Soft ethics, the governance of the digital and the General Data Protection Regulation

Article *in* Philosophical Transactions of The Royal Society A Mathematical Physical and Engineering Sciences · November 2018

DOI: 10.1098/rsta.2018.0081

CITATIONS
2

READS
10

1 author:

Luciano Floridi
University of Oxford
390 PUBLICATIONS **6,754** CITATIONS

SEE PROFILE

PHILOSOPHICAL TRANSACTIONS A

rsta.royalsocietypublishing.org

Research

Cite this article: Floridi L. 2018 Soft ethics, the governance of the digital and the General Data Protection Regulation. *Phil. Trans. R. Soc. A* **376**: 20180081.
http://dx.doi.org/10.1098/rsta.2018.0081

Accepted: 20 July 2018

One contribution of 9 to a theme issue 'Governing artificial intelligence: ethical, legal, and technical opportunities and challenges'.

Subject Areas:
artificial intelligence

Keywords:
data governance, digital ethics, General Data Protection Regulation, soft ethics

Author for correspondence:
Luciano Floridi
e-mail: luciano.floridi@oii.ox.ac.uk

Soft ethics, the governance of the digital and the General Data Protection Regulation

Luciano Floridi[1,2]

[1]Oxford Internet Institute, University of Oxford, 1 St Giles, Oxford OX1 3JS, UK
[2]The Alan Turing Institute, 96 Euston Road, London NW1 2DB, UK

LF, 0000-0002-5444-2280

The article discusses the governance of the digital as the new challenge posed by technological innovation. It then introduces a new distinction between *soft ethics*, which applies after legal compliance with legislation, such as the General Data Protection Regulation in the European Union, and *hard ethics*, which precedes and contributes to shape legislation. It concludes by developing an analysis of the role of digital ethics with respect to digital regulation and digital governance.

This article is part of the theme issue 'Governing artificial intelligence: ethical, legal, and technical opportunities and challenges'.

1. The mangrove society: from digital innovation to the governance of the digital

Today, in any mature information society [1], we no longer live online or offline but *onlife*, that is, we increasingly live in that special space, or *infosphere*, that is seamlessly analogue and digital, offline and online. If this appears confusing, perhaps an analogy may help to convey the point. Imagine someone asking whether the water is fresh or salty in the estuary where the river meets the sea. Clearly, that someone has not understood the special nature of the place. Our mature information societies are growing in such a new, liminal place, like mangroves flourishing in brackish water. And in these 'mangrove societies', machine-readable data, new forms of smart agency and onlife interactions are constantly evolving, because our technologies are perfectly fit to take advantage of such a new environment, often as the only real natives. As a result, the pace of their evolution

THE ROYAL SOCIETY PUBLISHING

can be mind-blowing. And this in turn justifies some apprehension. However, we should not be distracted by the scope, depth and pace of digital innovation. True, it does disrupt some deeply ingrained assumptions of the old society, which was exclusively analogue, for example, about competition, customization, education, entertainment, health, logistics, politics, production, security or work, just to mention some crucial topics, in a merely alphabetic order. Yet that is not the most consequential challenge we are facing. It is rather how we are going to design the infosphere and the mature information societies developing within it that matters most. Because the digital revolution transforms our views about values and their priorities, good behaviour and what sort of innovation is not only sustainable but also socially preferable—and governing all this has now become the fundamental issue. Let me explain.

To many, what digital innovation will throw up next may seem the real challenge. The question itself is recurrent and trite: What is the next disruption? What is the new killer app? Will this be the year of the final battle between virtual reality versus augmented reality? Or is it the internet of things that will represent the new frontier, perhaps in some combination with smart cities? Is the end of TV as we know it coming soon? Will healthcare be made unrecognizable by machine learning, or should our attention rather be focused on the automation of logistics and transport? What will the new smart assistants in the home do, apart from telling us what the weather is like, and allowing us to choose the next song? How is military strategy going to adapt to cyber conflicts? Behind similar questions lies the unspoken assumption that digital innovation leads, and everything else lags behind, or follows at best: business models, working conditions, standards of living, legislation, social norms, habits, expectations, even hopes. Yet this is precisely the distracting narrative that we should resist. Not because it is wrong, but because it is only superficially right. The deeper truth is that the digital revolution has already occurred. The transition from an entirely analogue and offline world to one that is increasingly also digital and online will never happen again in the history of humanity. Perhaps, one day, a quantum computing gadget, running artificial intelligence (AI) apps, may be in the pocket of your average teenager, but our generation is the last one that will have seen a non-digital world. And this is the really extraordinary turning point. Because that landing on the infosphere and the beginning of onlife happen only once. What this new world will be like, as we create it, is both fascinating, in terms of opportunities, and worrisome, in terms of risks. But the 'exploration' of the infosphere, to indulge in the geographical metaphor a bit longer, no matter how challenging, prompts a much more fundamental question, which is socio-political and truly crucial: What kind of mature information societies do we want to build? What is our *human project* for the digital age? Looking at our present backwards—that is, from a future perspective—this is the time in history when we shall be seen to have laid down the foundation for our mature information societies. We shall be judged by the quality of our work. So, clearly, the real challenge is no longer good digital *innovation*, but the good *governance* of the digital.

The proof that this is the case is all around us, in the mushrooming initiatives addressing the impact of the digital on everyday life and how to regulate it. It is also implicit in the current narrative about the unstoppable and unreachable nature of digital innovation, if one looks just a bit more closely. Because in the same context where people complain about the speed of digital innovation, and the impossible task of chasing it with some normative framework, one also finds that there is equal certainty about the serious risk that the wrong legislation may kill digital innovation entirely or destroy whole technological sectors and developments. You do not have to be Nietzsche ('Was mich nicht umbringt macht mich stärker'—'What does not kill me makes me stronger' [2]) to realize that the inference to be drawn is that updating the rules of the game is perfectly possible—after all, everybody acknowledges that it can have immense consequences—but that reacting to technological innovation is not the best approach. We need to shift from chasing to leading. If we then like the *direction* in which we move, or *where* we are going, then the *speed* at which we are moving or getting there can actually be something very positive. The more we like our destination, the faster we will want to get there. It is because we lack a clear sense of socio-political direction that we are worried by the speed of our technological travelling. We should be. Yet the solution is not to slow down, but to decide together where we want to go.

For this to happen, society needs to stop playing defence and start playing attack. The question is not whether, but how. And to start addressing the how, some clarifications are helpful. This is the contribution made by this article.

2. Ethics, regulation and governance

On the governance of the digital, there is much to be said, and even more still to be understood and theorized, but one point is clear: the governance of the digital (henceforth *digital governance*), the ethics of the digital (henceforth *digital ethics*, also known as computer, information or data ethics [3]) and the regulation of the digital (henceforth *digital regulation*) are different normative approaches, complementary, not to be confused with each other, but to be clearly distinguished, in the following sense (see figure 1 for a visual representation).

Digital governance is the practice of establishing and implementing policies, procedures and standards for the proper development, use and management of the infosphere. It is also a matter of convention and good coordination, sometimes neither moral nor immoral, neither legal nor illegal. For example, through digital governance, a government agency or a company may (i) determine and control processes and methods used by data stewards and data custodians in order to improve the data quality, reliability, access, security and availability of its services; and (ii) devise effective procedures for decision-making and for the identification of accountabilities with respect to data-related processes. A typical application of digital governance was the work I co-chaired for the British Cabinet Office in 2016 on a 'Data science ethical framework' [4], which was '[. . .] intended to give civil servants guidance on conducting data science projects, and the confidence to innovate with data.'[1] Despite the title, many recommendations had nothing to do with ethics and concerned only reasonable governance.

Digital governance may comprise guidelines and recommendations that overlap with *digital regulation*, but are not identical to it. This is just another way of speaking about the relevant legislation, a system of laws elaborated and enforced through social or governmental institutions to regulate the behaviour of the relevant agents in the infosphere. Not every aspect of digital regulation is a matter of digital governance and not every aspect of digital governance is a matter of digital regulation. In this case, a good example is provided by the General Data Protection Regulation (GDPR, more on the GDPR presently).[2] *Compliance* is the crucial relation through which digital regulation shapes digital governance.

All this holds true of *digital ethics*, understood as the branch of ethics that studies and evaluates moral problems relating to *data* and *information* (including generation, recording, curation, processing, dissemination, sharing and use), *algorithms* (including AI, artificial agents, machine learning and robots), and corresponding *practices* and *infrastructures* (including responsible innovation, programming, hacking, professional codes and standards), in order to formulate and support morally good solutions (e.g. good conduct or good values) [3]. Digital ethics shapes digital regulation and digital governance through the relation of moral evaluation of what is socially acceptable or preferable.

Digital governance in figure 1 is just one of the three normative forces that can shape and guide the development of the digital. But it is not uncommon to use that part for the whole and to speak of digital governance as referring to the entire set. This is 'governance' as a synecdoche, a bit like using 'coke' for any variety of cola. It is what I did at the beginning of this article, when I stated that the real challenge today is the governance of the digital. By that I meant to refer not just to digital governance but also to digital ethics and digital regulation, i.e. to the whole normative map: E + R + G. And this is also how I interpret the report 'Data management and use:

[1]Available from https://www.gov.uk/government/publications/data-science-ethical-framework.

[2]Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), OJEU L119, 04/05/2016.

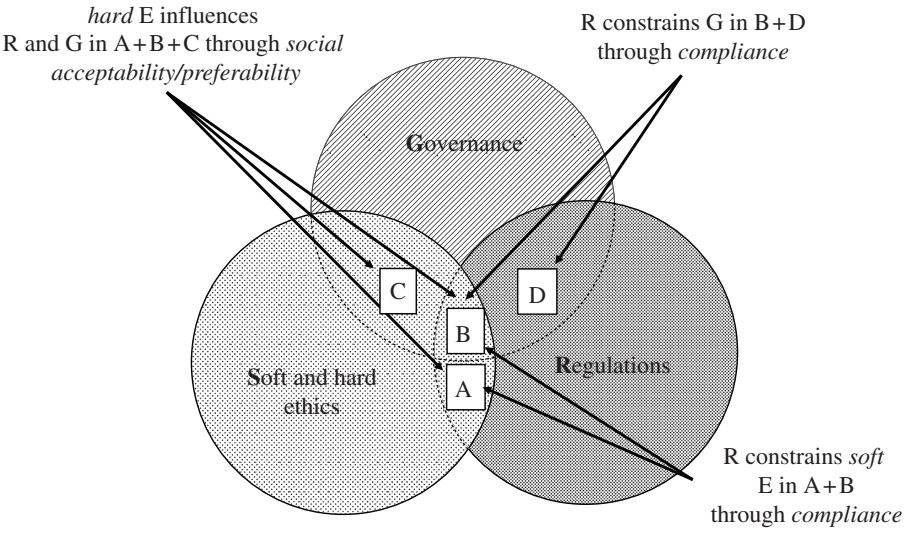

Figure 1. The relationship between digital ethics (E), digital regulations (R) and digital governance (G).

governance in the 21st century' that we published in 2017 as a joint British Academy and Royal Society working group [5]. As long as the synecdoche is clear, there is no problem.

Once the map is understood, some important consequences become clear. Let me discuss each of them in a separate section.

3. Compliance: necessary but insufficient

When policy-makers, both in political and in business contexts, wonder why we should engage in ethical evaluations when legal compliance is already available (this is a recurring topic in the discussion of the GDPR, for example), the answer should be clear: compliance is necessary but insufficient to steer society in the right direction. Because digital regulation indicates what the legal and illegal moves in the game are, so to speak, but it says nothing about what the *good* and *best* moves could be, among those that are legal, to win the game, that is, to have a better society. This is the task of both digital ethics, on the side of moral values and preferences, and of good digital governance, on the side of management. And this is why, for example, the European Data Protection Supervisor (EDPS, the EU's independent data protection authority) established the Ethics Advisory Group in 2015, in order to analyse the new ethical challenges posed by digital developments and current legislation, especially in relation to the GDPR. The report we published [6] should be read as a contribution to a normative governance of the infosphere in the European Union (EU), and a stepping stone towards its implementation. So what kind of digital ethics should we adopt, to complement legal compliance?

4. Hard and soft ethics

If we look at figure 1, digital ethics may now be understood in two ways, as *hard* and *soft ethics*. The distinction is above all a matter of theory—it is logically possible and often useful to distinguish soft and hard ethics and discuss each separately—not so much a matter of practice, because in reality soft and hard ethics often come intertwined inextricably.

Hard ethics (see A + B + C in figure 1) is what we usually have in mind when discussing values, rights, duties and responsibilities—or, more broadly, what is morally right or wrong, and what ought or ought not to be done—in the course of formulating new regulations or challenging existing ones. In short, *insofar* (and it may not be very far) as ethics contributes to making, shaping or changing the law, we can call that *hard ethics*. For example, lobbying in favour of

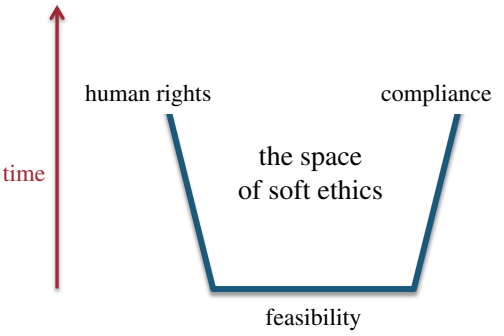

Figure 2. The space of soft ethics. (Online version in colour.)

some good legislation or to improve that which already exists can be a case of hard ethics. Hard ethics helped to dismantle apartheid legislation in South Africa and supported the approval of legislation in Iceland that requires public and private businesses to prove that they offer equal pay to employees, irrespective of their gender (the gender pay gap continues to be a scandal in most countries). It follows that, in hard ethics, it is not true that 'one ought to do x' (where x ranges on the universe of feasible actions) implies 'one may do x'. It is perfectly reasonable to expect that 'one ought to do x' may be followed by 'even if one may not do x'. Call this the Rosa Parks Principle, for her famous refusal to obey the law and give up her bus seat in the 'coloured section' to a white passenger, after the whites-only section was filled.

Soft ethics covers the same normative ground as hard ethics (again, see A + B + C in figure 1), but it does so by considering what ought and ought not to be done *over and above* the existing regulation, not against it, or despite its scope, or to change it, or to by-pass it, e.g. in terms of self-regulation. In other words, *soft ethics is post-compliance ethics* because, in this case, 'ought implies may'. This is why in figure 1 I wrote that regulations constrain software ethics through compliance. Call this the Matthew Principle, from Matthew 22:15–22: 'Render to Caesar the things that are Caesar's'.

As already indicated above, both hard and soft ethics presuppose *feasibility* or, in more Kantian terms, assume that 'ought implies can', given that an agent has a moral obligation to perform an action x only if x is possible in the first place. Ethics should not be supererogatory in this specific sense of asking for something impossible. It follows that soft ethics assumes a *post-feasibility* approach as well. Add that any ethical approach, at least in the EU, accepts, as its minimal starting point, the implementation of the Universal Declaration of Human Rights, the European Convention on Human Rights and the Charter of Fundamental Rights of the European Union. And the result is that the space of soft ethics is both partially bounded, and yet unlimited. To see why, it is easy to visualize it in the shape of a trapezoid (figure 2), with the lower side representing a feasibility base that is ever-expanding through time—we can do more and more things thanks to technological innovation—the two constraining sides, left and right, representing legal compliance and human rights, and the open upper side representing the space where what is morally good may happen in general and, in the context of this article, may happen in terms of shaping and guiding the ethical development of our mature information societies.

I already mentioned that hard and soft ethics often go hand in hand. Their distinction is useful but often logical rather than factual. In the next section, I shall analyse their mutual relation and their interaction with legislation by relying on the specific case provided by GDPR. In this section, a final clarification is in order.

When distinguishable, soft digital ethics can be more easily exercised the more digital regulation is considered to be on the good side of the moral versus immoral divide. Thus, it would be a mistake to argue for a soft ethics approach to establish a normative framework when agents (especially governments and companies) are operating in contexts where human

rights are disregarded, e.g. in China, North Korea or Russia. In other contexts, when human rights are respected, hard ethics may still be necessary to change some current legislation that is perceived to be ethically unacceptable. The Irish abortion referendum in 2018 is a good example. In a digital context, hard ethics arguments have been used to contrast the decision by the US. Federal Communications Commission (FCC) (December 2017) to rescind the rule about net neutrality (the principle according to which all Internet traffic should be treated in the same way, without blocking, degrading or prioritizing any particular legal content). The outcome is that, in March 2018, Washington became the first state in the USA to pass legislation mandating net neutrality. Within the EU, soft ethics may rightly be exercised to help agents (including individuals, groups, companies, governments, organizations) to take more and better advantage, morally speaking, of the opportunities offered by digital innovation. Because, even in the EU, legislation is necessary but insufficient. It does not cover everything (nor should it), and agents should leverage digital ethics in order to assess and decide what role they wish to play in the infosphere, when regulations provide no simple or straightforward answer, when competing values and interests need to be balanced (or indeed when regulations provide no guidance), and when there is more that can be done over and above what the law strictly requires. In particular, a good use of soft ethics could lead companies to exercise 'good corporate citizenship' within a mature information society.

Time has come to provide a more specific analysis, for which I shall rely on the GDPR. The choice seems reasonable: given that digital regulation in the EU is now determined by the GDPR, and that EU legislation is normally respectful of human rights, it may be useful to understand the value of the distinction between soft and hard ethics and their relations to legislation by using the GDPR as a concrete case of application. The underlining hypothesis is that, if the soft/hard ethics analysis does not work in the case of the GDPR, it probably won't work anywhere else.

5. Soft ethics as ethical framework

To understand the role of hard and soft ethics with regard to law in general and the GDPR in particular, five components need to be introduced (figure 3).[3]

First, there are the ethical, legal and social implications of the GDPR, e.g. on organizations. This is the impact of the GDPR on business, for example. Then there is the GDPR itself. This is the legislation that replaces the Data Protection Directive 95/46/EC. It is designed to harmonize data privacy laws across Europe, to protect and empower all EU citizens' data privacy, independently of geographical location, and to improve the way organizations across the EU approach data privacy. The GDPR comprises 99 Articles; this is the second element. As it is often the case with complex legislation, the Articles do not cover everything, leave grey areas of normative uncertainty even about topics that they do cover, are subject to interpretations and may require updating when applied to new circumstances, especially in a technological context where innovation develops so quickly and radically; think for example of face recognition software, or so-called deep fake software. So, to help understand their meaning, scope and applicability, the Articles are accompanied by 173 Recitals. This is the third element. Recitals, in EU law, are texts that explain the reasons for the provisions of an act, but are not legally binding, and are not supposed to contain normative language. Normally, Recitals are used by the Court of Justice of the European Union (CJEU) in order to interpret a Directive or a Regulation and reach a decision in the context of a particular case.[4] But in the case of the GDPR, it is important to note that Recitals can also be used by the European Data Protection Board (the EDPB, which replaces the Article 29 Working Party), when ensuring that the GDPR is applied consistently across Europe.

[3]In a previous version of this article the text read as if I argued that ethics shapes and interprets the law. This is simply untenable and I am grateful to one of the anonymous referees for highlighting this potentially erroneous reading.

[4]See for example 'C-131/12 Google Spain SL, Google Inc. v Agencia Española de Protección de Datos, Mario Costeja González', http://curia.europa.eu/juris/document/document.jsf?text=&docid=152065&pageIndex=0&doclang=EN&mode=req&dir=&occ=first&part=1&cid=257894. Or Domestic CCTV and Directive 95/46/EC (European Court of Justice (ECJ) Judgment in Case C-212/13 Ryneš), http://amberhawk.typepad.com/amberhawk/2014/12/what-does-the-ecj-ryne%C5%A1-ruling-mean-for-the-domestic-purpose-exemption.html.

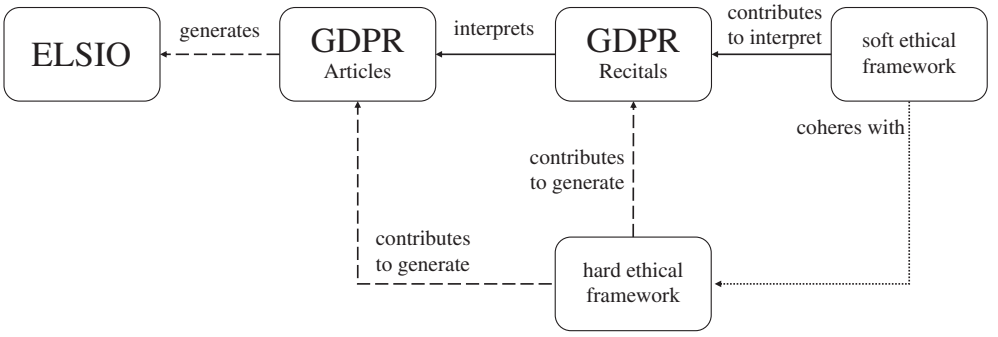

Figure 3. Soft and hard ethics and their relation to regulation. Note that the diagram is simplified by omitting references to all the other elements that contribute to the various frameworks.

The Recitals themselves will require an interpretation, and this is the fourth element. Part of this interpretation is provided by an ethical framework, which contributes, together with other factors, to understand the Recitals. Finally, the Articles and the Recitals were formulated thanks to a long process of negotiations between the European Parliament, the Council of Europe and the European Commission (the so-called Formal Trilogue meeting), resulting in a joint proposal. This is the fifth element, namely the perspective that informed the elaboration of the GDPR. This is where hard ethics plays a role, together with other factors (e.g. political, economic, etc.). It may be seen in action by looking at a comparative analysis of drafts from the European Parliament and European Commission and the amendments to the Commission's text proposed by the European Council.[5] So here is a summary of what we need to consider (figure 3):

(1) The ethical, legal and social implications and opportunities (ELSIO) generated by the Articles in (2). The distinction between implications and opportunities is meant to cover both what follows from the GDPR (implications) and what is left uncovered (partially or completely) by the GDPR. The reader who finds the distinction redundant (one may argue that opportunities are just a subset of the implications) should feel free to drop the O in 'ELSIO'. The reader who finds the distinction confusing may wish to add to the diagram another box, labelled 'opportunities', and another arrow, from the GDPR to it, labelled 'generates'. In figure 3 I adopted a compromise: one box double label. Note that opportunities need not be necessarily positive, they can be negative, also in the ethical sense of possible wrong-doings, e.g. the GDPR may enable one to exploit an ethically wrong opportunity.
(2) The Articles of the GDPR that generate (1).
(3) The Recitals of the GDPR that contribute to interpret the Articles in (2).
(4) The soft ethical framework that contributes to interpret the Recitals in (3) and the Articles in (2), that is coherent with the hard ethical framework in (5), and contributes to deal with ELSIO in (1).
(5) The hard ethical framework that contributes to generate the Articles in (2) and the Recitals in (3).

Hard ethics in (5) is the ethical element (together with others) that motivated and guided the process leading to the elaboration of the law, in this case the GDPR. Soft ethics in (4) is part of the framework that enables the best interpretations of the Recitals in (3). For soft ethics in (4) to work well in interpreting the Recitals in (3) it must be coherent with, and informed by, the hard ethics in (5) that led to their formulation in the first place.

[5]European Digital Rights, Comparison of the Parliament and Council text on the General Data Protection Regulation, https://edri.org/files/EP_Council_Comparison.pdf.

Another very good example is offered by the recent House of Lords Report on AI [7]. The argument developed in the report is that the USA has abandoned moral leadership altogether, and Germany and Japan are too far ahead on the technology side to make competition possible, but that this creates a vacuum where the UK should position itself, as a leader on ethical AI both as a socially desirable goal and as a business opportunity. This is part of the justification in the recent creation of the Centre for Data Ethics and Innovation (the Centre actually focuses quite strongly on AI as well). The fundamental lesson is that, instead of promoting a set of new laws, it may be preferable, within the current legislation, to foster an ethical approach to the development of AI that would promote social good.

Clearly, the place of ethics is both before (hard) and after (soft) the law, as what contributes to make it possible first and may complement it afterwards. In this, the position I am defending about the relationship between ethics and law is close to (and may be seen as the ethical counterpart of) Dworkin's when he argued that the law contains not only rules but also principles [8]. Especially in difficult, unclear or uncovered cases (Dworkin's 'hard cases'), where the rules fail to be applicable in full or unambiguously to a particular situation or offer an unacceptable approach, legal judgment is and should be guided by principles of soft ethics. These are not external to the legal system and used just for guidance (a position defended by Hart) but they are implicitly incorporated in the law as some of its ingredients (they are baked in), and help the exercise of discretion and adjudication.[6]

6. Ethical impact analysis

Given the open future addressed by digital ethics, it is obvious that the *foresight analysis* of the ethical impact of digital innovation, or simply ethical impact analysis (EIA), must become a priority [9]. Today, EIA can be based on data analytics applied strategically to the ethical impact assessment of digital technologies, goods, services and practices (figure 4). It is crucial because the task of digital ethics is not simply to 'look into the [digital] seeds of time/And say which grain will grow and which will not' (*Macbeth*, I.3, 159–162), it also seeks to determine which ones *should* grow and which should not.

Or, to use a metaphor already introduced above, the best way to catch the technology train is not to chase it, but to be already at the next station. We need to anticipate and steer the ethical development of technological innovation. And we can do this by looking at what is actually feasible, privileging, within this, what is environmentally sustainable, then what is socially acceptable and then, ideally, choose what is socially preferable (figure 5).

7. Digital preferability and the normative cascade

We do not yet have, for the infosphere, a concept equivalent to *sustainability* for the biosphere, so our current equation is incomplete (see figure 6).

In figure 5, I suggested that we interpret the x in figure 6 as social 'preferability' but I am aware that this may be just a placeholder for a better idea to come (note that, of course, digital technologies also have an ecological impact, so sustainability is relevant, but may also be misleading). This may take a while, given that 'the tragedy of the commons' was published in 1968 but the expression 'sustainable development' was only coined by the Brundtland Report almost 20 years later, in 1987 [10]. Yet the lack of conceptual terminology does not make the good governance of the digital any less pressing or a mere utopian effort. In particular, digital ethics, with its values, principles, choices, recommendations and constraints, already influences the world of technology significantly, and sometimes much more than any other force. This is so because the evaluation of what is morally good, right or necessary shapes public opinion—hence the socially acceptable or preferable—and the politically feasible, and so, ultimately, the legally enforceable, and what agents may or may not do. In the long run, people (as users,

[6]I am very grateful to one of the anonymous referees for calling my attention to this link with Dworkin's legal theory.

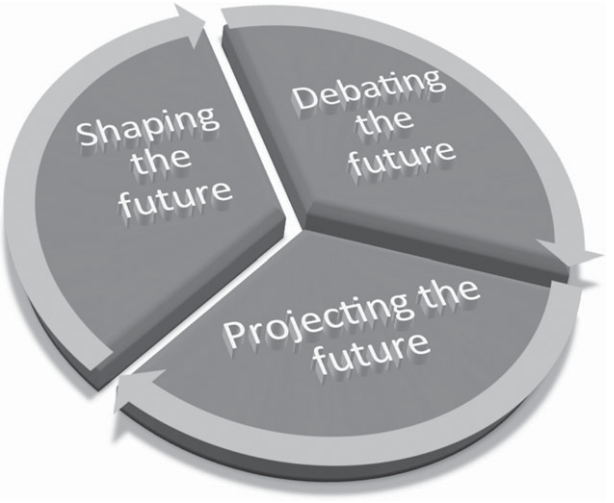

Figure 4. Ethical impact analysis (EIA): the foresight analysis cycle.

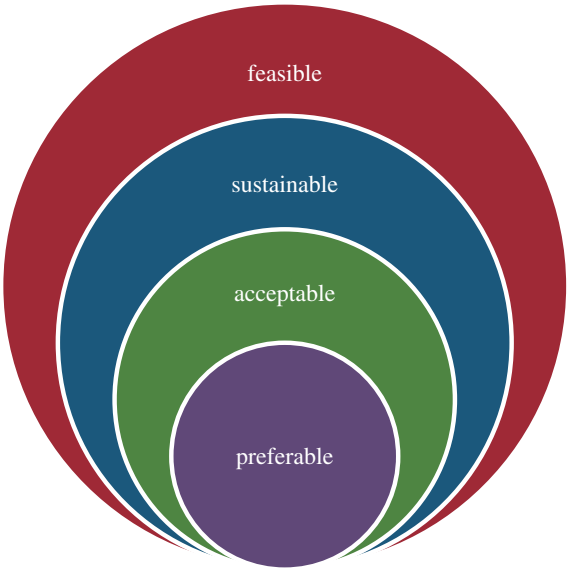

Figure 5. Digital ethics impact assessment. (Online version in colour.)

$$\text{biosphere} : \text{sustainability} = \text{infosphere} : x$$

Figure 6. A difficult equation to balance.

consumers, citizens, patients, etc.) are constrained in what they can or cannot do by the goods and services provided by organizations, e.g. businesses, which are constrained by law, but the latter is shaped and constrained by (also, although not only) ethics, which is where people decide in what kind of society they want to live (figure 7). Unfortunately, such a normative cascade becomes

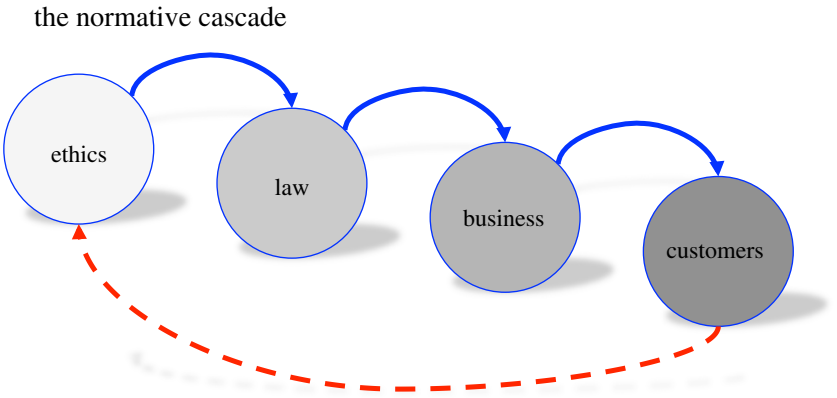

Figure 7. Example of a normative cascade, with business as agent and people as customers. Business could be replaced by government and people by citizens. (Online version in colour.)

obvious mainly when backlash happens, i.e. mostly in negative contexts, when the public rejects some solutions, even when they may be good solutions. A normative cascade should instead be used constructively, to pursue the construction of a mature information society of which we can be proud.

8. Digital ethics' dual advantage

Digital technologies offer many opportunities but also associated challenges and potential risks. Ensuring socially preferable outcomes means resolving the tension between incorporating the benefits and mitigating the potential harms, in short, promoting these technologies while avoiding their misuse, underuse and harmful use. This is where the value of an ethical approach becomes obvious. I argued above that compliance is merely necessary, but significantly insufficient. Adopting an ethical approach to digital innovation confers what may be defined as a 'dual advantage', echoing the 'dual use' terminology popular in philosophy of technology at least since the debate on civil and military uses of nuclear power. On the one hand, soft ethics can provide an *opportunity strategy*, enabling actors to take advantage of the social value of digital technologies. This is the advantage of being able to identify and leverage new opportunities that are socially acceptable or preferable, balancing any precautionary principle with the duty not to omit what could and ought to be done, e.g. to take advantage of the wealth of data accumulated, or the forms of smart agency available. On the other hand, ethics also provides a *risk management solution*. It enables organizations to anticipate and avoid costly mistakes (the Cambridge Analytica scandal involving Facebook data is by now a classic example). This is the advantage of prevention and mitigation of courses of action that turn out to be socially unacceptable and hence rejected. In this way, ethics can also lower the opportunity costs of choices not made or options not seized for fear of mistakes.

Soft ethics' dual advantage can only function in an environment of public trust and clear responsibilities more broadly. Public acceptance and adoption of digital technologies, including artificial intelligence, will occur only if the benefits are seen as meaningful and risks as potential, yet preventable, or minimizable, or at least something against which one can be protected. These attitudes will depend in turn on public engagement with the development of digital technologies, openness about how they operate, and understandable, widely accessible mechanisms of regulation and redress. The clear value to any organization of the dual advantage of an ethical approach amply justifies the expense of engagement, openness and contestability that such an approach requires.

9. Conclusion

Ethics in general, and digital ethics in particular, cannot be a mere add-on, an afterthought, a late-comer, an owl of Minerva that takes its flight only when the shades of night are gathering—once digital innovation has taken place, and possibly bad solutions have been implemented, less good alternatives have been chosen, or mistakes have been made. Nor can it be a mere exercise of *questioning*. The building of *critical awareness* is important, but it is also only one of the four tasks of a proper ethical approach to the design and governance of the digital. The other three are *signalling* that ethical problems matter, *engaging* with stakeholders affected by such ethical problems, and, above all, *providing sharable solutions*. Any ethical exercise that in the end fails to provide some acceptable recommendations is only a timid preamble. So ethics must inform strategies for the development and use of digital technologies from the very beginning, when changing the course of action is easier and less costly, in terms of resources and impact. It must sit at the table of policy-making and decision-taking procedures from day one. For we must not only think twice but, most importantly, we must think *before* taking important steps. This is particularly relevant in the EU, where I have argued that soft ethics can be properly exercised and where a soft ethical approach to SETI (science, engineering, technology and innovation) developments is acknowledged to be crucial. If soft digital ethics can be a priority anywhere, this is certainly in Europe. We should adopt it as soon as possible.

Data accessibility. This article has no additional data.

Competing interests. I declare I have no competing interests.

Funding. This article is part of research on data governance funded by Microsoft.

Acknowledgements. I am most grateful to the two anonymous referees for their detailed and constructive comments. The article is much better thanks to their helpful feedback.

References

1. Floridi L. 2016 Mature information societies—a matter of expectations. *Phil. Technol.* **29**, 1–4. (doi:10.1007/s13347-016-0214-6)
2. Nietzsche FW. 2008 *Twilight of the idols, or, how to philosophize with a hammer*. Oxford, UK: Oxford University Press.
3. Floridi L, Taddeo M. 2016 What is data ethics? *Phil. Trans. R. Soc. A* **374**, 20160360. (doi:10.1098/rsta.2016.0360)
4. Cabinet Office, Government Digital Service. 2016 *Data science ethical framework*.
5. British Academy, and Royal Society. 2017 *Data management and use: governance in the 21st century—a joint report by the British Academy and the Royal Society*.
6. EDPS Ethics Advisory Group. 2018 *Towards a digital ethics*.
7. House of Lords, Artificial Intelligence Committee. 2017 *AI in the UK: ready, willing and able?* Report of session 2017–19 HL Paper 100.
8. Dworkin RM. 1967 The model of rules. *Univ. Chicago Law Rev.* **35**, 14–46. (doi:10.2307/1598947)
9. Floridi L. 2014 Technoscience and ethics foresight. *Phil. Technol.* **27**, 499–501. (doi:10.1007/s13347-014-0180-9)
10. Brundtland GH. 1987 *The Brundtland report, World Commission on Environment and Development*. Oxford, UK: Oxford University Press.