freedom lab

THEME 01

Today's class of uncontrollable technology

AI   SYNTHETIC BIOLOGY   TECHNOLOGY

How can we understand the rising complexity and uncontrollability of technologies? Here, we explore and compare the cases of synthetic biology and artificial intelligence, two disruptive technologies that produce outcomes that are not fully controllable or predictable and whose impact on society will only grow in the following decades. These disruptive technologies will furthermore challenge basic aspects of human self-understanding, including our notion of autonomy.

Our observations

- "The world is getting more and more complex." Although it is rather a non-starter, the expression is widely used in different contexts today. As we are trying to get a grip on everyday changes that we witness, the expression needs more specification. In our research, we explore uncertainties in the geopolitical, socio-cultural and technological realm and how they are influencing and reinforcing each other. For instance, how the internet is influencing global power dynamics.

- When looking at the rising complexity in new technologies in particular, the challenges and fears concerning their complexity are often related to the feeling of losing control over our own technological inventions and the consequences this would have for our society. Science fiction often tells us stories of technological innovations getting out of hand (*Frankenstein*), computers that are controlling us (*The Matrix*), or human-made viruses that threaten the entire world population (*The Walking Dead*). The rising complexity of technology, and more specifically, the uncontrollability and unpredictability of today's technology is explored here by introducing philosopher Jan Schmidt's concept of "late-modern technology". Instead of trying to explain the uncontrollability and unpredictability of individual technologies, the concept helps us to see them in a wider class of technologies showing the same characteristics, such as the seemingly different innovations in AI and in synthetic biology (the scientific domain that involves redesigning organisms for specific uses by engineering them to have new abilities, such as cell factories).

- According to the classic-modern view of technology, uncontrollable and unpredictable outcomes of technology are undesirable. Man gains control over his environment by making use of technology. Constructability and controllability, including a clear input-output relation, are key in this regard and technology was traditionally equated with and defined by stability. Think of cars that are made in a production line.

- By contrast, late-modern technologies are a class of technology in which this idea of stability is abandoned. Late-modern technologies confront us with our ideas about autonomy and control over our own inventions. Autonomy can be regarded as the most celebrated outcome of the Enlightenment and makes up the foundation of moral philosophy that is still dominant in today's moral theory.

- An entire class of "autonomous" technologies is in the making or has already been deployed, from autonomous vehicles to autonomous weapons. These increasingly guide our behavior at a time when human human autonomy is challenged by the distraction and information overload in our digital age. As we described before, technological decisionism confronts us with the fact that our decisions will increasingly be supported, if not steered, by artificial intelligence. As non-living or non-human things are increasingly actively participating in and shaping our environment, we cannot ascribe autonomy to humans only anymore, as is acknowledged in the theory of new materialism.

Connecting the dots

When thinking or talking about technology, we often use words that describe the mechanical characteristics of technology. Not seldom is technology in books or movies depicted as machines or robots. Indeed, in our language this machine image is also widely present. The machine metonym is closely connected to an ontological assumption: a machine is assembled by humans, built up from parts to a whole, it has controllable and predictable characteristics. This is a classic-modern view of technology.

However, when turning to present cases of technological advances such as synthetic biology, this becomes problematic. Even if the goal was to create synthetic organisms as controllable and predictable entities, a living organism, whether "natural" or a product of human intervention, by definition evolves and interacts with other organisms and the environment in multiple ways. These characteristics do not fit the part-whole view and make organisms less controllable and predictable than machines. This complex interaction of technology with other technological or living systems creates complexity. In addition, organisms reproduce and grow, something that the machine metonym does not imply either. As a result, using machine metonyms might blind us from the implications of creating new life forms, such as synthetic organisms, as happens in synthetic biology. In the case of Artificial Intelligence, similar problems arise when using the machine metonym. AI, and more specifically machine learning, is confronting us with a case of technology that shows more autonomy than the machine metonym suggests. So, what are these cases of technological innovation showing us? How are they different from technologies that better suit our more mechanistic and predictable view of technology?

Already in 1985, philosopher Hans Jonas envisioned a historically new technoscientific era when technologies would show different characteristics than the previous class of technologies, such as a certain degree of autonomy and limited predictability. In current philosophy of technology debates, scholars differentiate between modern technology, or classic-modern technology, and late-modern technology. We can understand synthetic biology and AI as cases of the latter. Late-modern technologies differ from classic-modern technologies in two fundamental ways.

First, they show self-organization, autonomous behavior or agency properties. In the case of AI, an autonomous system goes beyond the behavior programmed in the initial algorithm, as it can learn by itself from data and environment, its behavior transgresses the initial objectives and conditions set by its creators (i.e. human engineers, computer scientists) and therefore gain a lower degree of predictability. Similarly, an organism created by means of synthetic biology, starts to interact with and "learn" from its environment in a way that makes it hard to predict its behavior. In both cases, the technology autonomously interacts with an open-ended and uncertain context, the real-world environment, and is thus less predictable than technological systems that merely react to human input and are otherwise passive. In that sense, technologies are sometimes regarded as "black boxes", as insight into their input and output processes is difficult to acquire.

Second, in the case of late-modern technology, the technology no longer appears in its modern way, rather, technological traces are disappearing. Culturally established borders and modern dichotomies such as "natural" vs "artificial" are becoming blurred. For instance, a synthetic cell has an artificial pathway, but shows no traces of technology: it cannot easily be distinguished from "natural" cells. Similarly, the thinking of AI can sometimes hardly be separated from human thinking or decision-making. In 2018, Google gave a demo of its voice assistant calling a hairdresser to make an appointment and shocked the audience when the hairdresser did not notice that she was not talking to a human. Indeed, this novel kind of technology appears human or natural to us. This is what is called the naturalization of technology. However, moral debates about these sorts of technology, such as the debate about acceptance of GMOs, are often still framed in modern terms, with a strict distinction between us humans, the technology we use, and the natural environment. Late-modern technology is thus difficult to predict and control, difficult to separate from the context and environment of its application, it can be said to "have a life of its own". The fact that human beings are surrounding themselves with more and more technologies that are less controllable and show autonomous features, inevitably gives us the sense that we are facing greater technological complexity, losing control over our technology and that our notion of autonomy, which we regard as a fundamental human trait, is being challenged. Late-modern technologies such as AI could even undermine our autonomy, as its ubiquitous deployment could steer us implicitly and explicitly in our behavior. As is often the case with new technological developments, late-modern technologies force us to define and reframe values and views that used to be implicit and unchallenged.

Implications

- **Seeing advances in AI and synthetic biology in a wider class of technologies is also helpful in discussing the challenges for both. For instance, in both areas, a centralization of knowledge can lead to negative consequences for society, e.g. that not everyone can benefit from or even be involved in their creation. In AI and synthetic biology, there are efforts to organize knowledge and IP in open-source governance structures, such as the OpenAI initiative and open source seed initiatives for (GMO) seeds.**
- **The rise of artificial intelligence or technological decisionism might teach us something about our human thinking. Similarly, synthetically created organisms might tell us something about living organisms. In a sense, late-modern technology can give us insights into fundamental concepts.**