THE LOOMING SHADOW: TAKING SERIOUSLY POTENTIAL EXISTENTIAL THREATS BROUGHT ABOUT BY ARTIFICIAL INTELLIGENCE

BORIS D. GROZDANOFF

Institute of Philosophy and Sociology, Bulgarian Academy of Sciences

QAISEC / DISI / RAISON

grozdanoff@gmail.com

Abstract

The explosive development of Artificial Intelligence (AI) is poised to become a watershed in human progress, capable of radically transforming our existence. In this paper I examine a potential existential threat space, enabled by future AI systems. I discuss a spectrum of existential risks, including genetic modification, warfare, cyber attacks, as well as the emergent threats posed by Artificial General Intelligence (AGI), like the emerging global erosion of truth. I critically discuss the recent LeCun - Bengio debate on AI safety and I argue, contra LeCun, that existential risks brought about by AI systems are a non-trivial threat that is currently in the period of transcending from the realm of conceivability to the domain of possibility. In order to contain such threats as much as practically possible, I argue that the best strategy for a democratic benevolent state-level agent is to develop an AGI system first as this would enable us to achieve superiority that could conceivably contain and block the existential threats. I examine a number of recent influential arguments against the possibility of existential AI risks, I provide counterexamples to each of them and I argue that all of them fail.

Keywords: Artificial Intelligence, AI safety, Existential Threats, Truth, AGI

The first Global AI Safety Summit, [1] held in the historical site of *Bletchley Park* in England, concluded a week ago with the international acceptance of a declaration that has already become historical, *The Bletchley Declaration* on AI Safety. [2] The event gathered political leaders as well as managers and AI architects of the world's leading companies that spearhead the evolution of the artificial intelligence technologies, such as OpenAI, Meta, Google DeepMind, Inflection and Microsoft. The declaration announces an international commitment to AI safety and provides a detailed list of threats in different domains, e.g. finance, news, bioengineering, cyber security and others. Yet, the text is quite careful in avoiding an explicit formulation of the *existential threats* brought about by future AI systems and the careful reader will only see a masking expression, such as "unforeseen risks" or "…

serious, even catastrophic, harm, either deliberate or unintentional, stemming from the most significant capabilities of these AI models." Given the quite intense emerging debate in the public space, often led by leading figures like the Turing award winners Hinton, Bengio and LeCun, it is quite clear that the matter of potential existential threats for humanity brought about by future AI systems is a theme considered serious enough to deserve the time of some of the most brilliant minds in artificial intelligence today.

This marks the quite fascinating transition of the problem from the realm of science fiction to the domain of scientific and political debate, no matter how masked. In this paper I will try to begin addressing it, and for this purpose I will discuss a number of recent arguments formulated in social media by Yann LeCun, perhaps the most influential expert proponent of the position that AI systems do not and will not pose existential threats to humanity. Instead of merely repeating counter-responses, such as the ones by Bengio, [3] in an interview as well as in the same exchange, and by Hinton, both of which I mostly agree with, I will offer a few of my own that directly respond to the justification behind the claims. I have extracted and reformulated, as best as I could, the formal argumentation behind a number of reasons why should we not accept existential threats by AI as a serious enough possibility. I show that each of the arguments considered here fails in the light of mere logical reasoning and counterexamples from history of science, technology, politics and warfare.

The social media debate between LeCun and Bengio uses the modal instruments of *conceivability* and *possibility* [4] to argue pro and contra AI existential threats. Conceivability of threat scenarios steps on current advances in ML technologies and modally justified expectations of further technological progress. Its directions are largely unanticipated in detail today, but given the speed of the progress made in the last 5 years alone, it seems quite inevitable to expect the fast emergence of much more *capable AI models* advancing towards generality. Together with the parallel progress of *quantum computers* which harness the unchartered but immense power and parallelizability of computation via the more informationally powerful qubits (compared to classical bits, which encode 1 of only two possible states), as well as novel logical operations that harness the power of quantum entanglement, we are facing imminent technological breakthroughs that have already begun to transform the global society.

In the social media exchange Yann LeCun puts forward the following theses that contain a number of distinct arguments, each of which aims (explicitly or tacitly) to defend the position that AI threats receive far more attention that they deserve:

"The heretofore silent majority of AI scientists and engineers who:

1. - do not believe in AI extinction scenarios or
2. - believe we have agency in making AI powerful, reliable, and safe and
3. - think the best way to do so is through open source AI platforms "

(1) it's not about worrying about risks nor imagining catastrophe scenarios, it's about *designing* AI systems for safety. I'll repeat the turbojet analogy. Speculating about the potential dangers of turbojets is pointless. Doing all the careful engineering to make them reliable is what matters. How to do it is a question for specialists, and scaring the public about it is counterproductive.

(2) that's just completely false and, in fact, quite naive. There is a huge amount of investment going into making AI systems safe and reliable. There is also a huge amount of investment in *using* AI to make various systems safer and more reliable, including cars, medical diagnosis systems, communication platforms, and social networks.

(3) AI systems are not weapons any more than a car or a computer is a weapon. AI systems are designed to make *people* smarter. Assimilating them to weapons, particularly nuclear weapons, is downright ridiculous. For the record, I'm all in favor of more restrictive gun laws in the US, and I have no issue with regulating AI *products*: obviously, a driving assistance system or a medical image analysis system should go through some sort of certification. But such regulations are already in place.

Today's systems do little more than approximately regurgitating information they have been trained on from the public Internet. In other words, they can't give you actionable information that is not already available through a search engine. Future AI systems may have more powerful capabilities. But they will almost certainly be very different from current Auto-Regressive LLMs. Until we have a design, it's premature to speculate about ways to make them safe, just like speculating on the safest design of turbojets in 1925. [5]

We can unpack LeCunn's arguments against the possibility of AI-related existential threats as the following distinct arguments:

1. (some) AI experts do not believe in AI extinction scenarios, because of their expertise of the field, which justifies this conclusion
2. (some) AI experts in virtue of their understanding of the field see very high probability that AI designers would manage to make (a) powerful AI systems, (b) reliable AI systems and (c) safe AI systems
3. The best way to reach (2a), (2b) and (2c) is through the method of open source AI systems.
4. If we design AI systems to be safe (which we can) we will not face AI extinction scenarios.

The turbojet analogy shows that a reliable design and implementation of AI systems can avoid AI extinction scenarios.

5. Immense investment is being injected in AI reliability and safety, therefore we will not face AI extinction scenarios.
6. AI systems are not by its architecture weapons and should not be assimilated to weapons. Due to them not being architecturally weapons we will not face AI extinction scenarios where AI systems function as weapons (because they are not weapons by design).
7. Today's systems cannot give actionable information that can be used for AI extinction scenarios.
8. Future AI system might be pore powerful, but it is premature to speculate about their safety until we have their design. Lack of information of design of AI future systems does not provide grounds (today) to argue that they can bring about AI extinction scenarios.

These arguments present a variety of perspectives on the potential risks of AI and the probability of it bringing about existential threats to humanity. I propose a short analysis for each argument, and as I identify the weak premises, I offer a counterexample from history of science and history of technology and shortly discuss why the argument fails, in terms of the formal logical validity or soundness of its premises, conclusion, or both.

In Argument (1) the thesis is that AI experts do not believe in AI extinction scenarios due to their expertise. The argument assumes that most – if not all – AI experts have a consensus, and that their expertise can guarantee accurate predictions regarding future AI risks. Unfortunately, the history of science is full of instances where experts underestimated risks (e.g., early assumptions about the safety of nuclear energy). As a recent significant example we can take the *Fukushima Daiichi nuclear disaster*, which occurred in 2011 and can serve as a stark illustration of the limits of experts' consensus. Prior to the disaster, nuclear experts widely considered the safety measures at the Fukushima Daiichi Nuclear Power Plant in Japan as technologically sufficient. However, an unforeseen combination of a massive earthquake and a subsequent tsunami caused catastrophic failures in the plant, leading to nuclear meltdowns and releases of radioactive materials, thus leading to one of the worst nuclear accidents in history. It was later revealed that the disaster was primarily due to the plant's inability to withstand a tsunami wave, a risk that was obviously underestimated by the experts. The Fukushima disaster serves as a quite apt counterexample to the claim in (1) that expert consensus on AI can accurately predict and mitigate all risks. Just like the unanticipated disaster that tore down the nuclear plant's defenses, unforeseen developments in AI could potentially lead to catastrophic outcomes that current experts have failed to predict. This tragedy illustrates that even expert consensus can fail to anticipate and mitigate complex risks in advanced technologies. For what is worth, given the discrepancy between the complexity level of this incident and that of the AI systems expected in the near future, we have serious reasons to worry that the limits of what AI experts can foresee would be much greater than those of Fukushima experts. Expert consensus is not infallible at all. Predicting the long-term trajectory of a rapidly evolving field like AI is fraught with uncertainty, thereby making argument (1) unsound.

In Arguments (2) and (3) AI experts are confident that AI designers can create powerful, reliable, and safe AI systems. The argument assumes current understanding and capabilities regarding AI will continue to improve linearly and without unforeseen challenges. Here, too, we can provide a historical illustration how despite software engineering advancements, critical software systems still fail: the Boeing 737 MAX crashes in 2018 and 2019 highlight the challenges in designing complex systems; the same type of challenge lies at the heart of the assumption of the AI Design Capability Argument. The incidents in question involved two Boeing 737 MAX airplanes which crashed in Indonesia and Ethiopia, respectively, due to malfunctions in their flight control systems. The malfunctions consisted in a flawed automated flight control feature (MCAS) that repeatedly pushed the plane's nose down. These flaws were a result of design and regulatory oversights, despite Boeing's expertise in aircraft design and overall strong safety record. They demonstrate that building reliable and safe systems can be fraught with unforeseen challenges, even when experienced designers are involved. Likewise, even with advanced technology and expertise, the design of complex systems like AI may be characterized by critical flaws which are practically impossible to foresee. This counterexample challenges the assumption behind argument (2) that AI systems can be designed to be inherently powerful, reliable, and safe. This assumption effectively *overlooks the complexity and unpredictability of AI development*, especially at advanced levels, thereby making the argument potentially unsound.

In Argument (4) the thesis is that open-source AI systems are the best method for ensuring safety and reliability. The argument assumes that open-source development inherently leads to safer and more reliable systems. However, it is a well known fact that open-source projects can suffer from lack of oversight, inconsistent quality control, and security vulnerabilities, as was the case with the "Heartbleed bug" discovered in 2014 in the OpenSSL cryptography library, which highlights the non-hypothetical and often rather significant role of vulnerabilities in open-source projects. OpenSSL, a widely used open-source tool, had one simple yet critical programming bug that went unnoticed for two years and compromised both the reliability of the library and the security of millions of websites by exposing significant volumes of sensitive data. The Heartbleed case shows how open-source projects, while beneficial in many ways (including, non-trivially, to the development of AI discovered vulnerabilities), are not immune at all to serious oversights and vulnerabilities. To further stress the latter point: it is not unjustified to affirm that perhaps the best tools today to discover bugs, errors and vulnerabilities unknown to human programmers are AI systems themselves. The potential of AI models to discover zero day exploits has increased immensely and we can expect the total number of AI-discovered zero day exploits to exceed the that of human-discovered ones soon. This counterexample challenges the assumption of argument (3) that open-source development is the best method to ensure AI safety and reliability by showing that open-source development does not inherently guarantee safer or more reliable systems. The argument fails because it overlooks the human limits and the practical challenges in open-source project management and quality assurance.

Argument (5) claims that designing AI systems for safety can prevent AI extinction scenarios, as shown by the turbojet analogy. The argument likens the complexity of AI systems to that of turbojet engines, severely oversimplifying AI's unique risks. Complex systems, however, provide much more fertile ground for failure, than analogue smaller systems. Thus, as a counterexample illustration, we can consider how the financial markets, designed with state of the art multilayered safeguards in place, have still experienced catastrophic failures, like the notorious 2008 financial crisis [6] which triggered a global collapse of the housing market and financial institutions. Despite all measures and constantly improving regulations designed to ensure the stability of financial systems, a combination of complex financial instruments and systemic vulnerabilities led to a *world-wide economic meltdown*. The chaos which ensued highlights the fact that complex systems – even those designed with safety in mind – can still experience unforeseen failures. This illustration serves as a counterexample to the frivolous idea that AI, like turbojets, can be made entirely safe through careful design alone. The very analogy is overly simplistic, does not account for AI's distinct complexity and inevitable social effects, and thereby clearly makes the argument unsound.

Argument (6) claims that current and expected significant investments in AI safety and reliability will prevent AI extinction scenarios. The argument assumes investment alone can guarantee success in achieving safety and reliability in digital technologies. This claim, again, is heavily opposed by a constant multitude of digital challenges which remain despite continuously growing investments in the domain. As a general illustration, we can consider the persistent challenge before cybersecurity. Governments and corporations invest billions annually in a broad spectrum of cybersecurity measures, yet significant breaches and cyber threats continue to occur globally. What is more, they appear to be ever-growing. Technologies such as IotT, Big Data, AI and quantum computers only open up novel avenues for attacks (for example, AI opens the door for a much larger number of zero day attacks whereas powerful 4000+ qubits quantum computers will gravely endanger the all prevailing asymmetric encryption), albeit some of them also help us develop novel methods of cyber protection. The cyber threat example shows that while investment is by all means crucial, it does not automatically solve complex technical challenges – significant breaches and cyber threats continue to impact states and organizations globally. These incidents often result from sophisticated attack techniques, human errors, and, naturally, the evolving nature of technology. The existence of such cyber security challenges demonstrates that investing in AI safety and reliability does not guarantee success at all. This provides a counterexample to the claim in argument (6) that financial investment alone is enough to prevent AI-related existential threats.

Argument (7) claims that AI systems are not weapons by architecture, and thus won't lead to AI extinction scenarios. The argument seems to assume that only systems explicitly designed as weapons can pose existential threats. This, again, is far from true. Plenty of disasters have been caused by the use – sometimes unintentional, other times deliberate – of non-weapon technologies as a weapon. Consider the ongoing aggressive war led by the Russian Federation in Ukraine, the blowing up of the Kakhovka dam, the shelling of the Zaporizhia nuclear power plant, and the weaponisation of the global grain deliveries. The Bhopal disaster in 1984 also comes to mind, where a leak of methyl isocyanate gas at a pesticide plant in India led to thousands of deaths. It illustrates the potential dangers of non-weapon technologies, as it was caused not by a weapon but an industrial accident resulting from poor maintenance, safety system failures, and inadequate emergency response measures. The Bhopal tragedy presents a significant challenge to the argument that AI systems are safe from causing existential threats simply because they are not designed to be used as weapons. The claim in (6) fails because it overlooks the potential (as well as actual history) of non-weapon systems causing harm through misuse or unintended consequences, and the multitude of examples render the argument unsound.

Argument (8) claims that today’s AI systems cannot facilitate AI extinction scenarios. The argument assumes that current limitations of AI will persist indefinitely, and that – given their early artificial complexity – they cannot be used as a stepping stone to a next generation of AI models that will lack these limitations. The second assumption is blatantly wrong, as we already see the potential of current models for discovering novel types of cyber attacks, deep fake propaganda and erosion of truth, which shows it can affect the global population. As instruments that are constantly and speedily improving, today’s AI models serve an unforeseen multitude of purposes, and no restraints have been proven to contain malicious use – the ingenious ways to hack the restrained responses, by role playing and for research purposes, for example, demonstrate this quite clearly. The first assumption can be confronted with a number of counterexamples, one of the weaker ones being that the rapid evolution of technology, such as the progression from early computers to modern supercomputers, shows that current limitations can be quickly overcome. Let us recall the room-sized ENIAC in the 1940s and compare it to today's powerful handheld devices: the speed and trajectory of the technological evolution shows that current limitations in AI are unlikely to be indicative of the capabilities of future models. The rapid development of computing technology, from early analogue computers to today's advanced smartphones, supercomputers and quantum computers, illustrates the exponential growth in technological capabilities. This growth is driven by continuous innovation, investment, and the increasing integration of computing into various fields. The transformative progress in computing technology serves as a strong counterexample to the notion that the limitations of current AI systems are and would remain static. It provides serious reasons to expect that present AI limitations will very soon be overcome, and thus makes it unwise to discount future AI risks based on current capabilities.

Argument (9) claims that speculations about the safety of future AI systems are premature as we currently lack a sufficiently clear idea of their design. In fields like bioengineering, however, potential risks are often *discussed and regulated before a particular application is fully developed*. As a counterexample, consider the discussions about the potential risks of CRISPR, which have been in progress since long before these gene editing technologies were in use. The capability to edit human genes raises a multitude of highly non-trivial ethical concerns regarding their unintended consequences, including ecological impacts and existential threats scenarios. This proactive approach to risk assessment in a rapidly developing field clearly demonstrates the importance of *speculating* and *regulating* future technologies in order to mitigate potential threats. The case of gene editing technologies openly challenges the idea that it's too early to speculate about the safety of future AI systems. Instead, it emphasizes the importance of discussing and regulating potential risks in advance, even before being aware of all relevant design details. Argument (9) fails because it neglects the importance of foresight and risk assessment in the development of new technologies, thus exposing itself as unsound.

A side remark about the relation between violence and artificial intelligence ought to be made. We need to stress that an event or action *x* can be considered an instance of violence if and only if there are both conscious and intentional components to it. Thus, we should not qualify as violence any harm brought about by a wild animal, but also, any harm brought about by inanimate objects, like a cliff falling in the mountains and hitting a tourist, or a sudden wave that flips over a surfer in the ocean. In order for violence to take place, be it physical, emotional or psychological, the incident must be recognized as something that was brought about consciously, i.e. something that was desired and intended. Of course, so far none of these conditions have been met at the current stage of AI development, as existing AI models possess neither consciousness nor desires or intentions. Even from a semantic point of view, they have been deprived of such capabilities, as they do not – at least in my understanding – appear to have any human-like understanding of terms or situations. Thus, we can only begin to research hypothetical AI-cause violence once we have been presented with an AI model that understands, intends, desires and, most of all, has a consciousness. Until then, the harm caused by an AI would not be any different from any harm brought about by a falling cliff.

These arguments present overly-optimistic viewpoints on AI safety and reliability, and they fail to fully consider the complexities and unpredictability involved in the development and deployment of advanced AI systems.

An additional, less obvious threat with a potential existential reach is *the erosion of truth as a main instrument of societal adaptation and evolution*. Although we may tend to ignore the phenomenon of truth as a practical instrument in our daily lives, in an increasingly information-based society truth is not an exotic luxury preserved for philosophers, but a *vital instrument of* *social and potentially even individual survival*. We have cultivated our concept of truth at the cost of much suffering throughout the history of humanity, and now we are confronted by a new reality that would have been difficult to find even in the most original science fiction novel. We are facing *the prospect of never being able to distinguish a truth from a falsity and then confidently act upon that distinction*. The global information space, pulsing with the beat of the Internet, is increasingly governed by enormous data flows and AI is the supreme technology that can work with them as it deems (computes) right. Given the enormity of information and the ever more sophisticated models of data (such as deep fakes), but also given propaganda and fake science, the better part of humanity might eventually find that *not believing virtually anything at all* provides a better chance of social and individual adaptation. AI is a technology that can both develop perfect fake realities and expose them. We are at the beginning of an *existential epistemic transformation* where we would need to find novel ways of reaching truth and proving it. The artificial erosion of truth has the potential to disintegrate the functional fabric of human society as we know it. Much like prey that fails to recognize a cleverly disguised predator, humanity might fall into the trap of an incomprehensible fake version of reality. Unexpectedly, it is on epistemology to begin helping humanity with this challenge, which does have considerable existential threat potential.

In conclusion, I would like to suggest that, given the lack of sufficiently good grounds (logical, technological, societal, economic and political) to expect that future AI models would not pose an existential threat to humanity, we can be confident that the containment of the threat would not come directly from "guaranteed safe AI architectures" (the development of which we have virtually no reason to expect at all). And even if we, somehow, miraculously, manage to develop provably and objectively safe models, the current global political map shows that we have virtually *no rational reasons to expect that non-democratic states or other malevolent actors would refrain from purposefully developing harmful models*. Thus we are left with a single instrument that may contain harmful AI systems. Paradoxically or not, this could only be *a superior system of artificial intelligence*, superior in its intellectual capacities, but also in its *generality*, which is what opens the space of its functionality. In this functionality we would need to embed our hopes and tasks to contain all possible effects of the operation of (less general) AI systems that may harm humanity. Any direction of a generally harmful model could only be prevented, contained or – hopefully – eliminated by a model with superior generality.

And *this model needs to come first and before any potentially harmful AI model* of lesser generality. In other words, my view is that humanity has no better chance to defend itself against potential existential AI threats than ensuring that a democratic state or a union, such as the one gathered in Bletchley (China being a controversial invitee), either develops the first AGI model or at least exercises effective control over the corporation that achieves that first; a task now burdened not only with technological challenges, but also with the weight of humanity's hopes, as dramatic as that might sound. The possibility that this AGI turns harmful itself or implements actions deemed harmful by human society clearly remains.

NOTES

[1] https://www.gov.uk/government/topical-events/ai-safety-summit-2023.

[2]https://www.gov.uk/government/publications/ai-safety-summit-2023-the-bletchley-declaration/the-bletchley-declaration-by-countries-attending-the-ai-safety-summit-1-2-november-2023.

[3] Interview with Yoshua Bengio by Susan Dagostino for WIRED magazine, published at https://www.wired.com/story/ai-godfather-yoshua-bengio-humanity-defense/?fbclid=IwAR1ZBVIyhLYaluFLxb8qiNnrguuiso8Tn_jvbk-z23OEFcoMj9hbCFPzY50. His theses on existential threats by AI are summarised here: "But over the winter, it dawned on me that the dual use nature and the potential for loss of control were very serious. It could happen much earlier than I had projected. I couldn't continue ignoring this. I had to change what I was doing. Also, I had to speak about it because very few people—even in the AI community—took these questions seriously until the past few months. In particular, I and people like Geoffrey Hinton and a few others were listened to more than those folks who were talking about this and had realized the importance of those risks much earlier than we did … In the future, we'll need a humanity defense organization. We have defense organizations within each country. We'll need to organize internationally a way to protect ourselves against events that could otherwise destroy us. It's a longer-term view, and it would take a lot of time to have multiple countries agree on the right investments. But right now, all the investment is happening in the private sector. There's nothing that's going on with a public-good objective that could defend humanity".

[4] For a discussion on conceivability and possibility as an instrument of logical reasoning, see the comprehensive volume by Gendler and Hawthorne [2002] "Conceivability and Possibility", Clarendon Press.

[5]https://www.facebook.com/yann.lecun/posts/pfbid02We6SXvcqYkk34BETyTQwS1CFLYT7JmJ1gHg4YiFBYaW9Fppa3yMAgzfaov7zvgzWl

[6] https://www.economist.com/special-report/2017/05/04/how-the-2007-08-crisis-unfolded

BIBLIOGRAPHY

Gendler and Hawthorne. (2002). *Conceivability and Possibility*. Clarendon Press.

Lochbaum, D., Lyman and Stranahan. (2014). *Fukushima: The Story of a Nuclear Disaster*. The New Press.

Gates, D. (2019). Aerospace reporter Dominic Gates talks about the 'huge issues' in the Boeing 737 MAX's flawed design. – In: *The Seattle Times*, May 8, 2019, Updated May 9. https://www.seattletimes.com/business/boeing-aerospace/dominic-gates-talks-about-the-huge-issues-in-737-maxs-flawed-design/

Robison, P. (2021). *Flying Blind: The 737 MAX Tragedy and the Fall of Boeing*. New York, Doubleday.

Wise, J. (2019). Where did Boeing go wrong? – In: *Slate*, March 11. https://slate.com/technology/2019/03/ethiopian-air-crash-where-did-boeing-go-wrong-with-the-737-max.htm

Wu, H. (2014). Heartbleed OpenSSL Vulnerability: a Forensic Case Study at Medical School., NJMS Advancing Research IT (Report). Rutgers University.

Durumeric, Z., F. Li, et al. (2014). The Matter of Heartbleed. – In: *Proceedings of the 2014 Conference on Internet Measurement Conference*. New York, NY, USA: ACM. pp. 475–488.

Glanz, J., et al. (2023). Why the Evidence Suggests Russia Blew Up the Kakhovka Dam. – In: *The New York Times*, https://www.nytimes.com/interactive/2023/06/16/world/europe/ukraine-kakhovka-dam-collapse.html

Moloney, M. & E. McGarvey. (2023). Ukraine war: Russian air strikes cut power at Zaporizhzhia nuclear plant. – In: *BBC News*, https://www.bbc.com/news/world-europe-64897888

Parker, E. (2023). When a Quantum Computer Is Able to Break Our Encryption, It Won't Be a Secret. – In: Lawfare, https://www.lawfaremedia.org/article/when-a-quantum-computer-is-able-to-break-our-encryption-it-won-t-be-a-secret

Lapierre, D. and J. Moro. (2002). *Five Past Midnight in Bhopal: The Epic Story of the World's Deadliest Industrial Disaster.* Grand Central Publishing.

Ceruzzi, P. (2003). *A History of Modern Computing.* MIT Press, Second Edition.

Doudna, J. and S. Sternberg. (2017). *A Crack in Creation: Gene Editing and the Unthinkable Power to Control Evolution.* Mariner Books.

Sosa, E. (2018). *Epistemology* (Princeton Foundations of Contemporary Philosophy, 18) Princeton University Press.

Huemer, M. (Ed.). (2002). *Epistemology: Contemporary Readings* (Routledge Contemporary Readings in Philosophy). Routledge.