

Information, science and democracy, for an ethics of scientific knowledge¹

Giuseppe Longo

Centre Cavallès, CNRS and Ecole Normale Supérieure, Paris,
and School of Medicine, Tufts University, Boston

<http://www.di.ens.fr/users/longo>

Abstract The modern theories of information opened to way to extraordinary new tools for communication and knowledge construction. When identified instead with the world (DNA is a complete information carrier and a program for ontogenesis, atomic particles and planets exchange information ...), a new form of scientism affects knowledge and trust in science. Vague and implicit assumptions, with no precise references to the sciences of information, force strong consequences, in biology in particular. Similarly, the powerful optimization methods that make intelligible the dynamics of inert matter are expected or used to direct also human activities and replace human government by automatic governance. We are then meant to follow the formal, mechanical rule or the unique optimal path, instead of inventing new forms of sense in human interactions. Yet another fantastic tool, the newly available Big Data, is being considered to substitute scientific invention by correlations among numbers: fortunately, mathematics a priori shows that a deluge of spurious correlations may be found in any very large set of numbers. An ethics of knowledge may then be proposed in order to reconstruct the trust in science, as critical thinking at the heart of democracy.

1 - The sciences of Information

The concept of information has been dealt with by at least two rigorous and important scientific theories: elaboration of information, after Alan Turing, and communication or transmission of information, after Claude Shannon. These two theories are based on fundamental properties of mathematical invariance, which is to say these authors singled out some key mathematical

¹ Extended and revised synthesis of an interview by P. Bartolini in Italian (ROARS, september 2016), translated into English by M.R. Doyle and S. Savic. In A. Fabris (ed.), *Trust*, Studies in Applied Philosophy, Epistemology and Rational Ethics 54, Springer 2020.

notions or properties that can be transferred from one context to another, while preserving that which matters. In both cases, typically, the characteristics of information do not depend neither on the code chosen (apart from the negligible price of coding, be it by 0s and 1s or by numbers from 0-9 or any other sequence of signs) nor on the material support: one can elaborate information on transistors, chips, silicon, and can transmit signals by wires, drums, smoke... This ancient invention, formalized in a revolutionary manner by Turing in 1936 and by Shannon in 1948, has enabled us to distinguish between *software* and *hardware* and to propose an autonomous theory of programming, i.e. of elaboration of information, or of transmission of information (Shannon) — both independent of the specific coding and of the material support (the hardware), a fantastic theoretical and practical feat. In Manchester, in 2012, at a conference marking 100 years from the birth of Turing, the students constructed a Turing Machine (1936), the mathematical prototype of modern computers... out of Legos. This machine worked, although of course a little bit slowly....

More precisely, in the 1930s, Gödel, Church, Turing ... each of these mathematicians gave his own, very different definition of the notion of “computable function”— these turned out to be all equivalent. In short, a function is computable when it can be described by an algorithm – a set of instruction or formal rules in a sufficiently powerful language. The consequence of that non-obvious equivalence results is that this notion does not even depend on a formal language or system of calculation used. Independence or invariance of a notion, typically from the formal system used, the coding and the material support, it is the force and the scientific relevance, in particular, of the notions of information and program.

2 – Mysticism and Scientism

As it often happens, the mystics have taken the beautiful mathematical invariance in the two theories of information as an absolute and have applied it to all phenomena, including natural ones. Thus, when one of the leaders of the Human Genome Project writes that, soon, we will be able to transfer genetic information on a compact disk and claim/assert "Here is a human being, this is me!" (Gilbert 1992), he applies the distinction between software and hardware, fundamental in computer sciences but catastrophic when it is about understanding the living. Gilbert's statement, a rather common attitude in molecular biology, betrays an alphanumeric, linguistic vision of biology as just an analysis of software, independent of coding, of its material realization and, of course, even physical dimensions – in discrete data types, phenomena in any finite physical dimension be encoded in one dimension, a string of 0s and 1s. This “linguistic” approach is very well described by François Jacob and Jacques Monod, two great scientists, protagonists of molecular biology, rare among those who have made explicit a broad theoretical reflection on the subject: "The surprise is that genetic specificity is written out, not with ideograms like in Chinese, but with an alphabet" (F. Jacob, opening lecture, Collège de France, May 1965).

All this would reassure us that God is on our side, as confirms Francis Collins, director of the National Human Genome Research Institute, in June 2000, publicly announcing the decoding of five human genomes together with Venter, Clinton and Blair: "we have caught the first glimpse of our own instruction book, previously known only to God". Mysticism and theology, or

researchers' enthusiasm about an extraordinary technical performance? Which philosophy and which practice of biology are they proposing? Which vision of the human?

It is firstly the sense of radical materiality of life that one loses, that within which one cannot distinguish the software and the hardware: The living state of matter is made of this DNA, RNA, only, of these membranes and their physics-chemistry and of nothing else Synthetic biology in fact transfers fragments of these from one to the other: The DNA or fragments of DNA in a membrane taken from another cell; it does not re-write neither transfer software onto another hardware, a practice common in computer science. When the colleagues from molecular biology will know how to transfer information and biological programs in a DNA on Lego and make a cell that works—then we will be able to affirm that they have finally "extracted" and separated information and instructions, the software of the living, written by God or not, from its hardware, in the mathematical and informational sense of the words. Are these only innocent metaphors? No: The characteristic of a metaphor is that it transfers and enriches the meaning of a theoretical proposition. In our case, it is a reference to commonsensical, vague, poorly defined notions of information and program (where is the operating system, the compiler in a cell? Which are the proper invariant properties of computing and information mentioned above?), but from which strong consequences have been derived, heavy weight onto knowledge, as it may be shown in reference to cancer research (Longo, 2018a). This is unacceptable in science, because this imposes a bias without rigorously specifying its origin and forces research onto tracks whose principles remain implicit, a form of "scientism" that affects science and trust in science. For example, as molecules, DNA in particular, are the obvious locus for discrete/digital coding of information, it follows that "everything is molecular", or even "everything is genetic", the most organized chemical structure in an organism. Thus, the claim that "any phenotype has a causal origin in the genotype", including, say, the gene for monogamous behavior (Nature, 1999), as, for many, genes should broadly encode also behavioral information (Plomin 2018)². Instead, DNA is the extraordinary physico-chemical trace of Evolution, a fundamental *constraint* in ontogeny, as I will hint below.

3 – Strong consequences from vague hypotheses

As for an analogy, the geocentric/informational approach in biology strictly parallels the geocentric approach to the Universe. They are both based on commonsense: Everybody knows what information and programs are, as it is evident that the Earth does not move ... Then, from both vague "evidences", complicated theories and strong consequences have been derived, based on ad hoc epicycles on top of epicycles for every planet, genes controlling gene expression for every phenotype. Science instead is made by putting forward strong and rigorous hypotheses, which go against common sense, against the idea that the Sun raises on motionless Earth, that light travels along a Euclidian line, or that a measurement remains the same if one changes the order between position and momentum (the non-commutativity of quantum physics) ... From vague hypothesis on information and programs, which have for too long governed molecular biology, an understanding of how transmission and elaboration of information may work on

² Unfortunately, these simplistic views of genetic determination reach the general public: "Being Rich and Successful Is in Your DNA" (*Guardian*, July 12, 2018); "A New Genetic Test Could Help Determine Children's Success" (*Newsweek*, July 10, 2018); "Our 'Fortunetelling Genes' make us" (*Wall Street Journal*, Nov. 16, 2018).

macromolecules has been *deduced*. That is, the strong idea of the “*Necessarily stereospecific* molecular interactions explain the structure of the code ... a Boolean algebra, like in computers” (Monod, 1970) was derived. This prevented observing the stochastic character of these interactions, already known in physics-chemistry, whose probability depends on the context - the cell, the organism, the ecosystem - for a classic reference, see (Elowitz, 2002) and for more references (Bravi & Longo, 2015). In this way, little attention has been paid to the possible role of endocrine disruptors of finite combinations of the 82,000 molecules that we have produced and thrown into the ecosystem in a little more than a century (Diamanti-Kandarakis et al. 2009), as most of them are far from presenting exact affinities with bio-molecules. Yet, they do affect hormonal cascades as they modify the *probabilities* of interactions, in particular with cellular receptors. Or, they may even replace hormones, in low, but sufficient, probabilities due to lower chemical affinities or by changes of affinity constants due to modifications of the chemical context, (Vandenberg et al. 2012). Of course, stochastic macromolecular interactions and stochastic gene expression, in a cell, a tissue, an organism constraining those interactions, are incompatible with the current notions of information and program, their determinism and their focus on DNA, elaborating and transmitting information in the exact Boolean fashion we find in computers.

In this way, by affirming that "the organisms are simple vehicles of genetic information" (to take a 2002 reference text in molecular biology of evolution) in the service of the "selfish gene" (as in a popular philosophy), it has been believed, for example, that we could completely pilot the development of plants by genetically modifying them (GMOs). To be more precise, GMOs are the children of the "central dogma" of molecular biology (Crick, 1958), according to which information propagates from DNA to RNA to proteins in a one-way and linear manner (except for some retroactions RNA-DNA). To the contrary, today, we understand the strong interactive and retroactive eco-systemic effects, acting on a number of levels. For example, GMOs heavily affect the microbiome and roots' mushrooms, essential symbionts of plants. As a consequence, they gradually transform humus into a sort of sterile sand (Bizzarri, 2012). Again, the exact macromolecular interactions (stereo-specificity, the key-lock paradigm) were *derived* from the informational/programming myth, which justified or gave a pseudo-rigorous frame to the genocentered approach, against empirical evidence. And this is massively affecting our lives and, in many, the general trust in science.

When we speak of the consequences of this "programmable/informational" and DNA-centered vision in cancer research, the organismal perspective in (Sonnenschein, Soto 1999) may be an alternative reference. It was confronted by harsh criticisms for at least a decade. But a paradigm shift is taking place (Baker, 2014). Recent articles by leading researchers in the informational perspective acknowledge the failure of the search for the “signal” from the carcinogen “de-programming” the genetic program. So, R. Weinberg, a biologist from the MIT, is among those working for long on search for genetic therapies for reprogramming DNA, promises made over and over since 1971, the beginning of "War on Cancer" by Nixon, and still today. In (Weinberg, 2014), he acknowledges the failure of the programming/informational, genocentered approach to cancer. Similarly, Gatenby, another leading personality in the field, in a 2017 paper in *Radiology*, acknowledges the little or absence of *causal* relation between DNA mutations and the various forms of cancer, see (Longo, 2018a) for this and more references. Where has information gone? Should it be just diluted in epigenetics? Ok, then one should define the invariant properties that make the sciences of information robust and that should be used in an epigenetic frame: is

the proteome and its interactions that support the biologically relevant information? Or use another word and concept or another *explicit* theory.

But, careful: Molecular biology has given us an immense quantity of data and has shown evidence of fundamental mechanisms. It is extraordinary to see the experimental fineness and originality of methods that enabled Changeux, Jacob, Monod and Lwoff to discover the mechanism of chemical allostery and of the lactose operon, a regulation of the genetic expression by DNA itself. However, on the basis of the operon specific case, they leave aside the epigenetics, to say the least: genes regulate gene expressions... So, they do not cite Barbara McClintock, the first author to analyze the *epigenetic* regulation of gene expression. In fact, since the 1950s, the studies of major geneticists like McClintock and Waddington, suspected of "Lamarckism" (by giving a role to external influences on DNA), have been abandoned, for decades. They had always studied the chromosomes, an extremely important component of the cell, in a context: the proteome, the cell, the organism. Evelyn Fox-Keller tells this story in a book on Barbara McClintock. In this way, for decades, molecular biologists have researched regulators of genetic expression in DNA, assigning it the role of the complete informational description of an organism, like an Aristotelian *humunculus*, very small indeed, but completely codified in chromosomes and therefore "programmed" (yes, we are modern). On the contrary—some geneticists have known this for 60 years—every relevant molecular cascade, from DNA to RNA, to proteins, depend causally on the context and on its material, physico-chemical, realization. The activity of these macromolecules is context driven or used by the context - the cell, the tissue, the organism – in its irreducible tri-dimensionality and materiality: No software as a mathematical invariant, no information as independent of coding, hardware, dimensions etc. And since several years we know that the individual cells, taken in small groups from cancerous tissues and transferred into healthy tissues take up normal function, without necessarily "reprogramming" their DNA.

4 – Randomness and bio-resonance

The concept of randomness in science is a delicate one. It is not an absolute. It depends on theories. It should be understood as the unpredictability with reference to the intended theory (Calude & Longo, 2016): it is what the theory shows/assumes to be unpredictable. In physics, classical and quantum dynamics propose two concepts of randomness, epistemologically and mathematically distinct. The first depends upon the joint role of non-linearity and of the physical measurement, namely upon the interactions in a dynamic and upon the access to the world given by measurement. A fluctuation/perturbation below measurement is the hidden cause of a phenomenon observable and measurable after some time – non-linearity may amplify the hidden fluctuation (Poincaré wrote in 1902: "and we then have a random phenomenon"). It is about a chaotic determinism of throwing dice, of a double pendulum or the solar system (luckily for us, in quite different temporal scales!). Quantum mechanics proposes another definition, which is intrinsic to the theory and starts from the indetermination (and non-commutativity) of measurement. Moreover, the phenomenon of "entanglement" distinguishes quantum randomness, mathematically, from the classical one. In biology, or at least in the cell, these two forms of randomness overlay each other. There is significant evidence of quantum phenomena with phenotypic effects (Buiatti & Longo, 2013) and these may be amplified or dumped,

according to the context, by classical phenomena. Moreover, the interaction of different levels of organization (molecules, tissues, organism) produces effects of bio-resonance, as we defined it, which may destabilize and stabilize an organism. But the most important thing is that this randomness, in biology, is not "noise" as in (Monod, 1970), a notion related to information, and like many claim still today (Bravi & Longo, 2015). It is instead an essential component of the variability, adaptability and diversity of the living and thus of its structural stability on the level of species, of populations and also in an organisms. They are all "more stable" when adaptive and diverse and this is partly due to the peculiar form of randomness one has to describe in biology, as we hinted above (for more about unpredictability in Evolution, see (Longo, 2016)).

To take an example, the immune system is a paradigm for the "functional generation" of randomness, essential to the stability and adaptability of an organism. But even an organ of such a uniform appearance such as the liver contains up to 50% of cells with a "wrong" number of chromosomes or with all kinds of mutations. By the diversity of enzymes' production, this contributes to the resilience of the organ (and the organism) to a variety of unforeseeable toxic shocks. These biological features have nothing to do with "noise" which destroys information. Instead, adaptability, exploration of diversity, changing niche constructions ... are at the core of biological dynamics. They simultaneously modify organisms and ecosystems, phenotypes and functions, and therefore the "observables" themselves in the evolutionary processes. Following a fundamental chapter of Darwin's Origin of Species, organism's "correlated variations" and its "extreme sensitivity" to internal and external conditions make it, but also the ecosystem, a conceptual challenge that cannot be envisaged merely as elementary "stacking" of elements. Machines are constructed by stacking constitutive elements, ones on top of others: We construct them by association, assemblage of components, nuts and bolts, chips and bits, simple and elementary. Occasionally, some noise may affect the construction or its functioning.

Also, computers and networks of computers are superpositions and a very complicated blend of simple and elementary components, which elaborate and transmit *bits* of information. In the natural sciences, on the contrary, it is not necessarily the case that the fundamental is *elementary*, nor that elements are *simple*. Galileo's or Einstein's theories are fundamental but do not concern the elementary, such as Democritus's atoms or quanta. Moreover, these last ones, as much as the cells that compose the elementary of the living, are not at all simple. But, even more crucially, a multi-cell organism is not made by adding tissue to it, by attaching organs to it, putting eyes in holes ... like in all machines. On the contrary, it is made by differentiation from one cell, a zygote, which is already an organism and which, at each cellular differentiation, keeps its organismal unity. There, we have an original complex object, a cell, which further complexifies by constructing a new unity by differentiation, and not a stacking of elementary and simple bricks, glued together. All this occurs in a permanent construction of diversity, radically unpredictable, during phylogenesis: organisms co-construct niches in changing ecosystems that "enable" them (Longo, et al. 2012). The changing space of possibilities singles-out biological evolution from physical dynamics (Longo, 2016).

5 – Information everywhere, an abuse

Besides the reinforcement of reductionists approaches in biology due to the notion of digital, programmed/transmitted information, I do not see how that sort of analyses can be improved by saying that “information is everywhere”. Which are the principles of this extended and even more vague notion? Is this a reference to the remarkable new field of “Geometry of information”, a new mathematical area rigorously dealing with homotopy theory and continuous symmetries and their breaking (Barbaresco, Mohammad-Djafari 2015)? Instead, in an attempt to make the commonsensical reference to digital information more rigorous, Maynard-Smith, in a 1999 article we discuss in Perret-Longo’s paper in (Soto et al., 2016), confuses, in the examples, Turing-Kolmogorof and Shannon-Brillouin dual approaches to entropy and complexity as amount of information on discrete data types – in one case they are covariant in the other contravariant. In (Longo, Seno, 2018), (Longo 2019) and (Montévil 2020), we tried to be more specific as for the use of delicate concepts such as negentropy, anti-entropy etc..

More generally, we do use, as humans, *gradients* (variations) of energy or matter to transmit and elaborate information (modulation of frequency, of electric flows, drums, smoke ... switches). Gradients of flows, in organismal and cellular exchanges typically, are very important. The claim that one has information, in a natural phenomena, any time a gradient matters more than the *quantity* of a flow, is an amazing abuse. It is an anthropomorphic-magic projection on nature and it blurs more appropriate scientific analyses of the role and nature of those gradients, those flows and their specific material structure. “Information” is too important today, to be left to a vague reference to an un-specified notion and then diluted on nature: robust, well adjusted theories to empirical evidence should be given, in search for bridges, dualities, unifications These is how the remarkable theories of information mentioned above have been constructed, their relations are still to be fully developed. Physics, since Galileo's inertia (momentum conservation) keeps proposing new observables and pertinent parameters: Force, energy, entropy ... new quantum observables. A change of scale forces a change of theory: Quanta, thermodynamics, hydrodynamics, Relativity ... all *incompatible* theories. Unity is very difficult to achieve: Newton, Maxwell, Boltzmann and Einstein and no others proposed remarkable unifications - each a scientific revolution. A book by three physicists who work on the “borderline” of theories explains this beautifully, a review is in (Longo, 2016a). In which way, claiming that both atoms and planets exchange information would help in unifying the quantum and relativistic fields, an open issue since a century? Diversity of tools and concepts, presented with rigor and facing the variety of phenomena (and scales!), is at the core of the scientific knowledge we may trust: The endeavor towards unity is a further, very difficult achievement. Scientism, against the history and the practice of science, let many believe that science is the progressive “occupation” of reality with familiar tools and concepts, possibly just one – information for example. This is a way to claim the full control on nature, by handling information, like the information in the DNA of plants (the GMOs) or the incredible fuzz made on “genome editing”. A fantastic discovery, the way some viruses modify DNA, has been transformed in financial support for actions on DNA, which result in plenty of uncontrolled side-effects (Sikandar 2019). These explorations can be a component of scientific work, what is wrong is its mythical handling and the resulting distortion of research and its funding by phantasmagoric promises, see (Audédat 2015) for a survey. Finally, this hides a mechanistic view of nature, which should be controlled like our piecewise constructed machines, with no appreciation of the peculiar organismal unity and historical dynamics in animal, plants and ecosystems.

6 – Optimal governance vs government

Besides “mechanisms”, scientism also claims to understand and to govern the world by optimization methods. For example, the notion of "Human Resources", which analyzes work in the same way as it analyzes material resources, have been invented in the ex-USSR. Its aim was the optimization of all components of production, including human work, by the mathematical methods of "linear programming". This is how the modern idea of purely technical government is born, put in the hands of engineers of the soviet Gosplan; this is what we call, today, “governance”. These methods have since found their place in our major private companies in the West and, then, also in the public and national sectors, from health to government. Shouldn't we have to govern them like we do govern (big) business? The government, should it be entrusted to the "best" entrepreneurs, like it was in Berlusconi's Italy, with a party as centralized and authoritarian as the worst Stalinist party of our democracies. Trump is a further example of this identification: Business governance rules government. This is how the circle closes. In the State Enterprise, Stalinist or neoliberal, scientism and governance blend wonderfully. Even the murdering effects may be compared: for one or two decades “only”, instead of 70 years like in USSR, in Latin America neo-liberal governance has been accompanied by jailing, murders and torture (“it works!”), with the explicit economic counseling and guidance by Milton Friedman and the Chicago boys. So far, in our countries, the democratic components of the liberal tradition, Constitutions issuing from revolutions or anti-fascist resistance, allowed to restrain the authoritarian tendency of neo-liberal governance, so far.

Alain Supiot in his book *Governance by Numbers* (2017) more closely explains the difference between "governance" and government by law. The former is about an "objective" management, according to formal rules, potentially mechanizable, independent of all contexts. It uses methods of optimization, like in the “linear programming techniques” invented in URSS, yielding one single possible and optimal path, a “geodesic”. Geodesics are produced also by the laws as equations of Walras equilibrium economy: These spontaneously determine a unique possible paths, an optimal one - societies have just to follow it. To the contrary, the law of man, the “government”, is motivated by its social or human meaning, when first discussed in the agora, in the moment of its being voted, and afterwards, when it is *interpreted* by a government or a judge who apply it in their domains, who give it a possibly new contextual sense.

In fact, the term "natural law" has been historically articulated into the "law of man" (and of the gods) in a very interesting way. In different cultures, it is interpreted in very different ways (Longo 2017). The formal rule, based on numeric or formal writing, is instead manageable by automatism independent of interpretative ambiguities, of which we make in the best case a "fine tuning". Governance may be largely implemented by machines elaborating information; it is claimed to be “objective”. In the Merkel-Sarkozy agreements, for example, it is written that the punishment for the States that transgress the rules about the deficit will be implemented in an "automatic way", potentially mechanizable. But the essential question is that of democracy, that of law which has to be discussed and interpreted, which gives meaning to the life in a community, with the possibility of a real and organized disagreement: The analysis of Stalinism and of neo-liberalism, European or American is subordinated to it. So, similarly to the huge Soviet industrial conglomerates, American corporations, since the Securities and Trust Act of 1933, have the status of a medieval fiefdom or of a legal person, a State within the State.

When we talk about science, within large, once enlightened corporations, like in renaissance principalities, such as IBM, ATT-Bell Labs or Digital, with which I collaborated at the end of the 1980s, there were spaces of extraordinary freedom for research: Large groups of very diverse researchers could think freely, protected by the master/corporation, thanks to a long term, enlightened vision of industrial research. Apple, Google and Microsoft offer still today remnants of this, autonomous research centers with about a few dozens people free to think, but these are no longer hosts of hundreds of researchers who, until the beginning of the 1990s, explored all possible directions of American industry: research for immediate profit dominates. Apple and Google have begun with great innovative ideas, at the end of 70s and 80s, but today they mostly apply, in computer memories all more miniaturized, powerful and faster, about the same ideas – or transfer and develop new ideas coming from public research, like the new Artificial Intelligence (the so called Deep Learning, a major technical innovation). When we talk of Microsoft, Roberto di Cosmo, professor of computer science in Paris, explains its growth in *Planetary Hold-up*, as a moment of disruption without research content in the 1980s. More generally, the significant reduction in industrial research in the past 20 years has been described by a number of authors, see (Seno, 2014), (Rafols et al. 2014), (Audedat 2015) and many others. It is one of the main consequences of the transformation of industrial property into a shareholders scheme, managed by universal managers, interchangeable, who have only to increase the value of a company's shareholders in a short or very short period of time. The permanent hype for new gadgets is transforming the public's trust in science: It is an instrument for advertising more than for knowledge construction. In this frame, note that the value of shares, in view of volatility and the evaluation of risks, is a matter of information, including on insurance assets, and on insurance based on information, far away from actual products and productivity. The continual flow of announcements of new technological applications manipulates this information and the associated shares values: It is at the heart stock market activities. This is the cultural hegemony that, in my view, dictates the universal role of the notion of information. No scientific argument nor evidence may be strong enough to oppose to a hegemony which is first economic and political.

As a matter of fact, all of this is of little help in facing the problems of our world's complexity: This requires first discussing the method and processes of research. Any research activity has its own method and concepts to be spelled out explicitly, as hinted above as for the notion of information, possibly on the grounds of an *ethics*: questioning one's own principles and motivations of knowledge, by making a “step aside”, look at them at a distance. This requires giving oneself time to think, without knowing exactly where one is going, open to being judged, and severely so, in order to be hired, promoted, closely evaluated on the basis of one's *past* creative work, beginning with a PhD. Thesis work and so on, say, and not on promises. The computational governance of research by numbers, by “bibliometrics” typically, does not work either, nor financing (only) huge projects. Bibliometrics reinforces dominant fashions, kills diversity and critical thinking, an essential component of science and ... democracy. The identification of democracy with the majority vote, disregarding the division of powers and the formation of alternative views, is a major danger in all our countries. It similarly endangers science by the “vote of the majority” of all colleagues on Earth, counted by the number of quotations, by machines searches on Internet, in the short term (the impact factor concerns at latest two or five years from publication, which is nothing in fundamental science). If we do not appreciate the risky and isolated exploration of new paths, by analyzing contents, and we do not respect the ethics of individual or team research, the passion for knowledge, no institutional

engineering will allow us to correctly evaluate our activity, nor to finance scientific projects, which is also necessary; big problems of the ecosystem, for example, would really need it.... But if we do only Big Science, we kill science: No project whose final application is told in advance can never be very innovative. The real novelty, even technical, is always achieved from research that does not imagine it - and often long after or as an indirect finding.

7 – Following the mechanical rule vs. the imagination of configurations of sense

We are confronted by an important crossroads: extraordinary instruments of interaction and exchange of information can enrich our knowledge and our scientific practices, or they can be employed to normalize us and to make us “follow the rules”, to render us all homogeneous. Networks put everyone in an unprecedented position of being able to meet at a distance, of accessing the knowledge of all of humanity in its diversity. The exchange of cultures, ideas, objects, was at the heart of major moments in our history, such as ancient Greece and Renaissance Italy, to mention only two examples. But with this new speed of interaction we could do much better. Or, to the contrary, information networks can be used like "average fields" in physics: With too many neighbors, there are no more singularities, we cannot any more have individuality, we become all grey. We have discussed above the monochromatic images of the world of this type: DNA, the brain, human law, they are all seen as sets of formal rules, instructions and programs, of the same type, like in computers, but enriched - in the best case - by some network “noise”. An instrument, a computer, an excellent assistive tool used by science to understand the world, in mathematical modeling for example, may also be employed to flatten the world, or, worse, to be identified with the world, a new, common sense of mechanicity. It happens that it also organizes our activities, by computational evaluations of work, bibliometrics being only one example. Rigid medical protocols are yet another terrifying example, as sets of formal instructions. This is the world of pure information, elaborated and transmitted by machines, with no interpretation, nor material body, the flesh at the origin of meaning. Living organisms are always acting, moving, in a protensive gesture: They interpret by this any “friction” on the world, any hit arriving – this is the bodily and material origin of meaning.

The modes of life are profoundly changed by the “everything is information” approaches. We are constantly under the pressure of computing devices that, as many promised since the 1960s, were going to replace man in everything. This was the dream of Classical Artificial intelligence (AI), based on deductive machines, that is on Turing’s Logical Computing Machine, as he called it, but well beyond Turing’s lucid analysis, (Longo 2018b). Yet, today, the paradigm has changed: The connectionist approach, based on continuous deformations of multi-layer neural nets (“Deep Learning”) prevailed, far away from the old “deductive” AI. It is no longer an “imitation” of human intelligence, in Turing’s sense (his 1950 paper on the Imitation Game), but a tentative “model” of the morpho-dynamics of the brain – that is of changes of forms, including electric variations and wiring, a model that may implement a highly simplified “learning” process. This is a much more effective approach, much closer to the way animal brains may seem to work. So, not a day passes without hearing: "Attention, do accept no matter which working condition, accept to lose your rights because if not, you will be replaced by machines!" – a voluntary construction of a collective imaginary. This substitution already started since decades, and it continues to happen in offices, and above all, in production chains, since the ‘70s, in storage

centers, today, thanks to old style digital control machines - that is to say in professions which require *reiteration of identical gestures* - the original competence of the digital, Turing-like, machine. Since then, we observe a construction of an imaginary adapted to the subordination to the rules, to mechanical evaluation, to the governance which replaces the government, as observed above. Because also the neural nets machines just elaborate information, as they also are input-output devices. They somehow model the brain, somehow learn, but they are remote shadows of an actual animal brain. This is an always active organ, whose “default” state is a chaotic dynamics. In absence of sensations and perceptions, for example by torturing someone by complete silence, immersed in body-temperature water, in the dark, the brain gets “crazy”, its activity goes out of control. The brain is not an input-output machine, but an always super-active organ, *constrained* by a changing context. Its continual activity is canalized by and works only in its preferred ecosystem: The skull of an animal as part of a sensing body in an ecosystem - and in history, as for humans. Its material flesh, the only one we can witness, is essential for this.

Intelligence is not just “elaboration of information”, but it is first production of sense, within a biological and historical body in an active friction with the world. Our humanity probably started when we singled-out by interpolation and named non-existing forms of stars’ constellations and thus gave “meaning” to the non-sense of the little lights in the sky. Or, when we interpreted phenomena by inventing and naming ancestors and gods, by establishing the law, very differently in different communities. This is not elaboration of information but the “imagination of new configurations of sense”, in the stars, in the human community. Our specific, material body and its gestures, in a common history of the communicating community, are essential to this. And these invention are often historical: only the Greek geometers, say, invented the “line with no thickness” (Euclid’s definition beta). An abstraction from movement, a trajectory, the invention of the notion of border, which originated the a-logos of the diagonal or the infinity of the Greek π . The extraordinary Nine Chapters of the Chinese geometer Lu Hui, III century, do not imagine this limit construction: The construction of mathematical sense then took a different path. This what I mean by a “historical formation of sense”, its reduction to “elaboration of information” is a parody of knowledge construction.

Today, when we speak of constant novelty, certainly, we are surrounded by a million of new fantastic gadgets exchanging bits and bytes of information, but publicity on this matter is often more important than scientific innovation, as hinted above. If you read newspapers from the end of 1990s, you will find similar announcements to those of today's Self-Driving Car ... and what has become of the Google Glass in the past two years? It was presented as a revolutionary way to elaborate visual information ... So many promises are continually formulated whose realization ends up as gadgets that decorate our cars, typically. In Pittsburg (PA, USA), a couple of years ago, Uber has launched a big project of driverless taxis, driven by robots that elaborate input information. For the moment, they declare that there will still be an "employee" in the vehicle – as the experiment has been a major failure. Is it possible that such a smart company that earns so much money by just selling information, with no “hardware” (no cars’ ownership), is wrong? But no, the primary objective is the increase, in short term, of stock market value of the concerned companies, by the effect of advertising. As for driving automation, we will perhaps channel, by means of electronic tracks, the driving done by robots and humans as well, so that their coexistence on the road can be possible. As long as we are channeled as car drivers, it is great. But the myth of mechanization, by all kinds of algorithms from genetics to evaluation, the necessity of interacting by subordination to machines elaborating digital information – at first

theoretical, conceptual and then practical, can have as a goal or as an effect to completely channel our human and social behavior. The too much computational modeling of various forms of “meteorology of social dynamics” seem to go in the same direction: Social information is extracted for the historical context and elaborated by optimization techniques – more than a botched analysis, it seems a method for guidance.

8 – Spurious Correlations and non-sense in Big Data. Taking a step aside.

This is the social pressure underlying and justifying the “everything is information” myth. The reduction of human activities and intelligence to “elaboration/transmission of information” by input-output devices continually compares human activities to machines, a fantastic tool when soundly used, but often overestimated in their performances and presented as threatening the human intellect. The worst example is perhaps the new fashion of un-scientific Data Mining on Big Data, as immense reservoirs of information. Big Data and their sound statistical analysis are an opportunity without precedent if they are used to produce hypotheses, to (in-)validate theories and to propose new ones. But, to the contrary, and in a truly viral manner, some think today that we can optimize thought, by reducing it to zero: sufficiently powerful algorithms could in the end “replace scientific knowledge” (sic!). This is supposed to predict all kinds of dynamics and to direct action, without the need of a hypothesis, a theory or knowledge. The bigger a database, yotta of yotta bites, the more, they claim, we can avoid thinking: “machines will discern regularities that the science does not see”, but which are sufficient to predict and act (Anderson, 2008). “We kill based on Data Mining on metadata” declared the former CIA director, M. Hayden, in a debate (<https://www.youtube.com/watch?v=UdQiz0Vavmc>).

Luckily, mathematics allows to demonstrate the absurdity of these perspectives. In (Calude & Longo, 2016a), by a simple application of classical, not trivial results, we pointed out the “Deluge of Spurious Correlations in Big Data”. In short, for any given “correlation between numbers”, one can compute a number of elements, let's say m , such that *any* set of data (of numbers) having at least m elements satisfies the pre-given correlation. It is therefore also the case for a set of numbers, a database, produced by a random process, by throwing of dice or quantum measurements: The correlation will appear there as well and will then be “spurious”, because it occurs by chance. In other words, the presence of correlations may only depend on the *size* of the database and do not grant themselves in any case prediction nor taking action. In this way, in order not to think, these authors of algorithms who say they in principle ignore theories, including the ergodic theory, theories of algorithms and of finite combinatorics, that we applied for our result, are confronted with internal limits demonstrated by and within these theories: Spurious correlations inevitably infiltrates very large sets of numbers. This makes any kind of prediction risky when this one is not based on a thought that gives meaning and that enables us to choose what counts, beginning with the choice of what and how to measure and thus to associate it to a number. This thought is necessary to understand and theorize and, if possible, predict. Moreover, the power of scientific knowledge resides also in the fact of revealing the limits of the intended theory; this helps in better understanding, delimiting but also reinforcing the perspective that enables us to do science. Those who pretend to understand everything and to be able to make everything from a single object or concept, such as the DNA in biology or information, algorithms in all sciences ... they are certainly wrong. But I insist: the DNA,

information, algorithms... are all very important, the *science* around them is essential. The former is a fundamental chemo-physical trace of evolution, continuously employed by cells in order to produce proteins, largely using the Brownian motion of the proteome, which is the first functional random dynamics in biology. The latter, the algorithms, are about to change our lives, potentially for the better, if we keep in mind the limitations of any monomaniacal approach to knowledge (Longo, 2005, 2018).

In this “everything is information” or even “instructions and formal rules” craze, with the peak in the use of Big Data with no meaning, many often make reference to Alan Turing, inventor of the Logical Computing Machine (1936) – the mathematical foundation of computers – but with a dangerously incorrect interpretation of his work, see (Longo 2018b). We need to go beyond a collective imaginary in order to think the “next machine”, just as Turing was able to show the falsity of common hypotheses by his result of “non-computability”. In fact, between us, this digital machine and its powerful networks are perhaps a little boring, with their always identical iterations: If one initiates over and over a digital simulation of the wildest turbulence, with the same initial data, or if one opens a web page in Japan, they are always the same. But when faced with a non-linear physical dynamics such as a hurricane or a linguistic structure, all this is suddenly meaningless because no hurricane or human being will ever repeat anything identically. When digital information is networked, the sophisticated methods of “interleaving” and “semaphores” render the randomness, proper to the spatial-temporal continuum and of networks, ineffective. This randomness is considered as a “do not care”, as the experts—the real ones—say. In the areas of concurrency and network analyses in computer science, experts are able to make them function with computational certainty, according to the rules, eliminating noise. In fact, against computing’s common-sense view, which belongs more to molecular biology and to formally regulated governance, computer science today analyses and makes large use of randomness, including in continua, by stochastic calculus, analyses of networks dynamics, in techniques of coding, although in a way different from biology. Networks’ noise is thus analyzed and used in a very useful way, in cryptography for example. Yet, this has nothing to do with evolutionary “bricolage” of rare events, the unpredictable changes of the phase space (the space of observable and pertinent parameters), which are proper to biological dynamics and functional to them (Longo, 2016). The historical processes in natural and human sciences as well, economics in particular, require a different insight far away from the analysis by and the imposition of the optimal path in pre-given spaces of possibilities, see (Koppl et al., 2015).

It is the job of science to single-out the role of diversity of the human within the human, starting from the biological. Evolutionary diversity and adaptability is the result of the unexpected variant, the “hopeful monster” that is not noise which needs to be removed, like in information theories, or averaged out as in stochastic approaches, but it is that which makes life possible, according to Darwin's great intuition. This may also allow us to better use this fantastic discrete state machine for the elaboration of information, and its networks, and to invent new machines. Human beings will certainly invent yet another machine, if we do not identify this one and its algorithms for elaborating information with the world, if we abandon the necessity to work in the same computational modalities, evaluated, or with our actions predicted, by normalizing techniques without thought, like in the science handled by Big Data and bibliometrics. Making the founding principle explicit, in order to re-think them critically, is at the core of the scientific approach and its ethics. Science is not “elaboration of information”, even less the extrapolation from Big Data, but a historical construction of sense of the world, by a continual invention of

new concepts and structures. In order to trust science, its ethics requires a critique of one's own principles of knowledge, making a “step aside”, understanding different viewpoints. A good scientific practice is a major democratic training by this critical attitude, while contributing to democracy also by knowledge construction. It may only be the result of a difficult dialogue between remote or contrasting perspectives, a collaboration between different knowing subjects, based on ideas that may be unlikely but must be profound, proposed by small groups of thinkers who see things differently. Disagreement and diversity are not noise to be eliminated by the absolute dominance of the majority vote, such as the short term quotations in bibliometrics. They are instead at the core of science, since its invention by the democratic and the scientific debate on the Greek Agora. To do science, we need democracy, and science is an essential component of democracy, the only way to trust both.

Minimal References (more in the papers and references in <https://www.di.ens.fr/users/longo/download.html>)

- Anderson, C. (2008). The End of Theory: The Data Deluge Makes the Scientific Method Obsolete. *WIRED*. Retrieved from <https://www.wired.com/2008/06/pb-theory/>
- Audétat M. (ed) (2015). *Sciences et technologies émergentes: Pourquoi tant de promesses ?* Hermann, Paris.
- Baker, S. (2014) “Recognizing Paradigm Instability in Theories of Carcinogenesis”, *British Journal of Medicine & Medical Research*, 4(5): 1149-1163.
- Barbaresco, F and Mohammad-Djafari, A (Eds.) (2015). *Information, Entropy and Their Geometric Structures*, MDPI, Basel and Beijing.
- Bravi, B., & Longo, G. (2015). The Unconventionality of Nature: Biology, from Noise to Functional Randomness. In C. S. Calude & M. J. Dinneen (Eds.), *Unconventional Computation and Natural Computation* (Vol. 9252, pp. 3–34). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-21819-9_1
- Bizzarri M., *The New Alchemist. The Risks of Genetic Modification*. MIT Press, Boston, 2012.
- Buiatti, M., & Longo, G. (2013). “Randomness and multilevel interactions in biology.” *Theory in Biosciences*, 132(3), 139–158. <https://doi.org/10.1007/s12064-013-0179-2>
- Calude, C. S., & Longo, G. (2016). “Classical, quantum and biological randomness as relative unpredictability”, *Natural Computing*, 15(2), 263–278. <https://doi.org/10.1007/s11047-015-9533-2>
- Calude, C. S., & Longo, G. (2016a) “The Deluge of Spurious Correlations in Big Data”. *Found. of Science*, 1-18, March, 2016.
- Diamanti-Kandarakis E., Bourguignon JP, Giudice LC, Hauser R, Prins GS, Soto AM, Zoeller RT, Gore A.C. (2009). “Endocrine-disrupting chemicals: an Endocrine Society scientific statement”. *Endocr Rev*, 30, 293-342.
- Elowitz, MB, Levine, A, Siggia, E & Swain, P (2002) “Stochastic Gene Expression in a Single Cell”. *Science*, 297.
- Giuliani, A. (2010). “Collective motions and specific effectors: a statistical mechanics perspective on biological regulation” *BMC Genomics*, 11(suppl 1):S2.
- Koppl, R., Kauffman, S., Felin, T., & Longo, G. (2015). Economics for a creative world. *Journal of Institutional Economics*, 11(01), 1–31. <https://doi.org/10.1017/S1744137414000150>
- Longo, G. (2005). On the Relevance of Negative Results. *Intellectica*, (40) (in English: <http://www.di.ens.fr/users/longo/files/PhilosophyAndCognition/neg-resCE.pdf>).

- Longo, G. (2016) "How Future Depends on Past Histories and Rare Events in Systems of Life", *Foundations of Science*, pp. 1-32.
- Longo, G. (2016a) "A review-essay on reductionism: some reasons for reading *Reductionism, Emergence and Levels of Reality. The Importance of Being Borderline*", a book by S. Chibbaro, L. Rondoni, A. Vulpiani. **Urbanomic**, London, <https://www.urbanomic.com/document/on-the-borderline/>, May 8.
- Longo G. (ed.) (2017). *Lois des dieux, des hommes et de la nature*. Spartacus IDH, Paris.
- Longo, G. (2018) "Interfaces of Incompleteness" in Minati, G, Abram, M & Pessa, E (Eds.) *Systemics of Incompleteness and Quasi-systems*, Springer, New York, NY.
- Longo, G. (2018a) "Information and Causality: Mathematical Reflections on Cancer Biology", *Organisms. Journal of Biological Sciences*, vol. 2, n. 1, 2018.
- Longo, G. (2018b) "Letter to Turing", *Theory, Culture and Society. A Posthumanities Special Issue*, June 7, 2018 DOI: 10.1177/0263276418769733.
- Longo, G. (2019). "Information at the Threshold of Interpretation, Science as Human Construction of Sense". In Bertolaso, M. & Sterpetti, F. (Eds.) *Critical Reflection on Automated Science – Will Science Remain Human?* Springer, New York.
- Longo, G., & Montévil, M. (2014). *Perspectives on organisms: biological time, symmetries and singularities*. Berlin: Springer.
- Longo, G., Montévil, M., & Kauffman, S. (2012). "No entailing laws, but enablement in the evolution of the biosphere". ACM proceedings of the *Genetic and Evolutionary Computation Conference, GECCO'12*, Philadelphia (PA, USA), July 7-11.
- Longo, G. & Seno, L. (2018) "Digital networks, knowledge and "political" biases in their understanding and use" in *Philosophy of Internet and of the Hermeneutic Web*, (Stiegler et al., eds), M.I.T. Press, in preparation.
- Montévil M. 2020. Entropies and the Anthropocene crisis, *AI and Society*, to appear.
- Plomin R. 2018. *BLUEPRINT How DNA makes us who we are*, Penguin.
- Rafols I. et al. (2014). Big Pharma, little science?: A bibliometric perspective on Big Pharma's R&D decline. *Technological Forecast. Social Change*, 81, 22-38
- Seno, L. (2014). "Why the development engine broke down", *IEA Nantes*. Downloadable from: <http://www.acustica.org/publicat.htm#2014>
- Sikandar H. (2019). "Genome-Editing Technologies: Concept, Pros, and Cons of Various Genome-Editing Techniques and Bioethical Concerns for Clinical Application" *Molecular Therapy: Nucleic Acids* Vol. 16.
- Sonnenschein, C. and Soto, AM. (1999). *The society of cells: cancer and control of cell proliferation*. Springer.
- Soto, A. M., Longo, G., & Noble, D. (2016). "From the century of the genome to the century of the organism: New theoretical approaches". *Progress in Biophysics and Molecular Biology*, 122(1), 1–3.
- Vandenberg LN et al. (2012). "Hormones and endocrine-disrupting chemicals: low-dose effects and non-monotonic dose responses". *Endocr. Rev.* 33, 378–455.
- Weinberg, R 2014 "Coming Full Circle - form endless complexity to simplicity and back again", *Cell* 157, 3/27.