Discover Artificial Intelligence

Discover

Comment

The urgency of an algorethics

Paolo Benanti[1]

Received: 2 February 2023 / Accepted: 21 March 2023
Published online: 27 March 2023

Abstract
The following contribution was presented during the event, "Ethics of Artificial Intelligence: the Engagement of Abrahamic Religions in the Rome Call." The event was sponsored by the Vatican RenAIssance Foundation, the Abu Dhabi (UAE) Forum for Peace and the Commission for Interreligious Dialogue of the Chief Rabbinate of Israel. The Rome Call for AI Ethics (www.romecall.org), finalized in February 2020 and also signed at the time by some of the world's largest tech companies (Microsoft and IBM), along with the FAO and representatives of the Italian government, commits signatories to follow what its principles call for in terms of transparency, inclusion, accountability, impartiality, reliability, security and privacy. Religions have played and will continue to play a crucial role in shaping a world in which human beings are at the center of the concept of development. For this reason, an ethical development of artificial intelligence must be approached from an interfaith perspective. The potential of an interfaith event lies in the impact this message communicates. In the face of the radical transformations that digital and intelligent technologies are producing in society, the three Abrahamic religions together provide guidance for humanity's search for meaning in this new era.

1 Salutation

Most Reverend Excellencies, distinguished Authorities, Ladies and Gentlemen, dear Sisters and Brothers, "*Assalâm(u) 'aleikum*", "*shalom aleichem*", "*pax vobis*", peace to you. I am glad to be here this morning and to have this opportunity to share with you some reflections on the urgency of algorethics for the development of artificial intelligence.

2 What is technology?

Technology is present in our daily lives in numerous ways. We often make use of technological artifacts in such a spontaneous way that we can describe them as almost invisible. However, this "habit" in using technology may in fact make us like David Foster Wallace's famous fish:

> There are these two young fish swimming along and they happen to meet an older fish swimming the other way, who nods at them and says "Morning, boys. How's the water?" And the two young fish swim on for a bit, and then eventually one of them looks over at the other and goes "What the hell is water?"

This little story, which circulated well before Foster Wallace's retelling, carries with it a predictable, but not trivial, moral: sometimes we do not pay attention to or realize the most obvious things. Many aspects of daily living are unknown to us, precisely because they have always been present in the background of our existence. They become almost invisible.

✉ Paolo Benanti, benanti@unigre.it | [1]Pontifical Gregorian University, Rome, Italy.

This sort of filter on *the-continuously-present* in our experience has biological roots and characterizes our cognition. There is a fundamental mechanism, evolutionary in origin, which is necessary for our survival: we are open to perceiving the unusual and the unexpected, and we tend to shift the usual and the continuous into a band of lesser attention, even to the point of making it disappear. Some speak of this phenomenon as a true human faculty: a capacity to incorporate technological artifacts into our lives.

3 Technology as a sign of human surplus

Human beings live in a condition that elsewhere I have called *techno-human* [1]. What does all that mean?

When we talk about the *techno-human condition,* we do not mean to point only to a recent period in the history of civilization or to a particular set of technologies that could in some way threaten human beings and their existence today. The *techno-human condition* refers more generally to the way humans have always understood and lived their existence: an interaction with the environment mediated through tools, technological artifacts. It seems clear, then, that it is not possible to separate the history of human beings and civilization from the history of the tools they made.

Suffice it to say that archaeology calls *Palaeolithic* the era in which *Homo erectus* and *Homo sapiens* appeared because this is also the era to which the earliest human industry dates, namely that of the ancient Stone Age. Humans, chipping stone, made a tool out of it.

Philosophical anthropology, developed in the first half of the twentieth century, addresses the question of technique: since humans are *biologically deficient beings*, they are incapable of surviving in any natural environment and are therefore forced to act to construct their *place in the world*, making use of technique. According to this perspective, humans would be physiologically inferior to animals in that they lack specialized organs. Their survival therefore depends on their ability to compensate, by means of tools, for their natural deficiencies: only humans are living beings capable of transcending their own biological condition, starting from a marked limitation of the same [2].

The final product of such techniques represents something completely new, which not only has no counterpart in nature, but even stands as an anomaly, being an artificial human creation. This leads to the conclusion that "technique is also, like human beings, *nature artificielle*" [3].

But is it possible that it is a deficiency that sustains the inventiveness and specificity of who we are? If we look closely at humans and their place in history, it seems to us that we need to turn perspectives upside down.

There is a surplus to human beings' biological constitution. This surplus condition is, for example, what we experience when we need to take notes during a lecture. Our biological condition, our memory, is insufficient to retain the experience we live, and we need some technological artifacts, the pen and the notebook, to recall, express and transmit what we experience. Human beings then do not relate to reality in a merely biological way, but through mediations offered by technological artifacts.

Technology is human beings' way of channelling and expressing their surplus vis-à-vis their biological condition. It is thanks to technological artifacts that, as a species, we have become a global phenomenon.

In other words, what in all other species is provided by the genetic code, in our case is borrowed from technological artifacts. The capabilities that other animals have, are given to them through genetic competence and can generally change through great efforts of adaptation, which can lead to the mutation of their DNA. This is not the case for humans. Through technology we change the world and ourselves to inhabit the world. Technology is therefore a great gift that human beings possess and should be seen and understood in all its potential.

4 Language as technology

Of all the technologies that humankind has developed, I would now like to focus on a particular technological artifact: language. Today we have come to recognize that language is in fact a technology, human beings' invention, or rather the great invention that gives shape and substance to every other invention: the invention of inventing and communicating the invisible.

Over time, humans have gradually improved this new communication technology. Language paved the way for symbolic and hypothetical thinking. Language allowed humans to learn things they had not experienced first-hand. It enabled more than that. By combining linguistic elements in different ways, speakers could provide instructions to

imagine an unlimited range of things—not just things their listeners had not experienced, but things no one had: they could instruct them to imagine what could happen, what should happen, even what could not happen. Gradually, they discovered that they could use this ability creatively, to tell stories, create myths, or even deceive each other. And, crucially from our point of view, they discovered that they could use it to propose hypotheses.

Scholars agree that we may share a basic imagination capacity with other animals. What other animals seem to lack though is creative imagination, that is, the ability to combine different kinds of knowledge and ways of thinking to create new ideas and insights. [1]

Language, to instruct the listener's imagination, dissects the analog *continuum of* emotional experience, naming certain points of this experience that have been identified by community members: happiness, sadness, anger, jealousy, trust.[2]

Since we invented the technology of language, it has continued to develop uninterruptedly. However, from the moment our ancestors began to systematically build language bridges—always fragile and uncertain—over the experiential gaps that separated them, a new era in human life began: the instruction of imagination changed everything.

We learnt to imagine events that happened to others, and we learnt to take them into account in our decisions. Memories of the past gradually turned into new objects of communication. But then, if language is a technology we have developed, the greatest invention that has enabled us to communicate the very act of inventing, our question must now be: what kind of social entity is language?

Language is essentially a communication technology. This means, to a first approximation, that language must be classified along with the other communication technologies invented by humans, such as the book, the fax machine, the telephone, computer games, and social media—not with social institutions (such as government or the family), or cognitive capacities (such as vision or rationality).

Like other technologies, language must be built before it can be used, and it constantly develops and changes as a result of use. It spreads and propagates like other technologies, and like them it has its experts, role models, innovators, guardians, rebels, users, and abusers—active and passive at different levels.

Language is the first communication technology we invented. It revolutionized human life and effectively changed us as a biological species. It is still the most powerful technology we use today.

However, language cannot be understood as a *general-purpose* technology. We simply do not have—nor do I believe we ever will as a species—such a technological capability. If language were *general-purpose,* we would never need to invent anything else.

On the contrary, we must remember that every communication technology is functionally specific. Each technology employs a specific functional strategy, and the specificity of the strategy determines the functional envelope of the technology: what it can do with high levels of efficiency, how and where its efficiency decreases, where it becomes collaborative, and where it cannot go.

Thanks to language we have, as humans, a *place of sociality* that remains inaccessible to other living things. Thanks to language, our species, despite limited natural endowments compared to others, i.e. strength, speed, and instincts, has been able to climb the food pyramid and become the species that can dominate over others. This functionality of language is key to understanding its nature and role in our species and how it actively participates in the construction of the human individual as a social being.

Abrahamic religions look at language with extreme interest. The word is something that speaks of the divine, the invisible, the beyond. I do not wish to probe beyond this aspect here, but only to say that the technology of language and the experience of believers meet and intersect.

[1] We should point out that we will never be sure that other animals have the faculty of basic imagination. However, we can *imagine* with some degree of plausibility that, for example, when a chimpanzee produces a stick for termite extraction, it may be engaged in farsightedness – imagining the process of sticking the stick into the hole, extracting it, and then licking the termites out – and may also contemplate alternative choices, regarding both the size of the stick to be used and whether to eat termites or fruit. On this topic, current research and the plausibility of the presence of basic imagination in other living things we refer to [4] and [5].

[2] The idea of words identifying a point in the uninterrupted emotional experience is as fascinating and illuminating as ever in understanding the specifics of language understood as a technology. We should note that, as much as this strategy greatly aids mutual understanding between humans, it is not immediately to be understood as the highest or most effective form; in fact, it cannot be argued that any language, however rich and complex it may be, can replace a shocked expression, a muffled laugh, or a stream of silent tears, when it comes to communicating some of the most important messages of human social relations [6], [7].

Thanks to language, we have constructed a world: something that does not coincide with reality as such but is the result of our reading and interpreting reality.

Several scientific studies agree that other animals simply cannot communicate anything about their past experiences.[3] In comparing language with animal communication, scientific literature defines this human capacity as *displacement ability*: humans are apparently unique in being able to talk about things that are remote in space or time or both.[4] This is the meaning of Bertrand Russell's famous quote repeatedly found in the books of numerous linguists: "however eloquently a dog may bark, he cannot tell you that his parents were honest though poor."[5] Thanks to the technology of language, we have a story. Thanks to language we can also communicate something invisible but very important: a sense of moral duty. Again, thanks to language we can put limits on our freedom for the good of all: law is an artifact of language.

However, language is not a neutral tool. With language we can access the deepest and innermost thoughts of our neighbour. We can enter into dialogue. But, with language, we can also tell stories that never happened, make the other person imagine something that is not there. We can construct the enemy, we can lie, we can change the opinion of our fellow human beings. Language, like any technology, is deeply ambiguous.

5 The technology of language in its evolutions

But why are we looking at language? The key to this answer is found in the evolutions this technology has undergone throughout human history. Spoken language soon merged with written language. This transformation allowed an extension of the length of our words. The Abrahamic religions have also reflected on language, and their ancient wisdom is the bearer of great wealth.

Written language was followed by printed language, and the invention of printing meant that our words could be reproduced easily and smoothly. As we shall see in a moment, the invention of printing coincided with a real transformation in the way we understand the world. Today language experiences a new transformation. Language has become computed. Computers and artificial intelligences are the new tools of language and are the tools with which we educate and modify our understanding of the world and our fellow human beings. Before addressing these challenges, let us see what happened when printing was invented.

With the advent of printing and the dominance of sight over hearing, the world of thought and expression was also transformed. Movable type printing relentlessly places words into space because while writing simply transfers words from the world of sound to the world of visual space, printing fixes them within this space. "In printing, control of position is everything: "composing" the typeface by hand (which is the original form of typesetting) means placing already prepared characters by hand" [13].

Typography and its movable typefaces have definitively changed writing. The printable area is no longer organized as an ornamental and ornate surface, as in calligraphy, but is sharp, with perfectly regular lines, aligned to the right. Everything also appears clear visually, without the help of guidelines or margins that often appear on manuscripts.

If writing transfers the word into a visual space, making the invisible visible and not just audible, it is movable type printing that permanently roots the word in space.[6]

When printing produced its cultural effect, when the new technology of language transformed the sense through which we made the invisible visible, the book began to be seen as a kind of object that *contained* scientific information, narratives and other. It was no longer, as was the case previously, the recording of oral expression. And since, from a physical point of view, all printed books were the same, identical objects, while handwritten ones were not, this led to the use of labels.

The printed book, being an object made of letters, had a label in letters: the title page.

[3] See in this regard the famous article by Hockett [8], but also more recent ones [9] and [10].

[4] Even Dor agrees that the displacement ability seems to be decidedly lacking in the vocal signaling of primates, the closest *biological relatives* of humans, although it does occur in the dance of bees which grounds the hypothetical exceptions about bees that we noted earlier [11].

[5] The original expression appears in the first chapter of Part II of the celebrated *Human Knowledge*: "a dog cannot relate his autobiography; however eloquently he may bark, he cannot tell you that his parents were honest but poor" [12]. I wish to point out that this expression is quoted by both Chomsky and Dor constituting a kind of golden thread that intellectually binds all these authors.

[6] We will then see that with the digital, the written word is free from space as it is able to exist without any materiality.

Elizabeth Eisenstein, a U.S. historian, pointed out that when one was able to visually reproduce a message exactly and at the same time give an exact verbal description of physical reality, a *permanent Renaissance* was produced in the West [14, 15].

In fact, printing, with its way of proposing space and lines geometrically and visually, had, more or less directly, countless effects on the cognitive economy, that is, on the *mentality* of the Western world.

In other words, printing moved the ancient rhetorical art—orally based—from the core of academic education and encouraged and made possible a broad quantification of knowledge, through both mathematical analysis and graphs and diagrams. It is precisely this change of mindset that we must dwell on to understand the scope of this transformation [1].

For Alfred W. Crosby, U.S. academician of History, Geography, and American Studies, Europeans owe the fact that their culture became dominant in the modern age to science and technology.[7] The advantage acquired by the Westerners, Crosby argued, did not lay in these forms of knowledge, but in the ability to use thought patterns that could ensure rapid progress in both science and technology. At the same time, such thought patterns provided decisive skills in administrative, commercial, maritime, industrial, and military fields.

In other words, the characteristic that enabled Europeans to become *producers of technology* and to make science a technological tool is what French historians have called, speaking of other situations, a matter of *mentalité*.

What happened was that thanks to this transformation induced by printed language, a new way of understanding reality by modelling it emerged in Europe. There was the spread of a quantitative model of reality, capable of replacing its ancient qualitative predecessor, an idea also shared by Crosby.

> Copernicus, Galileo, the craftsmen who became experts at building cannons, the cartographers who mapped the coastlines of the newly discovered lands, the bureaucrats and businessmen who administered the new empires and companies in East and West India, the bankers who ordered and controlled the new monetary flows—all these people thought about reality in quantitative terms, with a consistency that was unmatched by men [15].

While it is true that these are the initiators of a real revolution, and while they should certainly be credited with great merit and innovation, we agree with Crosby in recognizing that they were the heirs of the changes in this *mentalité*, which had been fermenting for centuries.

It was only because of this profound and widespread change in their understanding of reality that their genius was able to express itself in the way we know.

With the invention of printing, we were able to change the way we understand reality. Space from quality becomes quantity. Air, water, earth, fire are no longer different types of space. Maps teach us to look at space as a distance between matter considered homotropic, as a quantity. The mechanical clock transforms time from something qualitative (before Christ and after Christ) and elastic (the length of hours in the winter and summer months) into something quantitative and fixed.

Visualizing the world as quantities, spaces and numbers is inextricably linked to the mutation that language has undergone in its transition from an oral to a printed artifact. In other words, in transforming itself, the linguistic artifact transformed the way we perceived the world we constructed with it: the linguistic artifact, a product of our intellect, in becoming printed language changed the way we perceived and defined our horizon.

The West has thus educated its mind to consider the universe in terms of minimal units of measurement, often thought of in the form of lines, squares, circles and other symmetrical figures. Subdividing things, energies, activities, perceptions into uniform parts and counting them quantified the world. Even today, we still aspire to conquer physical reality, grasping it, as it were, by that part that we see as the core, after removing the unnecessary.

The outcome of this revolution in mentality is illustrated in Newton's first principles of physics, which hypothesized the existence of absolute time and space. This transformation of *mentalité* was also considerably influenced by the Abrahamic religions: not only the contribution of the Mediterranean Muslims to the spread of the new numbers, but also the Jewish Kabbalah reading and the reflections of the Franciscans Raymond Llull and Luke Pacciola are signs of this role.

[7] We are greatly indebted to Crosby's work, which in our opinion presents an understanding of the Western scientific and technological phenomenon that is academically unparalleled. We will try in this part of the text to *vulgarize* his ideas by trying to make them more understandable and popularizable. We warn the reader that we will not quote in the footnotes punctual references to Crosby's work but point out in appropriate notes where we depart from it or where we should supplement it with other sources.

6 Information technologies as a sign of the greatest human power: the moon and the atomic bomb

In the early twentieth century, the human community was wired by the telegraph and then the telephone. Today, global connections take place via the computer: trading at the stock exchange, exchanges of money and goods, control of air and rail traffic, etc. take place through computers. The same route allows millions of people to exchange messages without the limits of time or space.

The revolution produced by computers and the information technology in the scientific-technological field has been radical to the point that today we are often convinced or create the illusion that by turning human problems into statistics, graphs, equations, these problems are solvable by calculation only. The advent of digital research, where everything is transformed into numerical data, leads to the ability to study the world according to new paradigms. The change is of the same magnitude as that introduced by printing.

Computation and digitization have confronted us with great new capabilities. We must not forget that if we got to the moon, it was because of a computer's ability to perform a series of calculations correctly and quickly, and to solve a series of differential equations that otherwise would have been very difficult.

Likewise, unfortunately, the explosion of the first atomic bomb is also related to the ability to solve differential equations with a machine calculating the impact of a uranium projectile on the mass of fissile matter, again, thanks to the computer.

Both events—one making us the only living species to leave planet Earth and the other making us potentially capable of destroying the human race and the entire planet—were made possible by the advent of this new language technology. If in the modern age we recognized nature as written in mathematical language, today we are able to use this new language with a machine and thanks to machines.

Beware, however: since their appearance, computers have shown two limitations, which we have been constantly trying to overcome over the past seventy years.

After writing their programmes, coders would use punch cards to load them onto their machines. If the programmes worked as expected, they would take the results and pass the machine to the next person. If, however, there was an error, they had to modify the programme and reload it later. Thus, computers were simply tools to perform the calculations presented to them. In addition, the central processing unit (CPU) was active only part of the time.

To improve computing performance and efficiency, computer scientists devised time-sharing, in which machines performed several tasks simultaneously, thus giving all users the impression that they were using the computer independently.

To give an idea of what the use of the computer meant in terms of time for writing and copying data in the first decade of its existence, we report what Herman Heine Goldstine—a mathematician and computer scientist who worked at Princeton University's Institute for Advanced Study and contributed to the development of ENIAC, the first electronic automatic digital computing machine—enthusiastically wrote on September 2nd, 1944: "to solve a quite complex partial differential equation of von Neumann's," it would take only 30 min out of which 28 would be needed "just in card cutting and two minutes for computing. The card cutting is needed simply because the solution of partial differential equations requires the temporary storage of large amounts of data. We hope to build a cheap high-speed device for this purpose" [16].

Cloud computing, IoT, Big Data, and the most sophisticated artificial intelligence algorithms are nothing but the result of our attempts to perfect these machines and be able to do more by improving computation execution time and data availability. This has led to the spring of artificial intelligence. Today, just as printing did, artificial intelligences have highlighted a new linguistic challenge. This, however, occurs at the boundary between humans and the machine: in the process of mutual questioning involving humans and the machine, projections and exchanges arise, hitherto unthought of, and as the machine becomes humanized, humans are becoming machinated [17, 18].

If in the second half of the last century, thanks to computers we took control over matter by becoming capable of going to the moon or destroying the world with the atomic bomb; today, thanks to this new season of machines, to this new way of managing the most powerful of our technologies, i.e., language, we can control people and their destinies.

Along with great promise, also great risks arise, which demand even greater responsibility. Thanks to technology we have saved humanity from a pandemic with catastrophic outcomes by accelerating research; but because of technology we are also making language which conveys a truth indistinguishable from that which simply spreads or produces hatred, causing a change in society and the world of work that may lead to the end of the middle-class and democracy itself.

In the face of these risks, an ethics based on precautionary principles is not enough. In other words, an engineers' ethic that only questions the process of producing a technology by ensuring standards is not enough. A well-designed weapon is no less dangerous by virtue of this care in design. It is not enough simply to make the production process transparent or to ensure that what is produced keeps the promises made. If we were to use the automobile metaphor, it is not enough to say that the brakes can stop the car in a given space and time. Nor is it enough, as has been done since the patent era, to make the production process transparent. Knowing what a brake looks like still tells me nothing about how I should drive the car. Faced with the challenges before us, the biggest question is not how to drive the car. The real question is what is our destination, where do we want to go and what is the direction to take at the next fork along the road? What is our destination? What is the purpose of our journey? These are sapiential questions.

These kinds of questions again challenge us. We can give our contribution to such questions by answering them not through ready-made answers, but through a wealth of timeless sapiential riches that have already accompanied humanity in the epochal transformations it has gone through.

Our contribution goes in the direction of helping humankind decide for itself. All of us here represent the generation entrusted with the ability to make this planet live in peace or in a worst-case scenario.

What will happen depends on how willing we are today to address what needs to be put in place to prevent this car, in keeping with our metaphor, from flying off the road. We need ethical guardrails. We need algorethics. We need each other, in a sincere dialogue capable of recognizing the good that technology can bring to humanity, but also of reminding human beings that how that technology will be used depends on their hearts and decisions.

Since the first day a member of our species picked up a club in a cave, he or she has had to question whether that was a tool to produce good or a weapon to offend. We need a new alliance among the women and men of our planet to question the ends of our actions.

Today, together we build a public space, the *Rome Call for AI Ethics*, where these questions can urgently resonate within our global society.

Today we say, together, forcefully: it is important for us to ask how we want to direct technology; it is essential for us to seek to build a tomorrow of peace and prosperity. We do not want to leave any woman or man behind along this path toward a future that is rich in promise but also fraught with risks.

Our human condition convinces us that technology is a gift. But our sapiential knowledge—this is what the Abrahamic religions are here today to remind us—tells us that our existence is always marked by possibility: for good or evil.

To choose good and avoid evil, we need ethics. To do this today, with the help of artificial intelligence, we urgently need algorethics.

Author contributions Paolo Benanti wrote the manuscript. The author read and approved the final manuscript.

Data availability We do not analyse or generate any datasets, because our work proceeds within a theoretical and philosophical approach.

Declarations

Competing interests The authors declare no competing interests.

References

1. Benanti P. Homo faber: the techno-human condition. Bologna: EDB; 2018.
2. Pansera MT. L'uomo progetto della natura. L'antropologia filosofica di Arnold Gehlen, Studium, Rome, 1990.
3. Gehlen A. L'uomo: la sua natura e il suo posto nel mondo. Milan: Feltrinelli; 1983. p. 33.
4. Mithen S. Seven steps in the evolution of human imagination. In: Proceedings of the British Academy. 2007;147: 3–29

5. Scruton R, et al. Imagination. In: Davies S, et al., editors. A companion to aesthetics. New York: Wiley-Blackwell; 2009. p. 346–50.
6. Aitchison J. Language change: progress or decay? Cambridge: Cambridge University Press; 2001.
7. Deacon TW. La specie simbolica: coevoluzione di linguaggio e cervello. Rome: Giovanni Fioriti Editore; 2001.
8. Hockett C. The origin of speech. Sci Am. 1960;203:88–96.
9. Bickerton D. Adam's tongue: how humans made language, how language made humans. New York: Hill and Wang; 2009.
10. Mufwene S. Language as technology: some questions that evolutionary linguists should address. In: Lohndal T, editor. In search of universal grammar: from old Norse to Zoque. Amsterdam: John Benjamins; 2013. p. 327–58.
11. Dor D. The instruction of imaginations: language as a social communication technology. Oxford: Oxford University Press; 2015.
12. Russel B. Human knowledge: its scope and limits. New York: Routledge; 1948. p. 74.
13. Ibid, 178–179.
14. Eisenstein EL. Le rivoluzioni del libro. L'invenzione della stampa e la nascita dell'età moderna. Bologna: Il Mulino; 1995. p. 55–102 and 123–160.
15. Crosby AW. La misura della realtà. Nascita di un nuovo modello di pensiero in Occidente. Dedalo: Bari; 1998. p. 7.
16. Goldstine HH. The computer from Pascal to von Neumann (ed. or. 1973). Princeton: Princeton University Press; 1980. p. 198.
17. Benanti P. The Cyborg. Corpo e corporeità nell'epoca del postumano, Cittadella, Assisi, 2012.
18. Benanti P. Digital Age. Teoria del cambio d'epoca. Persona, famiglia e società, San Paolo, Cinisello Balsamo, 2020.