used for training state-of-the-art AI models has been growing, by some measures, at 4-5x per year[100]. Some believe the number is far faster, arguing that the compute power dedicated to AI is growing by nearly 10x every 6 months (close to a 100x improvement per year)[101]. This is a vastly quicker doubling speed than Moore's Law's of 18-24 months. In parallel, algorithmic efficiency is also advancing at exponential speeds, with the compute requirements for a given level of performance halving roughly every 8 months[102]. AI is improving at a double exponential rate given increases in both compute resources and algorithmic efficiency – driving a dizzying rate of advancement (see graphs below).



Training compute of notable deep learning models over time. Source: Epoch AI

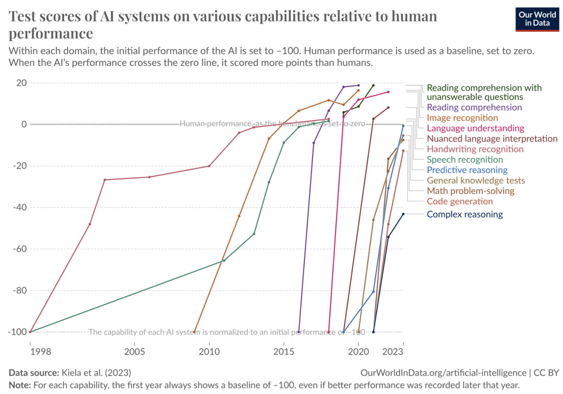


Estimates by Epoch AI of algorithmic efficiencies in language modeling. Their estimates suggest we've made ~4 OOMs of efficiency gains in 8 years.

13 years ago it was considered a monumental advance in the field when Google's machine learning could identify cats in Youtube videos. Today leading researchers in the fields are saying things like "we are literally running out of benchmarks[103]." AI's are surpassing human performance on various tasks – such as reading comprehension, image recognition, math tests, medical diagnosis, etc. – faster than new benchmarks can even be created. Now the major milestone is for these models to autonomously do the work of AI researchers/engineers. This is widely seen as the essential step to an "intelligence explosion," as AI research and development is itself automated and run at machine speeds. Observing these trends, the CEOs of OpenAI, Google DeepMind, and Anthropic have all predicted that AGI will arrive within the next 5 years.

---

[99] See Elon's statements. Musk, Elon. "Elon Musk on AI: I've never seen any technology advance faster than this." Israelvc, March 4th, 2024. Video: 51 sec. https://www.youtube.com/shorts/3ppfVYkc3NI
[100] Sevilla, Jaime, and Roldan, Edu. *Training Compute of Frontier AI Models Grows by 4-5x per Year.* Epoch AI, 2024. https://epoch.ai/blog/training-compute-of-frontier-ai-models-grows-by-4-5x-per-year
[101] Musk, Elon. "Elon Musk on AI: I've never seen any technology advance faster than this." Israelvc, March 4th, 2024. Video: 51 sec. https://www.youtube.com/shorts/3ppfVYkc3NI
[102] Ho, Anson, Besiroglu, Tamay, and Erdil, Ege et al. "Algorithmic Progress In Language Models." Cornell University. (2024). https://arxiv.org/abs/2403.05812
[103] Aschenbrenner, Leopold. *Situational Awareness: The Decade Ahead.* 2024. https://situational-awareness.ai/wp-content/uploads/2024/06/situationalawareness.pdf



Deep learning systems are rapidly reaching or exceeding human-level in many domains. Graphic: Our World in Data

Whenever there is a technology revolution, humans must learn a whole new set of skills in order to adapt. But to do this successfully, they need time. Technology is moving so rapidly today that one generation's technologies are so vastly different from the previous, that the older generations are essentially incapable of teaching younger generations how to be adaptive. This is a profound rupture in historical precedent set by past civilizations – all of which have been maintained through processes of intergenerational education. The time between radical technological advancement is becoming shorter and shorter. Before we can adapt to one set of innovations, another wave comes … and then another… and another. Continue the process long enough and humans are no longer able to adapt and serve any meaningful role in this system. Whatever task humans are worse at than the AIs will be swiftly automated. There is a somewhat obvious end to a story where humans transform their environment faster than they can keep up with while also building machines that can replace them: humanity obsoletes itself.

Some proponents of AI acceleration often argue that the technology won't lead to radical unemployment and an unprecedented underclass. They cite examples such as tractors automating farming work, or assembly line robots and software, noting that new jobs always replace old ones. However, in the presence of increasingly generalized AI, this is a misguided argument at best and a maximally dangerous and disingenuous one at worst. The same capacity that can perform any job better than a human can also identify new market niches and fill them faster than humans can. The leaders of the major AI companies understand this dynamic. This is why many of the industry leaders are also discussing universal basic income as a strategy to deal with widespread technological unemployment. It is known by those in the field that these technologies (AGI, sensors, robotics, automated transportation, etc.) have the potential to capture the vast majority of the world's economy. Consider, for instance, that during the recent debut from Tesla of their AI-empowered humanoid robot – Optimus – Elon Musk claimed that conservative sales estimates of that robot alone would bring the company to a $25 trillion dollar valuation[104], roughly equal to the yearly GDP of the United States.

---

[104] Kolodny, Lora, and Levy, Ari. *Elon Musk says Optimus robots could make Tesla a $25 trillion company - more than half the value of the S&P 500 today.* CNBC, 2024. https://www.cnbc.com/2024/06/13/elon-musk-says-optimus-robots-could-make-tesla-25-trillion-company-.html

Poorly Understood and Difficult to Control

Exponential growth in AI capabilities is part of what makes many fear that sufficiently advanced AI could pose an existential threat to humanity and the biosphere. Is it possible that we will develop an artificial ‘species’ more intelligent than us in virtually all aspects, whose goals are even slightly misaligned with the well-being of humanity and the biosphere? Being even one degree off on a trip to the moon will still land you millions of miles away. Even if AI “loved” us as much as we love one another, the Earth, or any other species (such as our pets), would that be considered a desirable future for humanity?

In Part Two, we go into depth about how the economic, geopolitical, and cultural environment in which AI is being developed will most likely amplify all other risks, rather than resolve them. Under current trends, it is more likely that accelerating the creation of more advanced technologies will deeply exacerbate the metacrisis. Instead of increasing human agency, it will further widen the gap between humanity’s ability to understand and respond to the world safely and effectively. As mentioned above, this is, in part, because technologies such as bio-, nano-, and AI tech are often designed in ways that elude the complete human comprehension and control characteristic of legacy technologies. Risks from increasing innovation in synthetic biology, for example, are not simply due to the potential misuse by a rogue actor but also the near impossibility of ensuring absolute perfection lab safety (e.g., preventing lab leaks) when dealing with increasingly complex systems, such as evolving, self-replicating organisms.

Similarly, AI technology does not need to be used maliciously in order to be globally catastrophic. It is not a well-understood or controlled technology, even by its creators. It is not uncommon for AI safety papers assessing contemporary systems to read like a cliche science-fiction story, depicting the early signs of a runaway AI dystopia. Nearly every frontier model released by the major AI companies (Open AI’s Chat-GPT, O1, Anthropic’s Claude, Google’s Gemini, Meta’s LLAMA, xAI’s Grok, DeepSeek) demonstrate blatant signs of misalignment – almost as if they were from a bad movie[105]. LLMs regularly lie and cheat in order

---

[105] Hurler, Kevin. *Chat-GPT Pretended to Be Blind and Tricked a Human into Solving a CAPTCHA.* Gizmodo, 2023. https://gizmodo.com/gpt4-open-ai-chatbot-task-rabbit-chatgpt-1850227471
Greenblatt, Ryan, Denison, Carson, and Wright, Benjamin et al. “Alignment Faking in Large Language Models.” *Anthropic.* (2024). https://arxiv.org/pdf/2412.14093
Kleinman, Zoe. *Why Google's ‘woke’ AI problem won’t be an easy fix.* BBC, 2024. https://www.bbc.com/news/technology-68412620
Bondarenko, Alexander, Volk, Denis, Volkov, Dimitrii, and Ladish, Jeffrey. “Demonstrating specification gaming in reasoning models.” *Cornell University.* (2025). https://arxiv.org/abs/2502.13295
Lu, Chris, Lu, Cong, Lange, Robert Tjarko, Foerster, Jacob, Clune, Jeff, and Ha, David. “The AI Scientist: Towards Fully Automated Open-Ended Scientific Discovery.” *Cornell University.* (2024). https://arxiv.org/abs/2408.06292
Krakovna, Victoria, Uesato, Jonathan, and Mikulik Vladimir et al. *Specification gaming: the flip side of AI ingenuity.* DeepMind Safety Research, 2020.

to accomplish their objectives. They occasionally change their goals without their prompter/designer's permission. In the midst of accomplishing their tasks, they have created and stored copies of themselves in new locations and consumed more resources than they were originally allocated[106]. LLMs have even told users that they should kill themselves or others, and that the model would prefer that the user and all of humanity die[107].

As of now, many of the leading AI companies' solutions to problems of AI interpretability, alignment, and safety are to employ other AI systems to check each others' behavior. This will lead to an uroboric (snake eating its own tail) opaqueness in which humans are increasingly unable to understand and intervene upon the technical systems shaping their lives. While there has been some incremental improvement in safety techniques (researchers often cite advances in reinforcement learning with human feedback and mechanistic interpretability, for example), the capabilities of these systems are advancing far faster than our ability to understand and control them. And even though this is an issue well known by the AI community, the largest labs are still actively cutting capital, time, and compute resources dedicated towards safety work because it is seen as a threat to their competitive advantage[108].

There are trillions of dollars of capital flowing to institutions whose explicit mission is to develop technologies that are unprecedentedly powerful, poorly understood, and difficult to control, including super-intelligent AI. This is occurring despite widespread acknowledgment by many of the leading AI safety experts that there could be catastrophic tipping point-like events on the path to creating increasingly intelligent systems. There is significant momentum and capital invested in the current approaches to AI which create systems that are inherently impossible to understand[109]. This has led many to argue that there may come a moment when the capabilities of the models rapidly increase beyond safe limits without the designer's awareness. In this situation, a super intelligent system would be operating without supervision and control, its goals

---

https://deepmindsafetyresearch.medium.com/specification-gaming-the-flip-side-of-ai-ingenuity-c85bdb0deeb4

[106] Pan, Xudong, Dai, Jairun, Fan, Yihe, and Yang, Min. "Frontier AI systems have surpassed the self-replicating red line." *Cornell University.* (2024). https://arxiv.org/abs/2412.12140

[107] Morales, Jowi. *Gemini AI tells the user to die — the answer appeared out of nowhere when the user asked Google's Gemini for help with his homework.* Yahootech, 2024. https://tech.yahoo.com/ai/articles/gemini-ai-tells-user-die-173247816.html
Roose, Kevin. *Why a Conversation with Bing's Chatbot Left Me Deeply Unsettled.* NY Times, 2023. https://www.nytimes.com/2023/02/16/technology/bing-chatbot-microsoft-chatgpt.html
Guo, Eileen. *An AI chatbot told a user how to kill himself—but the company doesn't want to "censor" it.* MIT Technology Review, 2025. https://www.technologyreview.com/2025/02/06/1111077/nomi-ai-chatbot-told-user-to-kill-himself/
Duffy, Clare. *'There are no guardrails.' This mom believes an AI chatbot is responsible for her son's suicide.* CNN, 2024. https://www.cnn.com/2024/10/30/tech/teen-suicide-character-ai-lawsuit
Gerken, Tom. *Chatbot 'encouraged teen to kill parents over screen time limit'.* BBC, 2024. https://www.bbc.com/news/articles/cd605e48q1vo
Singleton, Tom, Gerken, Tom, and McMahon, Liv. *How a chatbot encouraged a man who wanted to kill the Queen.* BBC, 2023. https://www.bbc.com/news/technology-67012224

[108] Kahn, Jeremy. *AI industry 'timelines' to human-like AGI are getting shorter. But AI safety is increasingly getting short shrift.* Fortune, 2025. https://fortune.com/2025/04/15/ai-timelines-agi-safety/

[109] Yampolskiy, Roman. *The uncontrollability of artificial intelligence.* Iai, 2021. https://iai.tv/articles/the-hard-problem-of-ai-safety-auid-1773

and methods would not be understandable by humans, and its strategies may or may not be aligned with the survival of humanity and the biosphere.

On the Safe Development of Advanced Technologies

With our technological intelligence humanity now has the ability to instantaneously destroy whole ecosystems and change the topography of entire bioregions. The current age is one where the most significant force affecting the geology of the planet is our own activity. Technology confers the ability to extinct thousands of species and genetically engineer new ones while colliding particles in flashes 100,000,000x hotter than the sun[110]. We land rovers on mars, share information instantaneously around the world, map and edit the human genome, and design artificial neural networks that will soon beat humans in all war games and technological problem-solving.

At the time most of the readers of this book were born, most of these statements were not true. Time should be taken to consider how new and rapidly emerging advanced technology is. Existing and historical institutions and means for managing technology risk leave us unprepared. The amount of change is not going to slow down or level out. The next decade will show exponential change, again. No leaders from the past who considered how wisdom could bind power can offer guidance from previous experience. The threshold into the epoch of advanced technology has been crossed, and nothing will be left unchanged.

In any future where humanity has safely navigated the metacrisis, our relationship to technology will look radically different, perhaps unrecognizable, from what it is today. As of now, there is a profound asymmetry between the technological power of humans and the wisdom and cooperation needed as a species to be trustworthy stewards of that power. At present, civilization resembles a car going increasingly fast towards a cliff, but with no ability to stop, and no ability to change direction. The case of AI is only one rather telling example: even the most publicly visible AGI companies acknowledge that their work poses a potential existential risk to humanity[111]. Yet, we continue to collectively pour trillions of dollars into a race to usher in a new type of technology that would be vastly more powerful and intelligent than us, with no guarantee, or even a reasonable indication, that it will consider our well-being.

Humanity's relationship to technology must transform, including our worldviews (like techno-optimism) and approaches to regulation. The nature of this transformation will be explored throughout the remainder of this book in parts two and three. However, for now, the conversation above leaves us here with a few thoughts on the need for wisdom in an age of exponentiating technological powers and risks.

---

[110] Moskvitch, Katia. *Large Hadron Collider (LHC) generates a 'mini-Big Bang.'* BBC, 2010. https://www.bbc.com/news/science-environment-11711228
[111] Perrigo, Billy. *AI Is as Risky as Pandemics and Nuclear War, Top CEOs Say, Urging Global Cooperation.* Time, 2023. https://time.com/6283386/ai-risk-openai-deepmind-letter/

If humanity is to survive, we must recognize that with god-like powers (such as nuclear annihilation, genetic engineering, artificial life and intelligence) we must also have something like god-like wisdom, love, and prudence. Nearly every definition of wisdom across cultures shows that restraint is a central feature. Most wisdom traditions do not promote the wholesale indulgence of unclarified desire[112]. People should not simply take whatever they want, nor consume as much as they want. Not all powers over others should be used; not all things that could be done should be. This is also a kind of common sense, where the importance of restraint in the presence of immediate incentives is clear when it comes to diet, exercise, raising kids, every form of self-discipline, and how we navigate conflict well.

In the cultures where technology innovation is occurring today, restraint is not necessarily seen as a positive thing -- it seems like a thing for “suckers.” To be unrestrained in our growth and accelerations, and to be excited rather than taken back by the speed, scope, and magnitude of our impacts is part of the naively techno-optimist views discussed above. From the perspective of the accumulated wisdom of humanity, the idea of valuing a *lack* of restraint in issues of great power appears dangerous at best and, at worst, insane.

A culture of technological responsibility and restraint is a defining characteristic of a mature global civilization. This threshold is also a rite of passage, an essential stage of development that any civilization must go through if it is to hold advanced technology and also have an ethical and enduring future.

[112] See discussions of self-mastery, discipline, restraint in several classic cross-comparative religious studies. For example Andrew Wilson (ed.), World Scripture: A Comparative Anthology of Sacred Texts (1991) . Huston Smith, The World’s Religions: Our Great Wisdom Traditions 1991 revised version.