Momentum of the Future

I

In January 1942, the United States of America was diving into the most expensive and uncertain technological research ever undertaken. Refugee physicists from Europe and the finest American minds had an open checkbook and a deadline: six months to determine the possibility of a nuclear chain reaction, another six months to produce it, and then a year to apply it. The purpose of the deadlines was clear: succeed before the Germans did. No one knew how close Germany was, but it was assumed that the U.S. was behind.

Knowing that the project succeeded, it seems money well spent. But the first stage of the research reminds us that, at first, no one knew that it was possible.

In this headlong race into the theoretical, there is a certain sense of inevitability: someone *will* develop the bomb; what matters is who does it first. Promising technology seems to take on a life of its own, both in the past and today. The momentum of this great project would sweep along the reluctant scientists who proposed it in a wave of fear and ambition.

So the Manhattan Project began based on a kind of Pascal's Wager. It seemed wiser to work and sacrifice, imagining punishment was imminent, than to hope for the best. Not because the evidence pointed towards certain doom, but because the consequences of ignoring the problem were potentially dire.

Before the bomb was proven possible, the world could have perhaps remained without it. But a successful demonstration of the theory would make certain the consequences. If the chain reaction was possible, in a sense, the bomb *had* to be built. Alice Kimball Smith, who was acquainted with those working on the project, recalls that some "found comfort in the hope that some insuperable obstacle might demonstrate the impossibility of an atomic weapon" (Boorstin 587). These hopes were dashed when the project reached its first goal.

On December 2, 1942, Einstein's theory was confirmed in a repurposed squash court under the University of Chicago. The energy he claimed was hidden in all matter turned out to be accessible. A year into the project, the wheels were very much in motion and began turning up deep concerns even as the physicists inched closer to their goal. More and more of the leading scientists were beginning to openly oppose using the bomb. But the prospect of a military trump card was too attractive, and the very same individuals who had called on the President to begin this project were now powerless to stop its movement.

II

Long after the Second World War, the promise of revolutionary new technology is still pulling all of us along inevitably towards the future. In the twenty-first century, it is not atomic fission but artificial intelligence (AI) that captures and haunts our world's collective imagination.

AI already exists in what is called a *weak* form. In short, there are programs that can learn and develop along certain design parameters. The great achievements of weak AI are programs designed to defeat even the best players at highly complex and nuanced games like Go and no-limit poker. Others can analyze and act on investment opportunities in a fraction of a second, or recognize faces and objects in photographs and videos. The machines have learned, and they have developed novel practices beyond what they were originally programmed to do, and in this sense they are intelligent.

Already in weak AI, computers are making decisions their creators do not understand. Google created an AI called AlphaGo Zero that taught itself how to play Go in three days by virtually facing off against an earlier version of itself countless times. This hidden learning process has led to an unconventional and highly effective playing style: no one on earth can tell you why AlphaGo plays exactly the way it does. But it always wins.

The Holy Grail however, remains elusive: *strong AI*.

Strong AI would be able to branch out from a specialized program, learning and acting in a flexible, multipurpose way, as humans do. Concepts from one area could be abstracted and applied to others, unprompted. If AlphaGo were a strong AI, it might be able to make equally perfect decisions in chess, or traffic, or combat. But much like the atom bomb in its early phases, there are doubts about the feasibility of this technology. It is difficult to say if logical, circuit-based processes can ever achieve original thought. After all, even the most interesting AI is still, in a sense, doing what it was programmed to do.

But despite these doubts, a similar momentum to the Manhattan Project's has begun. The wheels are turning and the money is flowing. "AI systems will eventually have been developed for all human intellectual domains," writes Andrew Sheehy of Forbes.com, in his article entitled "Superintelligence is not just possible, but inevitable." Governments seem to agree. Yearly budgets for US research into artificial intelligence have been in the hundreds of millions of dollars for the past decade, and have risen sharply in the last five years.

The race is on, and once again it is not a goal that is tirelessly pursued, but a force that seems to push unstoppably forward.

III

In 1945, the Manhattan Project achieved its goal. Whether or not the world is better off for the existence of atomic weaponry is, of course, another question.

The fear of a nuclear Nazi Germany had motivated many of the founders of the project. But once it was evident that the U.S. would be the first to have the bomb, the overwhelming attitude was that it should not be used. When it seemed a bomb was going to be produced, horror swept over those who understood it best. Many of the project's pioneers worked to prevent its deployment, trying to convince their coworkers and those in the administration to seek alternatives. A poll showed only fifteen percent of the atomic scientists in the Chicago laboratory supported full military use against Japan.

During this crucial time, seven of the most important Manhattan Project scientists wrote a letter to the Roosevelt administration, proposing that the bomb only be demonstrated as a peacemaking measure, not to kill. They had hoped that a world shown the potential of warfare including such weapons would be frightened and cowed into peace. And to drop the bomb on the enemy, they said, would have consequences that outweighed the benefits of a swift end to the war. The end was not as swift as it could have been—the bomb was dropped not once but twice, and war has continued between states, involving those with nuclear capability, since the end of WWII.

Fears of nuclear armageddon circulated widely, and sometimes intensely, for fifty years after the first successful test of the bomb. As the United States and USSR built more and stronger nuclear arms, other nations eventually joined in and developed their own. Today, denuclearization is a priority for the international community, but the inertia of the bomb remains. Once it got moving, it was too heavy to stop. Now that it's here, it is too heavy to move.

We will never know what effect nuclear restraint would have had on the rest of the Twentieth Century. It does seem that any hopes about the bomb's existence preventing or quickly ending conflict were misplaced. A reporter called the first test detonation "fascinating and terrifying, uplifting and crushing, ominous and devastating, full of great promise and great foreboding" (Boorstin 586). Unlike the reporters and the government, the men and women who understood best what they had created were the most appalled at it. Physicist J. Robert Oppenheimer famously recalled watching the Trinity test and thinking of a verse from the Bhagavad-Gita: "Now I am become Death, the destroyer of worlds."

IV

Today, a similar story is unfolding. In January 2015, Stephen Hawking, Elon Musk, and dozens of artificial intelligence experts released several open letters detailing their fear of unregulated and uncontrollable artificial intelligence development. Many high-level employees at some of the top tech companies signed. Yet Google, Apple, Facebook, Microsoft, and IBM are all racing forward with secretive AI projects.

As with the Manhattan Project, the 2015 letter came from experts who wanted to temper the optimism of research with self-control. They emphasize the positive applications for AI in the medical field, and towards ending poverty and hunger worldwide. As our civilization's triumphs are a result of human intelligence, superhuman intelligence promises even greater works. But our civilization has not only produced good.

The cautionary parts of the letter are strikingly similar to the letter from the atomic scientists: if this thing must be made, above all else restraint is necessary in its development and application. The fear is not necessarily of extreme dystopian, *Terminator*-style robot revolt; there are more subtle ways in which AI could irreversibly change our world.

The priority of one AI project, called Project Maven, was object recognition in video. Google was developing Project Maven for some time along with the Pentagon. It could "watch" countless hours of drone camera footage, in hopes of developing a system that could one day fly alone, only reporting when something of interest appeared. Strong AI could come into play here because the category target is flexible, intangible. There is no single criterion on which the program could decide to categorize structures. Appearance, behavior, and context all come into play when a human is assessing something unknown, all of which are nuanced and difficult to concretely identify. For now, humans are still far better at these kinds of tasks.

Teaching a computer to differentiate and recognize things in general, however, is an application to which AI is well-suited. China is currently developing city-wide security camera networks, which can track individuals over time, and it is far from perfect, but the process has begun—and if the bomb story is a good analog, it will only accelerate.

In other laboratories, face recognition is a priority as well. Voice recognition, as we all have in our smartphones, is often powered by AI which learns the idiosyncrasies of human speech and adapts to better understand and imitate it. Both of these are getting better every day, and will soon be seamless.

Images from NVIDIA's GAN Image Generator AI

AI not only recognizes the faces in the above photo, it created them. None of the images above were taken with a camera, or correspond to any single person. They are complete fabrications, created using a generative adversarial network, or GAN. In the creator company NVIDIA's GAN, two AIs go head to head, much like AlphaGo's learning process. In the GAN, the first AI creates a fake image, the other trying to determine where mistakes were made. Pixel by pixel, the two computers improve the image until the adversary machine can no longer find any mistakes. And, as you can see, neither can the human eye.

Consider the consequences of such perfect fakes for journalism. Consider also the application of the same technique to audio and video data. Soon, perfectly convincing videos known as *deepfake* videos will be produced that will call into question what has been for decades the most trustworthy source of information. .

This example emphasizes a second important distinction between these parallel stories. The bomb was a spectacle; it could not be ignored. Since before its first detonation, there were calls to cease development. But AI has no such obvious danger. It is thought to hold great potential to improve the world, and this makes the momentum it carries much harder to slow.

Maybe, as in the case of the bomb, there is no stopping the forward march of AI development, despite the risks. However, the letter from the AI experts can give us hope. There is always the possibility that we will stop before it is too late, or that the consequences are all imaginary and overly pessimistic.

Both these stories are of a quest for the impossible. It was long believed that the atom was impregnable, but once it seemed possible, the promise of the idea took on such force that no one

could stop its realization. In the same way, a machine that could outthink a human began as the dream of science fiction writers, but it is quickly becoming fact.

Just as in 1942, the momentum is here again, born from gigantic forces beyond the individual: vast amounts of money and effort, an international pressure to succeed first and most dramatically, and the allure of accomplishing the impossible. The AI explosion will be silent: the first AI able to design better AI could be mankind's last invention. Perhaps then we will all feel something of the powerlessness and ambivalence of those men and women watching the future arrive in the New Mexico desert. ¤

By Daniel Affsprung

This article draws on details and concepts from Daniel Boorstin's Pulitzer Prize-winning book The Americans, which introduced me to the idea of the momentum of great projects.

Works Cited

"Generating Photorealistic Images of Fake Celebrities with Artificial Intelligence." *NVIDIA Developer News Center*, 30 Oct. 2017, https://news.developer.nvidia.com/generating-photorealistic-fake-celebrities-with-artificial-intelligence/. Accessed 1 October, 2018.

Boorstin, Daniel J. *The Americans: The Democratic Experience*. New York; Vintage Books, 1974.

Good, Irving John. "Speculations Concerning the First Ultraintelligent Machine" *Advances in Computers*, ed. Franz L. Alt and Morris Rubinoff. vol. 6, 1966, pp. 31–88. *ScienceDirect*

Mozur, Paul. "Inside China's Dystopian Dreams: A.I., Shame and Lots of Cameras." *The New York Times*, 15 Oct. 2018. *NYTimes.com.* https://www.nytimes.com/2018/07/08/business/china-surveillance-technology.html Accessed 1 October, 2018.

Sheehy, Andrew. "Superintelligence Is Not Just Possible, But Inevitable." *Forbes.com*. 22 June, 2016 https://www.forbes.com/sites/andrewsheehy/2016/06/22/superintelligence-is-not-just-possible-but-inevitable/#af47f39225c0. Accessed 7 May 2019.