

Constructing a mutually supportive interface between ethics and regulation

Delacroix, Sylvie; Wagner, Ben

DOI:

[10.1016/j.clsr.2020.105520](https://doi.org/10.1016/j.clsr.2020.105520)

License:

Creative Commons: Attribution (CC BY)

Document Version

Publisher's PDF, also known as Version of record

Citation for published version (Harvard):

Delacroix, S & Wagner, B 2021, 'Constructing a mutually supportive interface between ethics and regulation', *Computer Law and Security Review*, vol. 40, 105520. <https://doi.org/10.1016/j.clsr.2020.105520>

[Link to publication on Research at Birmingham portal](#)

General rights

Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

- Users may freely distribute the URL that is used to identify this publication.
- Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
- User may use extracts from the document in line with the concept of 'fair dealing' under the Copyright, Designs and Patents Act 1988 (?)
- Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

Take down policy

While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact UBIRA@lists.bham.ac.uk providing details and we will remove access to the work immediately and investigate.



Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/CLSR

Computer Law
&
Security Review



Constructing a mutually supportive interface between ethics and regulation

Sylvie Delacroix^{a,1}, Ben Wagner^{b,1,*}

^a Law and Ethics at Birmingham Law School, a Fellow of the Alan Turing Institute and a Mozilla Fellow

^b TU-Delft at the Faculty of Technology, Policy and Management

ARTICLE INFO

Keywords:

Technology ethics
Regulation and governance
Healthcare ethics

ABSTRACT

When the word 'ethical' becomes synonymous with specious, you know that something is amiss. With each data governance scandal, with each creation of a corporate 'ethics board', 'ethical standards' seemingly lose a few more feathers, to the point of generating instant suspicion when invoked in any official report. We argue that a key challenge in this regard is to more precisely define the ethics-regulation interface. In order to do this, we first provide an overview of recent endeavours to develop ethical frameworks around technology. We then look at a successful process of refinement of the ethics-regulation interface: the case of healthcare ethics in the UK. The third section develops an account of what a more robust ethics-regulation interface could look like, which would support a process of cross-fertilisation between the political, ethical and legal approaches. Finally, the fourth and last section critically examines a 'live' implementation of such ethics-regulation interface, as put forward in Quebec's 'Bill 29'.

© 2021 Sylvie Delacroix and Ben Wagner. Published by Elsevier Ltd.

This is an open access article under the CC BY license

(<http://creativecommons.org/licenses/by/4.0/>)

1. Introduction

When the word 'ethical' becomes synonymous with specious, you know that something is amiss. With each data governance scandal, with each creation of a corporate 'ethics board', 'ethical standards' seemingly lose a few more feathers, to the point of generating instant suspicion when invoked in any official report. How did 'ethics' come to acquire such a bad name in the domain of AI and data governance? In large part, due to an insidious confusion between ethics and politics. This confusion is far from new. Almost forty years ago, Mouffe de-

nounced the tendency to 'turn to ethics' to gloss over the power conflicts characteristic of politics.²

The contemporary 'ethics boards fashion' is no different: as the scope and depth of the regulatory power yielded by pervasive data collection technologies becomes apparent, the cor-

² Since Mouffe's critique, there has been an ongoing shift towards corporate 'self-regulation' as a strategy to avoid regulation. While ethical approaches such as data ethics have attempted to go beyond such 'strategic' self-regulatory initiatives, they are in danger of being misused as mere tools to avoid regulation unless their relationship to existing legal and institutional frameworks is clarified. Such strategic, self-regulation approaches also need to be distinguished from existing compliance frameworks in professional contexts, such as academic research ethics boards or corporate compliance with an ethical code of conduct for employees. We hope that this article can contribute to clarifying the relationship between different kinds of ethical approaches and existing institutional legal frameworks.

* Corresponding author.

E-mail address: ben@benwagner.org (B. Wagner).

¹ Both authors have contributed equally to this work

porations deploying such technologies seek to avoid regulation. While regulators are still grappling for answers to problems whose full scope and nature are as challenging to delineate as they are to address, powerful corporations hurry to occupy as much of what currently remains an investigatory space as possible. The true *raison-d'être* underlying those shiny ethics boards and guidelines is political, not ethical: if they can prevent (or at least shape or reduce the scope of) legal regulation, they will have done their job.

The above state of affairs is unfortunate on more than one level. If these so-called "ethics-washing" (Metzinger, 2019) strategies do bear any fruit, they will have compromised some much-needed regulatory process. If, on the other hand, these strategies are thwarted, but we fail to question their underlying, antagonist understanding of ethics -whereby ethics and regulation vie for control in a zero-sum game- we will have lost something too.

Ethics is first and foremost a question: 'How should I / we live?'.³ There are no ready-made answers, no magic formula allowing us to get our 'moral sums' right.⁴ As the I / We at the root of the ethical question evolves, that question must be continuously asked anew. A wide range of de facto constraints can narrow down the range of available answers to that ethical question: one of the reasons why we often need legal regulation is precisely to rein in those de facto restrictions, just as the constraints imposed by legal regulation itself must be questioned too. To that extent, ethics is and should remain legal regulation's 'critical best friend'. In contrast to those models which describe the relationship between ethics and regulation in terms of at least partially competing spheres of influence, this paper emphasises the possibility of constructing a functional, mutually supportive relationship between ethics and regulation.

To understand the types of institutional frameworks that would foster such a mutually supportive relationship, it is helpful to look at the debate on the usage of data-intensive technologies such as AI through the prism of the evolving relationship between professional ethics and regulation within healthcare, since the latter has well-established tradition of cross-fertilisation between ethics and regulation. Considering professional ethics and how the latter 'interfaces' with legal regulatory frameworks can help to problematize some overly simplistic assumptions - including the idea that ethics can merely be a matter of getting one's 'moral sums' right.

In order to do this, the first section will provide an overview of recent endeavours to develop ethical frameworks around technology. We will also highlight some of the challenges with these approaches and how they may knowingly or unknowingly end up undermining the deployment of a legal framework that is capable of upholding fundamental human rights. The second section looks at a successful process of refinement of the ethics-regulation interface: the case of healthcare ethics in the UK. The third section develops an account

of what a more robust ethics-regulation interface could look like, which would support a process of cross-fertilisation between the political, ethical and legal approaches. Finally, the fourth section critically examines a 'live' implementation of such ethics-regulation interface, as put forward in Quebec's 'Bill 29'. The latter was proposed by the Government of Quebec in an endeavour to modify the Quebec Professional Code to include computing as part of the domain of engineering, with associated licensing requirements. In conclusion, we argue it is high time for community-based, professional institutions in the field of data-intensive technologies to be further developed and relied on in a bid to complement traditional, top-down regulation (in a way that is not dissimilar to the model of healthcare ethics in the UK).

2. The turn to ethics in data-intensive technologies such as AI

References to 'ethics' are frequently used as an attempt to preempt the political processes leading to regulation. This is most frequently the case when the relevant values within these ethical discourses are presented as 'self-evident' or widely agreed upon, when in fact every 'consensus' should be understood as the temporary result of a provisional hegemony, as a stabilization of power: as such, it always entails some kind of exclusion. The problem is that an 'ethics' discourse can be misused to mask such exclusionary politics.⁵

Section 1.1. reviews existing manifestos and ethical initiatives in the field of 'AI' with a view to highlighting the lack of responsibility attribution frameworks (and corresponding institutions), while section 1.2. emphasises the extent to which ethical debates of various kinds increasingly distract from such regulatory lacunae. Section 1.3, for its part, highlights the worrying way in which the EU Commission *Draft Ethics Guidelines*' focus on the 'ethical purpose' of key fundamental rights leads to their being watered down.

2.1. The institutional gap: overview of existing 'AI' ethics manifestos

Perhaps the two most notable aspects of the current state of 'data ethics' is the extent to which it has spread so widely in a such a short time and the extent to which it has become heavily politicised. Within little more than a week of setting up an 'AI ethics' board, Google decided to dissolve it following extensive criticism from its staff and the general public regarding its composition and remit (Statt, 2019). Similar heavy criticism has been levelled at the European Commission's AI ethics guidelines, with some claiming they are a case of "ethical white-washing." (Metzinger, 2019) The development of ethical guidelines has become a hotly debated political topic, which can lead to both ethics washing and ethics bashing alike. (Bietti, 2020) Inventories and broad analyses of ethical approaches to artificial intelligence have acknowledged a strong shift towards 'ethical' governance ap-

³ This definition of ethics by reference to the Socratic question is most famously articulated in (Williams, 2011).

⁴ The 'moral sums' image is meant to be provocative: it targets reductive understandings of ethics, whereby there is no distinction between ethics as an effort and morality as a positive set of standards.

⁵ Mouffe argues that the failure of liberalism to grasp the agonistic character of political life means that political conflicts get translated into moral terms (Mouffe, 1999).

proaches as well as a general lack of regulatory frameworks (Clarke, 2019; Daly et al., 2020; Hagendorff, 2019; Raab, 2020; Wagner, 2018), with one author describing the world of AI ethics as the "realm of paper tigers." (Haas and Gießler, 2020)

Such a degree of controversy is hardly surprising. There have been numerous technology manifestos over the past years, which have raised serious issues concerning the integration of norms and values in technology. Examples include the Castlebridge Report of 2016 which explicitly references the "new paradigm of ethics washing" (O'Keefe and Brien, 2016), the 2017 ethical design manifesto by Aral Balkan which is explicitly founded on human rights and calls for technology that "protects your civil liberties, reduces inequality, and benefits democracy," (Balkan, 2017) and the 2017 Copenhagen letter which argues it is: "time to put humans before business. Time to replace the empty rhetoric of 'building a better world' with a commitment to real action." (Rand-Hendriksen, 2018) Even in such everyday technical tasks as web design, there are increasingly extensive debates about what constitutes ethical practices (Rand-Hendriksen, 2018).

All of these initiatives can be seen both as a call to action and as a cry for help by respective communities given the status quo they perceive. What all these documents have in common is a considerable concern with misuse of power, whether this be power over data, human behaviour or emotions. It is noteworthy that even Facebook (Wattles and O'Sullivan, 2019) and Microsoft (Smith, 2018) have been calling for a more significant role for regulators, as well as more regulation in areas as diverse as facial recognition, "harmful content, election integrity, privacy and data portability." (Wattles and O'Sullivan, 2019) While it is easy to dismiss such statements by large, powerful actors, it also suggests that there is a considerable regulatory and institutional gap. Some actors are trying to fill this gap with ethical guidelines, others with regulatory initiatives. Initiatives that explicitly seek to build a mutually reinforcing interface between the two -ethics and regulation- are still few and far between.⁶

The rapidly growing debate surrounding the development of ethical approaches to artificial intelligence is also noteworthy, as instantiated in The Montreal Declaration for a Responsible Development of Artificial Intelligence, which was led by the University of Montreal in 2018 and signed by the Conseil national de recherches Canada (CNRC) (Else, 2018). Another interesting example is the 2017 Asilomar conference, organised by the Future of Life Institute in California, which developed 23 AI principles (Marchant, 2019). Notably, the "State of California recently adopted legislation 'expressing support' for the Asilomar AI Principles," (Marchant, 2019) which could be seen as developing a form of soft law ethical AI principles. As with all soft law principles, it should be noted that "soft law measures are very imperfect governance tools because of their lack of enforceability and accountability" (Marchant, 2019). Soft law mechanisms may not be helpful in specific realms of international governance (Abbott and Snidal, 2000) and these forms of governance also come with considerable challenges in the area of implementation (Abbott and Snidal, 2000).

⁶ This is unpacked in sections 3 and positively addressed in section 4.

To provide just a few examples of these soft law implementation mechanisms, the Council of Europe MSI-AUT Committee of experts on Human Rights Dimensions of automated data processing and different forms of artificial intelligence has also developed a recommendation on human rights impacts of algorithmic systems which was adopted by the Council of Europe Committee of Ministers on 8 April 2020. Finally, the Charlevoix Common Vision for the Future of Artificial Intelligence was developed during the Canadian chairmanship of the G7 in 2018, which is notably separate from the G7 Innovation Ministers' Statement on Artificial Intelligence. What is noteworthy about all of these declarations and statements is that they do not develop institutional frameworks or clear assignment of responsibility or accountability for individual groups or sectors.

In this context, the Institute of Electrical and Electronics Engineers (IEEE) has developed an entire Ethical Initiative with numerous different strands and working groups, focussing on ethical system design, transparency, autonomy, data privacy and algorithmic bias to name just a few areas.⁷ The International Standards Organisation (ISO) will soon publish ethical design standards for technology. At the same time, both the Association for the Advancement of Artificial Intelligence (AAAI) and the Association for Computer Machinery (ACM) have organised a flurry of conferences on ethical artificial intelligence and are in the process of developing their internal standards further (Stahl and Wright, 2018). There were even suggestions to 'professionalise' Machine learning as an occupation.⁸ This flurry of standard-setting by professional organisations is indicative of a willingness to consider better ways of responding to the need to develop ethical frameworks that are suited to the unique challenges inherent in data-intensive technologies.

2.2. When ethical debates distract from regulatory lacunae

For a relevant (and potentially important) ethical phenomenon to serve as a distraction from an actual policy issue is sadly all too common in the current policy debate on ethical AI. Rather than asking why Uber is allowed to implement poorly developed self-driving car systems on public roads (Marshall and Davies, 2018), the public is instead encouraged to think about how they would respond to numerous different variants of the trolley problem (MIT Media Lab, 2018). Rather than discussing serious challenges when machine learning systems are implemented within the criminal justice system (Angwin and Grassegger, 2017), policymakers appear more concerned that Artificial General Intelligence will soon take over the world (Cellan-Jones, 2014), with both general and specific applications of artificial intelligence continuing to be seen as a central challenge for policy-making in upcoming decades (Simon, 2019; Torres, 2019). In all three examples listed above, the purported flexibility of ethics is praised for its ability to easily and quickly respond to chal-

⁷ <https://standards.ieee.org/develop/project/7000.html>

⁸ Neil Lawrence for instance points at the "widespread social effects" of machine-learning models to suggest that the latter ought to be "validated" according to standards set by machine learning as a professional community (Lawrence, 2018).

lenges,⁹ when in fact this ‘turn to ethics’ masks a very real unwillingness -or inability¹⁰- to both address and implement ways of tackling thorny policy issues.

While there have been numerous calls for transparency of AI and more broadly automated systems in recent years, there are currently no general legal provisions that put this into practice. However, some limited provisions for this exist in certain areas, for example in articles 13–15 of the GDPR and the non-operative recital 71. This disconnect is most evident when the typical travel from ethics to standards to regulation is portrayed. For instance, there is no clear link between either the ethics or standards mentioned by Winfield and Jirotka and actual regulatory practices (Winfield and Jirotka, 2018). As both authors note, “there is little evidence that those principles have yet translated into practice, i.e. effective and transparent ethical governance” (Winfield and Jirotka, 2018: 9). While both authors mention the necessary “strong institutional frameworks” (Winfield and Jirotka, 2018: 10), they stop short of actually specifying relevant regulatory bodies which should be involved. It is unclear what strong institutional frameworks mean in this context, as regulatory question marks replace meaningful external oversight (Veale, 2020: 202).

2.3. Watering down fundamental rights

The tension between current regulatory and ethical approaches becomes particularly apparent when the Draft Ethics Guidelines for Trustworthy AI (18 December 2018) of the European Commission’s High-Level Expert Group on Artificial Intelligence are studied in greater detail. Not only do the Draft Ethics Guidelines prolong the confusion between ethics and fundamental rights at a European level, they also serve to deepen the confusion by enveloping the concept of fundamental rights within the concept of ‘ethical purpose.’ In the latter conception of ethics, ethics serve as a driver for the design of technical systems on an equal footing with fundamental rights. While fundamental rights are described as the “bedrock for the formulation of ethical principles”¹¹ the resulting operationalization process reduces human rights to very basic principles such as “do good”¹² or “do no harm,”¹³ which are extraordinarily far removed from actual fundamental rights.

The five principles developed in this context (‘do good’, ‘do no harm’, ‘preserve human agency’, ‘be fair’ ‘operate transparently’) are strange interpretations of fundamental rights

⁹ The slowness of law – it is argued – is not compatible with the short development cycles and agile sprints that are common in digital technologies (Gürses and Hoboken, 2017).

¹⁰ Raymond Geuss rightly emphasises the extent to which ‘Individual “ethics” comes to be presented as a purportedly separate subject area and topic for a freestanding treatment under certain political and social circumstances; particularly when the world of politics seems to have moved completely out of the control of individuals’ (Geuss, 2020).

¹¹ https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=58477, p.6.

¹² https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=58477.

¹³ https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=58477.

or their underlying foundations like human dignity. Human dignity or freedom of expression cannot be simplified within a framework of ‘do good’ or ‘do no harm’: this ‘operationalization’ of fundamental rights indeed strips them of much of their legal and normative meaning - to the point of making them unrecognizable. By focusing on their so-called ‘ethical purpose’,¹⁴ this ‘flexible’ operationalization of fundamental rights leads to their watering down: the baby has been thrown out with the bathwater.

The dangers inherent in operationalizing human rights as ‘ethics’ are apparent when the final EU Ethics Guidelines for Trustworthy AI (8 April 2019) are considered.¹⁵ The authors develop ethical principles based on the EU fundamental rights framework, which defines these principles as “respect for human autonomy [...] prevention of harm [...] fairness and [...] and explicability.”¹⁶ As with any attempt to subsume fundamental rights under supposedly univocal and abstract ethical principles, the authors thereby open the door to the possible reduction of fundamental rights to simple concepts like ‘do good’ or ‘operate transparently’. Given the considerable appeal of such reductive, simplistic translations for those focused on encoding such constraints within automated technical systems, the likelihood of fundamental rights thereby getting ‘lost in translation’ is dauntingly high.

The result of this process is a very weak institutionalisation of ethical frameworks, in which neither the institutional structures nor the actual accountability mechanisms are meaningfully defined. Yet the authors can claim to have ‘integrated’ EU fundamental rights into their decision-making process and thus to have responded to the frequent critique that these are often ignored (Veale, 2020).

The approach to ethics described above has understandably led to the backlash described in the introduction: if ‘ethics’ has come to acquire such a bad name in the domain of AI and data governance, it is in large part due to its being confused with a set of simplistic, ready-made answers (Bietti, 2020), in contrast to the Socratic, open-ended effort inherent in an understanding of ethics that focuses on the ‘how should I/we live’ question. This confusion is neither inevitable nor impossible to remedy. The next section invites the reader to consider the gradual refinement of the ethics-regulation interface within healthcare as a case study that may prove helpful in tackling the challenges currently undermining the governance of data-intensive technologies.

3. From data governance to healthcare: constructing a mutually supportive ethics/regulation interface

This section considers a domain where the ethics argument has long been invoked to rebut or limit the need for legal reg-

¹⁴ This ignores the fact that such ‘purpose’ is necessarily the result of a value-loaded interpretation (hence necessarily contested and contestable).

¹⁵ https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=60419.

¹⁶ https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=60419 p.12.

ulation: professional ethics. Because the latter argument has been articulated most convincingly in the domain of health-care, section 2.1. analyses the shifts that are currently taking place in the construction of the ethics / legal regulation interface within healthcare. These shifts are of particular interest for our purposes, since they reveal a tension between two very different ways of understanding the merits of (and rationale for) legal regulation. Section 2.2. Considers what, if anything, can be learned from that discussion when translated to the field of data and machine ethics.

3.1. The interface between ethics and legal regulation in healthcare

If one were to gauge the types of expertise yielding most regulatory power, medical knowledge is hard to beat, especially when it is considered in combination with its access to powerful drugs. Given the extent to which it impacts upon our most fundamental capabilities, the provision of healthcare cannot but be very high upon a regulator's list of priorities. Today medical practice is at the heart of multiple types of regulatory structures, from top-down legislation to professional bodies-based and court-based regulatory interventions. The latter are not only tasked with interpreting the relevant legislation - they are also frequently called upon to assess clinical negligence claims. In such cases, the Court has to assess whether a given healthcare provider has acted with the required degree of care. In this respect, UK Courts have long been criticized for what is often seen as their excessive deference to clinical discretion. Rather than allowing themselves to form a view as to whether a clinician acted with reasonable care, the continuing prevalence of the 'Bolam Test'¹⁷ (at least for matters other than risk disclosure) means that the Courts have to take into account 'prevailing practice' within the profession in question. A clinician's practice will be deemed reasonable as long as she can demonstrate that she acted in line with some accepted medical practice or opinion (even if the latter is a minority opinion).

Before turning to the factors that have led to restricting the scope of the Bolam test (hence restricting the Courts' deferral to what may be called 'professional ethical' judgment), it is interesting to note the language used by Foster and Miola to criticise law's 'abdication of responsibility for ethical issues to professional medical ethics' (Foster and Miola, 2015). Their phrasing indeed directly echoes the concerns that were touched upon in the introduction: Foster and Miola, for instance, deplore the fact that 'medical ethics has been allowed to take over from medical law' (Miola, 2004: 262–263). For Fos-

¹⁷ Under the 'Bolam test' (Bolam v W Friern HMC 1957), a doctor is not guilty of negligence if he (or she) has acted in accordance with a practice accepted as proper by a responsible body of medical men skilled in that particular art. As (Montgomery, 2011, p. 19) puts it, 'this test fixes the required standard of care not by reference to the judge's assessment of a reasonable balance of risk but by reference to peer review. Provided that a responsible body of the practitioner's peers accept that their practice was 'proper', then, even if the judge were minded to disagree, the practitioner would not be liable to pay compensation'.

ter and Miola, 'the more ethical in nature a decision is, the less justification there is for allowing medical ethics to become the arbiter.' (Foster and Miola, 2015).

The above is structured around an antagonistic understanding of the relationship between law and professional ethics, whereby the two respectively vie for 'control' in a mutually exclusive, rather than supportive fashion. In contrast, (Montgomery, 2011) points out that there is room for a mutually supportive relationship between law and professional ethics. Unlike the negative conception of professional liberty - as 'freedom from interference' - that implicitly structures Foster and Miola's account, Montgomery argues that 'the reason for protecting clinical freedom through the law is not to protect professionals from interference (a negative claim) but to ensure that these professional values can be acted upon (a positive claim)' (Montgomery, 2011). Far from necessarily being the product of some individual evaluation exercise on the part of each practitioner, these professional values can be the outcome of some ongoing, collective re-articulation effort that is 'co-originated' by the professional and lay community.

This individual v. 'co-originated', collective dimension of professional ethics matters for several reasons. It matters, first and foremost, because the shift towards greater legal regulation within healthcare originates in large part from a backlash against the long-dominant, paternalist understanding of professional responsibility. This paternalist understanding was most famously criticised by Terence Johnson (Johnson 1972). The latter denounced the self-serving nature of a reference to altruistic values¹⁸ and/or some esoteric knowledge to justify professional self-regulation. Johnson rightly exposed the extent to which such a justificatory discourse allows professionals to define both the needs of the 'lay' population and the way in which such needs ought to be met. Against such a background, Foster and Miola's suspicion towards any 'regulatory vacuum'¹⁹ that ultimately empowers 'the conscience of the individual medical practitioner' -rather than patients- is understandable.

However, once one moves away from an individualistic understanding of professional ethics -once one conceives of the latter as a community-based forum for the ongoing (re)articulation of professional standards and aspirations- the anti-paternalist concerns informing Foster and Miola's 'abdication of responsibility' diagnosis become less significant. Several factors condition the successful move away from an individualistic understanding of professional ethics. First, the relevant professional body - in the English context the BMA- needs to develop institutional structures that are in a position to articulate a set of positive obligations.^{20 21} This approach to

¹⁸ . The latter would typically emphasise that 'delivering health care is an altruistic vocation, not a commercial enterprise tainted by sordid financial pressures' (Montgomery, 2006).

¹⁹ Miola points at the 'vacuum' or vicious circle occurring in situations where the law delegates decision-making responsibility to professional medical ethics which in turn abrogates responsibility back to the law.

²⁰ Section 3 will discuss the possible shape and genesis of equivalent, national professional bodies in the data ethics space.

²¹ Moreover, until 1995, 'the General Medical Council's guidance on professional conduct and discipline (known as the "Blue Book") was primarily concerned with setting out the type of misbehaviour

professional regulation takes into account -instead of glossing over- the inevitable disagreements and uncertainty that structure medical practice, given the combination of medical research advances and evolving needs and aspirations of those whom medicine is meant to serve.

In such a dynamic context, the ongoing (re)-articulation of professional standards has little chance of succeeding if it does not give a significant place to bottom-up, community-based fora. This bottom-up aspect is at the heart of recent work on what many see as a positive shift in the 'way in which the compact between professionals and the state is changing'.²² Along this line, Montgomery for instance highlights the importance of the procedures introduced by the Mental Health Act 1983 (as amended by MHA 2007) as well as Chapter 31 of the Mental Health Act Code of Practice (2008 edition). For Montgomery, the mandatory inclusion of lay members in the tribunals tasked with ensuring that prolonged detention is justified against the criteria set out in the Act, 'indicate the emergence of a new professionalism, providing a context for clinical freedom rather than either deferring to professional opinion or overriding it'.²³

3.2. From healthcare to data governance

One may query the relevance of the above, healthcare-specific regulation-ethics interface in the context of data-intensive technologies on at least two grounds.

First, It is clear that there are significant differences between healthcare and data governance: it would be facetious to claim that the two can be regulated in an identical manner or for the same reasons. However, there is a constructive root that is sufficiently similar between the processes that underpin the gradual construction of regulatory institutions in those two domains. Historically, just as the pace and appetite for regulatory interventions were gaining momentum in early to mid-20th Century medical practice, existing professional bodies (in the UK, the GMC) did not seek to exercise much in the way of regulatory power. Mostly focused on representing the views of the medical profession (and propagating medical knowledge), these bodies only came to acquire 'regulatory teeth' relatively recently. This situation is similar enough to contemporary AI and data professional governance structures²⁴ to argue that it is more productive to draw on the lessons learned in the process of refining the ethics/regulation interface within healthcare (as outlined in 2.1) than to start from scratch.

that might call into question a doctor's continuing registration. Since then the focus has changed, and a series of documents under the umbrella title of "Duties of a Doctor" introduced a very different approach that set out the content of "Good Medical Practice". This marked a fundamental shift in the nature of professional regulation; from resistance against external scrutiny of clinical judgement on the basis of claims to professional autonomy, to explicitly asserting the principles on which the legitimacy of clinical freedom was based' (Montgomery, 2011).

²² (Montgomery 2011).

²³ (Montgomery, 2011, p. 24).

²⁴ Even if the latter are a lot more international in nature (and hence less visible to the general public).

Alternatively, one may query whether the 'grassroots' articulation of values relevant to professional practice through a collaborative, lay-professional endeavour (as outlined in 2.1 in the case of healthcare) is at all realistic in the domain of data-intensive technologies. While a definite answer to this question would require empirical studies that are beyond the scope of this paper, section 1.1 has sought to outline the fast-evolving professional norm development efforts that are taking place today. Without a commitment to open up more widely those efforts to community-based fora (as well as an endeavour to back the latter with adequate institutions), the risk is that well-intentioned but relatively rigid top-down legal regulation may inadvertently end up disempowering those it was supposed to empower. Section 4 unpacks this argument within a critical appraisal of Quebec's recent 'Bill 29'.

4. Towards a constructive ethics/regulation interface

4.1. The need for 'granular' engagement on the part of professional communities

Having outlined existing practices around the development of ethical frameworks in the area of data governance, as well as in healthcare ethics, we next outline four points that are of particular relevance (while far from exhaustive) when considering the extent to which granular engagement on the part of professional bodies is key to the development of a constructive ethics/ regulation interface.

The first is acknowledging that human agency plays a crucial role in decision-making around socio-technical systems, particularly those that are data intensive (Wagner, 2019). This acknowledgement is essential, as it demands a reflection on the psychological and environmental factors that impact such human decision-making. Taking into account those factors entails designing and managing a dynamic system, rather than merely engineering a specific artefact.

Second, the decisions made by human beings, in particular around the management of personal data, can contribute to either mitigating or exacerbating a type of human vulnerability, that is distinct from the physical fragility we all share just in virtue of being human. When the systematic recording of our machine-readable past prevents us from being able to project the self we each aspire to be, seemingly confining us to the 'self' anticipated by each of our profiles, what is compromised lies at the heart of our liberal democracies. We call this type of vulnerability 'development-of-self vulnerability.' That commitment is best summarised as a commitment to safeguard every person's need to continuously (re)-define their sense of self in a way that commands a minimal degree of respect (both from herself and from others). That effort of self-definition is easily imperilled, as outlined in detail in (Delacroix, 2018). This type of vulnerability has much in common with the German constitutional law notion of 'freie Entfaltung der Persönlichkeit', which can be translated as self-development of the personality of a human being. The latter plays a critical role in the German constitutional concept of

human dignity ([Eberle, 1997](#)).²⁵ In the context of this paper, it is enough to say that this development-of-self vulnerability is one of the reasons data protection frameworks such as the GDPR attempt to limit access to and prevent misuse of personal data. However, data protection frameworks are only mitigation strategies for this kind of development-of-self vulnerability; they cannot entirely prevent it. Genuine prevention not only calls for individual and collective responsibility that is robust enough to overcome the limitations inherent in regulation's typical focus on 'identifiable' sensitive decisions. It also calls for much greater awareness of the specific challenges inherent in development-of-self vulnerability than exists at present (if vulnerability is at all considered, it is typically restricted to 'inherent' forms of vulnerability, such as epistemic imbalances and corporeal fragility).

Third, in contrast to the development-of-self vulnerability described above, automated socio-technical systems that are deeply embedded in societal environments can also create new and specific forms of 'rights vulnerability'. Access to and reliance on fundamental rights end up being insidiously compromised when a diffuse set of individual decisions, potentially made by the subcontractor of a subcontractor, or in a frequently used software library on which many other pieces of software depend, have far-ranging legal consequences for human beings. As noted above, this is not to claim that human beings do not have agency in enabling such automated decision-making. The point is instead to acknowledge that once such automation exists, human beings become particularly vulnerable to hard-to-remedy human rights curtailments. This type of 'rights vulnerability' is most easily observed in areas such as failures around systems which automatically disburse unemployment benefits, as seen for instance, in Sweden ([Wills, 2019](#)).

Fourth, the rationale underlying our drawing upon the healthcare context is that it contributes to refining our understanding of the regulation-ethics interface. It should be evident in this context that the call for bottom-up articulation of professional standards within data ethics is not meant to supplant top-down regulatory interventions. In the domain of healthcare ethics, such regulatory interventions impose, at a minimum, a normative structure which constrains professional discretion. They can also remove discretion altogether, conferring rights and responsibilities in a way that is similar to data protection law in the context of data governance.

4.2. Challenges to the development of a constructive ethics/regulation interface

Advocating an ethics-regulation interface that places professional ethics at its heart stands in contrast to many existing approaches to technology ethics. In line with "The Californian Ideology," ([Barbrook and Cameron, 1995](#)) many of these approaches tend to emphasise virtuous individual decision making rather than systematically building structures that enable professional communities to co-develop – frameworks. For example, the procedures around Google's well-known ethical

framework 'don't be evil' slogan were de facto operationalized for many years as Google not doing "what Sergey [Brin] says is evil." ([Schmidt in McHugh, 2003](#)) Similarly, many approaches to ethics in computer science focus on teaching computer scientists how to 'be' more ethical. Whether for the CEO of Google or an undergraduate computer science student, these approaches are not designed to produce systematic accountability that enables responsible behaviour across an organization or industry.

Instead, these approaches transfer the responsibility for ethical decision-making to virtuoso individuals who – even as the CEO of Google – are unlikely to be able to shoulder this burden, devoid as they are from systemic institutional support frameworks. Here the distinction made by Mark Bovens between "accountability as a virtue and a mechanism" ([Bovens, 2010](#): 946) is helpful in order to understand how the institutionalization of ethics can fail. Following Bovens, accountability should be seen primarily from a procedural or institutional perspective, focusing on accountability as a mechanism rather than as a virtue. The latter approach avoids many of the challenges associated with ethics washing ([Bietti, 2020](#)) and is well suited to acknowledging the need for degrees of individual discretion while at the same time ensuring adequate oversight by professional organizations (and by courts – see section 2.1).

Importantly, the current ethical frameworks of many professional organisations in the technology sector – such as IEEE or ACM – currently lack key building blocks, such as the specification of meaningful consequences for rule infractions by members of that professional body ([Bovens, 2010](#)). Meaningful consequences could for example include a willingness to exclude members from a professional body if they break the rules of that professional body. If professional associations are unwilling to provide for these kinds of sanctions, the responsibility for providing accountability mechanisms falls to the State. In this context, the State may choose to pursue a self- or co-regulatory approach that is conducive to its particular objectives ([Latzer et al., 2019](#)). The development of non-binding ethical frameworks potentially enable governments to go beyond what the law would typically allow them to do ([Keller, 2019; Mueller, 2010](#)). This challenge is particularly evident in the EU's Code of Conduct for Terrorist content, but also in other more recent non-binding instruments at an EU level. The legitimacy of such ethical frameworks needs to be questioned, as the process is typically heavily driven and managed by the private sector. Even if such ethical approaches are institutionalized at a later date, the selection mechanisms chosen to develop them lack the legitimacy of robust stakeholder selection mechanisms that ensure balanced and representative participation.

To avoid going down this path, regulators and other public sector organizations would do well to encourage relevant professional organisations to develop such ethical frameworks rather than endeavouring to develop them themselves. The European Commission, as the convener for an expert panel on AI ethics, is not the appropriate organization to engage in this endeavour. Such endeavour to articulate principles should instead be spearheaded by national professional standards bodies which themselves draw upon localised debate fora. This is essential if one is to avoid a fossilised list of principles that

²⁵ ([Neal, 2012](#)) defines human dignity as "a particular sort of ethical response to universal human vulnerability". In contrast, see ([Sangiovanni, 2017](#)).

becomes increasingly out of touch with rapidly changing realities. At the same time, public sector organisations should not 'copy-paste' industry standards that were developed by the ACM, IEEE or ISO in the context of technology-specific ethical frameworks. Instead, governments need to be cognizant of their responsibility to fundamental rights and develop regulatory frameworks accordingly. In doing so, they may draw on the extensive experience they have acquired in other fields, such as healthcare ethics. Ethical frameworks should be developed by professional standards bodies or public sector organisations, within a clear institutional framework that includes effective sanctioning mechanisms.

5. Institutionalising the ethics/regulation interface for artificial intelligence

A recent proposal to introduce licensing requirements for the development of specific types of computing systems in Quebec makes for an interesting case study. The proposal (made in June 2019 by the Minister of Justice of Quebec, Sonia LeBel, as part of National Assembly Bill 29) is an amendment to the Professional code, and includes licensing requirements²⁶ for "an autonomous electronic or computer system for the operation of works referred to in this paragraph, including software" if they have been "derived from engineering principles."²⁷

This licensing proposal was met with considerable push-back from existing professional computing associations. While Bill 29 to amend the Professional Code was welcomed by the main engineering trade associations GERLI and the Quebec Order of Engineers, a Coalition for the Future of Computing is lobbying heavily against it. This coalition includes key members of the Quebec artificial intelligence community like the company Element.AI, Microsoft and the Quebec Association of IT freelancers. The coalition recommends removing the licensing requirement and instead proposes certification requirement for "critical software."

The arguments underlying this push-back help illustrate some of the challenges associated with the construction of an ethics-regulation interface of the kind we are proposing in a domain as wide and rapidly changing as computing systems development. In the case of Bill 29, existing market entry requirements for engineers were simply copied across and applied to computing, while ignoring the specificities of this field, which has become so broad and multifaceted that it encompasses numerous aspects of human life. Moreover, the licensing requirements are wide in scope, encompassing both automated systems and any other systems that derive their principles from engineering. Almost any computing system

could potentially fall within the reach of this licensing requirement. A more limited and precisely defined scope would be helpful, as the extensive nature of the current licensing requirement would likely place an unnecessary burden on low-risk everyday computing technologies. Instead, defining a core set of high-risk use-cases that legislators see as problematic and then slowly and carefully expanding them over time, would seem to be a much more promising approach to the current scope of this licensing requirement. As an example, one particularly promising 'licensing' application would target data-intensive technologies used for political campaigning during elections.

Another key challenge within the context of a licensing requirement consists in ensuring that individual members of professional associations are accountable for their actions within the context of that licensing regime. By accountability, we mean a form means a procedural accountability mechanisms as envisioned by Mark Bovens: "a relationship between an actor and a forum, in which the actor has an obligation to explain and to justify his or her conduct, the forum can pose questions and pass judgement, and the actor may face consequences." (Bovens, 2010) The most effective way of ensuring accountability is for members of these professional associations to face appropriate sanctions for breaking internal association professional ethics standards. For such action to be possible, internal association professional ethics standards would need to be sufficiently specified that a breach of them could be clearly identified. At present the ethical standards of both the IEEE and ACM are not very fully developed on this point.

For example, the ACM Code of Ethics and Professional Conduct²⁸ suggests that authors should "Treat violations of the Code as inconsistent with membership in the ACM," making it unclear whether members will actually be excluded for breaching this code and the ACM Enforcement Procedures for the ACM Code of Ethics and Professional Conduct²⁹ do not provide any further guidance on whether this will be the case. IEEE, ACM and other similar professional bodies would need to develop more significant internal accountability mechanisms for their members. In contrast, the Professional Code of Quebec helpfully sets out a framework of publicly managed disciplinary councils, for which there are a well-defined set of rules established by law (Section 117). As such the Quebec Professional code can provide considerable additional guidance on how to organise and structure relevant frameworks to ensure accountability of the members of professional organisations.

One last and important challenge when studying the professional code of Quebec is the extent to which existing trade associations can become part of the framework. A key difficulty is the high barrier for entry to new professional associations, as all existing organisations that implement the professional code are government appointed for an entire field or sector. It is thus not possible at present for a trade association like the Quebec Association of IT freelancers or a re-

²⁶ In the GDPR context, Such licensing requirements could conceivably be developed within the framework of Article 40 of the GDPR, which allows regulators to approve codes of conduct developed by professional bodies for automated data processing techniques (Gasser and Schmitt, 2019).

²⁷ Explanatory Document, Bill 29 – Act to amend the Professional Code and other provisions in particular in the oral health and the applied sciences sectors, Ordre des ingénieurs, June 2019 https://www.oig.qc.ca/Lists/Publications/Attachments/220/DOC_expliquatif_PL29_EN_2019-06-25.pdf

²⁸ <https://www.acm.org/code-of-ethics#h-4>.

²⁹ <https://ethics.acm.org/wp-content/uploads/2018/07/2018-ACM-Code-of-Ethics-Enforcement-Procedure.pdf>

gional chapter of the IEEE or ACM to become part of this system, with computer technologies likely falling under the provision of the Quebec Order of Engineers. Given the breadth and depth of professional communities that have developed in computing, this one-size fits all approach is likely to hinder progress. A more effective approach would be to open up the Quebec Code of Professions to input and/or active participation from other organisations, both public and private. This approach is particularly important to ensure a bottom-up process which ensures that all relevant voices within computing can be meaningfully articulated. Power could still be situated at an institutional level not dissimilar to what is discussed in the Quebec proposal, but without a bottom-up process preceding it this power would lack any meaningfully legitimacy. Without such an inclusive approach to bringing in relevant organisations, both the quality of the norms developed and the likelihood of their being implemented in a meaningful way are questionable.

Having highlighted all the above challenges, the underlying approach taken in Bill 29 remains noteworthy for the purposes of this paper. Were it to draw on genuine bottom-up mechanisms for the continuous re-articulation of adequate professional standards, it would have the potential to strengthen professional communities in a specific national jurisdiction by developing a constructive ethics-regulatory interface. At the moment, many professional organisations exist only at an international level, and even if they exist at a local level in the form of chapters, they do not often have local legal entities. Professional ethics frameworks provide an incentive to international associations in computing to more strongly ground their work in numerous jurisdictions across the world, as doing so provides advantages related to market access. However, it does not seem that the authors of Bill 29 have sufficiently engaged with other jurisdictions in Canada or international counterparts interested in adopting a similar approach. Such an approach could even envisage a web of mutual recognition agreements that could be developed between different jurisdictions. Such an international approach is crucial to ensure that the development of software remains possible across borders.

Developing international best practices around what an effective ethics-regulatory interface for data-intensive technologies could look like in this area constitutes an important next step. While each country needs to develop and refine an ethics/regulation interface that is suited to data-intensive technologies, there is no doubt that some degree of international standardisation would prove helpful. In this context, internationally oriented professional engineering bodies have a particularly vital role to play, ensuring the interoperability of the standards developed by national bodies.

6. Conclusion

There is undoubtedly something heartening about placing ethics at the centre of ongoing efforts to build tools that promise to make our lives better. There is even talk of tech companies considering the need to appoint 'chief ethics officers' to sit alongside legal teams and chief compliance officers (Swisher, 2018). A sense of queasiness however takes

over as soon as one pauses to consider just how easily that signifier ('ethics') can metamorphose into something that is not even remotely connected to the Socractic question - 'How should I [we] live?'. For as long as the latter remains a genuinely open question, there will be a limitless number of answers, and each answer will reflect social, cultural and political conditions that may -and often do- clash in the needs and aspirations they give rise to.

We believe that an endeavour to develop an ethics-regulation interface that gives a greater role to a variety of professional institutions with strong bottom-up, community based components would be in a position to rekindle much-needed grassroots, political debates as well as (re)-articulate the positive obligations necessary to addressing the considerable vulnerabilities that are concomitant with the current data-subject / data-controller power imbalances. Among these vulnerabilities, our paper has sought to draw particular attention to development-of-self vulnerability and rights vulnerability.

The creation of a licensing requirement for specific usages of data-intensive technologies could also play an important role in the development of a more effective ethics-regulation interface. This licensing requirement could, for example, target data-intensive technologies used for political campaigning around elections. Bill 29 proposed in Quebec demonstrates that there are already attempts underway to achieve this, although there is considerable additional work needed if it is to enhance the articulation of a constructive ethics-regulation interface in the computing industry.

Finally, but perhaps most importantly: our emphasis on the need to further develop (and rely on) community-based, bottom-up regulatory institutions such as professional bodies both presupposes and reinforces the call for robust top-down regulation. Instead of serving as an excuse for self-regulation, the progressive refinement of the ethics/regulation interface that could be facilitated by strong professional regulatory bodies would not only contribute to improving the role of technology in society. It may also temper the prominence of the 'need to duck political rifts by turning to ethics instead' mentality that has done so much damage to all..

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

REFERENCES

- Abbott KW, Snidal D. Hard and soft law in international governance. *Int. Organ.* 2000;54(3):421–56.
doi:[10.1162/002081800551280](https://doi.org/10.1162/002081800551280).
- Angwin J. and Grassegger H. (2017) Facebook's secret censorship rules protect white men from hate speech but not black children. Available at: <https://www.propublica.org/article/facebook-hate-speech-censorship-internal-documents-algorithms> (accessed 4 July 2017).
- Balkan A. (2017) Ethical design manifesto. Available at: <https://2017.ind.ie/ethical-design/> (accessed 11 July 2018).

- Barbrook R, Cameron A. *Californian Ideology: A critique of West Coast Cyber-Libertarianism*. The Hypermedia Research Centre; 1995.
- Bietti E. From ethics washing to ethics bashing. *FAT* '20: Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, Barcelona, Spain, 2020, p. 10. ACM Conference on Fairness Accountability and Transparency (FAT* 2020), 2020.
- Bovens M. Two concepts of accountability: accountability as a virtue and as a mechanism. *West Eur. Polit.* 2010;33(5):946–67. doi:[10.1080/01402382.2010.486119](https://doi.org/10.1080/01402382.2010.486119).
- Cellan-Jones R. Hawking: AI Could End Human Race. BBC News; 2014 2 December. Available at: <https://www.bbc.com/news/technology-30290540> (accessed 2 October 2018).
- Clarke R. Principles and business processes for responsible AI. *Comput. Law Secur. Rev.* 2019;35(4):410–22 Elsevier.
- Daly A, Hagendorff T, Li H, et al. AI, governance and ethics: global perspectives. In: Pollicino O, de Gregorio G, editors. *Constitutional Challenges in the Algorithmic Society*. Cambridge: Cambridge University Press; 2020.
- Delacroix S. (2018) Taking turing by surprise? Designing digital computers for morally-loaded contexts. *arXiv preprint arXiv:1803.04548*.
- Eberle Ej. *Human dignity, privacy, and personality in German and American constitutional law*. Utah L. Rev. 1997:963 HeinOnline.
- Else H. Europe's AI researchers launch professional body over fears of falling behind. *Nature* 2018. doi:[10.1038/d41586-018-07730-1](https://doi.org/10.1038/d41586-018-07730-1).
- Foster C, Miola J. The relationship between medical law, medical ethics, and medical morality? *Med. Law Rev.* 2015;23(4): 505–530.
- Gasser U, Schmitt C. (2019) The role of professional norms in the governance of artificial intelligence. ID 3378267, SSRN Scholarly Paper, 25 April. Rochester, NY: Social Science Research Network. Available at: <https://papers.ssrn.com/abstract=3378267> (accessed 10 May 2019).
- Geuss R. *Who Needs a Worldview?*. Harvard University Press; 2020.
- Gürses S, Hoboken J. (2017) Privacy After the Agile Turn. In: Selinger (ed.) *The Cambridge Handbook of Consumer Privacy*. Available at: <https://osf.io/ufdvb/>.
- Haas L, Gießler S. (2020) AI ethics global inventory. Available at: <https://inventory.algorithmwatch.org/> (accessed 28 October 2020).
- Hagendorff T. The Ethics of AI Ethics: An Evaluation of Guidelines. *Mind Mach.* 2020;30:99–120. doi:[10.1007/s11023-020-09517-8](https://doi.org/10.1007/s11023-020-09517-8).
- Keller D. State and Platform Hybrid Power Over Online Speech. Hoover Institution Essay; 2019 Aegis Series Paper No. 1902 Available at: <https://assets.documentcloud.org/documents/5699593/Who-Do-You-Sue-State-and-Platform-Hybrid-Power.pdf>.
- Latzer M, Saurwein F, Just N. Governance-choice method: in search of the appropriate level of state intervention. *The Palgrave Handbook of Media Policy Research Methods*. Basingstoke: Palgrave Macmillan; 2019.
- Lawrence N. (2018) The three ds of machine learning. Available at: <http://inverseprobability.com/talks/notes/the-three-ds-of-machine-learning.html> (accessed 10 May 2019).
- Marchant G. (2019) "Soft law" governance of artificial intelligence.
- Marshall A, Davies A. (2018) Uber's self-driving car saw the woman it killed, Report Says | WIRED. Available at: <https://www.wired.com/story/uber-self-driving-crash-arizona-ntsb-report/> (accessed 2 October 2018).
- McHugh J. (2003) Google vs. Evil. Available at: <https://www.wired.com/2003/01/google-10/> (accessed 10 May 2019).
- Metzinger T. (2019) EU guidelines: ethics washing made in Europe. Available at: <https://www.tagesspiegel.de/politik/eu-guidelines-ethics-washing-made-in-europe/24195496.html> (accessed 10 May 2019).
- Miola J. Medical law and medical ethics-complementary or corrosive? *Med. Law Int.* 2004;6(3):251–74.
- MIT Media Lab. Moral Machine; 2018 Available at: <http://moralmachine.mit.edu> (accessed 2 October 2018).
- Montgomery J. Law and the demoralisation of medicine. *Legal Stud.* 2006;26(2):185–210.
- Montgomery J. The virtues and vices of professionalism. In: Bhugra D, Malik A, editors. *Professionalism in Mental Healthcare: Experts, Expertise and Expectations*. Cambridge: CUP; 2011. p. 17–31.
- Mouffe C. Deliberative democracy or agonistic pluralism? *Soc. Res.* 1999:745–58.
- Mueller M. *Networks and States: The Global Politics of Internet Governance*. Cambridge, MA: MIT Press; 2010.
- Neal M. Not gods but animals": human dignity and vulnerable subjecthood. *Liverpool Law Rev.* 2012;33(3):177–200. doi:[10.1007/s10991-012-9124-6](https://doi.org/10.1007/s10991-012-9124-6).
- O'Keefe K, Brien D.O. (2016) New paradigm or ethics washing? An analysis of Facebook's ethics report. Castlebridge. Available at: https://castlebridge.ie/wp-content/uploads/2018/07/new_paradigm_or_ethics_washing.pdf.
- Raab CD. Information privacy, impact assessment, and the place of ethics. *Comput. Law Secur. Rev.* 2020.
- Rand-Hendriksen M. (2018) Using ethics in web design — smashing magazine. Available at: <https://www.smashingmagazine.com/2018/03/using-ethics-in-web-design/> (accessed 10 May 2019).
- Simon CJ. Ethics and artificial general intelligence: technological prediction as a groundwork for guidelines. *Proceedings of the IEEE International Symposium on Technology and Society (ISTAS)*; 2019. p. 1–6 November 2019.
- Smith B. (2018) Facial recognition: it's time for action. Available at: <https://blogs.microsoft.com/on-the-issues/2018/12/06/facial-recognition-its-time-for-action/> (accessed 10 May 2019).
- Stahl BC, Wright D. Ethics and privacy in AI and big data: implementing responsible research and innovation. *IEEE Secur. Priv.* 2018;16(3):26–33.
- Statt N. (2019) Google dissolves AI ethics board just one week after forming it. Available at: <https://www.theverge.com/2019/4/4/18296113/google-ai-ethics-board-ends-controversy-kay-coles-james-heritage-foundation> (accessed 10 May 2019).
- Swisher K. (2018) Who will teach silicon valley to be ethical? - The New York Times. Available at: <https://www.nytimes.com/2018/10/21/opinion/who-will-teach-silicon-valley-to-be-ethical.html> (accessed 10 May 2019).
- Torres P. The possibility and risks of artificial general intelligence. *Bull. Atomic Sci.* 2019;75(3):105–8 Taylor & Francis.
- Veale M. A critical take on the policy recommendations of the EU high-level expert group on artificial intelligence. *Eur. J. Risk Regul.* 2020;1–10 Cambridge University Press. doi:[10.1017/err.2019.65](https://doi.org/10.1017/err.2019.65).
- Wagner B. Ethics as an escape from regulation: from ethics-washing to ethics-shopping?. In: Hildebrandt M, editor. *Being Profiling: Cogitas Ergo Sum*. Amsterdam University Press; 2018.
- Wagner B. Ensuring meaningful human agency in automated decision-making systems. *Policy Internet* 2019;11(1).
- Wattles J, O'Sullivan D. (2019) Facebook's Mark Zuckerberg calls for more regulation of the internet. Available at: <https://www.cnn.com/2019/03/30/tech/>

facebook-mark-zuckerberg-regulation/index.html (accessed 10 May 2019).

Williams B. *Ethics and the Limits of Philosophy*. Routledge; 2011.

Wills T. Sweden: Rogue Algorithm Stops Welfare Payments for up to 70,000 Unemployed. AlgorithmWatch; 2019 Available at:

<https://algorithmwatch.org/en/rogue-algorithm-in-sweden-stops-welfare-payments/> (accessed 10 May 2019).

Winfield A, Jiroka M. Ethical governance is essential to building trust in robotics and AI systems. *Philos. Trans. A Math. Phys. Eng. Sci.* 2018;13.