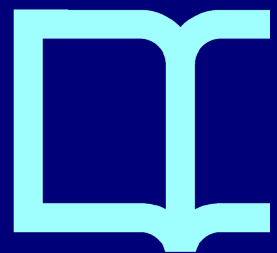**Document Version**

This is the Accepted Manuscript version.
The version in the Universitat Oberta de Catalunya institutional repository, O2 may differ from the final published version.

Enquiries

If you believe this document infringes copyright, please contact the Research Team at: repositori@uoc.edu

Ethical Idealism, Technology and Practice: A Manifesto

Joan Casas-Roma[1*]

[1*]Universitat Oberta de Catalunya (Department of Computer Science), Barcelona, Spain. ORCID: 0000-0003-1641-1325.

Corresponding author(s). E-mail(s): jcasasrom@uoc.edu;

Abstract

Technology has become one of the main channels through which people engage in most of their everyday activities. When working, learning, or socializing, the affordances created by technological tools determine the way in which users interact with one another and their environment, thus favoring certain actions and behaviors, while discouraging others. The ethical dimension behind the use of technology has been already studied in recent works, but the question is often formulated in a protective way that focuses on shielding the users from potential detrimental effects. Nevertheless, when considering collateral ethical benefits that the use of technology could bring about, virtue ethics and the notions of "practice" and "practical wisdom" present new opportunities to harness this potential. By understanding the combination of technology, its users and their interactions as a system, technology can be seen as the space where most of its users' daily practice happens. Through this practice, users can get the chance to collaterally develop and enhance their ethical awareness, sensitivity and reasoning capabilities. This work is shaped as a manifesto that provides the background, motivations and directions needed to ask a complementary question about the ethics of technology that aims towards the potentiality behind the use of technology. Instead of focusing on shielding the users, the proposed *ethical idealist* approach to the ethics of technology aims to empower them by understanding their use of technology as the space where the development of their practical wisdom, understood in the virtue ethics' sense, takes place.

1 Introduction and Motivations

With practice comes mastery. This is something that is often said about many skills –from skills belonging to scientific and technical disciplines, such as software programming, to artistic skills, like painting, or playing an instrument, as well as many other skills like cooking, or playing soccer. Even skills that rely almost entirely on cognitive faculties, such as performing mathematical operations, or memorizing a text, are also improved through practice. And even when talking about living, or, more specifically, about "living well", theories of virtue ethics argue that, in order to pursue a good life[1], in order to live well, practice and practical wisdom are needed in order to learn how to understand, act and live according to a series of virtues, principles, or values Vallor (2016).

Under this view, even the aspiration towards a good life requires repeated, conscious practice of the acts that, when built into a habit, allow us to live such kind of life. Conversely to the kind of practice related to honing specific skills, such as programming, painting, or cooking, which usually have a clear and delimited space and context where they can be trained, practice towards living well takes place almost continuously and in each and every act, decision, behavior and interaction we have in our day to day.

These acts and interactions can take place in a plethora of contexts and situations, and through many different channels and mediums –either by carrying out certain actions, through acts of speech, or even by interacting through a digital platform. In fact, a growing number of our daily interactions take place in digital environments and through technological tools of some sort – like working through internet-based platforms, buying products via our mobile devices, learning via online education platforms, or socializing through digital social networks. Considering this, technological tools, as well as the digital environments they create, can be understood as "spaces" where a large number of our daily interactions take place, and therefore where a large part of our daily practice (understood as the series of actions, choices and behaviors that we repeatedly do in order to build our habits) happens.

When a space changes in some way, the actions that can be carried out in it can change accordingly; this is as true about physical spaces, as it is about digital ones. Therefore, when new technologically-defined spaces appear,

---

[1]Some contemporary authors use different terms that roughly point to this notion. In (Garcés, 2021, pp. 19), the author refers to a "dignified", or a "decent life" (translated from the Catalan term "vida digne"); in (Han, 2017, pp. 17-28), the author introduces a sort of negative counterpart as the "bare life" (which corresponds more to a matter of existing, rather than to experiencing any sort of flourishing and fulfillment); in Vallor (2016), the author uses the term "flourishing" to refer to living a fulfilling life allowing both personal and collective growth –which, in a sense, is tied up to the existence of a possibility space enabling growth that certain ethical theories, like ethics of care Held et al. (2006), or relational ethics Metz and Miller (2016), aim to create. In the present work, these notions will be used interchangeably.

our interactions with, and in those spaces change as well. New technological advances are increasingly becoming more complex, more autonomous and more far-reaching than ever –specially when they involve the use of artificial intelligence (AI) techniques that partially, or fully automatize some decision-making procedures. Furthermore, the progressive, but rapidly-growing digitalization of almost every layer of society leads to the fact that almost every activity in our everyday life is mediated, manageable and accessible through technological tools of some sort and has its corresponding digital space.

These uses and interactions, however, often bring ethical considerations into the debate. There are several cases in different fields where AI-driven systems, despite having been initially thought to be objective and fair (Caliskan, Bryson, & Narayanan, 2017, pp. 1), have been found to exhibit biased decision procedures that have had a detrimental impact among some of their users Angwin, Larson, Mattu, and Kirchner (2016); Favaretto, De Clercq, and Elger (2019); Yapo and Weiss (2018). Due to this, the study of ethical implications in the design of new technology, and particularly in the design of AI-driven tools, is currently of paramount importance among public and private institutions alike Association of Nordic Engineers (2021); EU Parliament (2021); HLEG on AI (2019); IEEE (2016). Nevertheless, and even though most of these works focus on preemptively identifying their potential detrimental uses, AI-driven technologies can also bring numerous ethically-beneficial effects to our society Russell, Dewey, and Tegmark (2015). Technological advancements, therefore, always bring with them new uses and new interactions that present both risks and opportunities with a clear ethical dimension (in fact, and regarding to potential ethical risks, some works like Von Schomberg and Blok (2021) explore whether Responsible Innovation is even possible, when advancements are tied up to technological innovation).

The present work is a manifesto that explores and reflects upon the way the question about the ethics of technology (and, in particular, technological tools classified as *information technologies*, or IT tools, which are systems characterized by allowing to input, store, manipulate, retrieve and send information) is currently asked, and introduces a new approach towards how this question could be asked in order to identify and foster potential ethically-beneficial effects that the use of IT tools could bring to their users. This reflection, nevertheless, is not focused on reaping ethical benefits based on the *goal* that any particular technological tool might have (understanding the goal as the problem it explicitly tries to solve, the situation it intentionally aims to address, or the service it intends to offer by design; for instance, one of the main goals of social networks is to allow users to connect and interact with each other), but rather through the *way* technological tools might be used (understanding the way as how their users use such technology, which, continuing with the example of social networks, could refer to users connecting to help each other in a study program, or to provide support to other users suffering from a mental health disorder, or to organize an act of social protesting, or a social movement demonstration), regardless of their intended goal. The core of the approach

presented in this manifesto, therefore, focuses on the uses, affordances[2] and interactions created by technology, as well as on how its users could benefit from the actions, behaviors and reflections potentially prompted by those uses, affordances and interactions. The aim of this manifesto is not to answer the currently open question about ethics and technology, but rather to provide an alternative paradigm from which to ask and answer a different, but complementary question to the existing one: a question based on potentiality and focused not on the technology itself, but rather on its users and their flourishing through the way they use technology.

In order to guide this reflection, some preliminary notions, analogies and arguments are needed. This work will start by reflecting on the relation between ethical theories and ethical autonomy for the remainder of Section 1. Then, an analogy between two major theories in political philosophy and two alternative approaches to the ethics of technology will be presented in Section 2. The relationship between technology, ethical autonomy and the *ethical idealist* question introduced in this manifesto will be further explored in Section 3. Lastly, some concluding remarks and considerations will be presented in Section 4.

1.1 Ethical Theories and Delegated Autonomy

The notion of "autonomy"[3] is one of the key concepts that often appears when describing intelligent behavior, in general, as well as the human condition, in particular Chirkov, Ryan, and Sheldon (2010); Deci and Ryan (1995). Delegating our own autonomy is, nevertheless, something that happens quite often in particular tasks, or domains of expertise, where others might be deemed more competent to make a favourable choice in our place. However, delegating our autonomy by default could lead to a progressive loss of our own competence in that field and, if carried out in a systematic way, this delegation can end up hindering our capacity to understand, act and make informed choices, as one becomes progressively more detached with the tasks and skills that have been delegated. In (Garcés, 2021, pp. 44-45), the author points out how using AI as a means to make better decisions can be seen as a sort of "delegated intelligence". Under this understanding, we delegate our autonomy to make a choice to an algorithm-based system that does the choice for us. This not only aims for that choice to be, ideally, more fruitful, but it also delegates the responsibility and the reasoning behind that choice –but potentially at the expense of progressively hindering our own capacity to understand and make an informed choice on that matter.

---

[2]The term "affordance" was introduced by the American psychologist James J. Gibson in the following way: "the affordances of the environment are what it offers the animal, what it provides or furnishes, either for good or ill" (Gibson, 1979, pp. 127-137). In the field of human-computer interaction, an affordance is often understood in a similar way as the relation between the object and the user: "an affordance is what the user can do with an object based on the user's capabilities. As such, an affordance is not a 'property' of an object (...) An affordance is, in essence, an action possibility in the relation between user and an object" Interaction Design Foundation (n.d.).

[3]Different fields provide slightly different definitions of "autonomy", but they always have a common core involving the notions of agency and the capacity to make an informed choice devoid of any form of coercion.

A parallelism can be found between different ethical theories and the notions of ethical sensitivity, responsibility, and ethical reasoning –which, in the present work, will be grouped into the notion of *ethical autonomy*. Aiming for almost scientific-like objectivity, ethical theories based on weighing the consequences of an action, such as utilitarianism Mill (1987), try to define a sort of *calculus* to determine what is the best ethical choice –by bringing the most happiness to the greatest number of people or, alternatively, the least suffering to the smallest number Acton and Watkins (1963). Similarly, rule-based ethics (such as deontological ethics, and particularly deontological codes of conduct found, for instance, in private industries and associations, such as Association for Computing Machinery (ACM) (2018)) aim to distill the ethically-right actions through a set of rules to be followed Davis (1993). The goal behind such approaches is to detach ethical decisions from subjective judgments in order to ensure that the decision is as objectively optimal as it can be. The *subject* becomes less important, as the same situation should lead to the same choice, regardless of who is the one making it.

Regardless of any differences that there might be between these two approaches to ethics (as well as any challenges and shortcomings they might have), they do have one thing in common: the subject loses relevance, as the whole point is that the ethical dilemma should be "solvable" (in the sense of determining the best ethical choice) in terms of objective considerations, devoid of any personal involvement from the actors taking part in the scenario. Under this understanding, these ethical theories involve a *delegation of ethical autonomy* in which the actors that take part in the ethical dilemma resort to external rules and procedures in order to provide an objective justification for their choice. The actors' ethical sensitivity to identify ethically-relevant situations, their ethical reasoning capabilities to grasp the nuances behind their choices, and their overall autonomy as ethically-relevant actors is delegated to an external system that tells the actors how to act in each case and takes their cognitive responsibility out of the equation.

This might seem, at first glance, as something that is both positive and desirable. Ethically-relevant decisions can often have profound consequences that affect the well-being of potentially many people, and therefore it seems it would be desirable to have an objective way of knowing what the best choice would be. Nevertheless, this also has a side effect: it progressively alienates and detaches the actors involved in ethically-relevant decisions from the choices they make. These actors do not need to understand, empathize, or develop their ethical sensitivity: they only need to resort to the external rules forming the ethical theory and pick a minimal set that distills the nuances and complexities of the current situation at hand as best as possible in order to allow the actors to delegate their involvement and responsibility in it[4]. This not only tends to turn potentially complex scenarios into shallow and oversimplified caricatures, but it also progressively creates a detachment between the ethical

---

[4]This is expressed quite succinctly in (Pojman & Fieser, 2016, pp. 142) in the following way: "Deontological systems focus on an egoistic, minimal morality whose basic principles seem more preventative than positive".

agent making a decision, and the ethical patients that become affected by the outcomes; the only thing that matters in justifying such decision is being able to point at the rules that apply in that case. In other words, ethical actors no longer need to develop and train the ethical autonomy that would allow them to identify, reflect upon and reason about the particularities of each ethically-relevant decision. By delegating ethical autonomy to algorithmic systems based on rules and mathematical calculi, we alienate ourselves from our subjective understanding and empathy towards ethically-relevant situations and become progressively more detached from this kind of reasoning, which can slowly lead us to lose touch with the ethical dimension inherent to reality, as well as with the way our decisions affect the world and the others living in it.

Aside from the role that the actors involved in ethical decision-making might play, aiming for algorithmic-like precision in ethical reasoning presents further challenges. Even though many ethical rules and principles may be clear and easy to understand in the vacuum, or in hypothetical scenarios, ethical decisions taking place in our everyday lives are highly contextual and present a level of complexity that often goes beyond the simplicity that objective ethical theories seek to capture. Due to this, many of these systems need to resort to *ad hoc* exceptions (Vallor, 2016, pp. 24), additions and amendments to be able to deal with the myriad of contextual features that inevitably come with ethically-relevant scenarios –and, unfortunately, the need for these *ad hoc* amendments is often identified *after* the current system has already shown to be inadequate, and therefore after someone has already suffered this inadequacy.

Some rule-based conceptions of ethics, such as the theory of *prima facie* duties Dancy (1993), do establish a sort of "exception mechanism" in which certain rules can override other rules which, under normal conditions, would hold. After all, ethical dilemmas are often about the exception, rather than the norm. But being able to recognize such exceptions requires both awareness and understanding in order to evaluate when certain rules might prevail, or when the exception should take preference. At this point, ethical autonomy can no longer be delegated to a sort of algorithm-based decision system. Suddenly, an approach to ethics that aimed at keeping the subject "outside" of the ethical decision needs to bring the subject back into it in order to use their ethical autonomy to grasp the nuances that the rule-based abstraction cannot account for. But, precisely because the subject was deemed unnecessary when aiming for an objective and optimal ethical decision-making procedure, the subject has progressively lost the understanding and skill that would allow them to think, reason and act in an ethically-desirable way in this case: the subject was both doomed to progressively lose, as well as further deprived of developing and training, their *practical wisdom*.

1.2 Regaining Ethical Autonomy through Practice

The converse of the aforementioned ethical theories, in terms of subjective involvement, would be virtue ethics theories (such as Aristotle's, Confucius' and Buddhist theories (Vallor, 2016, pp. 36-42)). In a nutshell, these theories

identify a set of virtues, principles, or values that should guide one's actions and that should be pursued and practiced throughout one's life –thus swapping the question "how should I act", by the question "how should I live". The main distinction, with respect to theories based on rules, or calculi, is that those virtues are not prescriptions of what should be done, but rather directions one should (usually) aim for; nevertheless, the complexities of each decisions' context makes the subject central within the picture and requires that the subject uses their sensitivity, reasoning capabilities, responsibility and autonomy in order to decide what is best to do in each case –which might be in accordance with a virtue, or not. An example that is usually presented in this case refers to the virtue of "courage": while too little courage risks letting fear block almost every decision to be made, too much courage can lead to reckless behaviors resulting in stupid decisions. The key, therefore, does not lie on blindly sticking to the virtues for each and every decision, but rather on leveraging those virtues with a *practical wisdom* (from the Greek term *phrónēsis*, often translated also as "prudence", or "prudential reason" (Vallor, 2016, pp. 18-19)) built upon practice and experience. Virtues are not, therefore, "places" that one could reach, or precepts that could be "completed", but rather beacons to aim for throughout a lifelong process of cultivating one's own practical wisdom[5], but knowing that they are not categorical norms where one could permanently establish the justification behind their ethically-relevant choices.

Where is the catch? That this practical wisdom, alongside the way virtue ethics keeps the subject relevant to the ethical decision, does not guarantee that two different subjects would make the same decision in an otherwise equal scenario. The way those subjects identify, evaluate and understand the nuances of that particular scenario, their ethical sensitivity, their empathy, as well as their past experiences and future expectations can have an effect on the way they decide to act. This not only poses a challenge for algorithmic representations of such situations, but it can also lead to certain choices being considered correct, or optimal, whereas others might not be. In this sense, virtue ethics moves back part of the focus from a set of objective, measurable conditions that determine the best path in an ethically-relevant choice, to the subjective perception and judgement of the actors involved in such choice. One could argue that it "de-objectivises" ethical decision-making and brings the realm of subjectivity back into the equation[6].

---

[5]In fact, the way virtue ethics theories understand the cultivation of virtues as a lifelong process has many similarities with theories of heutagogy and lifelong learning Blaschke (2012, 2018); Carr, Balasubramanian, Atieno, and Onyango (2018).

[6]In (Han, 2018, pp. 44), and referring to an idea of Alain Ehrenberg, the author points out how the nowadays culture centered around performance and optimization avoids devoting effort into conflict, as conflict usually requires time –time that could be seen as "wasted", as it could have been used instead in a more productive way. Nevertheless, the author argues that conflicts are not destructive *per se*, as they also have a constructive side in which the subject grows and becomes more mature by facing them. Behind this reflection, one might find another clue regarding today's preference for "externally-driven" ethical theories, as they try to evade internal (subjective) conflict, whereas virtue ethics retain subjective conflict at its core –and as a prerequisite for the subject to develop their own practical wisdom.

What is, then, the benefit of virtue ethics[7], with respect to other ethical theories that rely on more objective approaches to ethical decision-making? That virtue ethics *inevitably requires* the subject to take an active role into the ethical decision-making. Without showing an awareness of the relevant context, without understanding the nuances it presents, and without building the ethical sensitivity needed to appreciate the potential consequences of such decision (in sum: without having developed ethical autonomy and practical wisdom), virtue ethics cannot work, as it does not try to reduce the sheer complexity of ethical decision-making into a set of simplified and alienating rules that could be mindlessly followed by anyone to reach (at least theoretically) an ethically-desirable result.

And why would it be desirable to take ethical autonomy back to the subject, instead of having an objective procedure to decide for us in ethically-relevant matters? Because current and upcoming challenges[8] will inevitably require both individual and collective involvement from an ethically-informed position that no rule, nor calculus-based approach, could fully replace. Ethical autonomy, therefore, is a capacity that is essential both in order to face emergent challenges from an ethically-informed position, as well as to identify and harness potential ethically-beneficial opportunities. Alienating the subject from ethically-relevant decisions and favoring a delegation of their ethical sensitivity and reasoning capabilities to an external decision procedure system will further hinder the development of the subject's ethical autonomy required to grasp the relevance of decisions they will face, as well as of outcomes that will affect them.

Taking this into account, the idea of regaining ethical autonomy through practice points towards the need to go beyond a conception of ethics fully described and governed by "objective" theories that are solely based on external rules and calculations (i.e.: those theories where the ethical subject becomes irrelevant), and instead pursue a conception of ethics based on enhancing and developing the ethical autonomy (i.e.: the ethical awareness, sensitivity and reasoning skills) of the ethical subjects by exposing them to actions, interactions and behaviors that can help "awake" and practice those relevant ethical skills. This idea of "practice" is a key difference to other approaches to the learning of ethical competences, which might instead be based on learning and sticking to a set of ethical rules and norms. Instead, the way to develop such ethical skills, in order to regain one's ethical autonomy, is through enacting actions, behaviors and interactions rooted in relevant ethical principles that can be grasped, understood and, ideally, replicated later in other contexts and situations.

---

[7]In fact, since the last decades of the 20th Century there has been a contemporary renewal of interest in virtue ethics theories from different authors (Anscombe (1958); MacIntyre (2013), (Vallor, 2016, pp. 20-23)), probably motivated in part by the shortcomings identified in other, more popular approaches to ethics –such as utilitarian and deontological approaches.

[8]It should be enough to recall how to recent COVID-19 global pandemic, or some of the current environmental challenges, require an active and responsible participation from practically every member of society and how, beyond practices that can be regulated, each own autonomous decisions are key in order to try and face such challenges with some guarantees.

For this practice to enhance the subject's skills, nevertheless, it would need to be seamlessly integrated as part of the daily actions and interactions that the subject carries out. This need leads to the question of where, then, this practice could be integrated... What would be a medium where the subject could be exposed to those actions, interactions and behaviors daily, and which could therefore become a medium where the subject could be repeatedly exposed to such practices throughout their day? What medium could we find that mediates most of our daily activities? In our present times, this medium are IT tools and devices –like computer programs and applications running on our computer, or on our own mobile device.

1.3 Information Technologies: A Medium for Practice

Currently, most of our everyday activities are carried out via interactions with technological tools –specifically, IT tools based on the input, processing, and output of some kind of information Frischmann and Selinger (2018). Whether we think about professional and work-related activities, public and private services, education and training, leisure activities (specially those that involve multimedia products), or even socializing, it is easy to see how information technologies can be usually found at some point (if not throughout) when carrying out those activities. In this sense, IT tools and devices can be understood as the medium through which most of our everyday activities take place, and with which we end up interacting with throughout our daily life.

Most of these technologies are *online technologies* (i.e.: that make use of the Internet in order to connect with databases, services, and other devices and users as part of their core functionalities). This use of the Internet not only brings access to all kinds of information and the chance to almost instantly use and update such information, but it often also allows users to interact with each other through the affordances offered by a particular tool[9]. For instance, the way social networks are designed might allow a user to interact directly only with those users that are already part of their "group", or might limit the kind of interactions available to different types of users, or may even determine the set of ways in which a user might interact with another user (i.e.: allowing text-based comment, reaction to certain posts, direct messages, citations, etc). Considering this, there is a huge subset of our daily interactions with other people that happens within a virtual space, and which is enabled, mediated and, at the same time, determined, by the IT tools and devices that grant us access to such virtual spaces. As such, when considering where our daily practice takes place, it is only natural to consider IT tools and devices as one of the main medium through which this kind of interactions take place.

IT tools and devices are, therefore, the medium where the present work will focus on as the place where actions, behaviors and interactions aimed at providing a opportunity to enhance their users' ethical autonomy can be

---

[9]In fact, and while some authors like Smart, Heersmink, and Clowes (2017) see the Internet as a scaffolding for our own cognition and cognitive processes, other authors such as Krueger and Osler (2019) go one step beyond and suggest that the Internet is a scaffolding also for our affective processes –including our moods, emotions and our ability to regulate the way we feel.

located[10]. In order to dive deeper into how these kind of tools and devices could be used for such purpose, we first need to ask ourselves what the current relation between ethics and technology (particularly, IT technology) is. Furthermore, we need to ask ourselves whether the currently existing relation, or even the way this relation is commonly being explored, provides us with a good starting point for the goal of the present manuscript... And, in case it is not, we then need to ask the following question: how could the relation between ethics, IT tools and users could be re-thought in order to devise IT tools as a medium that could incorporate actions, behaviors and interactions that form the daily practice allowing their users to enhance their ethical autonomy?

2 Two Approaches to the Ethics of Technology

The way a question is asked determines the direction towards which one looks for answers. When questions are asked in order to address a challenge, the answers found are usually intended to fill that gap, but potentially ignoring collateral matters that could be inadvertently left out just because they do not directly fit in the hole identified by the initial question. In these cases, it is not that the answer is not satisfactory, as it might indeed provide a solution to the challenge posed by the initial question, but the answer could end up being partial, or miss out on additional potential benefits due to its focused scope.

The question around the ethics of technology (more specifically, the question about the ethical effects that IT tools, specially AI-based and online systems, could have in their users, in particular, and in our society, in general) was built around the sudden and urgent need of dealing with the ethically-detrimental effects of systems that were initially thought to be "safe" (Caliskan et al., 2017, pp. 1). The urgency to "repair the damage", as well as to prevent further detrimental consequences, led to the question being posed in a *protective way* that aimed to safeguard the users against further potentially-adverse effects. In this sense, the question aimed at amending a damage that was inadvertently done by technological products in order to restore an ethically-neutral balance that was thought to exist, but which in fact did not. The question that has been mainly asked regarding the ethics of technology, therefore, is formulated in a way that looks to safeguard and protect the users, rather than in a way that aims to explore and exploit the potential opportunities for flourishing that technology could bring about.

---

[10]Although some offline technologies could also be relevant for this work, the focus will be placed on online technologies, as these are the kind of technologies where interactions among users is more common. Furthermore, and because some ethical theories like ethics of care and relation ethics, which strongly resonate with the motivations behind this work, are strongly based on interactions and relations between people, technologies that allow such interactions are a natural first step. Ideally, and even if it is carried out through an online medium, the development of one's practical wisdom, together with the enhancement of one's ethical autonomy, should be ultimately translatable to offline contexts: as one becomes more aware and sensitive towards ethically-relevant matters, one should be able to identify and react to them in an ethically-desirable way both online, and offline. In fact, and as it is pointed out in Floridi (2018), it can be considered that today we neither live offline, nor online, but *onlife*, as an almost inseparable mixture of the analogue and the digital, the offline and the online.

In order to better illustrate the differences between the current mainstream question about the ethics of technology and the alternative question (which will be introduced later in the present work), it is useful to draw an analogy between the notions of *human rights* and *laws*. Although both are aimed towards contributing to the quality of human life, the way in which they approach their goal comes from (and aims towards) different directions. In the Universal Declaration of Human Rights (as in United Nations (1948)), human rights are built upon the *potentiality* of human life and are articulated to express what human life could (and, ideally, should) become. The goal of human rights is to establish what human beings are entitled to have, in order to make a "fulfilling life" achievable and within reach; without granting those rights, without the potentiality that comes from them, the chance of flourishing and living such life gets seriously compromised, and living quickly becomes a matter of surviving, or simply existing.

This is where laws come into play. Human rights are *enabling*, in the sense that they are about what human beings should either have access to, or not be denied of. In a way, they can be seen as principles, or beacons towards which to strive for –similar to virtues in virtue ethics theories. However, an additional mechanism is needed in order to oversee that these rights are not hindered, arbitrarily restricted, or straight-on ignored. Among fulfilling other goals, laws are a mechanism created as part of our society in order to watch over the way human beings (as well as corporations, etc.) actually relate among themselves in order to guarantee, at least to a certain extent, that human rights are still respected. The way laws are articulated, however, is not about the entitlement, or the potentiality that human beings should have in virtue of certain principles, but rather about specific encoded norms that can be used to scrutinize particular cases and practices to determine whether they comply with certain requirements.

Under this understanding, human rights are about the potentiality, about setting the conditions that allow growth and flourishing, about the principles and ideals deemed to be pursued in order to achieve a fulfilling life; they set a direction towards which to walk to. Laws are about the actuality, about setting the mechanisms needed to ensure that the space needed for human rights to thrive is protected and not invaded by other interests, about the limits and boundaries that other activities have with respect to individual and collective life; they set up mechanisms to guarantee that the path towards rights is not blocked.

With this distinction in mind, we can draw a parallelism between two ways in which the question around the ethics of technology can be formulated: a *protective question*, based on a legalistic approach to safeguarding the user under the actual *status quo*, which we call the *ethical realist question* of ethics of technology; and the *expansive question*, based on a principle-oriented approach aimed towards setting and enabling the conditions for the user to flourish under a potentially new *status quo*, which we call the *ethical idealist question* of ethics of technology.

Aside from the analogy between human rights and laws, we can relate these two approaches to the question of technology ethics with two approaches from political philosophy, from which the names of both questions have been derived: political realism and political idealism Gilbert (1941); Roshwald (1971); Zuolo (2016). Whereas political realism is grounded on understanding the way society works in order to learn how to use these mechanisms (it focuses on understanding the *actuality* of human relations), political idealism focuses on discerning the potential that humans beings could achieve through a better (or ideal) organization of society (it focuses on envisaging the *potentiality* of human relations). In a certain way, laws can be related to the political realist perspective, as they are about the actual practices that do take place in human society, whereas human rights can be related to the idealist perspective, as they are about the potentiality of such society.

Ethical Realism: The Protective Question

The question of *ethical realism* is grounded on understanding the actuality of technological products, uses and trends, in order to safeguard its users from suffering ethically-undesirable[11] consequences. Similarly to one of the main core ideas in political realism, which focuses on understanding how society actually works in order to use this knowledge as a means to an end, ethical realism follows a similar logic: because technology works and is currently used in a certain way, an effort should be made in order to shield the users from potential detrimental effects stemming from such uses. In other words, the current *status quo* is taken for granted and the system (which, in this case, refers to the combination that emerges from technological tools, their users and other stakeholders, as well as from the interactions enabled by such tools) is taken as a model from which to identify the way actions, interactions and relations work[12]. In a way, the actual system is identified as the end-goal, and the mechanisms and rules derived from it are meant to keep the current system functioning –although with certain constrains to control undesirable effects.

Some of the terms that are usually heard under this perspective of technology ethics are notions like "privacy", "confidentiality", "fairness", "exclusion", or "transparency"[13], to name a few. Three main notions can be related to this question about technology ethics:

---

[11]The term "ethically-undesirable consequences" refers to those outcomes of certain decisions that have an effect that is detrimental for the user because it might be unfair, or biased, or be contrary to certain principle that would be beneficial for the user, among others.

[12]In this work, the term "system" is used to refer to the structure that arises from a particular way of organizing a matter, whereas the term "status quo" refers to the specific instantiation of elements that occupy a place in such structure.

[13]It is worth noting how "transparency" is often established as a requirement for "trust", in the sense that automated processes in AI-driven technologies should be explicitly monitorable by external users. However, and as pointed out by Byung-Chul Han in (Han, 2020, pp. 91), trust should not need transparency, as trust is built, precisely, under circumstances of asymmetrical information, when one party believes that the other one will do as promised without the need to supervise their work. Ironically, when one needs full transparency, trust is precisely what is missing.

- *Aristotle's "actuality"*: It refers to "the fact that is the case". Regardless of what could be, or could have been, the actuality concerns what *is*, at the current moment. In this sense, ethical realism is, just like political realism, about understanding what the actual relations between technology and their users, as well as between other stakeholders in the system, are.
- *Political realism and the law of prudence*: Political realism assumed, beforehand, that the person who held the power did not need to be *good*, in the sense of caring about other people's well-being. The system, therefore, needed to set up a series of mechanisms that constrained the power of the governing body in a way that they had to do some good, or, at the very least, that they could not do too much harm. Under this scope, political realism follows the "law of prudence": it should be assumed that everyone is ill-intended, and therefore the system should be built in such a way that limits what stakeholders in positions of power can do, in order to minimize detrimental consequences.
- *Law-like approach*: Following from the previous two considerations (i.e.: focusing on the actual state of affairs and setting up the system in a way that prevents things from going astray), the approach that ethical realism takes to the question about technology ethics follows a law-like approach: understand how the relation between technology and their users actually work, and create mechanisms to foresee and mitigate the detrimental consequences that technology could have on their users.

Considering this, the ethical realist approach to technology ethics can be characterized by the following terms:

- *Protective*: The "ethics" part within the question aims at shielding the users from detrimental consequences. The ethical layer is, under this view, defined in an *ad hoc* fashion that reacts to the uses and outcomes that the technology has over its users and society.
- *Realist*: The current social, political and economic state of affairs is taken for granted, as well as the current uses and trends in technology. It aims to answer the question "how things actually work" and to shape the ethical considerations around technology according to them.
- *Confronting*: Because the goals and interests behind companies and technological tools might not be aligned with individual and social interests, keeping the technology's potential within certain ethical boundaries supposes a confrontation between what *could* be done with that technology (in terms of the technology proprietor's interests) and what *should not* be done with it (in terms of ethical constrains). Ethical interests, therefore, can be often seen as conflicting with the (in most cases, economic) interests behind the technology.
- *Preserving*: The current state of affairs is not only taken from granted, but also assumed to continue being the way it is (within certain margins). For example, when considering how to address existing inequalities that might be caused by shortcomings in the current social and political organization,

the focus is seldom put on the *cause* of the problem (i.e.: the way society and political powers are organized), but instead it focuses on palliative approaches towards the *symptoms* shown by that problem. In other words, the roots of the current systemic organization are taken for granted, and ethical concerns are used not as a way to change those roots, but rather as a way of alleviating the existing negative consequences brought about by them.

- *Legalistic*: The ethical realist question to technology ethics follows a law-like approach focused on regulations constraining certain features and practices that technological companies must adhere to.

Ethical Idealism: The Expansive Question

Instead of focusing on understanding the current technological system in order to safeguard their users, the perspective of *ethical idealism* would instead look at how technology, as well as the way it can be used by their users, can create and modify affordances and interactions that enable, enhance and promote ethically-beneficial outcomes for their users –including chances for their users' empowerment and flourishing. Drawing a parallelism with political idealism, this approach to the ethics of technology sees the interrelations between technology, their users, and the way they interact with each other as a system that either enables, disables, promotes, or obscures interactions that set up the conditions that could enable ethically-flourishing practices in their users –both at an individual and a collective level.

In this sense, and conversely to what has been said about the ethical realist question, one should inquire not about how technology *currently* works, but rather about how it *could* work, and understand how potential changes could affect the uses, interactions and behaviors of its users. This follows one of the core ideas behind political idealism, which argues that the political and social system should be shaped in a way that creates the appropriate possibility space to enable the kind of actions and behaviors that can lead towards the ideal conception of what its citizens could *become*, rather than what they *actually are*. In other words, the system should allow, and ideally promote, those kinds of interactions that point towards that ideal state of affairs. In this case, the system acts not as an end-goal, or a model from which the set of existing interactions is derived, but rather as an *enabler* through which available interactions make it possible to get closer to a desired ideal conception. When, instead of the political and social organization, one thinks about the system formed by technology and its users, the political idealism approach amounts at shaping the system's available uses, affordances and interactions in ways that enable and point towards a potential desired state –thus providing the grounds for their users to become exposed to uses, behaviors and habits supporting the principles behind such ideal state.

Some of the terms that could be used under this perspective of technology ethics are notions like "opportunity", "empowerment", "flourishing", or "autonomy". In this case, the following three key notions can be identified:

- *Aristotle's "Potentiality"*: It refers to those facts that, given the current state of affairs, could *become* the case. Potentiality does not refer to counterfactual events (i.e.: what *could* have been the case, had things gone otherwise), but rather to future scenarios that, given the actual one, could indeed materialize. In this sense, potentiality is about visualizing states of affairs that could be within reach, and create paths leading to them.
- *Political idealism and systemic conditions*: According to political idealism, the structure of the social and political organization should be shaped in a way that promotes behaviors directed towards an idealized version of what people (and the system itself) could become. It is important to note how this ideal state might never be actually achieved: it is, just as in the case of the principles guiding virtue ethics theories, a state to strive for and walk towards. The important thing is that the systemic organization (be it of social and political structures, or of technological tools and users) is shaped in such a way that enables and promotes behaviors and interactions pointing towards the desired state of affairs.
- *Rights-like approach*: Following from the previous two considerations (i.e.: aiming towards what is possible and desirable, and setting up a system that enables practices leading towards it), the approach that ethical idealism takes to the question of technology ethics is similar to human rights: it identifies the desired principles to strive for, while envisaging a potential and ideal state of affairs that can be approached via practices and interactions supporting those principles.

Considering this, the ethical idealist approach to technology ethics can be characterized with the following terms:

- *Expansive*: Technology, their uses, and the affordances and interactions it creates should allow its users to engage in actions, behaviors and practices that lead towards their users' growth and potentiality. The ethical layer, under this view, is defined by shaping technological interactions in ways that enable and promote practices that enable their users to flourish and become more empowered.
- *Idealist*: The current social, political and economical state of affairs, as well as current uses and trends in technology, could be reshaped in ways that set up better systemic conditions to enable flourishing practices. It aims to answer the question "how things could be better" and to shape the ethical considerations around technology in ways that aim towards reaching that configuration.
- *Enabling*: Technology should be designed in a way that enables and promotes behaviors, interactions and practices that foster flourishing and empowerment in their users. In this sense, the affordances created by technological tools need to be carefully understood, explored and shaped in order to ensure not only that they do not restrict, or obscure flourishing practices, but rather that they enhance and favor them.

- *Mutable*: The idealist perspective should start by conceiving a better, ideal state of affairs that, although not being the actual case, could be within reach. In this sense, shaping technological uses to match the way things actually work would limit beforehand improvements based on how things could work differently. The ethical idealist perspective on technology ethics should keep, as part of its core, the idea that the current state of affairs should not be taken for granted and that the current social, political and economical configuration, as well as technology and the way it is used, could be reshaped in ways that enable to take steps towards a different ideal state of affairs[14].
- *Principle-based*: The way technological tools work, as well as the way they are used, should be rooted in sets of principles that can be "practiced" via the affordances and interactions enabled by those uses. In other words, the way technology is used should be seen as an opportunity to enable and foster behaviors, interactions and practices aligned with certain desired principles which lead to their practitioners' (i.e.: the users) flourishing and empowerment.

3 Technological Interactions as a System

It has been argued how the ethical idealist question of technology ethics aims at rethinking and reshaping the uses of technology in ways that allow to harness the ethically-desirable potentiality behind practices supporting certain principles. Similarly, the ethical idealist approach needs to be deployed upon a system built in a way that allows the necessary kinds of interactions to happen, in order for the potentialities behind an ideal state of affairs to be reachable and practicable. Whereas the notion of "system", within political idealism, refers to the configuration of political and social relations among the institutions and members of a society, its counterpart notion, in the case of technology, refers to the combination of technological tools and users, as well as the set of interactions generated by the affordances enabled by such tools.

Understood in such a way, certain actions, behaviors and interactions generated by the system are either enabled (and even fostered), whereas others are completely disabled (or obscured). As such, this system becomes the "possibility space" where their users' practice takes place. While some available actions, within this technological space, might highlight and foster behaviors supporting certain principles, others might not only completely disregard them, but can even support opposite ones. Therefore, technological tools become, through the affordances they create, *enablers* of certain practices, as well as *disablers* of others.

In the fields of behavioral economics, political theory and behavioral sciences there exists a technique that can be used in order to reinforce positive acts among, for instance, consumers. This technique, known as *nudging* Thaler

---

[14]There are already some works, like Mohamed, Png, and Isaac (2020), that combine theories about possible reconfigurations of power distributions in our society, such as decolonial theory, and the effects that such theories could have on technology and AI.

and Sunstein (2008); Weinmann, Schneider, and Vom Brocke (2016), aims to steer consumers towards buying products that are good for them –for instance, by placing healthy food on the most accessible and visible places in a supermarket, while leaving other less healthy options on the top, or the lowest shelves[15]. The idea is that, whereas consumers can still choose to buy the unhealthy product (they are free to do so, if they wish to), they are gently steered towards buying something that should be better for them (thus potentially even promoting healthier consumption habits in the long term).

In the same way, it should be easy to see how technology, through the interactions it creates as a system in where the users take part, can not only enable, but also potentially promote (i.e.: nudge) actions and habits that foster, into those users, behaviors that support certain ethically-desirable principles. By engaging in interactions that expose them to those principles, the users could get the chance to develop an awareness, as well as to progressively transfer behaviors inspired by such principles into other contexts. In other words, these interactions could create the core of the actions and behaviors that conform the users' everyday practice, which could in turn allow them to develop, train and regain, almost inadvertently, their own ethical autonomy (that is: their ethical awareness, sensitivity, and reasoning capabilities) and practical wisdom.

3.1 Examples of Actual and Potential Applications

In order to illustrate this idea, let us briefly go over a few examples of both actual, and potential ways in which a touch of the ethical idealist perspective could be integrated into different IT tools. These examples can be helpful to see either how available interactions and affordances can lead to particular behaviors and emotions, how certain general ethical principles can be integrated to prompt awareness and reflection into the user, how healthy habits can also be nudged through these IT tools, and how existing IT systems could be easily refurbished in order to integrate the aforementioned perspective.

3.1.1 Reactions in Social Networks

Consider social networks as one of the most straightforward cases of direct (by either commenting on, or reacting to, other users' posts) and indirect (simply by seeing other users' posts) interactions between users. For instance, the range of *reactions* the platform offers provides an easy way to interact with another user's post by offering a limited range of options to respond, but without prompting for a specifically-made answer –unlike a text comment.

The decision behind the range of available reactions, or even behind the way reactions are shown, is, of course, not a technical decision, but rather a design decision aimed at allowing users to react in some ways, while not in others. While the *Facebook* platform, for example, included a reaction allowing users to "like" another user's post, Facebook's CEO Mark Zuckerberg stated,

---

[15]Needless to say, this same approach can also be used as marketing and consumerism strategies without any intention of nudging healthy habits Wilkinson (2013).

in relation to some user group's demands for the inclusion of a "dislike" button, that "we need to figure out the right way to do it so it ends up being a force for good, not a force for bad and demeaning the posts that people are putting out there" Johnston (2014). Instead, and as a result of users demanding further options to quickly interact with posts, Facebook rolled out "reactions" standing for different emotions, rather than a reaction to dislike a post. Conversely, other platforms, such as *YouTube*, which had both a like and a dislike button that publicly showed both counters, decided to remove the public dislike count and only show it to the poster of the content Suciu (2021) –as a way for the user to know what type of content was most disliked by their viewers, but while avoiding the content to be publicly seen as disliked.

As it has been seen, the decision behind not including such "negative reaction" buttons, or at least not making their numbers public, follows the aim, as expressed by Facebook's CEO in the previous quotation, to avoid creating mechanisms allowing affordances that could lead to behaviors and interactions based on negativity and confrontation, which could risk dragging part of the interactions had within the social network platform to that set of negative emotions. These decisions, therefore, can be a example to show how certain affordances enabled by an IT platform can lead to behaviors that can be either based on positive interactions, or on negative ones, and how affordances that enable and promote negative relations are already identified as such and are, sometimes, already being intentionally omitted by the companies behind those platforms.

3.1.2 Overarching Ethical Principles

Some cases of digital games can encourage, through their game mechanics, behaviors and reflections connected to ethical principles that present their players with the opportunity to engage in actions that support such principles. As most games are usually related to achievement and to getting closer to the end of the game, a player's "saved game file", which stores the current progress within the game, can be seen as one of the most precious resources for the player within the game. Considering this, would a player be willing to sacrifice it in order to help another anonymous player within the game?

Square's Enix *NieR: Automata* Square-Enix (n.d.) explores this question in one of the possible endings of the game where the player is faced with an extremely powerful, incredibly unfair and almost undefeatable final boss. Eventually during that fight, the player starts to see messages from anonymous players who have previously managed to defeat the boss and encourage the player to keep fighting. Ultimately, one of those anonymous players will appear to grant a power-up (i.e.: something that makes the player's character more powerful within the game), thus making the player's victory possible. However, once the player has finally managed to defeat the final boss, the player will be told that those anonymous players who helped them had to make the sacrifice of completely deleting their saved game files (thus returning their own game at its original state, with no progress saved whatsoever) in order to provide the

same help they have received. At that point, the player will be presented with the same choice: are you willing to sacrifice all your game progress to help an anonymous player, just as you have just been helped now?

Although this example is not related to any particular affordance created by the technology's available interactions, it shows how some high level ethical principles, such as sacrifice, community and altruism, can be integrated in decisions included within the technology –in this case, included within the decisions available as part of the game's fiction, but which can still reach beyond the fiction itself and into the human player behind.

3.1.3 Nudging Healthy Habits

Although not directly related to fostering ethical principles, other IT tools can incorporate interactions aimed at contributing to the building of healthy habits. Some applications made for mobile devices, or even smart watches, would track the physical activity of the bearer and may suggest, at certain times, that the bearer engages in some sort of physical activity, such as going for a walk. Nevertheless, in those cases the nudging of doing something physical is actually part of the goal of the tool (it is designed and intended to be used as an assistant to track and nudge its user towards healthy habits involving physical activity). There are, however, other cases of IT products that, even if could be seen as unrelated, or even contrary to the goal and interests of the product, still choose to incorporate that kind of nudges.

For example, digital games like *Earthbound* Wikipedia (n.d.-a) and *Wii Sports* Wikipedia (n.d.-b) incorporated a system which, after a certain amount of time while the game had been running, suggested the players to take a break from the screen, catch some air, and come back later. Although this may often be seen as unrelated, or even contrary to the game's developers interest, which might want to make the game as immersive as possible and keep the players hooked at it for as much as possible, it instead chooses to encourage healthy habits, that go beyond the game, to their players.

Note that this is an example of an offline system that, although involves no relation with other users, can still incorporate a mechanism that has no relation with the goal of the game itself (i.e.: being enjoyable for the player), but which nonetheless can be incorporated within the game to benefit the player behind the screen by nudging habits that would likely be beneficial for them.

3.1.4 Cooperation and Community in Online Learning

Consider the field of education supported by digital platforms, like online learning courses. While the platforms where these courses are carried out always include many necessary functionalities (access to learning materials, communication, exercises, etc.), the *way* in which they are provided can significantly change how students interact with the learning environment and among themselves.

Take, for example, self-grading exercises within the virtual campus. These can be done individually, in the sense that student A performs an activity, gets the results, and that is everything there is to it. Now, how could a touch of the ethical idealist perspective be added into this already existing functionality? Say that, because in online learning environments students might feel somewhat isolated from the rest of the students McInnerney and Roberts (2004), we want to foster a sense of community and cooperation among them. How could the design of that feature's functionality be altered in order to shape interactions that nudge those principles to the students? Imagine that, for a hypothetical question X for which student A has gotten an incorrect answer, student B, whose answer on question X was right, gets the chance to share their own answer (once the activity has ended) and provide a brief explanation on the reasoning behind their answer, possibly while giving A the chance to ask back. Even though this does not, of course, replace what the module teacher might do to help that student, it can lead towards interactions that foster the creation of a sense of community and a cooperative mindset as a result of bringing students together in order to share (and help each other with) some parts of their learning process –which are the principles that were aimed to nudge through this modified functionality.

3.1.5 Distilling the Examples: Looking Beyond the Goal

Even though this are just some simple examples that are only commented briefly, they serve the purpose to illustrate how, given certain design considerations, the *way* in which technology is designed (regardless of the *goal* it is designed for, in the first place) can create a specific set of affordances and enable a set of interactions that can point towards certain behaviors, relations grounded on one, or another set of emotions, and even ethical principles that are enabled and nudged via the way the technology is used.

Nevertheless, the question that needs to be asked in order to identify and harness these collateral opportunities around the use of technology cannot be asked from an ethical realist perspective. For instance, the initial setting of the online learning case, which did not yet bring the interaction with other students into the picture, already did what it *had* to do (as a component of an online learning environment); furthermore, it did nothing that was ethically undesirable for the students themselves, while serving a purpose that is, in itself, ethically-desirable (namely: helping them with their learning). It is only when we switch to the ethical idealist perspective and start wondering how the system could be even better (in terms of potential ethical benefits), that these additional opportunities to use technology as a tool to foster ethically-desirable behaviors, to enhance flourishing interactions, emerge. A new and complementary question appears: how could this mechanism be enriched in such a way that it supports other principles that, although maybe not directly related to the goal that this technology aimed to address in the first place, are still ethically-desirable?

Still in the online learning example, those principles are related to community building and cooperative behaviors in online learning processes; all that is needed is to wonder not how self-graded exercises *currently* work in many learning environments, but rather how they *could potentially* work in order to allow (and nudge) interactions pointing towards those desired principles. Note that the hypothetical exercise itself has not changed: it is still the same exercise, asks the same content, and is carried out by the same student within the same online learning environment. Nevertheless, additional collateral features have been added into it in order to modify and alter affordances and interactions to enable ethically-desirable behaviors that were either not possible, or strongly obscured beforehand due to a design that focused solely on the technology's main purpose, and only looked at the ethical realist (protective) question to ensure that it did not have unintended detrimental consequences –but while disregarding the alternative (and complementary) ethical idealist question and the potential benefits behind it.

As it has been explained while presenting the ethical idealist approach, and as these brief examples has shown, the potential behind the ethical idealist perspective lies, precisely, in its radically idealistic roots: it is not about the goal of a particular technology, or about what it needs in order to be usable, but rather about going beyond those questions and asking oneself: how could this technology be used to incorporate an added value that exhibited, enabled and maybe even nudged their users towards some desired ethical principles and behaviors?

4 Towards Ethical Idealism in Technology

The present work introduces and merges different topics that could appear, at first glance, independent. It starts by considering some ethical theories and recognizing the need to regain ethical autonomy; it then shifts to a series of analogies related to human rights, laws and theories in political philosophy in order to understand how (and from where) the question about technology ethics is (or could be) asked; then, it argues how technological tools, together with their users and the interactions that exist between both, configure a system where a major part of the users' daily practice (understood as a combination of choices, actions, behaviors and relations) take place, and where some key ideas that characterize the previous political theories can be applied to. In this sense, technological interactions are understood as the space where their users carry out an important part of their practice.

At this point, it is argued how, by combining the notions of ethical autonomy and practical wisdom, by asking the ethical idealist question of technology ethics, and by understanding the system as the interrelation between technology, users and uses, the affordances created by technological tools can be shaped in ways that enable and promote actions, behaviors and habits supporting ethically-desirable principles. By becoming exposed to such principles and practices, their users get the chance to progressively develop the awareness,

ethical sensitivity and ethical reasoning skills required to identify, understand and potentially replicate similar behaviors supporting the same principles in other different scenarios; in other words, by shaping technological interactions in ways that nudge certain ethically-desirable principles and behaviors, the technological space becomes a place for the users to develop and train their ethical autonomy and their practical wisdom.

This way of proactively exploring collateral ethical opportunities behind the generalized use of technology, which has been introduced in this work as the *ethical idealist question*, requires an additional layer of consideration throughout the design, development, deployment and use of technological tools. However, and because the focus of the ethical idealist approach is on identifying and integrating subtle opportunities and nuances for ethically-desirable practices within the everyday uses of certain technological tools, their effects might seem unnoticeable at early stages of deployment in some cases. Nevertheless, regaining ethical autonomy through practice requires the progressive building of certain habits, behaviors and sensitivities that must be cultivated over time, and thus which may be challenging to clearly identify via evaluation techniques that focus on short-term results. This is not, however, something that should be seen as a shortcoming of this approach, or something that impedes it being put into practice, but rather as one of its defining features, and in accordance to how theories of virtue ethics, specially with regards to the life-long cultivation of practical wisdom, work.

It is important to remark that the two approaches distinguished in this work are not meant to be exclusive in any way. Their focus is on different facets of the use of technological tools, and therefore their contributions are aimed towards different goals. Whereas the ethical realist question aims to foresee and deploy mechanisms in order to ensure that technology will not be detrimental to their users, the ethical idealist question should look beyond the goal of that particular technology and ask what sort of actions, behaviors and interactions could emerge, from the use of that technology, that could be used to promote some desired ethical principle. Considering this, the question about the ethics of technology must categorically avoid the pursuit of a "theory of everything", with respect to whether the realist, or the idealist approach, is the right one. Both are complementary and ethically-beneficial in different ways within the design, deployment and use of technological tools. Considering only the ethical idealist approach would be naive and would likely lead to many unfair and detrimental uses of new technological tools, but considering just the ethical realist approach ignores opportunities to use those technological advancements in order to provide more, better and deeper flourishing opportunities for both individuals, as well as for communities and potentially the whole society in the long run. Under this new ethical idealist approach, technology should be seen not just as a self-contained tool aimed to achieve a specific goal, but rather as an *enabler* of a set of actions, behaviors and interactions that provide their users with much more than a mere tool to solve a problem, or to carry out an activity in the digital space: it provides their users with the space where,

through their everyday practice, they are given the chance to pursuit, regain and enhance their ethical autonomy and their practical wisdom.

This approach, nevertheless, has neither been introduced as a clear methodology, nor as a succession of steps to be followed, but rather as an active pull towards a change of paradigm in the way of thinking about the full range of potential ethically-relevant opportunities that technology can bring. Its main (apparent) weakness? That this approach is neither about any specific technology, nor concerns any specific field, nor it is related to any technology's end goal. The strength behind this apparent weakness? That, precisely because of it, this approach can potentially be applied to practically any IT tool, in practically any field of application, and that it can be applied to collaterally complement any technology's end goal. The ethical idealist question on the ethics of technology, which focuses on identifying the potential benefits of developing their users' ethical autonomy through the way they use technology, rather than because of why they use it, can be asked and answered as part of the design cycle of almost any technological product. It does require, from the stakeholders involved in such cycle, to take an exploratory and creative mindset in order to foresee and understand how this technology *could* be used (aside from how it is initially meant to be used) in order to shape those uses in ways that enable, foster and nudge actions, behaviors and interactions supporting ethically-desirable principles that could be practiced in the context where that technology will be deployed.

The ethical idealist approach is not meant to replace, but rather to complement. It is not about a method, but rather about a mindset. It is not aimed at how technology works, but rather at how it is used. It is about recognizing technology as one of the main spaces of interaction in our current society, about understanding how the users use IT tools and devices to carry out their interactions, about empowering the users through these uses, and about shaping the system of technological interactions in ways that enable the kind of practice required to cultivate our ethical sensitivity, awareness and reasoning. It is about technological interactions understood as a means to practice ethics, and about ethics understood as a practical discipline that opens up the path to pursue our own individual and collective flourishing, empowerment and regaining of ethical autonomy.

References

Acton, H.B., & Watkins, J.W. (1963). Symposium: negative utilitarianism. *Proceedings of the Aristotelian Society, Supplementary Volumes*, *37*, 83–114.

Angwin, J., Larson, J., Mattu, S., Kirchner, L. (2016). Machine bias. *ProPublica, May*, *23*, 2016.

Anscombe, G.E.M. (1958). Modern moral philosophy. *Philosophy*, *33*(124), 1–19.

Association for Computing Machinery (ACM) (2018). *Acm code of ethics and professional conduct.* https://www.acm.org/code-of-ethics/. (Accessed: 2022-07-27)

Association of Nordic Engineers (2021, January). *Addressing ethical dilemmas in ai: Listening to engineers.* https://nordicengineers.org/wp-content/uploads/2021/01/addressing-ethical-dilemmas-in-ai-listening-to-the-engineers.pdf. (Accessed: 2021-10-26)

Blaschke, L.M. (2012). Heutagogy and lifelong learning: A review of heutagogical practice and self-determined learning. *The International Review of Research in Open and Distributed Learning*, *13*(1), 56–71.

Blaschke, L.M. (2018). Self-determined learning (heutagogy) and digital media creating integrated educational environments for developing lifelong learning skills. *The digital turn in higher education* (pp. 129–140). Wiesbaden: Springer.

Caliskan, A., Bryson, J.J., Narayanan, A. (2017). Semantics derived automatically from language corpora contain human-like biases. *Science*, *356*(6334), 183–186.

Carr, A., Balasubramanian, K., Atieno, R., Onyango, J. (2018). Lifelong learning to empowerment: beyond formal education. *Distance Education*, *39*(1), 69–86.

Chirkov, V.I., Ryan, R., Sheldon, K.M. (2010). *Human autonomy in cross-cultural context: Perspectives on the psychology of agency, freedom, and well-being.* Dordrecht: Springer.

Dancy, J. (1993). An ethic of prima facie duties. P. Singer (Ed.), *A companion to ethics* (p. 219-229). Hoboken, NJ: John Wiley & Sons.

Davis, N. (1993). Contemporary deontology. P. Singer (Ed.), *A companion to ethics* (p. 205-218). Hoboken, NJ: John Wiley & Sons.

Deci, E.L., & Ryan, R.M. (1995). Human autonomy. M.H. Kernis (Ed.), *Efficacy, agency, and self-esteem* (pp. 31–49). New York: Springer.

EU Parliament (2021, April). *Proposal for a regulation of the european parliament and of the council laying down harmonised rules on ai and amending certain union legislative acts.* https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52021PC0206&from=EN. (Accessed: 2021-11-03)

Favaretto, M., De Clercq, E., Elger, B.S. (2019). Big data and discrimination: perils, promises and solutions. a systematic review. *Journal of Big Data*, *6*(1), 1–27.

Floridi, L. (2018). Soft ethics and the governance of the digital. *Philosophy & Technology*, *31*(1), 1–8.

Frischmann, B., & Selinger, E. (2018). *Re-engineering humanity.* Cambridge University Press.

Garcés, M. (2021). *Nova il·lustració radical.* Barcelona: Anagrama.

Gibson, J.J. (1979). *The ecological approach to visual perception.* New York, NY: Houghton Mifflin Harcourt (HMH).

Gilbert, F. (1941). Political thought of the renaissance and reformation. *The Huntington Library Quarterly*, 443–468.

Han, B.-C. (2017). *The agony of eros.* Cambridge, MA: MIT Press.

Han, B.-C. (2018). *The expulsion of the other: Society, perception and communication today.* Hoboken, NJ: John Wiley & Sons.

Han, B.-C. (2020). *The transparency society.* Stanford, CA: Stanford University Press.

Held, V., et al. (2006). *The ethics of care: Personal, political, and global.* Oxford University Press on Demand.

HLEG on AI (2019, April). *High-level expert group on artificial intelligence: Ethics guidelines for trustworthy ai.* https://www.aepd.es/sites/default/files/2019-12/ai-ethics-guidelines.pdf. (Accessed: 2021-11-03)

IEEE (2016, December). *Ethically aligned design: Prioritizing human wellbeing with autonomous and intelligent systems.* https://standards.ieee.org/content/dam/ieee-standards/standards/web/documents/other/ead1e.pdf. (Accessed: 2021-11-15)

Interaction Design Foundation (n.d.). *Interaction design foundation: Affordances.* https://www.interaction-design.org/literature/topics/affordances. (Accessed: 2021-11-29)

Johnston, C. (2014). *No dislike button for facebook, declares zuckerberg (the guardian).* https://www.theguardian.com/technology/2014/dec/12/no-dislike-button-for-facebook-declares-zuckerberg. (Accessed: 2022-07-27)

Krueger, J., & Osler, L. (2019). Engineering affect. *Philosophical Topics*, *47*(2), 205–232.

MacIntyre, A. (2013). *After virtue* (3rd ed.). London: A&C Black. ((First edition published in 1981))

McInnerney, J.M., & Roberts, T.S. (2004). Online learning: Social interaction and the creation of a sense of community. *Journal of Educational Technology & Society*, *7*(3), 73–81.

Metz, T., & Miller, S.C. (2016). Relational ethics. *The international encyclopedia of ethics*, 1–10.

Mill, J.S. (1987). *Utilitarianism and other essays.* New York, NY: New York: Penguin Classics.

Mohamed, S., Png, M.-T., Isaac, W. (2020). Decolonial ai: Decolonial theory as sociotechnical foresight in artificial intelligence. *Philosophy & Technology*, *33*(4), 659–684.

Pojman, L.P., & Fieser, J. (2016). *Cengage advantage ethics: Discovering right and wrong.* Wadsworth: Cengage Learning.

Roshwald, M. (1971). Realism and idealism in politics. *Social Science*, 100–107.

Russell, S., Dewey, D., Tegmark, M. (2015). Research priorities for robust and beneficial artificial intelligence. *Ai Magazine*, *36*(4), 105–114.

Smart, P., Heersmink, R., Clowes, R.W. (2017). The cognitive ecology of the internet. *Cognition beyond the brain* (pp. 251–282). Springer.

Square-Enix (n.d.). *Nier: Automata.* https://nierautomata.square-enix-games
.com. (Accessed: 2022-07-27)

Suciu, P. (2021). *Youtube removed 'dislikes' button – it could impact 'how to'
and 'crafts' videos (forbes).* https://www.forbes.com/sites/petersuciu/
2021/11/24/youtube-removed-dislikes-button--it-could-impact-how-to
-and-crafts-videos/?sh=766c3f4c5a53. (Accessed: 2022-07-27)

Thaler, R.H., & Sunstein, C.R. (2008). *Nudge: Improving decisions about
health, wealth, and happiness.* London: Yale University Press.

United Nations (1948). *Universal declaration of human rights.* https://www
.un.org/en/about-us/universal-declaration-of-human-rights. (Accessed:
2021-10-15)

Vallor, S. (2016). *Technology and the virtues: A philosophical guide to a future
worth wanting.* New York, NY: Oxford University Press.

Von Schomberg, L., & Blok, V. (2021). Technology in the age of innovation:
Responsible innovation as a new subdomain within the philosophy of
technology. *Philosophy & Technology*, *34* (2), 309–323.

Weinmann, M., Schneider, C., Vom Brocke, J. (2016). Digital nudging.
Business & Information Systems Engineering, *58* (6), 433–436.

Wikipedia (n.d.-a). *Earthbound.* https://en.wikipedia.org/wiki/EarthBound.
(Accessed: 2022-07-27)

Wikipedia (n.d.-b). *Wii sports.* https://en.wikipedia.org/wiki/Wii_Sports.
(Accessed: 2022-07-27)

Wilkinson, T.M. (2013). Nudging and manipulation. *Political Studies*, *61* (2),
341–355.

Yapo, A., & Weiss, J. (2018). Ethical implications of bias in machine learn-
ing. *Proceedings of the 51st hawaii international conference on system
sciences.*

Zuolo, F. (2016). Realism and idealism. *A companion to political philosophy.
methods, tools, topics* (pp. 75–85). London: Routledge.