Chapter 1

Introducing the Metacrisis: Civilization Emerging

Humanity's existence depends upon the relationship between two planetary-scale systems: the *biosphere* and *global civilization*. The biosphere is the complex and evolving system of all life on Earth and related physical processes. Global civilization, which depends upon the biosphere for its basic functions, is the total system of all human activity, including the built environment of interdependent technologies, governance processes, economic structures, communication systems, and cultural artifacts.[1] Together these vast entities constitute a life-support system for billions of humans. Civilization has been growing rapidly larger and more powerful and is for the first time in history fundamentally and irreversibly altering the dynamics of the biosphere. The relationship between the biosphere and global civilization has therefore reached a turning point.

By the middle of the last century, due to large-scale industrial technology and nuclear weapons, *homo sapiens* became the only species capable of both self-extinction and the destruction of the integrity of the biosphere as a whole. A growing global movement to safeguard the biosphere, arising in response to the effects of industrial activity on the natural world, underscored another unique attribute: we are the only species that can *conceive* of our responsibility for the fate of all life on Earth. These events signal epochal shifts in both the productive and destructive capabilities of civilization, and in the scope, urgency, and necessity of humanity's responsibilities to life and its own future.

Prior civilizations have shown patterns of rise, decline, and collapse (both gradual and sudden), and the current shift should be understood within this context. Many civilizations have failed

for the same set of reasons, including ecological overshoot and institutional breakdown.[2] Today, seemingly distinct civilizational blocs (e.g., Asian, European, Middle Eastern, etc.) have been woven into a global lattice of interdependencies, creating a single world system.[3] This is historically novel: no prior world system has been truly global, with technologies powerful enough to destroy planetary life support systems. This means that long-standing patterns that led to local catastrophes, such as wars or resource depletion, now have the potential to play out at a planetary scale. Moreover, these patterns are now amplified by new technologies that are increasing radically in both power and rate of change. There has never before been the possibility of *global* civilization breakdown or collapse. This is something we face now for the first time, precisely as a result of the successful growth of civilization beyond its planetary limits.

In the interest of all that is valuable in life, our current civilization must not succumb to these usual patterns of collapse. Instead, global civilization must transform from within while averting all catastrophes that could fundamentally undermine the potential for viable and humane futures. This constitutes the greatest challenge humanity has ever faced. In the face of potential collapse, some civilizations have succeeded in transforming to a higher order of problem-solving and wisdom.[4] Most, however, have collapsed and passed some of their knowledge on to adjacent, still-surviving civilizations.[5] This latter possibility is no longer an option in any meaningful sense, because there are no longer any adjacent societies. We are now one interconnected planetary society, and there is no "outside" left. In the broadest historical terms, what is discussed here as *the metacrisis* can be understood as a complex, dangerous, and ultimately terminal climax of the current world system – a sign that global civilization must transition into a new form if it is to survive.

The current world system enables the lives of more than eight billion humans at a higher-than-average standard of material wealth than any prior system.[6] It is precisely the success of this civilization that has led to the need for its fundamental restructuring. It is being forced to transform as a result of the destructive aspects of its productive capacities. As an example, it is evident that the extraction and pollution involved in the creation of material wealth cannot be continued at current rates without eventual global resource depletion and ecological despoliation.[7] Civilization must also transform its relation to its own technological power, as risks proliferate from engineers and scientists with the ability to access and manipulate the basic building blocks of matter, life, and mind.[8] As civilization has grown, so has the total complexity of the global risk landscape as well as its generative social dynamics. During the past hundred years, as catastrophic and existential risks crossed thresholds of global scope,

interconnection, mutual amplification, and intensity, these dynamics began to undermine the viability of the current world system. The metacrisis had begun.

If humanity survives, it will be because a world system is built that is capable of navigating the metacrisis. In practice, this means avoiding catastrophic global disorder, violence, and breakdown ("chaos") on the one hand, and global totalized control and authoritarian, dystopian order ("oppression") on the other.[9] Civilization must be rebuilt in response to the metacrisis, through a process explored here as a kind of navigation toward a "third attractor"—a place beyond the immediate and twin attractors of chaos and oppression. We conclude here with discussion of these probable near-term outcomes, alongside approaches to the problem and design space for building a viable future for humanity and all life on Earth.

Risks Pass Global Thresholds

Global catastrophic and existential risks—and the social dynamics that generate them—emerge as part of the natural development of civilizations working to solve increasingly complicated technological and social problems. Risks associated with industrial agriculture (for example, fertilizer runoff creating dead zones in the oceans) only arise as a result of successfully solving the problem of how to feed people. Nuclear weapons were created to end a global war and solve a problem of international security, and yet now the upper bound of global insecurity is radically higher as a result. There is a general principle in history that *successfully solving problems will often create new and sometimes more complex kinds of problems as a consequence.*

Understood in this light, the metacrisis— defined as the total state of all interconnected risks and social dynamics generative thereof—is the vast unintended outcome of the increasing impact of human problem-solving, mediated through ever more powerful technologies. This is another way of saying that the metacrisis is the accidental endgame—or *cumulative terminal negative externality*—of what is in fact, on its own terms, a largely successful, rapidly growing, and technologically advanced global civilization. The size, scope, and design of this civilization has led to the crossing of critical thresholds of global risk. The scale, scope, and interconnectedness of the risk landscape now creates fundamentally new conditions that set the context for decisions about the future.

In the decades after Hiroshima, world leaders and intellectuals spoke about the possibility of self-induced extinction as a threshold event, marking a new phase of history.[10] In many ways, the

so-called "post-war order" was built in response to this threat, including new global institutions like the United Nations and the International Monetary Fund. These major international projects recognized that humanity had crossed the thresholds of existential power, and that planetary responsibility impacted culture and civilization worldwide. The human rights and decolonization movements as well as the environmental, anti-nuclear, and global justice movements all arose downstream of this awareness.

While nuclear war has been avoided so far, other non-nuclear technologies with potentially catastrophic destructive powers have continued to develop since the middle of the last century. Beyond the domain of weaponry, technologies of extraction and production also advanced, enabling the exponential growth of economies and megacities.[11] As a result of the total accumulation of human impact since the Industrial Revolution, the biosphere and the chemical composition of the Earth have been fundamentally changed. This change marked the beginning of the Anthropocene.[12] The increasing scale and scope of risks generated by civilization is outpacing the development of the kind of global wisdom necessary to manage our new destructive capacities.

Today's civilization is crossing planetary boundaries of both extraction and pollution—a threshold beyond which human life support systems are under threat of failure.[13] Global civilization is taking more from nature than can be replenished and putting back into nature things that cannot be integrated or processed. Many prior civilizations have collapsed because of these same dynamics as they depleted their local ecosystems beyond repair.[14] Ours is the first civilization capable of depleting and despoiling Earth's biosphere as a whole. This threshold places our civilization in novel territory; seventy years ago, when the first concerns were expressed about planetary boundaries and limits to growth, these thresholds were a future problem. Now they have arrived.

Warfare and violence are another perennial cause of civilizational collapse. In today's global civilization, conflict is coupled to powerful new technologies such as AI, drones, cyber, and bioweaponry, all of which extend the destructive capacities of even small groups to previously unimaginable levels.[15] Arms races continue to result in new and more powerful catastrophic weaponry. Technologies developed for civilian application can be applied to military purposes while still remaining widely distributed within the civilian population (such as drones and AI). With the scale of violence now possible, conflicts can escalate both rapidly and unpredictably due to technological innovation. At a certain scale, volatility and violence can overwhelm

economic systems and governance capacities, leading to more conflict. A total war, waged at a global scale and involving our most powerful weapons, would end life as we know it. This is another threshold—the threshold beyond which violence becomes existential. Humanity first crossed this threshold when nuclear war became possible, and now there are more nuclear-armed nations and newer, non-nuclear means of civilization-damaging conflict than ever before.

Historically, horrific acts of violence have taken place both when warring peoples inflict it upon one another and when the ruling classes inflict it upon the ruled. Abuses of centralized systems of power are often as violent and dystopic as the catastrophic violence of war. In the past when such abuses were severe enough, the people would revolt. And while it may have been incredibly difficult to overthrow a corrupt government, it was not impossible. Today we are nearing this threshold of uncheckable power. Previously an autocratic ruler could not monitor everyone's behavior with widely distributed sensors and satellite technology. Nor could they use machine learning to process the information to inform real-time decision-making and immediately deploy swarms of drones for automated policing. Asymmetric power gained by newly emerging technologies could become so great that previous methods used to check corrupt systems of power become obsolete.

As these cases demonstrate, technologies structure civilization, shape humans, and impact nature. This has always been true, as technology has always been one of the most catalytic drivers of civilizational transformation. Today's civilization has expanded to a global scale with the adoption and use of a rapidly changing suite of technologies which will continue to change civilization radically, bringing many benefits while also increasing catastrophic risks.[16] Scientists, engineers, businesses, and governments are unleashing powerful new tools that will change the very foundations of life and our understanding of the universe. Humans can now perform work at all known levels of reality. Capabilities exist to alter some of the most basic components of the material world, to split atoms and change the composition of atomic nuclei.[17] Advances now allow for work at the most basic levels of life itself in the design and alteration of the human genome and how it expresses in body and mind.[18] Quantum computing has the potential to manipulate quantum mechanics—enabling work below the level of the classical physics of the universe.[19] The scale of industrial technology allows whole ecosystems to be destroyed, modified, or created—allowing for work at the scale of geoengineering.[20] Artificial intelligence is being developed across many kinds of substrates, and brain-computer interfaces are being created to augment human intelligence and the boundary of self—which amounts to work on the very nature of body, mind, and what it means to be human.[21]

The nature of these innovations will give humanity capabilities that challenge all existing legal and ethical frameworks, causing problems for all pre-existing mechanisms for technological governance and control. This is another threshold that places global civilization in novel territory—a threshold beyond which exponential technological power becomes ungovernable.[22]

Crossing these thresholds—global ecological limits, capacity for violence, uncheckable power and the ungovernability of technology —marks an historically novel state of global catastrophic risk. Understanding that these thresholds have been crossed is a necessary prerequisite for understanding the metacrisis. It is not just that there are many proliferating risks and crises; this has always been the case. The key difference now is that our civilization has crossed thresholds of global speed and complexity so that today's crises increasingly fall within the domain of global catastrophic and existential risk.

Generative Dynamics

As the metacrisis intensifies, perennial patterns of human behavior and motivation remain unchanged. The metacrisis is the result of the interplay between recent historical events and the amplification and cumulative impact of enduring human practices, including approaches to problem-solving, conflict, and the creation of new technologies. The basic social dynamics of civilizations unfold within *superstructures* (beliefs and worldviews), *social structures* (governance and institutions), and *infrastructures* (economies, technologies and material supplies).[23] The way these basic structures have functioned throughout history has usually resulted in self-terminating systems.[24] The metacrisis is the climax of a set of long-standing social dynamics that have driven technological progress, but that have also led to a range of atrocities, in addition to the current increase in global catastrophic risk.

In general, these fundamental roots of the metacrisis are referred to here as *generative dynamics*. These roots are less obvious than the risks themselves, but they may be found beneath all of the world's most significant and persistent problems. They are complex, recurring patterns of human activity that are incentivized by and embedded within legal, economic, and technological structures. Generative dynamics include issues like legally encoded economic incentives that encourage businesses to externalize their costs (e.g., officially accounting for production and consumption but not pollution or extraction).

Generative dynamics also include less explicit cultural norms and expectations that shape decision-making, such as what is taught in a culture about who is good or bad, successful or not. For example, whole populations will make different political and economic decisions depending upon whether their culture believes that the conspicuous consumption of material goods makes someone a good person or not.[25] The rules and agreements taken for granted in a culture, as well as the explicit codes of conduct and ethical commandments, combine to shape large-scale trends in decision-making that can change the trajectory of a civilization.

A key type of generative dynamic that is often found at the center of collective human activity is known as a "social trap." An arms race, such as those we have witnessed in relation to autonomous weapons or hypersonic missiles, is an example of a type of social trap (a *multipolar trap*).[26] When the various actors in a conflict see their only reasonable action as escalation, all sides race to increase their destructive capacity. And so begins a social dynamic that is difficult to escape, in which the long-term outcome for all actors (i.e., a total increase in destructive capacities for all) is not desired by any party, even as short-term gains in security are achieved in the process.

As the total landscape of global risks began to evolve beyond critical thresholds, many underlying patterns of human behavior and institutional design became untenable as the basic structures of future world systems. Understanding generative dynamics is the basis of understanding the metacrisis. It allows us to understand not only the fundamental connections between risks and crises, but also the ongoing amplification of the persistent social dynamics that give rise to such risks. These patterns of connection and amplification drive the total landscape of risk past critical global thresholds. At base, it is these patterns that now demand immediate innovations in global cooperation and governance.

Possible Futures: The Third Attractor

The dominant generative dynamics at work beneath a civilization set the possibilities for its future state. Continuing to run the current generative dynamics beyond a set of critical thresholds has two probable outcomes. These can be thought of as "attractors." When a complex system—like an organism or ecosystem—moves far beyond equilibrium, emergent patterns arise that restore order and establish new potential paths ahead in an otherwise chaotic situation.[27] The emergent probabilities that draw complex systems into the future are called *chaotic attractors*.

As noted above, there are two primary attractors looming over the future of our complex and increasingly chaotic world. Both can be considered as failure states for humanity. On the one hand, it is increasingly likely that the world system is *disordered by cascading global catastrophes*. On the other hand, as a means to avoid global catastrophe, the world system is likely to be *ordered through global autocratic control*, creating a system open to dystopian futures, either through capture, corruption, or breakdown. A third attractor must be found and made the shared point of social coordination, design, and orientation. Navigating the metacrisis requires extraordinary care to avoid two types of choices simultaneously: choices that could lead to intensifying and widening catastrophes, and choices that could lead to increasingly powerful centralized control systems (as a pathway toward dystopias).

The actions and prescriptions following from this insight suggest that the metacrisis cannot be "fixed" or "solved" (implying top-down solutions). Similarly, the metacrisis cannot be addressed exclusively through bottom-up, market, distributed, or peer-to-peer mechanisms. Instead, we suggest that the concepts of navigation and design provide a useful framework for thinking about how to address the metacrisis. Here, we discuss the need for efforts that are both top-down and bottom-up, while at the same time exhibiting a dynamic relationship between the two, at various scales and across domains. These interventions must focus upon the design, care, and steering of our built environment of interdependent technologies, governance systems, and cultural artifacts. This work has to be addressed simultaneously over short-term (referred to here as *triage*), medium-term (*transition*), and long-term time (*transformation*) horizons. The third attractor—and its related set of design principles—is the focus of the concluding chapters.

Overall Outline

The work begins with an overview of the risk landscape (Part One), offering a taxonomy of risk types and a description of the set of interrelated crises and potential catastrophic outcomes. The total risk landscape is the *explicit* domain of the metacrisis, in which the various visible aspects of the world system may be observed in a state of instability and fragility (see Figure 1). In Part Two, we will explore the generative dynamics of the metacrisis, which constitute the *implicit* domain of the metacrisis and include less obvious and often more difficult-to-perceive realities. The domain of the explicit (the cascading crises and expanding risks) is related to the domain of the implicit (the generative dynamics of human choice and design) in a feedback loop of self-creation, or autopoiesis. The concluding chapters of Part Three describe principles and

design criteria necessary to address the generative dynamics of the metacrisis, endeavoring to create what is needed to navigate toward a third attractor beyond chaos and oppression.
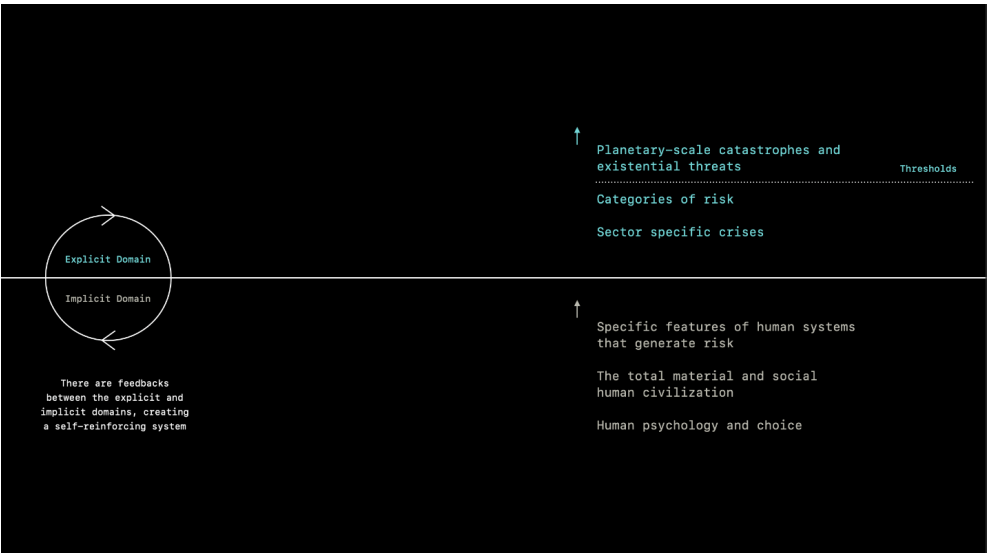
Figure 1

Figure 2.

Human relation to:		Threshold Approaching:
Autocratic Rule + Continuous increased technological power.	➡	Concentrations of power so great that all prior means used to overthrow authoritarian rule are rendered obsolete.

ENDNOTES

1. For an introductory overview concerning theories of civilization, see for example:
   - Joseph A. Tainter, *The Collapse of Complex Societies* (Cambridge University Press, 2011).
   - Peter Turchin and Sergey A. Nefedov, *Secular Cycles* (Princeton University Press, 2009).
   - Carroll Quigley, *The Evolution of Civilizations: An Introduction to Historical Analysis*. 2nd ed. (Liberty Press, 1979).
   - Jared M. Diamond, *Guns, Germs, and Steel: The Fates of Human Societies* (W. W. Norton, 1997).
   - Jared M. Diamond, *Collapse: How Societies Choose to Fail or Survive* (Penguin Books, 2005).

2. See note 1 above, especially Jared Diamond.
3. See for example: Immanuel Wallerstein, *The Modern World-System* vol. 1: *Capitalist Agriculture and the Origins of the European World-Economy in the Sixteenth Century* and vol. 2: *Mercantilism and the Consolidation of the European World-Economy, 1600-1750* (Academic Press, 1974).
4. See for example:
   - Michael Lecker, ed. *The 'Constitution of Medina': Muhammad's First Legal Document* (Darwin Press, 2004).
   - Samuel B. Payne Jr., "The Iroquois League, the Articles of Confederation, and the Constitution," *The William and Mary Quarterly* 53, no. 3 (1996): 605-620.
5. See Ronald K. Faulseit, ed., *Beyond Collapse: Archaeological Perspectives on Resilience, Revitalization, and Transformation in Complex Societies* (SIU Press, 2016).
6. This is a complex statement. While true in one sense (i.e., by some definitions of "material wealth"), it is also true that, in terms of total numbers, more people live in extreme poverty today than the entire global population before the Industrial Revolution. Some aspects of material wealth can lower quality of life (such as lack of access to nature, views, clean air, etc.), while many

aspects of immaterial wealth have declined in pursuit of its material manifestation. For an interesting discussion of these points see:

- Robert E. Lane, *The Loss of Happiness in Market Democracies* (New Haven: Yale University Press, 2001).
- Robert E. Lane, *The Market Experience* (Cambridge University Press, 1991).

7. For perspectives and contemporary research on planetary boundaries associated with industry and extraction, see for example:

- Johan Rockström et al., "Planetary Boundaries: Exploring the Safe Operating Space for Humanity," *Ecology & Society* 14, no. 2 (2009): 32, https://doi.org/10.5751/ES-03180-140232.
- Will Steffen et al., "Planetary Boundaries: Guiding Human Development on a Changing Planet." *Science* 347, no. 6223 (2015), https://doi.org/10.1126/science.1259855.
- Linn Persson et al., "Outside the Safe Operating Space of the Planetary Boundary for Novel Entities." *Environmental Science & Technology* 56, no. 3 (2022): 1510-1521, https://doi.org/10.1021/acs.est.1c04158.
- Lan Wang-Erlandsson et al., "A Planetary Boundary for Green Water." *Nature Reviews Earth & Environment* 3 (2022): 380-392, https://doi.org/10.1038/s43017-022-00287-8.

8. For perspectives on the risks associated with emerging technologies, see for example:

- Nick Bostrom, "The Vulnerable World Hypothesis," *Global Policy*, 10, no. 4 (2019): 455–476, https://doi.org/10.1111/1758-5899.12718.
- David Manheim, "The Fragile World Hypothesis: Complexity, Fragility, and Systemic Existential Risk. *Futures* 122 (2020), https://doi.org/10.1016/j.futures.2020.102570.
- Daniel Deudney, *Dark Skies: Space Expansionism, Planetary Geopolitics, and the Ends of Humanity* (Oxford University Press, 2020).
- Miles Brundage et al., "The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation" *ArXiv* (2018), https://arxiv.org/abs/1802.07228.
- Tucker Davey, "Benefits & Risks of Biotechnology," Future of Life Institute, Octover 28, 2022, accessed March 31, 2023, https://futureoflife.org/biotech/benefits-risks-biotechnology/.
- Jonathan B. Tucker, ed., *Innovation, Dual Use, and Security: Managing the Risks of Emerging Biological and Chemical Technologies* (MIT Press, 2012).

9. "Teetering Between Oppression and Chaos." *The Consilience Project*, August 3, 2022, https://consilienceproject.org/teetering-between-oppression-and-chaos/.

10. For this moment in global culture, see Johnathan Shell's works: *The Fate of the Earth* (Alfred A. Knopf, 1982) and *The Abolition* (Alfred A. Knopf, 1984).

11. For work on the exponential growth of urban landscapes, see for example: Geoffrey West, *Scale: The Universal Laws of Life, Growth, and Death in Organisms, Cities, and Companies* (Penguin Books, 2018).

12. See for example:

- Erle C. Ellis, *Anthropocene: A Very Short Introduction* (Oxford: Oxford University Press, 2018), https://doi.org/10.1093/actrade/9780198792987.001.0001.
- Will Steffen et al., "The Anthropocene: From Global Change to Planetary Stewardship." Ambio 40, no. 7 (2011): 739-61, https://doi.org/10.1007/s13280-011-0185-x.
- Carle Folke et al., "Our Future in the Anthropocene Biosphere," *Ambio* 50, no. 4 (2021): 834-869, doi:10.1007/s13280-021-01544-8.
- Colin N. Waters et al., "The Anthropocene is Functionally and Stratigraphically Distinct from the Holocene" *Science* 351, no. 6269 (2016): https://doi.org/10.1126/science.aad2622.
- *Human Development Report 2020, The Next Frontier: Human Development and the Anthropocene*, United Nations Development Programme (2020), https://hdr.undp.org/content/human-development-report-2020.

13. See note 7 above.

14. See note 1 above, in particular Tainter and Diamond.

15. See for example:

- "Artificial Intelligence and Arms Control," Center for a New American Security (2022), https://www.cnas.org/publications/reports/artificial-intelligence-and-arms-control.
- Melanie Sisson et al., "The Militarization of Artificial Intelligence," Stanley Center for Peace and Security (2020), https://stanleycenter.org/wp-content/uploads/2020/06/TheMilitarization-ArtificialIntelligence.pdf.
- Paul Scharre, *Army of None: Autonomous Weapons and the Future of War* (W. W. Norton & Company, 2018).
- Fillipa Lentzos, Catherine Jefferson, and Claire Marris, "The Myths (and Realities) of Synthetic Bioweapons," *Bulletin of the Atomic Scientists* (2014), https://thebulletin.org/2014/09/the-myths-and-realities-of-synthetic-bioweapons/.

16. For a range of commentary on the civilizational impacts of exponential technologies, see for example:

- Martin Ford, *Rise of the Robots: Technology and the Threat of a Jobless Future* (Basic Books, 2015).

CIVILIZATION
RESEARCH
INSTITUTE

- Jonathan Haidt and Eric Schmidt. "AI is About to Make Social Media (Much) More Toxic." *The Atlantic*, (May 5, 2023). www.theatlantic.com/technology/archive/2023/05/generative-ai-social-media-integration-dangers-disinformation-addiction/673940/.
- Yuval Harari, Tristan Harris, and Aza Raskin, "You Can Have the Blue Pill or the Red Pill, and We're Out of Blue Pills," *The New York Times* (March 24, 2023), www.nytimes.com/2023/03/24/opinion/yuval-harari-ai-chatgpt.html.
- Jennifer Kuzma and Todd Tanji, "Unregulated Genetic Engineering and the Future of Humanity: A Call for a New Cultural Biotechnology," *Biotechnology and Development Monitor* 48 (2001): 2-7.
- Toby Ord, "Future Risks: Unaligned Artificial Intelligence" in *The Precipice: Existential Risk and the Future of Humanity* (Hachette Books, 2020).
- Cathy O'Neil, *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy,* (Penguin, 2016).

17. The capability to split the atom was first demonstrated in a series of foundational experiments spanning the 1930s. Most of these developments were built upon Ernest Rutherford's work on the nature of the atom and subsequent characterization of α and β radiation. For a general overview of atomic physics see: Ernest Henley and Alejandro Garcia, *Subatomic Physics* (World Scientific, 2007). Foundational papers include:

- Ernest Rutherford, "The Scattering of α and β Particles by Matter and the Structure of the Atom." *Philosophical Magazine Series* 6, 21, no. 125 (1911): 669-688.
- Lise Meitner and Otto Robert Frisch, "Disintegration of Uranium by Neutrons: A New Type of Nuclear Reaction," *Nature* 143, no. 3615 (1939): 239-240.
- Otto Robert Frisch, Physical Evidence for the Division of Heavy Nuclei under Neutron Bombardment." *Nature* 143, no. 3616 (1939)": 276.
- Niels Bohr and John Archibald Wheeler, "The Mechanism of Nuclear Fission." *Physical Review* 56 no. 5 (1939): 426-450.

18. The potential for human genetic engineering became a reality following the discovery of the structure of DNA in 1953. Many further steps over the course of seventy years of genomic research have led to contemporary techniques such as CRISPR. For a contemporary overview, see: Jennifer A. Doudna and Samuel H. Sternberg, *A Crack in Creation: Gene Editing and the Unthinkable Power to Control Evolution*, (Houghton Mifflin Harcourt, 2017). Key research works include:

- James D. Watson and Francis H. Crick, "Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid," *Nature* 171, no. 4356 (1953): 737-738.
- Stanley Cohen et al., "Construction of Biologically Functional Bacterial Plasmids In Vitro," *Proceedings of the National Academy of Sciences* 70, no. 11 (1973): 3240-3244.
- Andrew Fire et al. "Potent and Specific Genetic Interference by Double-Stranded RNA in *Caenorhabditis elegans*," *Nature* 391, no. 6669 (1998): 806-811.
- Prashant Mali et al., "RNA-Guided Human Genome Engineering via Cas9," *Science*, 339, no. 6121 (2013): 823-826.

19. See for example:

- Congressional Research Service, "Defense Primer: Quantum Technology," Updated November 15, 2022. https://crsreports.congress.gov/product/pdf/IF/IF11836.
- Samo Burja, "Quantum Technology Appeals to World Powers," *Bismarck Brief*, July 27, 2022.
- Stephen Witt, "The World-Changing Race to Develop the Quantum Computer," *The New Yorker*, December 12, 2022.

20. See for example:

- Benjamin Bratton, *The Terraforming* (Strelka Press, (2019).
- Holly Jean Buck, *After Geoengineering: Climate Tragedy, Repair, and Restoration* (Verso, 2019).
- David W. Keith, *A Case for Climate Engineering* (MIT Press, 2013).
- Wake Smith, *Pandora's Toolbox: The Hopes and Hazards of Climate Intervention* (Cambridge University Press, 2022).

21. For a foundational text exploring the range of potentialities for AI, see: Nick Bostrom, *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/. For commentary on brain-computer interfaces, see: Simanto Saha et al., "Progress in Brain-Computer Interface: Challenges and Opportunities." *Frontiers in Systems Neuroscience*, 15 (2021): https://doi.org/10.3389/fnsys.2021.578875.

22. For a primary perspective on governance and emerging technologies see: David Collingridge, *The Social Control of Technology* (New York: St. Martin's Press, 1980). For recent developments and commentary in relation to the challenge of AI in particular, please see:

- UN Chief Executives Board for Coordination, "Report of the High-Level Committee on Programmes of the United Nations System Chief Executives Board for Coordination at its thirty-fourth session," 2017, accessed August 4, 2023, https://unsceb.org/sites/default/files/imported_files/CEB_2017_6%20%28HLCP%2034%29_0.pdf.

12

- Future of Life Institute "Pause Giant AI Experiments: An Open Letter," 2023, accessed August 4, 2023, https://futureoflife.org/open-letter/pause-giant-ai-experiments/.
- Eliezer Yudkowsky, "Pausing AI Developments Isn't Enough. We Need to Shut It All Down." *Time*, March 29, 2023, https://time.com/6266923/ai-eliezer-yudkowsky-open-letter-not-enough/.

23. See note 1 above.
24. See note 1 above.
25. See for example:
    - Thorstein Veblen, *The Theory of the Leisure Class: An Economic Study of Institutions* (Macmillan, 1899).
    - Robert E. Lane, *The Market Experience* (Cambridge University Press, 1991).
26. For an overview of core texts with relevance to social traps (and game theory more broadly), see for example:
    - William Poundstone, *Prisoner's Dilemma: John von Neumann, Game Theory, and the Puzzle of the Bomb*, (Doubleday, 1992).
    - Thomas C. Schelling, *Micromotives and Macrobehavior* (W. W. Norton, 1978).
    - Thomas C. Schelling, *Strategies of Conflict* (Harvard University Press, 1960).
    - Elinor Ostrom, Roy Gardner, and James Walker, *Rules, Games, and Common-Pool Resources* (University of Michigan Press, 1994).
    - Bruce Schneier, *Liars and Outliers: Enabling the Trust that Society Needs to Thrive* (John Wiley & Sons, 2012).
    - Bo Rothstein, *Social Traps and the Problem of Trust* (Cambridge University Press, 2005).
27. Ralph Abraham, and Christopher D. Shaw, *Dynamics—The Geometry of Behavior: Global Behavior* (Aerial Press, 1982).

Part I

The Risk Landscape