American Scientist

Digital Ethics Online and Off

LUCIANO FLORIDI

Molding the digital future as it simultaneously shapes us

ETHICS TECHNOLOGY

In 1964, the year I was born, Paramount Pictures distributed *Robinson Crusoe on Mars*. The movie described the adventures of Commander Christopher "Kit" Draper (Paul Mantee), a United States astronaut shipwrecked on Mars. Watching it on YouTube recently reminded me how radically the world has changed in just a few decades. The computer at the very beginning of the movie looks like a Victorian engine, with levers, gears, and dials—a piece of archaeology that Dr. Frankenstein might have used. The only truly prescient element of techno-futurism comes toward the end of the story, when the character Friday (Victor Lundin) is tracked by an alien spacecraft through his bracelets.

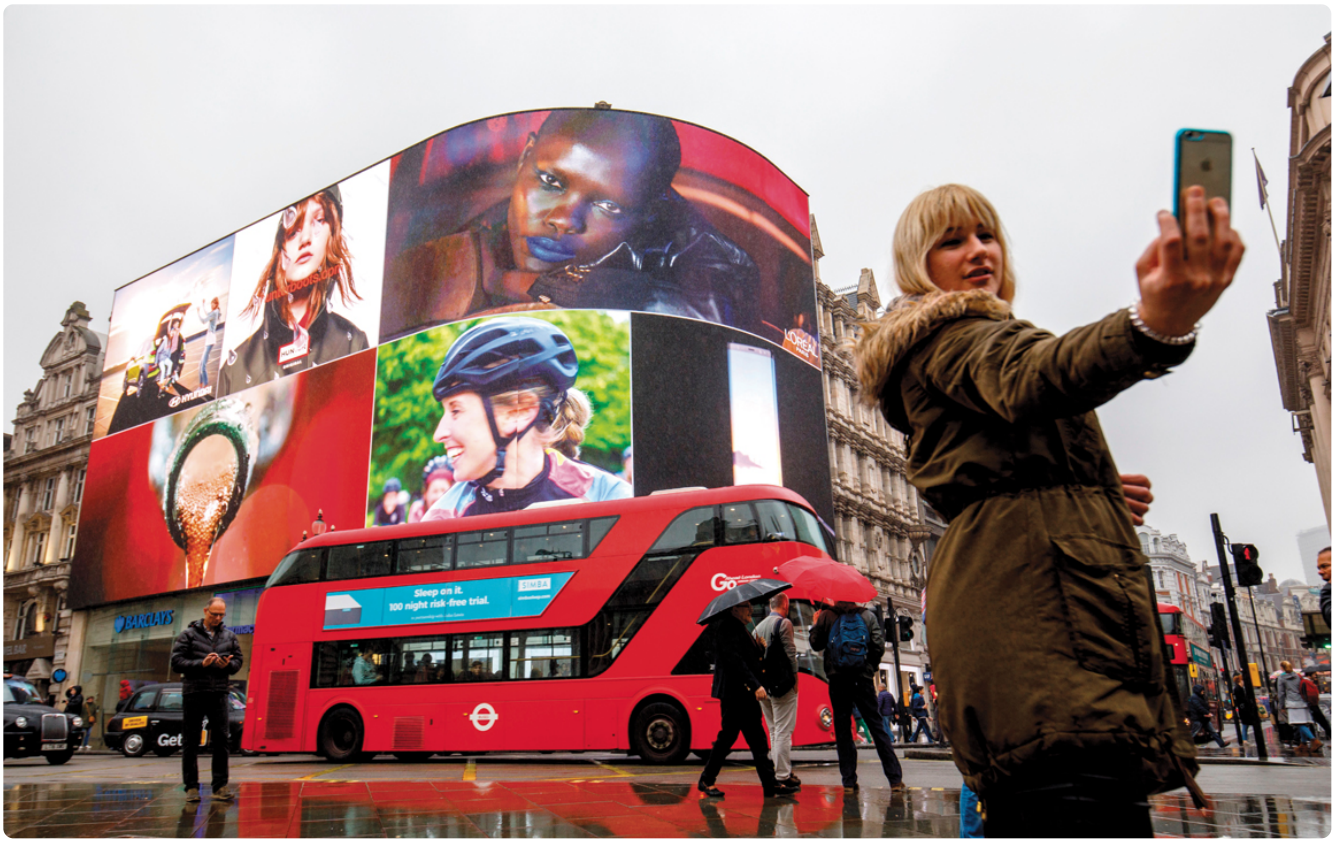
Robinson Crusoe on Mars belongs to a different age, one that was technologically and culturally closer to the previous century than to ours. The film describes a *modern* but not a *contemporary* reality, based on *hardware* not on *software*. Laptops, the internet, web services, touchscreens, smartphones, smart watches, social media, online shopping, streamed videos and music, driverless cars, robotic mowers, and virtual assistants were still to come. Artificial intelligence (AI) is mainly a project, not a reality. The movie shows a technology that is made of nuts and bolts, and mechanisms that follow the clunky laws of Newtonian physics. Even the iconic *Star Trek*, which debuted two years later, still imagined a future reliant on clunky mainframe computers, manual labor, and face-to-face social interactions.

QUICK TAKE

- Computers now generate so much information that data and algorithms shape our experience of the world. Systems have shifted from logic to parsing statistical trends.
- To navigate the digital world ethically, we must build a framework for weighing privacy, innovation, human rights, discrimination, and more.
- We now live in digital and physical spaces simultaneously. Grappling with these complex ideas and ethical dilemmas will help us to build a stronger future.

People born after the early 1980s have inhabited a totally different reality. To them, a world without digital technologies is like what a world without cars was for me: an abstract concept that I had only heard described by my grandmother. The social and ethical impacts of digital technology are now so deeply embedded that they can be difficult to perceive, much less comprehend.

Today's smartphone packs far more processing power in a few centimeters, and at an almost negligible cost, than NASA could put together when Armstrong landed on the Moon five years after *Robinson Crusoe on Mars*. The Apollo Guidance Computer on board Apollo 11 had 32,768 bits of random-access memory (RAM) and 589,824 bits (72 KB) of read-only memory (ROM): You could not have stored this issue of *American Scientist* on it. Fifty years later, your average phone comes with 4 GB of RAM and 512 GB of ROM. That is about 1 million times more RAM and 7 million times more ROM. As for the processor, the Apollo Guidance Computer ran at 0.043 megahertz. An average iPhone processor runs at 2,490 megahertz, about 58,000 times faster. To get a better sense of the acceleration, a person walks on average at 5 kilometers per hour, but a hypersonic jet travels slightly more than a thousand times faster at 6,100 kilometers per hour, just over five times the speed of sound. Only the most extreme spacecraft, such as the new Parker Solar Probe, can beat your walking speed by a factor of 58,000.



Pedestrians and passengers move through public spaces, such as London's Piccadilly Circus, with their smartphones, checking messages or taking selfies, and live in the digital and physical worlds simultaneously. Meanwhile, artificial intelligence algorithms are at work, using cameras to collect information about people and vehicles to display targeted ads on the large digital billboard.
AFP/Getty Images

Yet, today's world does not feel 58,000 times faster than it did when I was young. Where did all this speed and computational power go? The answer is twofold: feasibility and usability. We can do more and more in terms of applications, and we can do so in increasingly easy ways, not only in terms of programming, but above all in terms of user experience. Videos and operating systems software are computationally very hungry. Today's common forms of AI—such as the programs used to recommend videos on YouTube, set the price for an Uber ride, or select the ads you see online—are possible because we have the computational power required to run their software.

Today, thanks to this mindboggling growth in storage and processing capacities, at increasingly affordable costs, billions of people are connected and spend many hours online daily. Americans, for example, spend an average of 6.31 hours on the internet daily, just a bit less than the global average of 6 hours 41 minutes. The pandemic has surely driven these numbers even higher. AI is possible today also because we, humans, increasingly spend time in digital contexts that are AI-friendly.

Managing the Data Deluge

More memory, more speed, and more digital environments and interactions have generated immense quantities of data. We have all seen diagrams with exponential curves, indicating quantities that we do not even know how to imagine. According to market intelligence company IDC, in 2018, we reached 18 zettabytes of data created, captured, or replicated—that is 18×10^{21} bytes. This astonishing growth of data shows no sign of slowing down: According to IDC's projections, total data consumption will grow to 175 zettabytes in 2025. This number is hard to grasp, both in terms of quantity and significance.

Two consequences deserve a moment of reflection. The speed and the memory of our digital technologies are not growing at the same pace as the data universe, so we are quickly moving from a culture of recording to one of deleting. The question is no longer what to save but what to delete in order to make room for new data, transforming the idea of archiving the past. And most of the data available have been created since the 1990s, even if we

include every word uttered, written, or printed in human history and every library or archive that ever existed. Just look at any diagram illustrating the data explosion: It's not just the right-hand side, where growth skyrockets, but the left-hand corner, with so little data only a handful of years ago. Because almost all our data have been created by the current generation, they are also aging together, in terms of support and obsolete technologies, so their curation will be an increasingly pressing issue. Floppy disks from the 1980s are now obscure curios; compact disks from the early 2000s are quickly following them.

The real challenge is not innovation within the digital world, but the governance of the digital ecosystem as a whole.

More computational power and more data have made possible the shift from logic (if A then B) to statistics (A is related to B). Algorithms have progressed from library searches that could find a source by keyword to Amazon's tools that can parse your buying habits and recommend new books for you. Neural networks that were interesting only theoretically have become ordinary tools in machine learning and are used daily in medical diagnostics, to issue loans and credit cards, and in security checks. They are not really meant to help us understand cognition; they are just tools that make classifications. Early versions of AI were mostly symbolic and could be interpreted as a branch of mathematical logic, but in its current iterations AI is mostly connectionist,

seeking out complex patterns within datasets, and could be interpreted as a branch of statistics. AI's main war horse is no longer logical deduction, but statistical inference and correlation.

Computational power and speed, memory size, amount of data, algorithms, statistical tools, and online interactions have had to grow at breakneck speed to keep up with one another. The number of digital devices interacting with each other is already several times higher than the human population, but the causal connection goes both ways. So, most communication is now machine-to-machine, with no human involvement. We have computerized robots like Perseverance and Curiosity roving around on Mars, remotely controlled from Earth. Commander Christopher "Kit" Draper would have found them utterly amazing.

All of these trends will keep going, relentlessly, for the foreseeable future. The economic, cultural, and technological trends propelling them forward show no sign of abating. The expansion of the digital world has changed how we learn, play, work, love, hate, choose, decide, produce, sell, buy, consume, advertise, have fun, care and take care, socialize, communicate, and so forth. It seems impossible to find a corner of our lives that has not been affected by the digital revolution. In the last half-century, our reality has become increasingly digital, made of zeros and ones, run by software and data rather than by hardware and atoms.

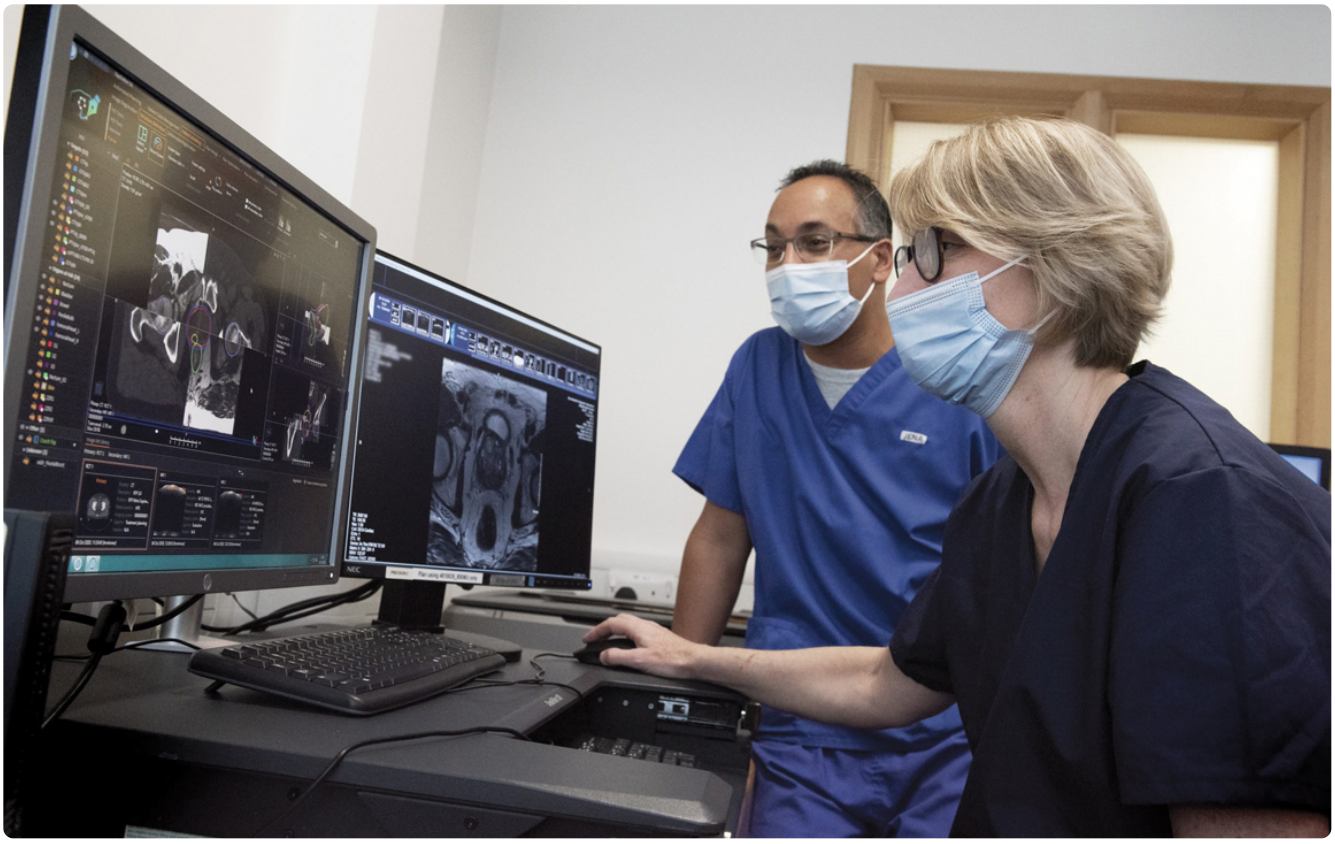
More and more people live increasingly *onlife*, both online and offline, and in the *infosphere*, both digitally and analogically. We use Instagram and WhatsApp to keep in touch with our friends, and we compare prices online even when we are standing in a shopping mall. This digital revolution is not merely technological. It affects how we conceptualize and understand our realities, increasingly interpreted in computational and digital terms. People now routinely refer to DNA as genetic "code," even though the idea is just a few decades old. The digital revolution has also fueled the development of AI. We now routinely share our onlife experiences and our infosphere environments with smart agents, whether they are algorithms, bots, or robots. They represent an unprecedented form of agency, one that does not need to be intelligent to be successful; for example, computers can play chess or Scrabble better than any of us. Netflix can guess what I want to watch next even before I know it.

Navigating Digital Ethics

What I have described so far, the digital revolution, provides huge opportunities to improve private and public life, as well as our environment. Consider, for example, the development of smart cities or the problems caused by carbon emissions. Algorithms can improve traffic flow and energy use; smart grids can reduce carbon impacts.

A recent report by Microsoft and PwC estimated that the use of AI in environmental applications could cut global greenhouse gas emissions by between 1.5 percent and 4 percent, while boosting global GDP by between 3.1 percent and 4.4 percent. Unfortunately, such opportunities are also coupled to significant ethical challenges. Examples include the extensive use of increasingly more data—often personal if not sensitive (Big Data)—the growing reliance on algorithms to analyze them in order to shape choices and to make decisions (including machine learning, AI, and robotics), and the gradual reduction of human involvement or even oversight over many automatic processes. These applications pose pressing issues of fairness, responsibility, and respect of human rights, among others—for example in predictive policing or when using facial recognition for monitoring behavior.

The ethical challenges posed by digital technologies and practices can be addressed successfully. Boston, Portland, and San Francisco are among the cities in the United States that have banned the indiscriminate use of facial recognition by law enforcement. Carefully and ethically deployed, however, the same technology has helped identify thousands of missing children in New Delhi.



Artificial intelligence can help speed the planning of complex medical procedures. Machine learning tools, such as those used with Microsoft's InnerEye technology, can sift through dozens of computed tomography images, finding data connections and marking the boundary between tumors and healthy tissue for a doctor's review. Oncologists sometimes need hours to go through such scans manually. AI medical tools can be more than 10 times faster.
Cambridge University Hospitals

Fostering the development and applications of data-based and software-based innovations, while ensuring the respect of human dignity, of the wellbeing of the environment, and of the values shaping open, pluralistic, and tolerant information societies is a great opportunity of which we can and must take advantage. Building such a robust alliance between the green of all our environments (natural as well as artificial) and the blue of our digital technologies will not be an easy or simple task. But the alternative—failing to advance and leverage ethically all digital technologies, their practices, and sciences—would have regrettable consequences.

On the one hand, overlooking ethical issues may prompt negative impact and social rejection, as was the case with the British National Health System's care.data program, a failed project to extract data from the offices of general practitioners into a central database. Social preferability and environmental sustainability must guide any digital development. On the other hand, overemphasizing the protection of individual or collective rights in the wrong contexts may lead to excessively rigid regulations, which may harm the chances to harness the social and ecological value of digital innovation. For example, the LIBE amendments, initially proposed to the European Data Protection Regulation and abandoned only after objections from the research community, would have made medical research more difficult by severely restricting scientific access to patient data records.

The demanding task of digital ethics is to maximize the value of digital innovation to benefit individuals, societies, and environments while navigating between social rejection and legal prohibition. To achieve this, digital ethics can build on the foundations provided by *computer ethics* since the 1950s, and on the later discipline of *information ethics*. This valuable legacy grafts digital ethics onto the great tradition of ethics more generally.

Within a few decades, we have understood that the most useful focus of our ethical strategies is not on a specific technology (computers, mobile phones, online platforms, cloud computing, and so forth) but on what is done with any digital solution. The shift from computer and information ethics to digital ethics

highlights the need to consider not only the technologies and sciences involved, but also the contexts and the applications (in business or in politics, for example), and the corresponding practices.

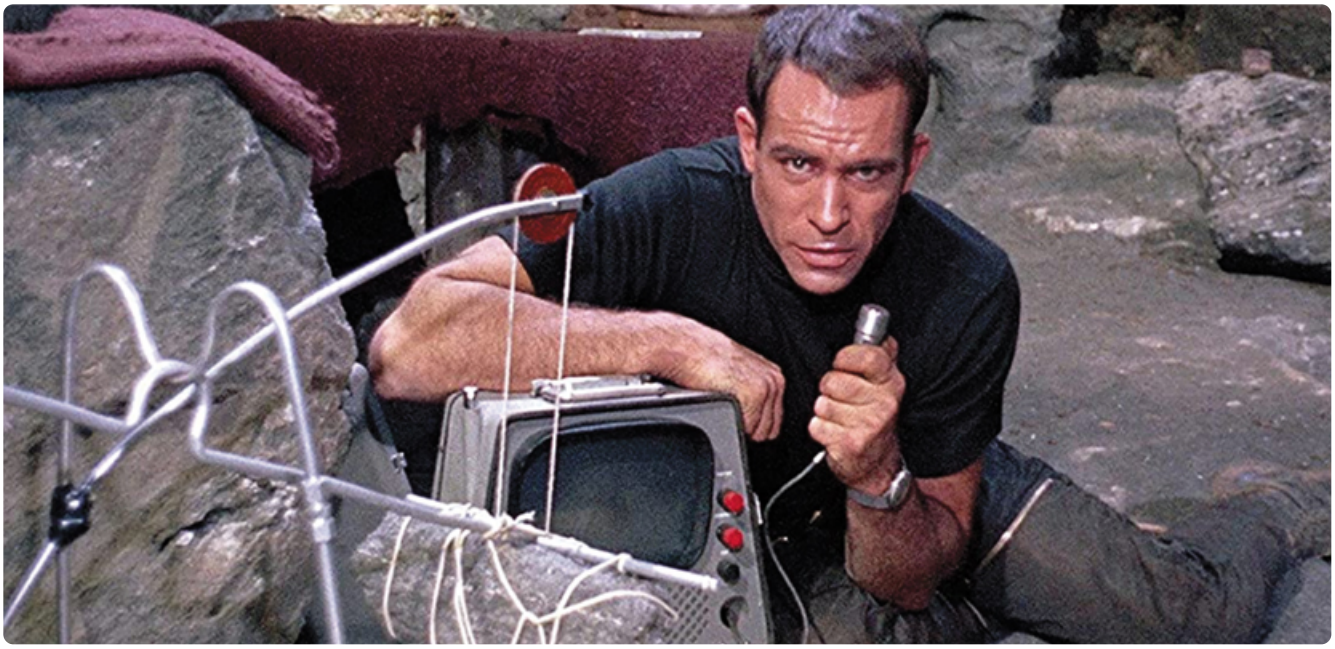
Digital ethics concerns the overall impact of the digital world, broadly construed, and narrower discussions of concepts such as “robo-ethics” or “machine ethics” miss the point. The ethical challenges brought about by the digital revolution—including privacy, anonymity, responsibility and accountability, transparency and explainability, and trust—concern a wide variety of digital phenomena, and hence they are better understood at an ecosystem level. The real challenge is not innovation within the digital world, but the governance of the digital ecosystem as a whole.

An Ethical Braid of Ideas

Today, digital ethics is a full branch of ethics that studies and evaluates moral problems related to *data* (including generation, recording, curation, processing, dissemination, sharing, and use), *algorithms* (including AI, artificial agents, machine learning, and robots), and corresponding *practices* (including responsible innovation, programming, hacking, and professional codes) in order to formulate and support morally good solutions (such as right conducts, values, and policies). There is growing interest in these fields but also a lot of confusion, which generates false hopes and ungrounded fears—for example, about what AI may really deliver. It won’t be a panacea and has nothing to do with the AI of Hollywood movies. Right now, there are many airport bestsellers

and too few PhD theses. In particular, we need more experts with the multidisciplinarity required by the field. But people are increasingly talking about the right issues in the right places: not only in academia, but also in companies and in governments.

The *ethics of data* focuses on problems posed by the collection and analysis of datasets and on issues relating to their use, such as Big Data in biomedical research and social sciences, social profiling and advertising, open data, and data philanthropy. Some of the key issues my colleagues and I debate concern the possible re-identification of individuals through the mining, linking, merging, and reuse of data. Risks extend to “group privacy,” when the identification of types of individuals, independently of the de-identification of each of them, may lead to serious ethical problems, from group discrimination (including ageism, ethnicism, sexism, and more) to group-targeted forms of violence. A classic example was provided by the algorithm running Amazon’s same-day shipping service. After it was launched, the company adjusted its algorithm to end discriminating against predominantly Black neighborhoods in many major cities. Content moderation, freedom of speech, and trust are also crucial topics of ethical debate.



Commander Christopher "Kit" Draper (played by actor Paul Mantee) finds himself shipwrecked on Mars, in the 1964 science fiction film *Robinson Crusoe on Mars*. The trailer boldly promises "Tomorrow come to life . . . Today!" However, yesterday's futurism, full of dials and knobs, looks quaint and far different from our digital present.
Everett Collection

These discussions can also help address the limited public awareness of the benefits, opportunities, risks, and challenges associated with the digital revolution. For example, politicians and technology developers often promote transparency as a measure that may foster trust. However, these proposals often leave the public unclear what information should be made transparent, to what degree, and to whom.

The *ethics of algorithms* addresses issues posed by the increasing complexity and autonomy of algorithms broadly defined, including AI and artificial agents such as internet bots, especially in the case of machine-learning applications—for instance, image-recognition software or automated decision-making systems. In this case, crucial challenges include the moral responsibility and accountability of both designers and scientists with respect to unforeseen and

undesired consequences as well as missed opportunities. The opacity, ethical design, and auditing of algorithms, and the assessment of potential, undesirable outcomes (such as racial discrimination or the promotion of misleading or incendiary content on social media) are attracting increasing research.

Finally, the *ethics of practices* addresses questions concerning the responsibilities and liabilities of people and organizations, such as governments that implement smart cities technology, in charge of digital processes, strategies, and policies. The goal is to define an ethical framework to shape professional codes—such as avoiding racial and gender bias in facial recognition—toward responsible innovation, development, and usage, which can ensure ethical practices fostering both the progress of digital innovation and the protection of the rights of individuals and groups. Three issues are central in this line of analysis: consent, privacy, and secondary use.

Although they are distinct lines of research, the ethics of data, algorithms, and practices are intertwined. For example, analyses focusing on data privacy must address issues concerning consent and professional responsibilities; and ethical auditing of algorithms often implies analyses of the responsibilities of their designers, developers, users, and adopters. Digital ethics must address the whole conceptual space, even if with different priorities. It must avoid narrow, ad hoc approaches but rather address the diverse set of ethical implications of digital realities within an inclusive, integrated, and consistent framework.

Molding Our Digital Future

Life today has become inconceivable without the presence of digital technologies, services, products, and practices. Anyone not perplexed by such a digital revolution has not grasped its magnitude. We have begun a new chapter in human history. Of course, other chapters have come before, all similarly significant. Humanity has experienced a world before and after the wheel, ironworking, the alphabet, printing, the steam engine, electricity, television, and the telephone.

Each transformation is unique. Some of them have changed our self-understanding, our reality, and our experience irreversibly, with complex and long-term implications.

We are still finding new ways of exploiting the wheel; just think of the infotainment controls in cars. By the same token, what humanity will achieve using digital technologies is unimaginable. Nobody in 1964 could have guessed what the world would be like only 50 years later, as abundantly documented by the predictions on display at the 1964 New York World's Fair. And yet, the digital revolution happens only once. To get it going the right way, the time to start is now. Future generations will never know an analog-only, offline, predigital reality. We are the last generation to have experienced it.

The price of living at this special time in history is that we have to confront worrying uncertainties, and do so with limited understanding. The mind-blowing potential transformations justify some confusion and apprehension.

You just have to look at news headlines to see the pervasiveness of those responses. However, our special position and perspective in history also bring extraordinary opportunities. Precisely because the digital revolution has just begun, we can shape it in positive ways that benefit both humanity and our planet. As Winston Churchill said, “We shape our buildings; thereafter they shape us.” The same is true of our built digital environment.

We are at the very early stages of constructing our new digital realities. We can get them right, before they start affecting and influencing us and future generations in the wrong way. The debate over whether digital technologies are more beneficial on balance or more damaging is pointless. The technologies are coming, and the really interesting question is how we can apply them ethically. The real answer to whether the glass is half empty or half full is that it’s fillable. To identify the best path ahead in the development of our digital technologies, we need more and better digital ethics.

We should not sleepwalk into the creation of an increasingly digital world. The insomnia of reason is vital, because its sleep generates monstrous mistakes. Understanding the digital revolution happening under our gaze is essential, if we wish to steer humanity toward a future that is both environmentally sustainable and socially just.

BIBLIOGRAPHY

- Floridi, L. 2014. T*he Fourth Revolution: How the Infosphere Is Reshaping Human Reality*. New York: Oxford University Press.

- Wacks, R. 2015. *Privacy: A Very Short Introduction*. New York: Oxford University Press.
- Wiener, N. 1954. *The Human Use of Human Beings: Cybernetics and Society*. New York: Da Capo Press.