RESEARCH ARTICLE

To Each Technology Its Own Ethics: The Problem of Ethical Proliferation

Henrik Skaug Sætra[1] · **John Danaher**[2]

© The Author(s) 2022

Abstract

Ethics plays a key role in the normative analysis of the impacts of technology. We know that computers in general and the processing of data, the use of artificial intelligence, and the combination of computers and/or artificial intelligence with robotics are all associated with ethically relevant implications for individuals, groups, and society. In this article, we argue that while all technologies are ethically relevant, there is no need to create a separate 'ethics of X' or 'X ethics' for each and every subtype of technology or technological property—e.g. computer ethics, AI ethics, data ethics, information ethics, robot ethics, and machine ethics. Specific technologies might have specific impacts, but we argue that they are often sufficiently covered and understood through already established higher-level domains of ethics. Furthermore, the proliferation of tech ethics is problematic because (a) the conceptual boundaries between the subfields are not well-defined, (b) it leads to a duplication of effort and constant reinventing the wheel, and (c) there is danger that participants overlook or ignore more fundamental ethical insights and truths. The key to avoiding such outcomes lies in a taking the discipline of ethics seriously, and we consequently begin with a brief description of what ethics is, before presenting the main forms of technology related ethics. Through this process, we develop a hierarchy of technology ethics, which can be used by developers and engineers, researchers, or regulators who seek an understanding of the ethical implications of technology. We close by deducing two principles for positioning ethical analysis which will, in combination with the hierarchy, promote the leveraging of existing knowledge and help us to avoid an exaggerated proliferation of tech ethics.

Keywords Ethics · Applied ethics · Technology · AI · Philosophy · History · Hierarchy

---

✉   Henrik Skaug Sætra
    Henrik.satra@hiof.no

1   Faculty of Computer Science, Engineering and Economics, Østfold University College, 1757-N Remmen, Norway

2   School of Law, NUI Galway, Galway, Ireland

⚛ Springer

1 Introduction

In a world of rapid development and dissemination of technology, ethics plays a key role in analysing how these technologies affect individuals, businesses, groups, society, and the environment (Sætra and Fosch-Villaronga, 2021), and in determining how to avoid ethically undesirable outcomes and promote ethical behaviour. While ethics is a staple in many academic fields, it is also gaining significant mainstream traction in in the tech industry and policy circles. In the 2021 version of Gartner's 'hype cycle' for AI, for example, *digital ethics* was placed at the peak of inflated expectations (Gartner, 2021), and terms such as *human-centred AI* and *responsible AI* are approaching the same stage.

Focusing on computer-based technologies, we know that computers in general and the processing of data, the use of artificial intelligence, and the combination of computers and/or artificial intelligence with robotics in particular are all associated with ethically relevant implications for individuals, groups, and society. This has given rise to a wide range of 'ethics of X' or 'X ethics' fields of inquiry and debate. Examples include computer ethics (Moor, 1985), data ethics (Hand, 2018), *big* data ethics (Zwitter, 2014), information ethics (Floridi, 1999), machine ethics (M. Anderson and Anderson (2011), robot ethics (Lin et al. 2011), and others that we describe in detail below.

In this article, we argue that while all technologies are ethically relevant, and studying the ethical implications of their development and use is crucial, we should not create a separate subdomain of ethical inquiry for each and every one of them. A frivolous proliferation of technology ethics is problematic for three reasons. First, the conceptual boundaries between the subfields are not well-defined. This creates problems for practitioners and regulators alike, as it becomes increasingly difficult to find historically established and valuable insight into the implications of technology. Second, it leads to a duplication of effort and constant reinventing the wheel. Third, there is a danger that participants overlook or ignore more fundamental ethical insights and truths. In general, historical efforts to build new domains of ethical inquiry risk burying and undermining historical insights, leading to a situation in which we increase the number of ethical domains and publications without increasing the actual ethicality of our decisions and practices.

We argue that the key to avoiding such outcomes lies in taking the discipline of ethics as moral philosophy seriously—acknowledging and pursuing it as a philosophical endeavour and not merely as a source of checklists and guidelines. We consequently begin with a brief description of what ethics is. We then proceed to present and review the main forms of technology-related ethics. Through this process, we develop a hierarchy of technology ethics—with a description of how certain forms of non-technology ethics are also relevant to this hierarchy. The hierarchy can be used by developers and engineers, researchers, or regulators who seek an understanding of the ethical implications of technology. It also shows how poorly defined some of the subdomains of technology ethics are. This process allows us to deduce two basic principles which will, in combination with the
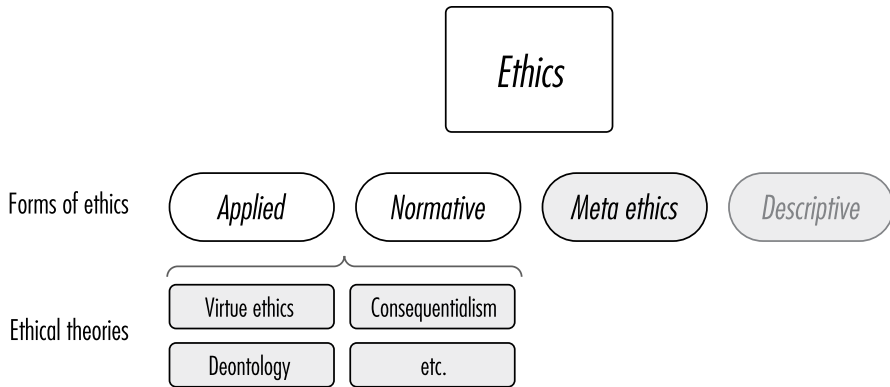
Fig. 1 The basic concepts in ethics

hierarchy, ensure that existing knowledge will be leveraged and that we avoid the proliferation of subdomains of ethics and a muddying of the waters of tech ethics.

While we perceive the proliferation of tech ethics as unfortunate, there are several seemingly plausible justifications for it. We end the article by discussing four such justifications and offering our replies.

2 The Core of Ethics

Ethics as a general object of study is originally positioned in the discipline of philosophy and more specifically in moral philosophy (Copp, 2005). 'Ethics' and 'morality' are often used interchangeably, and we follow Singer (2011) in conflating the terms, as ethics is fundamentally about making moral judgements. Ethics is a discipline as old as philosophy itself, but the structured approach to ethics as we use the concept today tends to be traced back to Aristotle (2014) and his *Nicomachean Ethics*. Ethics as a concept consists of a wide range of branches and theories, and we must be able to distinguish between these before we proceed to review the main types of technology ethics. The primary distinctions we focus on are the *forms* of ethics and the different ethical *theories*.

There are four primary forms of ethics, as shown in Fig. 1. *Meta* ethics is the most abstract form of ethics and deals with the origins and study of ethics, and whether or not there *are* moral truths at all (Copp, 2005). *Descriptive* ethics is about describing what a particular set of people believe to be right or wrong, without necessarily connecting this to any underlying theory or comprehensive conception of morality. *Normative* ethics, on the other hand, is about how people *should* act (Copp, 2005), and this is the domain of ethics where the three main ethical theories (utilitarianism, deontology, and virtue ethics) are debated and refined. Finally, there is *applied* ethics, which are normative ethical theories applied to particular circumstances. Applied ethics includes the different forms of technology ethics we will here pursue and all other kinds of *practical* ethics (Singer, 2011). There are many kinds of

applied ethics in addition to the ones we present below, such as medical ethics, business ethics, research ethics, care ethics, and migration ethics.

3 The Ethics of Science and Technology

While much of ethics is abstract, applied ethics focuses on concrete moral issues (Copp, 2005). Our focus in this article is on the need for different types of applied ethics theories or domains of inquiry, which we call 'domain-specific ethics'. Applied ethics tends to take the form of norms, guidelines, and frameworks. The Mertonian norms of science (Merton, 1973), for example, exemplify such an approach. Ethics, in applied form, often entails the codification and systematisation of what someone has discovered through philosophical analysis, which can at times be inaccessible to non-philosophers. There is often a division of labour between those who do meta ethics and normative ethics and those who do applied ethics, which draws upon the former to generate practical and often actionable insight for practitioners. There is also an important division of labour between the ethicist who develop norms and guidelines—codifies ethical considerations for a particular domain—and the ones who apply and must adhere to norms and guidelines (Sætra and Fosch-Villaronga, 2021). Developers need not necessarily deal with abstract ethics but can adhere to guidelines and checklists. Political institutions may or may not make the ethicist's proposed codified ethics law or support it in other ways (Sætra and Fosch-Villaronga, 2021). The key applied ethical questions arising from the use of technology tend to take the following form: will the use of an algorithm in a particular setting result in any harm/benefit to humans? What is the responsibility of a software developer to those affected by their technology? Is it ok to make machines that deceive humans? Can and should machines be designed to behave morally? And so on.

We limit our analysis to the overarching question of how to understand the ethical implications of our use of technology, and this includes both efforts to analyse these implications and to make sense of how people using technology—*and* the machines themselves—might act ethically. This sounds simple enough, but as we will show, many claims about the need for separate ethics for different types of technology have emerged. To make sense of this jungle of domain-specific ethics, we briefly review and analyse several of the most popular types of tech ethics. We describe the main features of each type and summarise their nestedness and relations to the other types. Our purpose, in doing so, is not merely descriptive. Through our review of these types of tech ethics we attempt to (a) highlight the conceptual confusion arising from the proliferation of subdomains and to (b) tentatively demarcate meaningful limits that could reduce the amount of overlap between the different subfields. We ultimately conclude that there are limits to this exercise in conceptual hygiene.

By beginning with engineering ethics, we omit the ethics of science. We acknowledge that this is one of the foundational types of applied ethics. Science is value-laden and political (Merton, 1972; Rollin, 2006; Sætra, 2018), and while not all technology-related activity involves science, much of it does and will partly be covered by science ethics, and the related research ethics. Nevertheless, our review of domain-specific

ethics is already quite extensive and the fact that the subfields discussed below could, on some occasions, be nested within science ethics merely serves to reinforce our larger argument. We also acknowledge that our presentations of the different domain ethics are necessarily simplified, and we are unable to describe in detail the nuances of and philosophical differences between the researchers and positions described as belonging to each domain. Such differences will certainly be important when approaching the analysis of the implications of a particular technology, but they are of less import with regard to understanding the potential problems relating to the proliferation of domain ethics and a fragmentation of the field of technology ethics.

3.1 Engineering Ethics

The first domain-specific ethic we consider is engineering ethics. This is a professionally oriented form of ethics, aimed at engineers with the purpose of promoting ethical practice. The aim is not primarily to promote a theoretical understanding of the ethical implications of engineering work. Harris et al. (2013), for example, present what they call their profession's perception of the primacy of the public good while discussing prohibited actions and the prevention of harm. They also discuss what they refer to as 'aspirational ethics' and the promotion of well-being through, for example, *design*. 'Ethics-by-design'—for example, in the form of value-sensitive design (Cummings, 2006)—is an important element of engineering ethics, which highlights its tight relation to design ethics (Costanza-Chock, 2020).

Engineering ethics is a vital part of the education of engineers, and 'teaching engineering ethics' is, according to Harris et al. (1996), seen as teaching *engineering*. They refer to professional ethics and engineering ethics as ethics for a particular *group* and state that 'engineering ethics applies to engineers (and no one else)' (Harris et al. 1996). This highlights how some domain-specific ethics are narrower and more specific than other forms of ethics. Science ethics, for example, is to a larger degree discussed by non-scientists, and AI ethics tends to be just as much a framework for understanding the implications of AI by non-practicing researchers and regulators as it is the codification of professional ethics for software developers.

That said, engineering ethics is a domain that, in theory, covers everything covered by the other domain-specific ethics discussed below. For example, the introduction to engineering ethics by Fleddermann (2004) begins by discussing how an incident involving a Ford Pinto car led to harm to humans, and that Ford was charged in a criminal court as they were held responsible for the design choices that determined the level and likelihood of harm. The Pinto was not autonomous, but the fundamental questions are the same as are asked in many fields of tech ethics: given that technology can cause harms and/or provide benefits (broadly defined) who is responsible for what in the design, production, and use of technology (Sætra, 2021a)?

3.2 Technology Ethics

Technology ethics is the highest-level technology-exclusive form of applied ethics. It is also informed by and heavily overlapping with work often referred to as

philosophy of technology (Ellul, 1964; Mumford, 1934; Winner, 1977). While seemingly similar to engineering ethics, it is less oriented towards professionals. It is much broader. Jonas (1982), in line with the argument proposed in this article, asked whether technology is sufficiently 'novel and special' to warrant its own brand of ethics, and answered in the affirmative. He gave four reasons for this. First, he argued that technology defies neutrality, as it is not only malevolent use of technology that is problematic, but also the long run effects of what we'd consider beneficial use. Second, he argued that technology has a tendency, after careful beginnings, to become 'an incessant need of life' (Jonas, 1982). Third, he highlighted technology's unique magnitude and ability to amplify human action, something that breaks historical *anthropocentric* ethics as the biosphere is increasingly affected by humans through technology. Fourth, he noted that technology promotes new and fundamental existential ethical questions, such as 'whether and why there ought to be a mankind?' However sound or unsound we find Jonas's argument, he exemplifies the question that should always be asked when one ponders whether a new ethic is necessary, namely, what makes a particular technology distinct from something already covered by an existing ethic?

Technology ethics is nowadays often portrayed as the discipline that contains lower-level applied tech ethics, such as machine, robot, and computer ethics (Gordon and Nyholm, 2021). A clearly defined ethics of technology is, however, relatively hard to come by, despite the fact that many combine the terms ethics and technology. Tavani (2016), for example, authored the book *Ethics and Technology*, but from the get-go decides to establish the term *cyberethics* instead of *technology ethics*. While cybertechnology—computing and communication devices—is certainly central to modern technology, this takes us closer to a specific form computer ethics and away from technology ethics more generally.

3.3 Computer Ethics

Computers are technology, and consequently encompassed by technology ethics. Some describe computer ethics as concerned with 'commercial behaviour involving computers and information', including issues of data security and privacy (Gordon and Nyholm, 2021).[1] If *information* ethics is defined as related to issues clearly linked to computer-mediated information, seeing it as a subdiscipline makes sense. However, issues of surveillance and privacy, as mentioned above, are clearly not restricted to digital information, and this generates certain challenges leading to some overlap with regard to which branch of tech ethics privacy and surveillance-related questions belong.

A more fruitful approach is found in Tavani (2016), where computer ethics is seen as the ethics of computing *machines*, unrelated to the issues of how such machines communicate. But since communication is now fundamental to much

---

[1] However, *information* ethics is often also described as a separate discipline, necessitating the need to determine whether issues of information are distinct from issues of computing, or more properly described as a subdiscipline of computer ethics (Capurro, 2006).

computing technology, it is difficult to maintain this distinction. Furthermore, computers are seen as the basic technological foundation of anything *digital*, and it will consequently be considered a high-level ethics related to how digital technologies are used.

In the broader landscape of ethics mapped in this article, a pertinent question is what sort of questions belong to computer ethics that are not related to its subdisciplines, such as AI ethics, and information ethics? We argue that for the concept to be useful alongside other form of ethics, computer ethics should in fact mainly relate to the materiality of computing—how machines are designed, built, their energy use, distributional effects, and accessibility. This would, if so, indicate that much of what is discussed in relation to the environmental sustainability of AI (Brevini, 2021; Sætra, 2022; van Wynsberghe, 2021), for example, is, in reality, more properly a question of the sustainability of *computing*.

Computer ethics can be taken to be the basic domain in which questions related to how computers change what human beings can do and how we do things are asked (Moor, 1985). Furthermore, if we follow the approach of Johnson (2004), and use the term to describe all examinations of the ethical issues related to an 'information society', this would then subsume professional ethics for computer scientists and engineers, issues of privacy, cybercrime, VR, and so on (Johnson, 2004). If such a definition stands, most of the forms of ethics described below would be superfluous and in reality a part of computer ethics. Many would count, primarily, as case studies in computer ethics.

3.4 AI Ethics

To build artificially intelligent systems, a precondition is the existence of computers. Indeed, the origins of computing technology and artificial intelligence are inextricably linked thanks to the pioneering work of Turing on computation and thought (Turing, 2009). As a result, AI ethics could be seen as a lower-level form of tech ethics (below computing ethics) with relatively high specificity. But since AI is currently a concept in vogue, the term 'AI ethics' garners a lot of attention, and this, in turn, tempts researchers to describe various challenges that more properly relate to other types of ethics as AI ethics. While in vogue, AI as a concept has a long history, with the term first used in 1956 (Russell and Norvig, 2014), often traced back to earlier work of researchers such as Turing (2009). The notion of autonomous technology is also relevant for demarcating AI ethics, as autonomy is tightly linked to various conceptualisations of intelligence (Winner, 1977).

Today, AI ethics is argued to encompass a wide array of issues, and there is a need to distinguish which questions should belong to AI ethics proper, and which questions belong to other ethics domains. AI ethics entails, according to Gordon and Nyholm (2021), issues including, but not limited to, the design and use of autonomous systems in general (both weapons and other systems), machine bias, privacy and surveillance, governance, the status of intelligent machines, automation and unemployment, and even space colonisation. According to Coeckelbergh (2020, p. 7), 'AI ethics is about technological change and its impact on individual

lives, but also about transformations in society and in the economy'. In his book, he includes challenges related to superintelligence, the difference between humans and machines, the potential for moral machines, issues related to data, privacy, bias, machine responsibility, policy, and even the meaning of life. Müller (2020) similarly discusses AI, but in combination with robot ethics, and includes specific discussions of bias, opacity, privacy and surveillance, machine ethics and machine morality, and the singularity. To top this off, some even argue that the scientific *communication* of advances in AI, and the selection of imagery and stock photos of AI or AI-related themes, is a part of AI ethics (Romele, 2022).

If we take one step back, and consider AI to be software capable of either thinking or acting humanly or rationally (Russell and Norvig, 2014), it seems pertinent to drastically reduce the number of topics seen to properly relate to AI ethics. While certain AI systems are based on machine learning approaches which entail analysing data, issues of privacy and surveillance still seem to be issues more properly conceived as belonging to data ethics, or even a form of ethics not restricted to privacy and surveillance as digital phenomena at all. Furthermore, robot ethics, machine ethics, and information ethics deal directly with subsets of the issues that are argued to belong to AI ethics.

The core topics remaining are those related to how intelligent systems allow us to do new things and to do things differently. This could relate to automation and employment, as mentioned by Müller (2020). However, it would be restricted to automation based not on replacing human force with animal or machine force, for example, but on systems performing tasks with a cognitive element that previously required humans. Other issues of automation belong more properly to technology ethics in general. Long-term existential or x-risk issues such as superintelligence and the singularity can be properly said to be part of AI ethics, although they equally branch out into discussions of other technologies (biotech, nuclear weapons, and so on). AI will in turn clearly be relevant for understanding issues of robotics, but AI ethics should only be concerned with issues relating to systems independent of embodiment, as embodied systems will create novel challenges best analysed through specific forms of robot ethics.

3.5 Robot Ethics

Robots—machines that sense and purposefully act with a certain degree of autonomy in a particular environment (Winfield, 2012)—are necessarily driven by *some* form of artificial intelligence. They are, however, always embodied. Does embodiment by itself make a difference such that we ought to distinguish *robot* ethics from AI ethics? Lin et al. (2011) argue that it does, as 'advanced robotics brings with it new ethical and policy challenges' divided into three main categories: safety and errors, law and ethics, and social impact. But are these really novel? Aren't they true of all forms of technology?

It is worth noting that the ethics of robots has been contemplated long before the contemporary fad for AI ethics and robot ethics (Winfield, 2012). While not a codified ethics, the science fiction literature is replete with analyses of the potential

ethical implications of both robots and AI. Isaac Asimov's three laws of robotics from the 1942 short story *Runaround* (Asimov, 2013) are perhaps the most famous example of this:

> "We have: One, a robot may not injure a human being, or, through inaction, allow a human being to come to harm."
> "Right!"
> "Two," continued Powell, "a robot must obey the orders given it by human beings except where such orders would conflict with the First Law."
> "Right"
> "And three, a robot must protect its own existence as long as such protection does not conflict with the First or Second Laws."

The question nonetheless remains: what distinguishes robots from 'regular' AI? In theory, Asimov's laws could easily be said to apply to AI as well, even if embodiment makes issues of harm and protection even more pertinent.

Lin et al. (2011) argue that a robot's ability to 'directly exert influence on the world' generates sufficient novelty, and their analyses encompass a wide array of robot applications, such as labour, service, military, medical, education, care and companionship, and transportation. This means that social robots, military robots and autonomous weapons systems (AWS), and autonomous vehicles are likely candidates for analyses under the robot ethics umbrella. Some have proposed separate ethics for specific types of robotic system, e.g. a specific ethics for AWSs (Horowitz, 2016) or 'autonomous driving' (Geisslinger et al. 2021), but we consider those to be encompassed in robot ethics.

The first category of ethical and social issues pertaining to robots relates to their safety and potential for error. But what Lin et al. (2011) list in this category reads quite similarly to what is often discussed under computer ethics—namely, considerations of what happens when computer scientists make errors or unforeseen consequences emerge when technology is applied in new settings and at large scales. They argue, however, that the magnitude of damage potentially done is larger when robots physically operate in our environment as opposed to software errors leading to the loss of data, for example. It could, however, easily be shown that a wide range of software errors also have fatal outcomes, so *direct* physical harm seems insufficient for creating a new ethic.

The second category relates to law and ethics and includes issues such as *responsibility* when robots, for example, cause harm. However, this applies just as much to computers in general, and in particular to AI, as it does to robots. There has been much debate over the presence of so-called *responsibility gaps* resulting from the unpredictable nature of modern AI (Matthias, 2004; Sætra, 2021a). Some connect this directly to robots (Gunkel, 2017), but if the issue stems from the nature of advanced AI (namely, its capacity for autonomous decision-making), then this topic belongs to AI ethics, and not robot ethics.

Third, there is the social impact of robots. This is an area in which it seems more likely that robots constitute a special ethical case, as their physical presence in human social environments can have various implications not directly comparable to the presence of computers or other non-autonomous devices. Some have

focused specifically on human likeness, anthropomorphism, and how social robots can change human beings and society (Danaher, 2020a; Sætra, 2021c), while others focus on the social implications of autonomous vehicles and weapon systems (Fleetwood, 2017; Horowitz, 2016). While social robots are often portrayed as particularly problematic, there is also an argument to be made that social *AI* is capable of generating many of the same challenges as social robots in interactions with human beings (Sætra, 2020). Having a relationship with an app on one's smartphone is perhaps similar to having one with a robot.

In short, robot ethics can be described as the ethic of how human beings 'design, construct, and use robots' (Gordon and Nyholm, 2021), where robots are understood as embodied AI systems. However, many ethical questions seemingly caused by robots can and should be treated as issues of AI, computer, or technology ethics.

3.6 Machine Ethics

Closely linked to, but perhaps separable from, AI and robot ethics is the field of machine ethics. The motivating question behind this field of inquiry is: To what extent can machines be ethical, or deal with ethical challenges? Answering that question necessitates a new field of inquiry, according to M. Anderson and Anderson (2011). The goal of this field, Anderson (2011, p. 22) states, is to make 'a machine that follows an ideal ethical principle or set of principles in guiding its behaviour'. Why is that important? In one of the originating articles in the field, Allen et al. (2006) use the trolley problem to say something along the following lines: since an *autonomous* machine might run into ethical dilemmas akin to the trolley problem, we must explore how we can make machines ethical agents capable of ethical decision-making.

This form of applied tech ethics is to be contrasted with other types of tech ethics. The target of most applied tech ethics is humans and human institutions: how can they be improved to address ethical challenges? In machine ethics, the target is the machines themselves: how can we codify ethics into autonomous machines or train these machines to act ethically? That said, this framing of machine ethics is potentially problematic. As argued in Sætra (2021a), presenting machines as autonomous entities partly beyond the control, and perhaps even beyond the responsibility, of the humans who make and deploy them is both controversial and potentially misleading.

While its proponents argue that machine ethics is about 'adding an ethics dimension' to autonomous machines (M. Anderson and Anderson, 2011), all autonomous machines arguably already have an ethical dimension, as all tasks performed by machines in a sociotechnical system have consequences of ethical value. Moor (2006) explores to which extent machine ethics even exists. He argues that it is reasonable to see computers as 'technological agents'. He perceives all computing technology to be normative by its nature: computers are designed to do certain things and, consequently, follow a certain ethical code, even if this is only implied. Still, Moor accepts that it is important to distinguish an inquiry into the ethical impact of machines (and technology more generally) from the agenda of putting 'ethics into a machine'. The former is *computer ethics*, while the latter is *machine ethics*.

With such an interpretation, machine ethics can be portrayed as a field in which the moral status of machines as potential moral agents is examined, with a particular emphasis on modelling and codifying existing human moral systems or new moral systems into autonomous machines, and so may have a distinctive identity as a sub-domain of technology ethics. Machine ethics also has branches of its own, with fields such as machine medical ethics emerging to focus on ethical machine behaviour in different domains (Kochetkova, 2015).

3.7 Information Ethics

According to Floridi (1999), 'standard ethical theories' cannot deal satisfactorily with computer ethics problems. Floridi argues that computer ethics thus needs a *new* foundational ethics on which to build. This new foundation is what he terms 'information ethics', which he sees as a "particular case of 'environmental' ethics or ethics of the infosphere" (Floridi, 1999). Any information entity could, in this ethics, be considered worthy of moral recognition and status. It is thus clear that it is a theory aimed at the expansion of our moral circles, which again explains why it is portrayed as a particular form of environmental ethics. What is much less clear, however, is the value of this form of environmental ethics. Is it really a necessary foundation for computer ethics and all its branches, or might 'standard' ethical theories and environmental ethics be capable of more than suggested by Floridi?

One example of why information ethics could be relevant is the evaluation of the moral status of an artificial agent (Capurro, 2006). Say that we are building a social simulation, using agent-based modelling to explore issues related to the emergence of social effects. In this process, we construct a number of artificial agents with rules to determine their actions, including 'goals' they might be coded to optimise. Do such figments of our imagination have any value? Can we do with them as we please, including setting them free in worlds we create, in which they might be attacked by other agents, and even 'killed' through commands such as 'If (energy = 0) [die]'? Would things change if we say that the agents in question are far more sophisticated AI agents living in some metaverse in a not-too-distant future? Even if such agents have no biological life, and cannot necessarily suffer or experience joy in a specific human or animal sense, they are indeed information entities, and thus *potentially* recipients of moral consideration.

In this form, information ethics connects quite directly with robot and AI ethics, which encompasses questions relating to the moral value and status of robots and AI (Gunkel, 2014, 2018). While embodiment could be said to matter, the basic cognitive capabilities of robots and the artificial agents just mentioned are exactly the same, which makes it pertinent to ask if questions related to the moral status of artificial agents most properly belong to information ethics or robot ethics or AI ethics. However, we could just as well ask whether information ethics is really necessary for asking these questions at all, or whether, for example, environmental ethics have already provided us with the required framework for analysing moral status and various forms of inclusion in moral communities (Nolt, 2014).

3.8 Data Ethics

Most attempts to define a domain-specific ethics entail attempting to highlight how special the domain is. So it is with data ethics. Hand (2018) states that ethical issues related to data are 'more challenging' than issues related to other technologies. This is, he states, because data and data science are ubiquitous, and because the issues involved are so complex. Right or wrong, data is central to modern society, and Hand (2018) attempts to capture a wide range of issues under the data ethics umbrella, including what data is, who owns it, consent, confidentiality and transparency, trustworthiness, and privacy.

There are ample principles, checklists, and guidelines for data ethics. Drew (2016, p. 4) presents a set of principles, such as 'use data and tools that have the minimum intrusion necessary' and 'keep data secure'. Such general principles are hopefully universally accepted, and many have also been codified in law, e.g. in the EU's GDPR. Hand (2018, p. 189) refers to the checklists of other unnamed domains of ethics, and makes his own, with entries such as 'store data securely' and 'be clear about the benefits of the analysis, and who derives the benefits.' One attempt at summarising the principles and demarcating data ethics is the *data ethics canvas* by the Open Data Institute, aimed at providing practitioners with the tools and questions required to avoid 'adverse impacts on people and society' (Open Data Institute, 2021). From an outsiders' perspective, there appears to be a constant scramble to present and be the originator of the best framework, and Franzke et al. (2021), for example, argue for the benefits of their Data Ethics Decision Aid (DEDA) over the data ethics canvas.

There are consequently many varieties of data ethics, and some, such as the 'data ethics of power' (Hasselbalch, 2019) is less prescription- and checklist-oriented and more focused on elucidating how data relates to power and changed power relations. All in all, however, data ethics might most fruitfully be understood as the practically oriented guide to practitioners and users of computing technology, AI, and robotics, as all the broader questions crammed into this umbrella are also analysed by other domains.

It is also worth noting that while data ethics purports to be the domain of privacy and surveillance, issues related to these phenomena are much older than modern data science (Westin, 1967), and the questions involved in understanding them should perhaps not be limited to 'data ethics'. The ethics of privacy (Moor, 1999; Siegel, 1979) and surveillance (Macnish, 2018; Marx, 1998) are established domains that seem able to serve data scientists with both historical and new insight into these phenomena without them having to be connected specifically to data ethics. Furthermore, attempts to brand even more specific data ethics have been made, such as *big* data ethics (Richards & King, 2014; Zwitter, 2014) and the even more specific 'ethics of biomedical big data analytics' (Mittelstadt, 2019a). However, these endeavours are in reality either specific instances of data ethics, or cases of attempting to understand how data is currently used in combination with new forms of analysis, or AI, and we consequently believe the latter is encompassed in data ethics and/or AI ethics.

3.9 Digital Ethics

Digital ethics is a term not used quite as often as many of the others, but it does capture a segment of ethical questions not directly belonging to the other domains detailed above. In particular, if we use the term in line with Capurro (2009), digital ethics, or digital media ethics, relates to the use of information and communication technology, and captures questions related to the use of, for example, mobile phones and navigation services. Digital ethics can consequently be seen less as a technical checklist-oriented ethics and more oriented towards the challenges 'raised by digital culture' and issues related to participation, equitable access, and the implications of our use of digital media (Luke, 2018). As seen in the *Oxford Handbook of Digital Ethics* (Véliz, forthcoming), the term has been understood to subsume *all* forms of ethics described in this article, including Internet ethics, AI ethics, and robot ethics. This broad use of the term is seemingly also found in Aggarwal (2020), who describe how advances in *AI* changes the ethical implications of digital technologies. Muddying the water further, Aggarwal proceeds to state that Intercultural digital ethics is a subfield of both digital ethics *and* information ethics.

Another usage of the term is found in Whiting and Pritchard (2018), where digital ethics is defined as 'the moral principles or rules of behaviour that govern and guide qualitative Internet research from its inception to publication and the curation of data'. Such a definition would, however, position digital ethics as a branch of science, and more specifically research, ethics.

3.10 Internet Ethics

The final domain-specific ethic we include is Internet ethics, which is seen as a low-level technology ethics nested in digital ethics. In Langford (2000), Internet ethics is presented through explorations of how the Internet relates to privacy and security, law and the Internet, the potential for fostering moral wrongdoing, information integrity, democratic implications, and professional responsibilities. This suggests that it is both used to describe the agenda of analysing Internet implications and Internet ethics as a professional ethic for engineers. The term is not as established as many of the other domain ethics here discussed, and Tavani (2016) suggests that 'cyberethics' is a preferable term that covers more than just the Internet, including other interconnected communication technologies.

While Internet ethics is already low level, this does not stop others from developing even lower-level domain ethics. Social networking ethics (Lannin and Scott, 2013; Vallor, 2012), for example, and even search engine ethics (Tavani, 2012), have been proposed.

3.11 The Great Chain of Technology Ethics and Neighbouring Ethics

The preceding considerations lead to the overview of technology-related domain ethics summarised in Table 1. This categorisation shows how one might think about

the hierarchical relationships between the different domains of technology ethics. It represents our attempt to make sense of the proliferation of subdomains. However, as noted in the preceding text, the way in which the subdomains are understood or applied in the philosophical, legal, and regulatory literature is not as conceptually pure or logical as we might like.

The relationships between the various types are also shown in Fig. 2, which presents the tentative distinction that has emerged between ethics aimed at directing human action and those directed at other entities, specifically the technologies themselves. We have shown how, for example, engineering ethics is aimed at guiding the conduct of engineers, while machine ethics is about the ethical behaviour of machines. This hierarchy coupled with a proper understanding of how the domains relate to each other and their goals can help identify which forms of domain-specific ethics are novel enough to warrant unique research agendas and which are already sufficiently captured by higher-level ethics.

While we argue that even some of these forms of ethics should have a more marginal importance than it might appear in the modern discourse on technology ethics, the real challenge is further exacerbated by the fact that we have already excluded a number of proposed domain-specific ethics, such as social network ethics, search engine ethics, cyberethics, and programming ethics.

We have chosen not to go into detail on the various types of adjacent or supporting ethics. Some of these relate directly to those detailed above. Business ethics, for example, can relate very closely to computer ethics since computing technology is widely deployed by businesses. Care ethics is closely related to robot ethics since one major potential application of robots is in care settings. Environmental ethics is arguably complementary to (and possibly foundational to) both robot and information ethics. We have already noted that privacy and surveillance ethics already covers many of the bases purportedly covered by data ethics, and social and distributive ethics arguably provides the foundational analyses so often foregrounded in various forms of data ethics and AI ethics.

While we are admittedly sceptical of the importance of many of these domain-specific ethics, we do not argue that they are all superfluous and that we only need one, or very few, types of general ethics. The complexities of new technologies and the business operations of those who use them will sometimes require analyses based on intimate knowledge of the technology in question and a certain degree of specialisation. Furthermore, case studies involving the application of higher-level ethical principles or theories to particular technologies will always be needed. The question, then, is how to evaluate the need for particular types of domain-specific ethical inquiries or theories, as opposed to allowing for more specialisation within the more foundational domains or case studies arising from them.

4 An Ethical Division of Labour

As the preceding section has shown, there is significant potential overlap between the various domains of ethics related to modern technologies such as social robots, AI, and big data. While many branches of ethics have emerged for a reason, we have

Table 1 Summary of domain-specific ethics

Domain-specific ethic	Topic	Target group(s)	One level up	One level down, example(s)
Engineering ethics	The responsibilities of engineers in the development and application of technology	Engineers	Science ethics	Technology ethics, computer ethics, design ethics, etc
Technology ethics	The use of and implications of technology on society	Policymakers, engineers, all professionals involved in designing and applying technology	Engineering ethics	Computer ethics, machine ethics, robot ethics
Computer ethics	Computing machines and unconnected computers, the materiality of computing	Computer/IT professionals	Technology ethics	AI ethics, information ethics, digital ethics
AI ethics	Artificially intelligent systems, embodied or not. Superintelligence, automation, and employment	Software engineers and regulators	Computer ethics	Machine ethics, robot ethics
Robot ethics	The use of and implications of robots	Roboticists, lawmakers	Technology ethics, AI ethics	AWS ethics, ethics of autonomous vehicles, ethics of sex robots, etc
Machine ethics	The design of moral machines	Roboticists, lawmakers	AI ethics and/or robot ethics	Machine medical ethics
Information ethics	The evaluation of what is recognised as morally valuable	Philosophers and lawmakers	Environmental ethics	-
Data ethics	The collecting, storage, and usage of data, including personal data	Data scientists	Computer ethics	Big data ethics
Digital ethics	Questions related to how digital media impact people and society	Researchers and media practitioners	Computer ethics	Internet ethics and other specific media-related ethics
Internet ethics	Questions related to how Internet-based technology impacts people and society	Researchers and digital media professionals	Digital ethics	Social networking ethics, search engine ethics

Directed at human action

Directed at other entities

Engineering ethics

Technology ethics

Computer ethics

Information ethics

Data ethics

Big data ethics

Digital ethics

AI ethics

Internet ethics

Social networking and ethics

Robot ethics

Social robot ethics
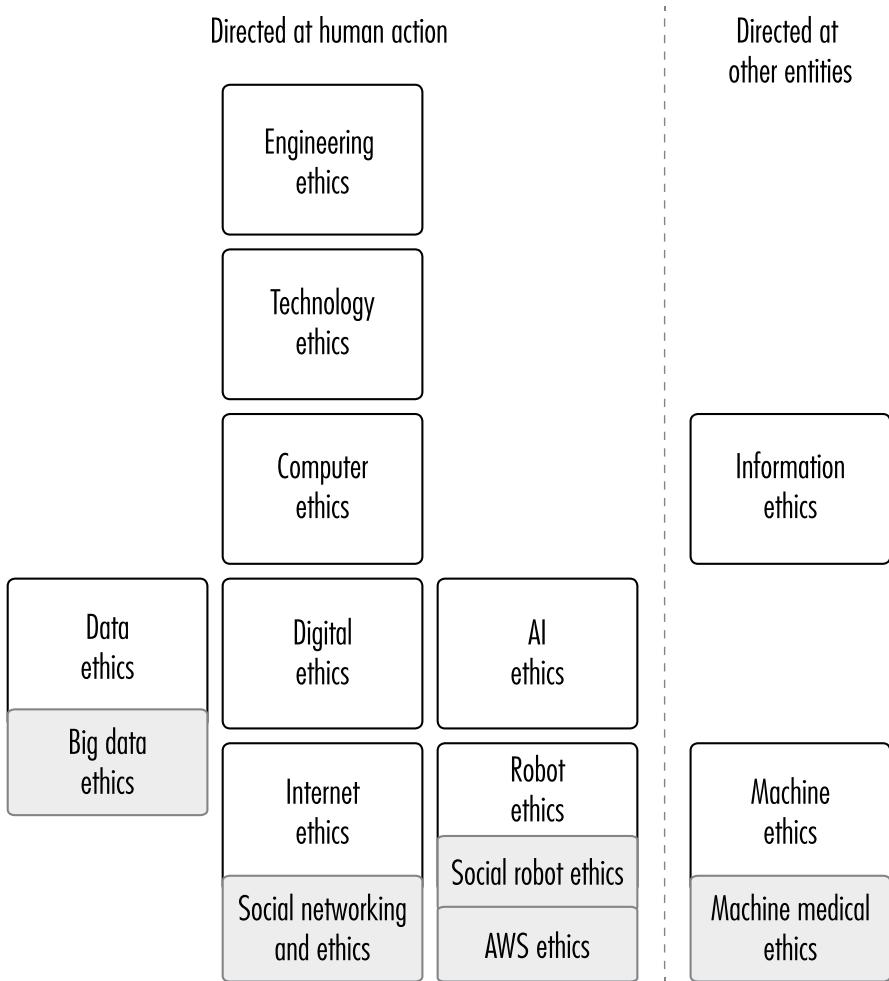
AWS ethics

Machine ethics

Machine medical ethics

Fig. 2 The great chain of technology ethics

also shown that there is much confusion and a lack of consistency in how the various terms are used, as seen, for example, in the Oxford handbooks on digital ethics (Véliz, forthcoming) and AI (Dubber et al. 2020). In addition, certain technologies are associated with significant hype, and this could easily lead ethicists who are unfamiliar with the higher-level traditional technology related ethics—or who seek attention and impact within and outside academia—to align their work with the hype terms in vogue at any point in time. Big data has been an obvious example for some years, now superseded by AI, which might in turn give way to various forms of virtual/extended reality and crypto, as the metaverse and web3 seem poised for prominence. Academic specialisation is not necessarily a bad thing, but we argue that the proliferation of technology ethics domains can be, and we should avoid excessive proliferation.

4.1 The Problems of Proliferation

Our position is that the negative consequences of proliferation of domains often outweigh the positive ones, and there are three main arguments in favour of such a position.

Firstly, the conceptual boundaries between the subfields are not well-defined nor respected. This leads to a general confusion and a lack of consistency, as people from purportedly different domain-specific ethics proceed to work on the same issues. Privacy ethics is a good example of how researchers and practitioners in different domains work on the same topic. People working in Internet ethics, for example, discuss privacy-related issues arising from the tracking of online information and the monetisation of this information by social media platforms. People working in AI ethics discuss the very same issues as they pertain to, for example, facial recognition technology and predictive analytics services. The discussions are similar, perhaps even equivalent. Part of the reason for this is because AI technology has become seamlessly blended into many online services. But problems arise as soon as people unduly characterise the challenge generated by AI as novel or specific and neglect to connect their discussions to foundational insight into the nature of privacy. If AI ethicists proceed to generate their own conceptions of privacy and surveillance, and the same occurs in, for example, data ethics and digital ethics, the risk of inconsistency emerges. Furthermore, AI ethics is, as we have shown, presented as a domain encompassing many topics arguably belonging to higher-level ethics, such as technology or computer ethics, and this creates a confusion as to what belongs where. While case-based analyses of problems related to issues such as privacy related to a particular service, autonomous vehicles, and robots in public spaces are clearly necessary, we take issue with the attempts to compartmentalise such questions in specific domains.

Secondly, it leads to a duplication of effort and constant reinventing of the wheel. This is related to the first point, as insufficient demarcation leads both new and old practitioners to create new foundations and approaches within lower-level forms of ethics that ignores what came before. This way of doing compartmentalised and siloed ethics is inefficient and wasteful, as similar and overlapping knowledge is produced without sufficient interaction. This is a problem even *within* domain-specific ethics, as shown by the various analyses done on the proliferation of guidelines and principles of responsible, ethical, and trustworthy, AI (Floridi and Cowls, 2019; Jobin et al. 2019), which tend to repeat many of the same points (Dotan, 2021). We extend this argument, because we have seen that not just within domain-specific ethics, but also between them, we find an even larger universe in which similar and overlapping topics become the subject of guidelines and principles rooted in a too low-level ethics to facilitate knowledge sharing and interdisciplinary debates about the consequences of technology.

Not only is it unnecessary to invent the wheel over and over—doing so will arguably also lead to the constant invention of poor wheels, rather than improvement on the basic concepts. One example could be from the domain of robot ethics, in which various authors, in a large number of different outlets, debate the potential for robots

to be friends, lovers, and romantic partners. To a large extent, the debates about those different robots work along the same basic lines: some people argue that robots currently lack the mental properties associated with human friends, sexual partners, and lovers; others argue that they do not or that they may acquire those properties in the future (Danaher, 2019, 2020b; Gunkel, 2018; Sætra, 2021b). Contributors to the debate simply list the same basic mental properties over and over again and debate their actual or possible instantiation in a robot. There is very little progress and much duplication of effort. Why does this happen? One possibility is that new researchers from different fields continuously stumble upon topics that seem novel from their perspective (love, friendship, sex, workplace relations, or even general sociology, philosophy, etc.), but that are already being dealt with in other disciplines, or, in the best of circumstances, in specialised interdisciplinary arenas. Journal editors and reviewers are unaware of these pre-existing literatures and thus give the green light to a new take on an old, well-debated issue. Approaching a problem from different angles, or multiple times, is not necessarily a bad thing, but to provide scientific value, it should be purposeful and based on extant knowledge. Cross-validation from different disciplines and philosophical perspectives, and replication in general, is immensely valuable for separating the valuable from the discardable in extant literature. However, while a fragmented field of technology ethics *might* incidentally have such positive effects, the benefits will be limited if such quasi-replication is performed without knowledge of that which is supposedly replicated.

The community of editors and researchers involved in applied technology ethics consequently have a shared responsibility to search for existing work and to use reviewers who knows the field of study. This will allow newcomers to better make use of existing knowledge, while also connecting the various disciplines that need be involved in the ethics of technology, which will often by necessity be interdisciplinary.

Thirdly, there is a danger that participants overlook or ignore more fundamental ethical insights and truths in their zeal for carving out a new domain-specific ethics. While ethicists constantly reinvent the wheel, they could instead choose to rediscover and apply foundational insight from higher-level ethics. By doing so, they would be adhering to the scientific ideal of accumulation, and by standing on the shoulders of giants, they would arguably be able to get much farther into what is truly unique about their lower-level case studies (Merton, 1942). AI ethics is once again an interesting example. Much of what is now labelled AI ethics has been expertly detailed by writers in the philosophy and ethics of technology, such as Mumford (1934), Ellul (1964), and Winner (1977). While their accounts may be wordy, and contain very few checklists, the key questions they address reveal how non-novel the challenges purportedly attributed to cutting edge AI really are. Winner (1977, pp. 326–327), for example, discusses principles related to the need to design autonomous technology in a way that makes it intelligible and accessible to those it affects, and that flexibility and mutability are crucial for avoiding 'circumstances in which technological systems impose a permanent, rigid, and irreversible imprint on the lives of the populace'.

Or consider the work on the problem of bias in AI systems. This has become a major topic of debate and concern in recent years, much of it stemming from a

landmark report by the public interest journalism platform ProPublica on bias in recidivism algorithms used by the US criminal justice system (Fazelpour and Danks, 2021). While there is value to the recent work on bias—in particular the so-called impossibility results derived by mathematicians and computer scientists (Kleinberg et al. 2018)—a lot of the conceptual terrain on bias and fairness was mapped long before the modern AI hype cycle. For instance, Friedman and Nissenbaum (1996), in an article entitled 'Bias in Computing Systems' published in 1996, addressed many of the basic forms of bias in computing systems, all of which overlap directly with concerns about bias in modern AI systems. The economist/philosopher John Roemer (1998) mapped in detail the incompatibility between different standards of fairness and non-discrimination in his work on equality. And many classic contributions to our understanding of sexism, gender bias, racial injustice, and other forms of discrimination harbour insights that are clearly relevant to the understanding of bias in AI (Benjamin, 2019; D'Ignazio and Klein, 2020; Noble, 2018). There is a danger that these insights are ignored and, again, reinvented because they are included in work on political philosophy/economics or associate themselves with technology (computers) that is overlooked by participants in the modern AI ethics debate. People working with a novel ethics silo are too busy debating with their peers to harness the more foundation insights from previous generations.

4.2 Two Criteria for Choosing Technology Ethics

The way out of the predicament generated by the proliferation of ethics consists of two simple criteria. Seeing how applied ethics is ripe with checklists and principles, we have made our own very simple principles for labelling and positioning ethical research:

1. If the questions you address are sufficiently addressed by higher-level ethics, do not align your work with lower-level technologies.
2. If you are in fact pursuing novel questions, consider if they are general and apply to other, more basic technologies and questions, and do not always rush to create a new domain-specific ethics attached to lower-level terms and technologies.

The first principle suggests that ethicists should always locate their work in the highest-level domain ethics that covers the questions they address. For example, if they research questions related to general problems facing computer and software engineers, these questions are most likely already partly answered in computer ethics or engineering ethics, and it is beneficial to continue the debate there, rather than to pretend that it belongs to AI ethics because AI is some ground-breaking and magical technology and not just new software made by software engineers.

The second principle opens up the possibility that genuinely new domains which answer new questions, potentially requiring new approaches, are discovered. While real novelty might be rarer that one imagines, it is clearly also conceivable that something new is found by new generations of ethicists working with new technologies in new societal contexts. However, when this occurs, the work should not

automatically be placed in low-level ethics such as AI ethics just because that is where the hype is currently strongest. Very often, the questions, despite being new, relate to fundamental issues of science, engineering, and technology, and if so, that is where the work belongs.

A shorthand for choosing one's domain might be to apply Occam's razor whenever a domain is chosen, as the highest-level ethic capable of explaining and describing a phenomenon is simpler in terms of not including superfluous specifications of particular technologies. This principle nicely captures our main point, which is that science and theoretical models should not be needlessly duplicated, and that precedence should be given to simplicity (Duignan, 2021). When having to choose between aligning with, for example, AI or more basic theories involving less complicated technological foundations capable of dealing with an issue, the simpler should be chosen when it works as well or better than the alternatives. The simplicity and non-complexity of ethical theories are important for fostering understanding and for pedagogical purposes. It is also important for furthering scientific progress, however, as testing and falsifying theories are easier the simpler a theory is (Popper, 2005).

4.3 Opposing Views

Before concluding, we must consider some objections to the argument we have made.

First, someone might argue that we need specialised, domain-specific rules to move ethics out of the armchair and into the real world. As we have mentioned, the complexities of technologies and business practices will potentially preclude effective analyses of ethical challenges at too high and abstract level. For example, trying to get to grips with the intricacies of algorithmic audits will be a tall order for the ethicist who insists that they will solve this with a general understanding of *technology* without an intimate understanding of how algorithms work and are applied. Abstract normative ethics is fun for philosophers, but it does not always connect with the real-world problems faced by designers and users of technological systems. For instance, one could argue that the specific models of 'fairness' in machine learning, that have been developed in the recent past, are valuable (perhaps) and you would not get those if you did not create a specific subfield of AI ethics.

In response, it is important to bear in mind that we are not arguing against the entire enterprise of applied ethics or the attempt to use case studies involving specific technologies to develop and understand ethical theories. We need to apply ethical principles to new technologies and new scenarios. This is an essential and beneficial practice. It is, rather, the creation of new tech-defined domains of inquiry to which we object. These, we argue, come with the risks of creating relatively sealed-off specialities that reinvent the wheel and ignore foundational insights. We submit that you can get the benefits of applied ethical insight without always seeking to carve out a new domain of ethical inquiry. Academic specialisation is both natural and beneficial, but mainly when it is based on a scientific approach of cumulative

knowledge building and the realisation that there will often be some giant's shoulders to stand on (Merton, 1965).

Second, and related to the previous objection, one might argue that subfields are needed to attract interdisciplinary expertise. You will not get engineers interested in meta ethics or normative ethics, but if you carve out a subfield of applied ethics that relates to their area of expertise, then, you might get them interested and that is what we need if ethics is to have real-world impact. In other words, the labels matter when it comes to motivating people to care. Creating a specific subfield of 'AI ethics' is just good marketing to the key stakeholders in that technology.

In response, there is probably some merit to this objection. Interdisciplinary work is hard. Different fields do not share the same conceptual foundations and assumptions. Building bridges to mutually beneficial collaboration requires a lot of work. That said, it is not obvious that repeatedly carving out new subfields of tech ethics is beneficial for interdisciplinary collaboration. Indeed, it may be counterproductive. If, as we have argued above, many digital and smart technologies work on the same basic technological foundation (computing machinery), then introducing new subfields simply risks perpetuating unnecessary disciplinary silos: AI engineers should be talking to data scientists and roboticists and vice versa (to give but one example). Stipulating that AI ethics is distinct from data and robot ethics may preclude the necessary interdisciplinary collaboration.

Third, one might argue that our position is too strong in that it applies to all subfields of ethics. Take medical ethics as an example. In the aftermath of WWII, this developed into a conceptually rich and rigorous subfield of ethics, generating its own checklists of principles, journals, conferences, specialists, research centres, and university degrees. Most people would accept that medical ethics is a valid and useful subfield of applied ethics. Could it not be argued that AI ethics (or whatever subfield of tech ethics we happen to be concerned with) is in a similar position? Indeed, some prominent AI ethicists have argued that the field should develop along the lines provided by the medical ethics model, albeit while noting significant differences between the two fields (Mittelstadt, 2019b; Véliz, 2019). A related approach could be to argue that AI ethics should look to and take inspiration from business ethics (Schultz and Seele, 2022).

In response to this, it is worth bearing in mind that our objection is not to the existence of subfields of applied ethics per se. Many subfields are valid and worth developing. Our objection relates, more particularly, to the proliferation of subfields of ethics with poorly defined conceptual boundaries and excessive reinventing of the wheel. We are arguing that we should create and accept subfields of ethics with caution and not zeal. In this respect, it is noteworthy that the field of medical ethics, for instance, has not undergone the same degree of ethical proliferation as technology ethics. While there are closely related fields of applied ethics (e.g. bioethics and neuroethics), medical ethicists have avoided the temptation to create distinctive subfields of applied medical ethics, such as keyhole surgery ethics, nephrology ethics, or oncology ethics. This may be because each branch of medicine shares the same basic ethical goal—to improve the health of patients—and so the ethical focus remains the same across all subspecialities of medicine. Whatever the reason, this is very different from the situation we find in technology ethics. As we have clearly

shown with our survey, the literature reveals an excessive number of poorly defined subfields. Some of these may be worth retaining, but only after an appropriate cull.

Fourth, and finally, one might argue that people (particularly academics and researchers) must chase the money and follow the hype cycle. We need grants, we need readers, and we need to attract students. By attaching our expertise to new technologies and generating ethical insights about their application—however generic or unoriginal they may be—we make ourselves relevant and can attract the requisite attention and financial input. This may be cynical and self-serving, but it is a practical necessity and cannot be overlooked.

We certainly sympathise with this objection. We feel the pull of these incentives too. But this is a poor reason for creating subfields of ethics if, as we have argued, this is counterproductive to ethical decision-making and insight. Indeed, if this is the motivation for the proliferation of tech ethics, then it seems we have even more reason to reject this practice.

5 Conclusion

The ethics of technology is garnering attention for a reason. Just about everything in modern society is the result of, and often even infused with, some kind of technology. The ethical implications are plentiful, but how should the study of applied tech ethics be organised? We have reviewed a number of specific tech ethics, and argued that there is much overlap, and much confusion relating to the demarcation of different domain ethics. For example, many issues covered by AI ethics are arguably already covered by computer ethics, and many issues argued to be data ethics, particularly issues related to privacy and surveillance, have been studied by other tech ethicists *and* non-tech ethicists for a long time.

We have proposed two simple principles that should help guide more ethical research to the higher levels of tech ethics, while still allowing for the existence of lower-level domain specific ethics. If this is achieved, we avoid confusion and a lack of navigability in tech ethics, ethicists avoid reinventing the wheel, and we will be better able to make use of existing insight from higher-level ethics. At the same time, the work done in lower-level ethics will be both valid and highly important, because it will be focused on issues *exclusive* to that domain. For example, robot ethics will be about those questions that only arise when AI is embodied in a particular sense, and not all issues related to the moral status of machines or social AI in general.

While our argument might initially be taken as a call to arms against more than *one* fundamental applied ethics, we hope to have allayed such fears. There are valid arguments for the existence of different types of applied ethics, and we merely argue that an exaggerated proliferation of tech ethics is occurring, and that it has negative consequences. Furthermore, we must emphasise that there is nothing preventing anyone from making specific guidelines for, for example, AI professionals, based on insight from *computer* ethics. The domains of ethics and the needs of practitioners are not the same, and our argument is consequently that ethical research should be more concentrated than professional practice.

To change the undesirable situation we have described, actions on several levels and in different sections are required. The most obvious target group of our efforts is researchers, who produce much of the research and generate the foundational ethical analyses used to guide practitioners and shape policy and regulation. Existing research can use the proposed hierarchy and the two criteria for choosing how to position their work to ensure that they search for ways to build on extant research in the higher-level ethics. Academia can also play a crucial role in changing this situation, and a key action relates to restructuring the way we teach ethics of science and technology to students. One way to ameliorate the current situation would be to split ethics education into introductory joint classes in science and technology ethics before splitting the student group into domain specific groups. In such groups, one should focus on (a) what can be learned from other domains and (b) what is novel for the technology of the specific domain. Industry and practitioners are a major cause of the proliferation, and hence, another potential target for reform, but we argue that it seems unlikely that the industry itself will see the need to solve the problem (Sætra et al. 2021).

The tech industry and practitioners have incentives to hype new technologies both to more effectively sell them and also to more effectively be able to avoid or shape regulation to suit its interests. Nevertheless, if incoming students are educated as just discussed, and researchers increasingly resist excessive proliferation, industry will inevitably find the grounds towards increased proliferation harder to tread. Finally, government and regulators are the consumers of tech ethics research and producers of tech regulation, and increased awareness of and knowledge about the challenges here discussed will help resist efforts to pursue potentially unnecessary laws that might be covered through more foundational and general—and thus potentially future-proof—regulation.

Declarations

References

Aggarwal, N. (2020). Introduction to the special issue on intercultural digital ethics. *Philosophy Technology, 33*(4), 547–550. https://doi.org/10.1007/s13347-020-00428-1

Allen, C., Wallach, W., & Smit, I. (2006). Why machine ethics? *IEEE Intelligent System, 21*(4), 12–17.

Anderson, S. L. (2011). Machine metaethics. In M. Anderson & S. L. Anderson (Eds.), *Machine ethics* (pp. 21–27). Cambridge.

Anderson M, & Anderson SL (2011) *Machine ethics.* Cambridge University Press.

Aristotle. (2014). *Nicomachean Ethics* (C. D. D. Reeve, Trans.). Indianapolis: Hackett Publishing.

Asimov, I. (2013). *I, robot*. Harpercollins Publishers.

Benjamin, R. (2019). *Race after technology*. Polity Press.

Capurro, R. (2006). Towards an ontological foundation of information ethics. *Ethics and Information Technology, 8*(4), 175–186.

Capurro, R. (2009). *Digital ethics.* Paper presented at the Global Forum on Civilization and Peace. http://www.capurro.de/korea.html. Accessed 2 July 2021.

Coeckelbergh, M. (2020). *AI ethics.* MIT Press.

Copp, D. (2005). *The Oxford handbook of ethical theory*. Oxford University Press.

Costanza-Chock, S. (2020). *Design justice: Community-led practices to build the worlds we need*. The MIT Press.

Cummings, M. L. (2006). Integrating ethics in design through the value-sensitive design approach. *Science and Engineering Ethics, 12*(4), 701–715. https://doi.org/10.1007/s11948-006-0065-0

D'Ignazio, C., & Klein, L. (2020). *Data feminism*. MIT Press.

Danaher, J. (2019). The philosophical case for robot friendship. *Journal of Posthuman Studies, 3*(1), 5–24.

Danaher, J. (2020b). Welcoming robots into the moral circle: A defence of ethical behaviourism. *Science and Engineering Ethics, 26*(4), 2023–2049. https://doi.org/10.1007/s11948-019-00119-x

Danaher, J. (2020a). Robot betrayal: A guide to the ethics of robotic deception. *Ethics and Information Technology, 22*, 1–12. https://doi.org/10.1007/s10676-019-09520-3

Dotan, R. (2021). The proliferation of AI ethics principles: What's next? Retrieved from https://montrealethics.ai/the-proliferation-of-ai-ethics-principles-whats-next/. Accessed 10 Dec 2021.

Drew, C. (2016). Data science ethics in government. *Philosophical Transactions of the Royal Society a: Mathematical, Physical and Engineering Sciences, 374*(2083), 20160119. https://doi.org/10.1098/rsta.2016.0119

Dubber, M. D., Pasquale, F., & Das, S. (2020). *The Oxford Handbook of Ethics of AI*. Oxford University Press.

Duignan, B. (2022, August 24). *Occam's razor*. Encyclopedia Britannica. https://www.britannica.com/topic/Occams-razor

Ellul, J. (1964). *The technological society*. Vintage Books.

Fazelpour, S., & Danks, D. (2021). Algorithmic bias: Senses, sources, solutions. *Philosophy Compass, 16*(8), e12760.

Fleddermann, C. B. (2004). *Engineering ethics*. Prentice Hall.

Fleetwood, J. (2017). Public health, ethics, and autonomous vehicles. *American Journal of Public Health, 107*(4), 532–537.

Floridi, L. (1999). Information ethics: On the philosophical foundation of computer ethics. *Ethics and Information Technology, 1*(1), 33–52.

Floridi, L., & Cowls, J. (2019). A unified framework of five principles for AI in society. *HDSR, 1*(1). https://doi.org/10.1162/99608f92.8cd550d1

Franzke, A. S., Muis, I., & Schäfer, M. T. (2021). Data ethics decision aid (DEDA): A dialogical framework for ethical inquiry of AI and data projects in the Netherlands. *Ethics and Information Technology, 23*, 551–567. https://doi.org/10.1007/s10676-020-09577-5

Friedman, B., & Nissenbaum, H. (1996). Bias in computer systems. *ACM Transactions on Information Systems (TOIS), 14*(3), 330–347.

Gartner. (2021, September 7th). Gartner identifies four trends driving near-term artificial intelligence innovation. Retrieved from https://www.gartner.com/en/newsroom/press-releases/2021-09-07-gartner-identifies-four-trends-driving-near-term-artificial-intelligence-innovation. Accessed 10 Dec 2021.

Geisslinger, M., Poszler, F., Betz, J., Lütge, C., & Lienkamp, M. (2021). Autonomous driving ethics: From trolley problem to ethics of risk. *Philosophy & Technology, 34*(4), 1033–1055. https://doi.org/10.1007/s13347-021-00449-4

Gordon, J.-S., & Nyholm, S. (2021). Ethics of artificial intelligence. In *Internet Encyclopedia of Philosophy*. https://iep.utm.edu/ethics-of-artificial-intelligence/

Gunkel, D. J. (2014). A vindication of the rights of machines. *Philosophy & Technology*, *27*, 113-132. https://doi.org/10.1007/s13347-013-0121-z

Gunkel, D. J. (2017). Mind the gap: Responsible robotics and the problem of responsibility. *Ethics and Information Technology*, 1–14. https://doi.org/10.1007/s10676-017-9428-2

Gunkel, D. J. (2018). *Robot rights*. MIT Press.

Hand, D. J. (2018). Aspects of data ethics in a changing world: Where are we now? *Big Data, 6*(3), 176–190.

Harris, C. E., Jr., Davis, M., Pritchard, M. S., & Rabins, M. J. (1996). Engineering ethics: What? why? how? and when? *Journal of Engineering Education, 85*(2), 93–96.

Harris Jr, C. E., Pritchard, M. S., Rabins, M. J., James, R., & Englehardt, E. (2013). *Engineering ethics: Concepts and cases.* Cengage Learning.

Hasselbalch, G. (2019). Making sense of data ethics The powers behind the data ethics debate in European policymaking. *Internet Policy Review, 8*(2), 1–19.

Horowitz, M. C. (2016). The ethics & morality of robotic warfare: Assessing the debate over autonomous weapons. *Daedalus, 145*(4), 25–36.

Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence, 1*(9), 389–399.

Johnson, D. G. (2004). Computer ethics. In L. Floridi (Ed.), *The Blackwell Guide to the Philosophy of Computing and Information* (pp. 65–75). Blackwell Publishing.

Jonas, H. (1982). Technology as a Subject for Ethics. *Social Research, 49*(4), 891–898.

Kleinberg, J., Ludwig, J., Mullainathan, S., & Sunstein, C. R. (2018). Discrimination in the age of algorithms. *Journal of Legal Analysis, 10*, 113–174.

Kochetkova, T. (2015). An overview of machine medical ethics. In S. P. v. Rysewyk & M. Pontier (Eds.), *Machine Medical Ethics* (pp. 3–15). Springer.

Langford, D. (2000). *Internet ethics*. Macmillan London.

Lannin, D. G., & Scott, N. A. (2013). Social networking ethics: Developing best practices for the new small world. *Professional Psychology: Research and Practice, 44*(3), 135.

Lin, P., Abney, K., & Bekey, G. (2011). Robot ethics: Mapping the issues for a mechanized world. *Artificial Intelligence, 175*(5–6), 942–949.

Luke, A. (2018). Digital ethics now. *Language and Literacy, 20*(3), 185–198.

Macnish, K. (2018). *The ethics of surveillance: An introduction*. Routledge.

Marx, G. T. (1998). Ethics for the new surveillance. *The Information Society, 14*(3), 171–185.

Matthias, A. (2004). The responsibility gap: Ascribing responsibility for the actions of learning automata. *Ethics and Information Technology, 6*(3), 175–183.

Merton, R. K. (1942). Science and technology in a democratic order. *Journal of Legal and Political Sociology, 1*(1), 115–126.

Merton, R. K. (1965). *On the shoulders of giants: A Shandean postscript*. The Free Press.

Merton, R. K. (1972). *The sociology of science: Theoretical and empirical investigations*. University of Chicago press.

Merton, R. K. (1973). The normative structure of science. In N. W. Storer (Ed.), *The sociology of science: Theoretical and empirical investigations* (pp. 267–268). University of Chicago Press.

Mittelstadt, B. (2019a). The ethics of biomedical 'big data' analytics. *Philosophy & Technology, 32*(1), 17–21. https://doi.org/10.1007/s13347-019-00344-z

Mittelstadt, B. (2019b). Principles alone cannot guarantee ethical AI. *Nature Machine Intelligence, 1*(11), 501–507.

Moor, J. H. (1985). What is computer ethics? *Metaphilosophy, 16*(4), 266–275.

Moor, J. (1999). Ethics of Privacy Protection. *Library Trends, 39*(1 & 2), 69–82.

Moor, J. H. (2006). The nature, importance, and difficulty of machine ethics. *IEEE Intelligent Systems, 21*(4), 18–21.

Müller, V. C. (2020). Ethics of artificial intelligence and robotics. In E. N. Zalta (Ed.), *Stanford Encyclopedia of Philosophy (Summer 2020 Edition)*. https://plato.stanford.edu/archives/sum2020/entries/ethics-ai/

Mumford, L. (1934). *Technics and civilization*. Routledge & Kegan Paul LTD.

Noble, S. U. (2018). *Algorithms of oppression: How search engines reinforce racism*. New York University Press.

Nolt, J. (2014). *Environmental ethics for the long term: An introduction.* Routledge.

Open Data Institute. (2021). Data Ethics Canvas. Retrieved from https://theodi.org/article/data-ethics-canvas/#1562602644259-1d65b099-ea7b. Accessed 10 Dec 2021.

Popper, K. (2005). *The logic of scientific discovery*. Routledge.

Richards, N. M., & King, J. H. (2014). Big data ethics. *Wake Forest L. Rev., 49*, 393.

Roemer, J. E. (1998). *Equality of opportunity*. Harvard University Press.

Rollin, B. E. (2006). *Science and ethics.* Cambridge University Press.

Romele, A. (2022). Images of Artificial Intelligence: A Blind Spot in AI Ethics. *Philosophy & Technology, 35*(1), 1–19. https://doi.org/10.1007/s13347-022-00498-3

Russell, S., & Norvig, P. (2014). *Artificial intelligence: A modern approach* (3rd ed.). Pearson.

Sætra, H. S. (2018). Science as a vocation in the era of big data: The philosophy of science behind big data and humanity's continued part in science. *Integrative Psychological and Behavioral Science, 52*(4), 508–522.

Sætra, H. S. (2020). The parasitic nature of social AI: Sharing minds with the mindless. *Integrative Psychological and Behavioral Science, 54*, 308–326. https://doi.org/10.1007/s12124-020-09523-6

Sætra, H. S. (2022). *AI for the sustainable development goals*. CRC Press.

Sætra, H. S., & Fosch-Villaronga, E. (2021). Research in AI has Implications for society: How do we respond? *Morals & Machines, 1*(1), 60–73.

Sætra, H. S. (2021a). Confounding complexity of machine action: A Hobbesian Account of machine responsibility. *International Journal of Technoethics, 12*(1). https://doi.org/10.4018/IJT.2021a0101.oa1

Sætra, H. S. (2021b). Loving robots changing love: Towards a practical deficiency-love. *Journal of future robot life, 3*(2), 109–127. https://doi.org/10.3233/FRL-200023

Sætra, H. S. (2021c). Social robot deception and the culture of trust. *Paladyn, Journal of Behavioral Robotics, 12*(1). https://doi.org/10.1515/pjbr-2021c-0021

Sætra, H. S., Coeckelbergh, M., & Danaher, J. (2021). The AI ethicist's dilemma: Fighting big tech by supporting big tech. *AI and Ethics, 2, 15–27.* https://doi.org/10.1007/s43681-021-00123-7

Schultz, M. D., & Seele, P. (2022). Towards AI ethics' institutionalization: Knowledge bridges from business ethics to advance organizational AI ethics. *AI and Ethics*, 1-13 https://doi.org/10.1007/s43681-022-00150-y

Siegel, M. (1979). Privacy, ethics, and confidentiality. *Professional Psychology, 10*(2), 249.

Singer, P. (2011). *Practical ethics.* Cambridge university press.

Tavani, H. T. (2012). Search engines and ethics. In E. N. Zalta (Ed.), *The stanford encyclopedia of philosophy*.

Tavani, H. T. (2016). *Ethics and technology.* Wiley.

Turing, A. M. (2009). Computing machinery and intelligence. In Robert Epstein, Gary Roberts, & G. Beber (Eds.), *Parsing the Turing Test* (23–65). Springer.

Vallor, S. (2012). Social networking and ethics. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*.

Véliz, C. (forthcoming). *Oxford Handbook of Digital Ethics*. Oxford: Oxford University Press.

Véliz, C. (2019). Three things digital ethics can learn from medical ethics. *Nature Electronics, 2*(8), 316–318. https://doi.org/10.1038/s41928-019-0294-2

Westin, A. F. (1967). *Privacy and freedom*. IG Publishing.

Whiting, R., & Pritchard, K. (2018). Digital ethics. In C. Cassell, A. L. Cunliffe, & G. Grandy (Eds.), The SAGE Handbook of Qualitative Business and Management Research Methods (pp. 562–579). Sage Reference.

Winfield, A. (2012). *Robotics: A very short introduction.* OUP Oxford.

Winner, L. (1977). *Autonomous technology: Technics-out-of-control as a theme in political thought*. MIT Press.

van Wynsberghe, A. (2021). Sustainable AI: AI for sustainability and the sustainability of AI. *AI and Ethics*, 1-6 https://doi.org/10.1007/s43681-021-00043-6

Zwitter, A. (2014). Big data ethics. *Big Data & Society, 1*(2), 2053951714559253.