

DIGITAL DRIVEN TECHNOSCIENCES: EPISTEMOLOGICAL AND ETHICAL QUESTIONING

THIERRY MAGNIN

*Université catholique de Lille,
ETHICS EA 7446, Chaire STFEC, France
thierry.magnin@univ-catholille.fr*

PIERRE GIORGINI

*Université catholique de Lille,
ETHICS EA 7446, Chaire STFEC, France
pierre.giorgini@univ-catholille.fr*

ABSTRACT

Technosciences such as biotechnologies aim to analyse, modify and reconstruct livings and complex biochemical components, using top-down or bottom-up approaches to simplify these complex systems. Because of the large number of involved parameters, these technosciences need more and more big data techniques (Artificial Intelligence with machine and deep learning) to find efficient correlations to predict the evolution of systems in their environments. This prediction is especially used to reconstruct livings or artificial pieces of livings in biotechnology and to give medical diagnostics in medicine. Two examples corresponding to these two different cases are analysed in this paper.

The aim is to highlight interests and limitations of such digital approach in terms of human free will and responsibility, in relation to provisional scientific truth. A difference is established between classical digital-assisted technoscience and digital-driven technoscience through an epistemological analysis of this new way to practice science. This allows to precise values which are favoured in a digital driven science, what we call an « ethics of knowledge ». Thus “capabilities to predict for making”, based on efficient correlations, look more important than “capabilities to know for making » and determine complex biological mechanisms or causes of evolutions. Consequences about technoscientific mentality are drawn and practical ethical questioning is consequently reassessed.

KEYWORDS

Technoscience, artificial intelligence, digital and big data techniques, epistemology, ethics.

1. INTRODUCTION

Artificial Intelligence (AI) corresponds to a set of digital techniques which allows the numerical treatment of a large number of data that the scientist himself cannot take into account alone. According to European definition (<https://www.europarl.europa.eu/doceo/document/TA-9-2020-0275>), AI corresponds to “a system that is either software-based or embedded in hardware devices, and exhibits intelligent behavior, including collecting and processing data, analyzing and interpreting its environment, and taking action, with a certain degree of autonomy, to achieve specific objectives”. Thus, big data analysis is used for complex biological systems, for instance in biotechnology and for biomedical research and health care.

One of the most important aspects of digital big data treatment is “a capacity to search, aggregate, and cross-reference” large and differently structured datasets (Boyd and Crawford, 2012). That is particularly interesting to increase the technical performance in the elaboration of new biochemical components or new livings and for the prediction of diseases related to complex biological systems.

Nevertheless, even if such digital approaches are already very efficient, one must be very careful to precise their limitations, showing their non-neutrality and risks. In a recent paper (Guillermin and Magnin, 2017) we criticized a kind of fundamentalist view conceiving big-data-based processes of knowledge production as a neutral scientific method, free of human subjectivity, allowing the highest forms of objectivity and rationality. Indeed, according to data fundamentalism, “massive data sets and predictive analytics always reflect objective truth,” and “unless you have really large datasets then you're not getting close to objectivity and truth” (Crawford, 2013a; Crawford, 2013b).

The aim of this new paper is to deepen the critical analysis of the digital-driven technosciences through two different examples. The first one is related to the prediction of gut microbiota evolution in human health, deepening the analysis of our previous paper already referred. The second is an analysis of the incidence of AI to predict the repair of a genomic sequence after a voluntary rupture due to CRISPR-Cas 9.

In both cases, we detail the different steps of the digital approach to be able to precise efficiency and limitations. We focus particularly on the different steps and degree of validation by biologists of the successive ‘patterns’ produced by deep learning processes. The objective is also to evaluate the mode of validation by biologists and the correlative free will that can be applied in practical ethics.

We will then analyse such results regarding the epistemology of technoscientific mode of knowledge. This analysis focuses on fundamental values which are involved in digital-driven technoscience in regard to digital-assisted science (i.e. what we call « ethics of knowledge»). Finally, consequences in practical ethics are emphasized.

2. PREDICTION OF GUT MICROBIOTA EVOLUTIONS

Gut microbiota composition is known to play a major role in the development of diseases such as cardio-metabolic disorders or diseases (CMD, ranging from obesity and diabetes to atherothrombotic diseases and heart failure). The main factors of evolution are the individual's genome, the environmental parameters (such as pollution or in utero environment) and lifestyle specificities (e.g. physical or dietary habits). Nevertheless, the biological mechanisms are not known today because of the complexity of such ecosystems, which are often named as 'symbiosis' (see the works of the Metagenopolis consortium (<http://www.mgps.eu>), INRAE). It is why only big data processes can operate to show possible evolutions of such complex systems.

Dietary habits are often tested because they are known to play a major role on the composition and diversity of gut microbiota and its microbiome (or metagenome), in relation with patient's metabolic functioning susceptible to cause CMD's symptoms (Dao et al., 2016).

Thus, the Metagenopolis consortium developed big data approaches to find efficient correlations related to gut microbiota evolution from broad systems of data related to both biological and lifestyle parameters. It involves different types of CMD patients at different stages, with different predisposing factors. Biochemical, clinical and anthropometrical data (weight, glucose tolerance status, or adiposity) are integrated with information about medical history and lifestyle (dietary habits and physical activity, but also socioeconomic status or educational level, and psychological stress or perceived quality of life). These various data are complemented with the collection of biosamples allowing high-throughput molecular sequencing to establish profiles of subjects involved in the study. Thereby, a metagenomic characterisation of the gut microbiota composition and diversity can be coupled with metabolomic analysis of molecules involved in metabolic processes, and with transcriptomic monitoring of host's genome expression.

Statistical and machine learning approaches are developed for patient stratification and correlation detection. Big data techniques permit the extraction of insightful correlations and "patterns" in the following sequence:

1. The computer is given a set of data and a set of rules: for example, the abundance of bacterial species that make up the microbiota (as a set of data) and the designation of a patient suffering from disease or a not healthy patient (as a rule).
2. With this dataset and these human-injected rules, the digital machine, which searches for correlations, produces a 'pattern' (a statistical model) associated with the injected rules. For example, a patient with cirrhosis shows a decrease in the diversity of his gut microbiota in terms of species unlike a healthy person. And the machine concludes that it is a specific

pattern of a patient and proposes this pattern. The expert biologist then checks the validity of the pattern in terms of biological and medical hypotheses on the underlying mechanisms. And he checks the assumptions made for these mechanisms when experiments are possible. In particular, he wonders whether the pattern obtained for one data set will be true for another population and another data set. Note that the assumptions made by the biologist cannot always give realistic experiences; however, he will have to validate them as "plausible". This (or these) validated patterns will thus make possible distinctions between the types of subjects (patients or healthy people). This is called the learning phase.

3. The machine is provided with a new set of data (e.g. species abundance) to distinguish patients from healthy people using patterns defined in training; a new validation from the expert biologists is requested.
4. If this step is validated, these patterns are proposed as a diagnostic algorithm.

In this process, we see that digital processing remains supervised and that the biologist validates the steps. The scientist seems to "keep control", even if limits appear as often in science: certain hypotheses are considered as plausible but not validated by experiments. This is the way by which technoscientists operate within the current framework of Metagenopolis with a very prudential posture.

Scientists can also rely on expert systems that accumulate past experiences. These systems can be entered into the machine to which part of the validation can also be entrusted. Thus, the diagnostic results here from a form of collaboration between the scientist and the digital machine: we can speak of "digital-assisted technoscience".

We note that the degree of validation of the patterns by the scientist depends in particular on the number of parameters in interactions and the number of injected rules. We can thus inject new rules and study, for example, a linear correlation between the abundance of species in the microbiota and the amount of physical activity of the tested people. Multidimensional patterns can then be supplied by the machine, which also needs to be verified by the biologist. The more heterogeneous the data (biological and behavioural for example), the more the pattern can become "complex", the more delicate the validation becomes. The biologist will first look for what could invalidate the proposed complex pattern, or whether this pattern "makes sense or not" in relation to the medical and biological context being tested.

Likewise, we can think of letting the machine find its own rules. The objective will remain for the scientist to validate the patterns provided by the machine, to find a plausible explanation for these patterns (why they could correspond to real mechanisms). Here we remain in the realm of explicability.

But one can conceive that with the complexification of the interactions, with the novelty of the situations analysed compared to the current scientific knowledge, it is

tempting to "leave to the machine" the care of indicating possible evolutions without being able to verify them in terms of explicability. There is an important "ethical threshold" (what we called "ethics of knowledge" coming from an epistemological analysis) when the criterion of explicability is abandoned. There is a tendency here to validate "since it works" without understanding and knowing the underlying mechanisms. We can clearly point out an ethical dilemma between the technical efficiency sought in the context of a diagnosis, for example, and the explicability which can validate it scientifically. Such a dilemma depends on the circumstances: one can thus have spotted a disease biomarker without knowing how to explain it at a given point in the process. But if this marker "works" we fully understand that it should be used, which does not mean that it should not be sought to clarify it.

Thus, when an unsupervised and unverifiable diagnosis at this moment by the scientist is proposed, we enter a form of "digital-driven technoscience" which radically changes the mentality of the scientific process in regards to scientific truth.

3. WHEN AI IS COUPLED WITH CRISPR-CAS9

The same kind of situation is encountered in studies where AI is used to predict organization of complex genetic systems.

Discovered in 2013 and recognized by the 2020 Nobel Prize for Emmanuelle Charpentier and Jennifer Doudna, the CRISPR-Cas9 method is a gene editing technique that is of great interest to gene therapy. This therapy consists of eliminating or repairing genetic mutations in a targeted manner, directly in living cells. This involves, for example, replacing a defective gene with a healthy artificial gene by using the "DNA scissors" represented by CRISPR-Cas 9. Gene therapies are evolving rapidly and are already used in clinical trials on human somatic cells to treat numerous pathologies such as cancers, neuromuscular diseases such as infantile spinal muscular atrophy, or haematological diseases such as haemophilia.

At the ethical level, modifying a human somatic cell by CRISPR does not pose any major problems, as long as its use is scientifically validated, when the benefit / risk ratio (particularly in terms of off-target effects ") is positive and when the patient was able to give informed consent. This is not the case with human germ cells, where the changes will evolve in a person and be passed on to his descendants. The situation is then complex, not only in terms of irreversibility of transmission but also in terms of future modifications by epigenetic factors (taking into account the effects of interactions between genes and other factors of the biological and psychic environment).

In recent years, AI tools have been developed to fight against the undesirable off-target effects mentioned above, to reduce the prices of research, of the production on an industrial scale of living materials making it possible to implement the corresponding therapies, and to lower the price of drugs. We give here two recent

examples, one on the prediction of the processes by which a sequence break introduced by CRISPR is repaired (which is an important element to know) and the other one relating to the reduction of the search time for the development of treatment processes in gene therapy.

The first example is taken from an article in the electronic journal "Direct Industry e-mag" (Celia Sampol, February 24, 2020) that recounts the work of Richard Sherwood (professor of medicine at Brigham and Women's Hospital at Harvard Medical School in Boston), creator of InDelphi, a computer program involving AI to predict the results of CRISPR repair. The process under review regards what happens in a cell after that cell's DNA is cut with CRISPR. The cell repairs CRISPR cuts in very specific ways, and it is scientifically very important to know the "genetic" result of this repair.

Richard Sherwood's team was able to develop a machine learning system to have a computer program capable of predicting what type of mutations will be "caused" (possible mutations, uncertain causes) by the repair of the cut. The corresponding algorithm was able to propose "repair models" without having access to basic biological and genetic mechanisms. The importance of machine learning is essential here. This involves entering as much data as possible, including hundreds of millions of actual DNA mutations known today. Using this large data set on the types of mutations caused by CRISPR, actors were able to "train the algorithm" to look for "patterns of possible mutations." These models are products of the AI machine and not as observed and experimentally reproducible mechanisms.

The second example is taken from a recent interview with David Del Bourgo, director of a French start-up, WhiteLab Genomics, in *The Conversation* magazine (Sylvie Gamet, 2020). He gives a convincing example of research linking AI to CRISPR. It shows that the complexity of the data to be processed when modifying a few base pairs of the human genome is such that scientists have to reckon with millions of possibilities and hypotheses to solve. They must then proceed by iterations (trial and error), which takes a lot of time. AI can reduce this time considerably by processing tens of millions of pieces of data much faster, while retrieving information from the field of research, particularly through scientific publications and data on disease-related combinations that one wants to treat.

Artificial intelligence engines thus collect all available data, analyse and classify them, using "educated algorithms". Various tools such as natural language processing, image analysis, data interpretation and machine learning are used and combined to make these algorithms ever more efficient.

The education of algorithms is a crucial step. A biologist specializing in gene therapy verifies the processing of information by machines. The machines thus learn how to classify the data of interest according to the assessment made by the

biologist during this process. The "educated" machines then complete the comprehensive collection of state-of-the-art data and their processing in seconds, where it would take a few years for a human to process only part of it. As in the previous example, the machine provides interesting digital solutions for researchers, in record time, without, however, explaining the mechanisms presiding over mutations.

These processes are still at the research stage. The next challenge is to develop a predictive model to advise research teams on the optimal exploration routes to take depending on the issue, to reduce the number of iterations, in collaboration with different biotechnology companies. But accessing data, verifying its validity and sharing it are not easy today.

These two examples of CRISPR conjugated to AI treatments pose the same questions as in the case of gut microbiota. On the one hand, AI makes it possible to obtain digital solutions that can enable fine technical performance and decision support. But let's not forget that the AI process gives correlations (not causes) and possible "patterns" and solutions for predicting the evolution of the studied systems (which is very valuable), without the researcher having access to the biological mechanisms presiding over these solutions. In many cases, a biologist often (but not always) validates the patterns proposed by machine learning, by experiments when possible, by reasoning on plausible hypotheses for other cases.

4. EPISTEMOLOGICAL AND ETHICAL QUESTIONING ON « DIGITAL-DRIVEN TECHNOLOGICAL SCIENCE »

AI is able to find solutions to complex scientific questions that are not elucidated in terms of mechanisms, due to its ability to detect possible solutions to problems when too many parameters are involved and humans alone cannot disentangle them or even imagine them. It is in this sense that we have shown how we move from digital-assisted technoscience to digital-driven technoscience. The performances of digital-driven technosciences are astounding and say something of a form of scientific truth that one could predict and simulate without understanding its complexity. Epistemological work is here necessary to refine the elements of "numerical reductionism", but also to clarify the meaning of truth for the technoscientist including AI with different kinds of algorithms (related or not to physical or biological laws). In the last case (algorithms not related to mathematized reality), predictability by correlations prevails on research of causes and mechanisms. Which is a questionable choice in terms of ethics of knowledge. Such an analysis is essential to work in a relevant way on the major ethical questions raised by digital-driven technosciences (both for "ethics of knowledge" and practical ethics).

The first questions relate to the inevitable "digital reductionism" which must be clarified at the following levels:

- sampling: the quality of the data collected, their classification, the question of amalgamating different databases,
- coding choices,
- validation of patterns by a scientific expert, or by an expert system incorporated into the machine: what power is left to digital machines?
- the transparency of the processing algorithms and the selection of correlations.

Then the posture of the “technoscientist using AI” must be explained. Some remarks in this direction need to be deepened, particularly when the criterion of explicability is abandoned.

If technique preceded science in human history, their union has contributed a great deal to the scientific experimental approach to understanding a mathematized reality. Technics allowed both increasing precise observations and the performance of detailed critical experiments. These interactions strengthened the approach of “understanding in order to predict and make” and the development of new techniques, understood then as the armed wing of science. Going hand in hand with this development, work in the epistemology of science has made it possible to better define the relations object-subject via the theory of measurement and uprose the question of absolute objectivity of sciences (e.g. history of quantum physics in the 20th century; Gödel's theorem; see Magnin, 2015).

Faced with the complexity of empirical reality, scientists have learned to live with uncertainty, unpredictability and incompleteness. In addition, applied sciences have allowed the development of an empirical approach, based more on “making to understand”. Technosciences then appeared as a new coupling between sciences and techniques, driven by making (in relation with techno-economical goals) and intended to simplify a complex nature in an attempt to reconstruct it (top-down and bottom-up methods). With digital as the driving force behind the development of technosciences and their convergences, “making or constructing-reconstructing natural / hybrid / artificial objects”, enabled and driven by digital technology, becomes a major asset for “predicting and acting”, for designing the real, more than for its “knowledge” (underlying mechanisms, explicability).

Thus, is there a new threshold of development linking science and technology through digital technology emerging, particularly when algorithms are not related to mathematized reality? It would constitute a formidable crossing from an ethical point of view which requires further study. It is linked to the nano-bio-neuro-technological convergence with the dazzling development of AI, all driven by a kind of autonomy from digitization and simulation vis-à-vis the mathematical conceptualization of reality. We speak here about an “ethics of knowledge”, i.e. the choice and corresponding values which preside to the digital-driven technosciences in regards

to classical scientific processes which are based on explicability and critical experiments to validate hypotheses and theories.

The first change is related to the choice to favour prediction in regard to explicability. If these trends are confirmed, we would then go from “scientific knowledge by making”, specific to technosciences (different from “knowledge for making” in usual sciences), to “predicting for making without explicability” and without the elaboration of a classical scientific knowledge. New relationships to scientific truth for which “intelligent machine prediction” would replace classical explicability occurs, inducing a break in the classical definition of scientific truth. For example, we would no longer be “in the falsifiable way of Popper” to define scientific truth (A. Chalmers, 1987), namely the possibility of considering and carrying out a critical experiment that could invalidate the hypotheses of a theory.

Thus “capabilities to predict for making”, based on efficient correlations, is more important than capabilities to “know for making” with the determination of biological mechanisms or causes of evolutions. The truth of “digital prediction” with algorithms not related to mathematized reality, which is based on correlations and not on causes, is then favoured. It would be a disruptive change in the frame of reference of the ethics of knowledge. The relation to truth evolves, the weight of the technical performance and techno-economic foundations is increasing while human free is decreasing!

In conclusion, one can ask: “what sense of responsibility and precaution should we put into action to preserve free will in digital-driven technosciences”? The evolution of the “ethics of knowledge” studied in this paper induces correlated impacts for the sense of responsibility in practical ethics.

On a practical level, this opens up a considerable amount of risk. For instance, it radically impacts jobs in different fields such as engineering, medicine, economy, insurance, law. As long as the real causality has not been differentiated from the usual correlation, it remains fragile in the extrapolation of its future.

We have to be very prudent considering numerical trends as “possible” solutions which can be pertinent but must be deeply analysed before applications (as already shown in the analysed examples of this paper). Don’t forget that the main question in practical ethics is: are all possibilities desirable? To answer with collective responsibility, we need “free will”!

REFERENCES

Boyd, D., and Crawford, K. (2012), “Critical Questions for Big Data.” *Information, Communication & Society*. 15.5: 662-679.

Chalmers, A., *Qu’est-ce que la science ?*, Paris, La Découverte, 1987, p. 76.

Crawford, K. (2013a), “Big Data Becomes 2013’s Scariest Fad.” *Future Tense*. 05- 13-2013, <http://www.abc.net.au/radionational/programs/futuretense/katecrawford-big-data/4686244>.

Crawford, K. (2013b), “The Hidden Biases in Big Data.” *Harvard Business Review*. 04-01-2013, <https://hbr.org/2013/04/the-hidden-biases-in-big-data>.

Dao, M.C., et al. (2016), “*Akkermansia Muciniphila* and Improved Metabolic Health During a Dietary Intervention in Obesity: Relationship with Gut Microbiome Richness and Ecology.” *Gut* 65.3 (2016): 426-436.

Gamet, S. (2020), “Pourquoi combiner IA et thérapie génique”, *The Conversation magazine*: march 26, 2020.

Guillermin, M. and Magnin T.,(2017), “Big Data for Biomedical Research and Personalised Medicine: an Epistemological and Ethical Cross-Analysis ”, *HSS VI.3* : 13-36

Magnin, T. (2015), “Le scientifique et le théologien en quête d’origine”, Paris, DDB: 72-74

Popper, K., in CHALMERS, A. (1987), “Qu’est-ce que la science ? ”, Paris, La Découverte : 76.

Sampol, C. (2020), *Direct Industry e-mag*, “CRISPR-Cas9: Artificial Intelligence Gets Involved in Gene-Editing”: february 24, 2020.