THEME 01

# The age of genetic information

BIOLOGY TECHNOLOGY GENETICS

**Great advances both in AI and in genetics have opened up a wide scope of hybrid fields and applications, such as DNA data storage and the prospect of precision medicine. The age of information is no longer merely about digital information, but increasingly about two types of information: digital data and genetic data. As genetic data are increasingly valuable in the light of new technologies and applications, questions over privacy and ownership are arising.**

## Our observations

- Complex systems like our body and the biological sphere around us are full of data. Our own genome alone amounts to 100GB. Genetic data entails the entire history of life and is a very efficient way of storing data, while digital data storage has only proved to have a limited lifespan. That is why DNA data storage is receiving more attention and new technologies use the four base nucleotides of DNA as digital bits to encode data instead of binary zeroes and ones. Project Lunar Library aims to become the biggest collection of data written in synthetic DNA. Synthetic DNA can serve as very efficient data storage, the founder of the project says: “a tiny liquid droplet (of synthetic DNA) could contain Amazon’s entire data center”.
- With their users’ consent, companies such as 23andMe, Ancestry.com and MyHeritage are allowed to sell users’ genetic information. Without genetic privacy protections, individuals do not have ownership over their data anymore. As a counteraction, companies are building platforms based on blockchain technology to let individuals control and even profit from their genetic information. A Russian non-profit, Zenome.io, has launched a decentralized “genomic internet”, aiming to deal with the challenges of storage, privacy, and sharing data.
- As the value of genetic information becomes more apparent in improving healthcare with precision medicine, among other things, countries are trying to complete national genome projects in order to get access to the insights that come from sequencing genetic data. The list of countries with such initiatives include the U.K., Japan, China, Australia, Saudi Arabia, the U.S., Estonia, France, the United Arab Emirates, and Turkey.
- Recently, the World Economic Forum published the top 10 emerging technologies of 2018. What marks most of these technologies is that they have resulted from advances in AI and biotechnology: with personalized medicine, it will be possible to analyze genetic data, AI-led molecular design will help identify and develop new medicines, implanted living cells can act like in-body drug factories, genetic engineering technology can permanently change the traits of a population, which is called “gene drive”, plasmonic materials are novel nanomaterials that will allow new applications in biology and medicine, lab-grown meat means that growing biological tissue in the lab is possible, and finally, electroceuticals are small devices in the body that can stimulate nerves.
- In developed countries, farms producing greens will increasingly change into automated factories controlled by AI. In the first automated farm, Iron Ox, cloud-based AI monitors the whole farm and directs the robots. The next step is for algorithms to teach robots to identify crop diseases.

>> see next page

# Connecting the dots

Over the past years, progress has been made in two fields – genetics and artificial intelligence – that now are increasingly intersecting and challenging the ways we look at basic concepts such as information and ownership. In the first field, the completion of the Human Genome Project (HGP) in 2003 was a milestone. This international research aimed at completely mapping and understanding all the genes of human beings. It took over a decade to reach this goal and cost more than a billion dollars. Progress in this field has accelerated to such a level that today, the HGP can be carried out in days and for a fraction of the costs. Gene sequencing is becoming ever-cheaper and genetic information is thus increasingly accessible. Genetic engineering increasingly allows us to reprogram basic building blocks of life, with the latest gene editing tools like CRISPR, for example, to make precise changes in DNA. Furthermore, we are heading towards the era of personalized medicine, in which treatments are tailored to an individual's genetic composition. In the second field, the domain of AI, we are witnessing the progress made with machine learning. While current machine learning systems are narrow AIs, neural networks show a (limited) degree of generality. As we wrote earlier, innovation at the intersection of the biological and the artificial domain is accelerating. These hybrid fields link biological data with computer science, as in synthetic biology and bioinformatics, and raise questions about information and ownership.
In the digital age, information is predominantly looked at as the digital binary code of zeroes and ones. Code and algorithms are man-made and thus subject to ownership. While we give away our personal data with relative ease by leaving a digital footprint on the internet, the algorithms based on these data are controlled by a concentrated number of parties. However, the merge of biology and technology further begs the question of ownership. Traditionally, patents have applied solely to inventions, granted as rewards for human innovation and discovery. Genetic code used to be exempt from such laws. In the 80s, this changed, as the American scientist Ananda Mohan Chakrabarty applied to patent a bacterium that he had modified genetically so that it could consume oil. Chakrabarty's application was at first rejected because his "innovation" concerned a product of nature. Only later, the Supreme Court ruled that it was Chakrabarty's inventiveness that added value to the bacterium and granted him the first-ever patent granted to a life-form. Ever since, genes considered to have been "isolated from their natural state and purified" have been eligible for patent protection, resulting in the entire human genome being nearly fully covered by patents. Doing research on a patented gene means paying for a license. The rise of genetics and the broadening field of applications, such as DNA data storage, genetic data for precision medicine, and AI-controlled processes to produce biological products, urge us as individuals but also as countries and regions to think about ownership and access to genetic information. While the U.S. and China have labelled biotech a priority for national security, investing heavily in research and development, Europe has adopted a more careful approach to biotechnology applications and is lagging behind in the biotechnology race. Meanwhile, decentralized alternatives to patenting life are on the rise. As the WEF top ten emerging technologies shows, the increased capacity to combine the developments in the field of digital information and genetic information leads to new possibilities in many domains, and this will only accelerate in the future.

# Implications

- **In the future, AI will be increasingly applied to discover patterns in genetic data. For instance, genetic puzzles that are now modeled as games to be solved, such as Fold It, will increasingly be solvable by AI. This could lead us to not only learn more about diseases, but possibly about behavior and mental processes as well.**
- **As we have written before, the combination of digital code and the biological code heralds a symbiosis between organisms, our bodies, our products and even our buildings. Rather than creating inventions assembled from parts, MIT professor and biodesigner Neri Oxman says that in the coming "biological age, designers are empowered to dream up new, dynamic possibilities, where products and structures can grow, heal and adapt".**

>> see next page