European Journal of Law and Technology, Vol 10, Issue 3, 2019

EJLT European Journal of Law and Technology

Regulating AI and Machine Learning: Setting the Regulatory Agenda

Julia Black * and Andrew Murray **

Abstract

Disruptive technologies arrive with regularity. Whether it is the first industrial revolution with steam powered factories and transportation, or subsequent revolutions which brought about chemical engineering, communications revolutions, aviation and eventually biotechnology and digitisation. We stand at the edge of the next revolution the AI revolution where methods of artificial intelligence and machine learning offer possibilities hitherto unimagined. How this revolution develops and how our society absorbs the potential of this new technology will be largely determined by the models of regulation and governance applied to the nascent technology. In this paper the authors examine lessons from history and propose a framework for identifying and analysing the key elements of regulatory regimes and their interactions which can form the basis for developing a new model for AI regulatory systems. Furthermore, it argues that the goals of such systems should be to manage the risks different models and uses of AI pose, not just the ethical issues they create.

Keywords : Regulation, Artificial Intelligence, Regulatory Models, History of Techno-Regulation, Decentred Regulation, Polycentric Regulation

Introduction

Successive periods of industrial revolution have been enabled by the creation of new technologies, and with that new risks, new concentrations of power and new coordination challenges. In each phase, states have, to varying degrees and in varying ways, sought to manage those risks and challenges through forms of regulation. The transition to new manufacturing and engineering processes in the 18th and 19th century prompted the rise of occupational health and safety regulation for example. The second industrial revolution in the late 19th and early 20th centuries was characterised by the new technologies of energy production, chemical processes, engineering and methods of communication. In each area, successive waves of regulation were introduced to manage risks and facilitate coordination, for example of transport, of the airwaves, and of airspace. The third industrial revolution was marked by the creation of, amongst other things, nuclear energy, electronics and computing, information technology, biotechnologies, and the plethora of technologies driven by the space race. By the early 1990s, sociologists Beck and Giddens were arguing that we were seeing the advent of a 'risk society',

preoccupied with the risks to health, safety and in particular the environment posed by the proliferation of these new technologies. [1] Regulation turned in part to managing these new risks, which also were starting to prompt significant ethical debates, for example on what limits should be imposed on the development and deployment of technologies of genetic engineering of humans, animals and plants. However, this was also the period when economic theories of neo-liberalism and the concomitant political philosophy of the limited state came into full force, restricting the legitimacy of state intervention into markets to being only that which was necessary to make markets function efficiently; an edict which can be in tension with the risk society thesis. The fourth industrial revolution is upon us now, characterised not just by increased digitisation but by technological innovation which cuts across the spheres of the biological, the physical and the digital. Each period builds on its predecessor, creating incredible opportunities but bringing its own risks, and each occurs in a shifting political context in which contesting political views of the legitimate role of the state shapes arguments as to what the role and purpose of regulation should be.

This sweep through history is necessarily partial and incomplete, but illustrates some of the historical regulatory context to current debates on the regulation of AI and machine learning (ML), and the related issues of the regulation of their raw material: data collection and use. The first and second parts of this paper outlines the varying state-based regulatory responses to disruptive technologies over the last hundred years or so, highlighting the different rationales for and modes of regulatory intervention and the contested political philosophies which underpin arguments about the appropriate role for states or firms in regulation. The third draws out some of the parallels which can be seen between early debates on the regulation of the internet and contemporary debates on the regulation of AI and machine learning. The fourth part argues that if we are to make progress in developing effective and accountable systems of regulation, we need first to stand back from the particularities of these debates to analysing regulatory systems as a whole. Building out from the decentred or polycentric analysis, in its final part we provide a framework for designing regulatory systems, for analysing deep rooted causes of failures, for thinking through the potential impacts of changes in any part of the system, and for helping us understand how each element would need to operate and be accountable if regulation of AI is to be both effective and trusted.

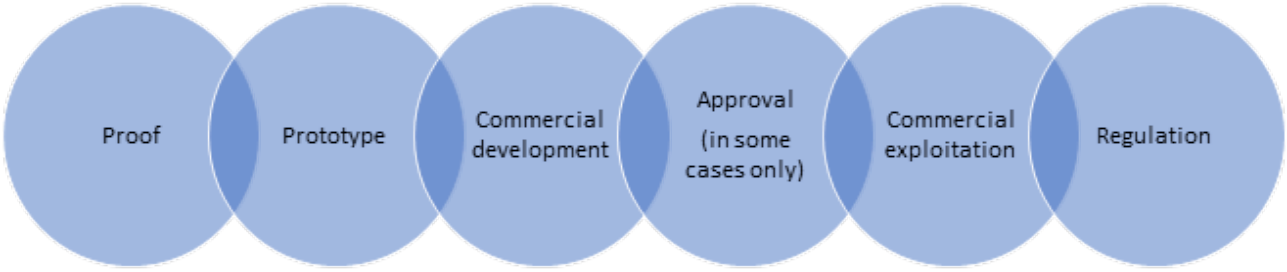
1. Regulation and Technology - an historical perspective

An important message that regulators of AI must learn from prior disruptive (and disrupting) technologies is that technological development and the disruption that it brings occur in the vanguard of socio-economic and socio-legal change. This has been seen repeatedly: from the introduction of the railway and later the motor vehicle to the development of the telephone system in the 19th century, through aviation, commercial radio, broadcast and film media, and the development of the internet in the 20th century.

Looking at the historical record we have identified and modelled six key stages that disruptive technology normally passes through from its concept to it becoming a regulated product or service. There are three initial stages: (1) proof of theoretical concept; (2) development of a prototype; and (3) development of a commercial manufacturing and distribution system. These initial stages are followed by one of two further stage pathways. For technologies that need to make use of centralised, scarce, or public resources, or which pose systemic risks or risks of 'deep regret', such as to life, a prior licensing or approval stage is usually required, so stage (4) is licensing or approval for development; (5) is commercial marketing and exploitation; and (6) is reactive regulation and control. For technologies which do not use centralised or public resources or it is judged that risks are diffused, or can be managed by individual

consumers or through remediation, the fourth stage may be avoided and direct movement from stage (3) commercial development to stage (5) commercial exploitation may take place:

Figure 1: Disruptive technologies - stages of development



Recognition of which form the disruptive technology takes is important for regulators as it determines whether ex ante or ex post exploitation (stage 5) regulation is possible (and appropriate). Historically factors such as access to, and use of, a shared public resource allowed for ante-exploitation (stage 4) regulation. If we look for example at the history of motor vehicle regulation, we see that while the first motor vehicles (in the modern sense) were developed in the 1880s, legal regulation of "horseless carriages" predates this considerably, with the Locomotives on Highways Act 1861 being the UK's oldest legal provision on the regulation of powered road vehicles. The early regulation of this nascent technology was possible as the technology was to be found on the public highway; a space shared with other users including horse riders, operators of carriages and carts, and pedestrians. The need to regulate this shared space predated the disruptive innovation of the powered motor vehicle which meant that although a disruptive technology it was also quickly a regulated technology. [2] When we look to other disruptive technologies of the late 19th and early 20th centuries we find that the need to use public, or scarce, resources, often meant early (stage 4) intervention into the nascent commercial market for the technology. We find, for example, early government regulation of radio communications. Originally demonstrated by Guglielmo Marconi in 1896, radio communications were quickly regulated, with the International Radiotelegraph Conference taking place in Berlin in 1906, and with the US Congress noting in 1910 that regulation was necessary as "the physical limitation on the airwaves or electromagnetic spectrum restricts the number of stations." [3]

These early interventions into areas with scarce and shared public resources can be contrasted with the development of telecommunications regulation and (a lack of) intervention into the market for the home telephone system. Here there was no early (stage 4) intervention. The (semi) modern analogue telephone was developed in the 1870s by a number of inventors but usually attributed to Alexander Graham Bell who patented the telephone in 1876. With Bell holding a monopoly (in the US market) take up was slow until 1894 when the technology entered the public domain. During the 18 years of Bell's monopoly the

average number of daily calls per 1,000 population grew relatively slowly, from 4 to 37. [4] Once Bell's legal monopoly was removed uptake grew quickly: thousands of competitors began connecting users, increasing the daily calling average per 1,000 people from 37 in 1895 to 391 in 1910. [5] It was only in 1910 (at the same point that Congress first regulated the much newer technology of radio) that legal regulation entered into the market through the Mann-Elkins Act of 1910, a now discredited attempt to regulate the market to create natural monopolies and the direct cause of the AT&T monopoly. [6] In contradistinction to radio therefore, where there was a limited supply of spectrum and as a result a stage 4 intervention, with telephone as copper wires were theoretically limitless (as long as someone paid for their construction and installation) regulation came much later as a stage 6 intervention.

What can be learned from this? We can see that the interaction of regulation with the development and deployment cycle of disruptive technology can follow the two paths identified above. Either: (1) proof; (2) prototype; (3) commercial development (4) approval; (5) commercial exploitation; and (6) regulation and control OR (1) proof; (2) prototype; (3) commercial development (4) commercial exploitation; and (5) regulation and control.

The distinction historically between the two forms of intervention in the context of telecommunication technologies was one based on resourcing and scarcity not, for example, on the nature of risks posed to health or life. More recently however the model has changed slightly. Moves towards market deregulation in the 1980s has seen public regulators less likely to get involved in approval or licensing where the issue is a market or scarcity one. Regulators tend now only to intervene prior to commercialisation in areas of public safety or security in order to manage risks; so prior approval for pharmaceutical products and medical devices is a vital part of their commercial development and deployment cycle. Interestingly development of autonomous vehicles (AVs) seems to be following the approval model. No state is permitting the unregulated testing of autonomous vehicles on public roads. The United Kingdom, one of the leaders in AV testing, has created a rigorous regulatory regime controlling the use of AV on public roads. [7] This is likely the result more of public safety concerns rather than the fact that public highways are a common resource. There have been a wide variety of vehicles licensed to use public highways in recent years, always at the point that safety concerns are met: thus, the primary reason for intervention at stage 4 today is safety not concerns over markets or resourcing.

2. Regulation and Technology Redux: Internet Regulation (and Failure)

The default approach of states today is not to interfere in nascent (likely disruptive) technologies but rather to let the market regulate unless there are public safety or security concerns. The model for late 20th century disruptive innovation and regulation is probably internet regulation and governance. [8] While the internet is parasitical upon commercial telecommunications networks (and so applying the early 20th century model seen in radio communications we may have expected it to be regulated) moves were made to deregulate telecommunications markets to ensure data carriage via local loop unbundling and shared access (rights of carriage) for commercial broadband providers. [9] Carriage requirements, partnered with the placing of network protocols such as TCP/IP into open source and the release of the WWW software protocols by CERN in 1993 created an marketplace not subject to prior approval despite the relative scarcity of network capacity at the time: the theory it appears was the market would regulate the nascent uses of this exciting new technology. Coordination on standards, domain names, and other key features of the internet was, and is, provided by non-state bodies, such as the World Wide Web Consortium (W3C), Internet Engineering Taskforce (IETF) and Internet Corporation for Assigned Names and Numbers (ICAAN).

However, the truth was the market was never quite as free as free market theorists imagined and the technology was never quite as regulated as utility regulators may have liked. We tend to map the regulation of the internet through clearly defined phases that in many respects are quite similar to those seen in the development of telecommunications regulation about a century before.

The first phase may be defined as the market regulation phase. It was marked by a strong libertarian ethos and the belief that only cybernauts or netizens (to use the language of the time) could set the limits of their own freedoms in this new space. This early cyber-libertarian movement attracted a number of prominent supporters including most famously John Perry Barlow. Cyber-libertarians were identifiable by their adherence to the belief that the incorporeal and borderless nature of the digital environment would render traditional law-makers powerless, and would empower the community within cyberspace to elect its own law-makers and to design its own laws tailored to that environment. The high point, in the public conscience, of this argument was Barlow's Declaration of Independence for Cyberspace. [10] Here he set out the cyberlibertarian argument that traditional governments had no moral authority in cyberspace as it was a space separate from the traditional post-Westphalian jurisdictions recognised in international law. In essence the argument may be boiled down to a simple claim that cyberspace was a space separate from analogous real-world spaces such as international air routes, the high seas, or even outer space, in that it could not be physically represented and existed only as a space made of protocols and data. Traditional sovereign governments could, according to cyberlibertarians, not exert any moral authority as the action of any state to control any part of cyberspace would have impact throughout the space beyond the sovereign limits of any government (or governments if acting in concert).

The legal framework to this argument was of course famously supplied in David Johnson and David Post's seminal 1996 paper Law and Borders. [11] There they argued that no state has authority to regulate activities which occur in cyberspace for four, interconnected reasons. Firstly, that lawmaking is the exercise of power over those persons whom the state can control. By asserting a claim to apply national law, the state in question is also asserting a right to control the cyberspace activities of individuals who reside in other states, and this conflicts with those other states' monopoly rights to exercise power over their citizens. Secondly, that while they recognise that some overlap in power claims is legitimate, via the effects test of private international law which is accepted and adopted by all states, the effects test should not apply at all to activities in cyberspace. This is because those activities have no greater effect in any one state than in the remainder of the world, and so no state can legitimately claim to apply national law in preference to any other national law merely on the ground of effects. Thirdly, the legitimacy of a state's law-making power derives from the consent of the governed and their participation in the law-making process. Claiming to apply national laws to cyberspace activities goes beyond the boundaries of that legitimacy because it extends the ambit of those laws to persons who have not so consented and who have no way to participate in the lawmaking process, for example through elected representatives. Finally, a lack of borders means that cyberspace users do not receive the notice to which they are entitled that their activities are now subject to a particular state's laws. The rule of law, they argued, requires notice of a law's claim to authority over one's actions. [12]

This may be seen as a unique claim, about a unique technology. Prior technologies did not allow for a creation of a space outside real space. It was clear that in areas of shared responsibility such as in aviation, shipping, or even in space law nation states shared both control and responsibility. This was necessary due to both the limited availability of resources (shipping lanes, aviation corridors, orbital paths) and to mitigate risks to both persons and assets. Cyberspace was, to cyberlibertarians, different as there were in theory an unlimited resource in bits, and in practice little risk to persons or (replicable) digital assets.

Of course, in truth little of the Cyberlibertarian argument was true. Resources, in the form or telecommunications bandwidth were scarce, while risks of harm were real in multiple forms from simple copyright infringement or abuse of personal data, through online fraud to online abuse, hate, and threats of violence (including death threats). This truth formed the foundation of the opposing cyber-realist movement spearheaded by the likes of Cass Sunstein, Lawrence Lessig and Jack Goldsmith. This movement examined the nexus between the real world and the digital. Much like maritime law they recognised the effectiveness of regulation at the margins. Rather than focusing on port authorities they focused on the access points of the internet and the code of the space itself. Henry Perritt pointed out that Town Hall Democracy, as proposed by the Cyberlibertarians, could not function in such a large and varied space as there was no form of self-governing community in the space. Instead, he pointed out that contractualism and control by the bodies that provide access would be the default for of regulation unless states intervened to protect individuals: [13] a point made forcefully by Lawrence Lessig in his book Code and Other Laws of Cyberspace . [14]

With the 1990s debate fractured between the market regulation position of libertarians and the contractualism and code position championed by digital realists such as Lessig, Jack Goldsmith, [15] and Tim Wu, [16] the academic debate began to centre on more practical questions of modelling effective regulation. In essence if the digital realists were right, how could effective regulation be modelled for a place which, as the libertarians had pointed out, had no clear borders and no government? Some such as Murray re-examined the role of the "citizens of cyberspace" pointing out that legitimacy is still drawn from the governed and as such a model of "symbiotic regulation", regulation "which affords all participants in the regulatory matrix an opportunity to shape the evolutionary development of their environment", [17] was to be preferred. Others like Brown and Marsden examined the role of co-regulation, [18] while Reed returned to the concept of a governable cyber-space by re-examining the authority and legitimacy (what he called respect-worthy) of laws in cyberspace, [19] a theme he would return to in his recent book with Andrew Murray. [20]

While this academic debate was ongoing what happened in the wider world of law and policy was a model of regulatory weakness and ultimately failure. Governments seemed paralysed, with no government (at least no Western government) wanting to be the first to be seen to be regulating this innovative, creative space where freedom seemed to create both economic and civic benefits. Governments interacted with this space in a piecemeal fashion: some laws on copyright infringement here, [21] some on hate and harmful speech there, [22] but no coherent strategy for regulation of the space emerged. By around 2010, and the emergence of the currently dominant model of regulatory thought for the online environment - intermediary or platform responsibility/liability, [23] it was clear that as predicted by proponents of digital realism, effective regulation of the online environment had been ceded through contractualism to a small number of key online platforms: platforms which act within their own spheres as private nation states. [24] There is the state of Facebook which controls much of our online social media experience. There is Alphabet which controls our search, and much of our mobile experience, Apple which controls the rest of our mobile experience and much content experience, Amazon which controls a large portion of our content experience and much of the Internet of Things, and Microsoft which essentially sweeps up everything else. Real world states are now rushing to bring forth legislation pell-mell to oblige these "gatekeepers" to regulate our online lives and experiences in line with the values of the state rather than in line with the corporate values of the gatekeepers. [25]

The experience of internet regulation from 1995 to today serves as a warning which is applicable to all emerging technologies. By failing to take early steps to structurally regulate the internet and instead focusing on individual harms, governments failed to appreciate that they had left markets to control a communications technology which relies on network effects and can create system wide risks and impacts. As a result, we ended up in a position similar to the natural monopolies of telecommunications

seen in the 20th century, however this time it was as a result of a failure to intervene rather than by design for as Lawrence Lessig had made clear in 1998, East Coast codemakers had the ability to control West Coast codemakers - they just required the will. [26]

3. Lessons from history for the regulation of AI and Machine Learning: bringing risks back in

Today parallels may be seen between the internet regulation debate of the 1990s and the debate surrounding the regulation of AI and Machine Learning. Firstly, AI and ML similarly to the internet, but distinctly from motor vehicle regulation, telecommunications regulation or radio regulation, does not seem to raise any scarcity issues. Like bits datasets are seemingly limitless, scalable and more valuable as their accumulation grows. Secondly, like internet regulation in the 1990s, risks of harms are being characterised as particular or individual risks rather than systemic or structural. There are clearly identified risks around bias and the attendant regulatory regime in the General Data Protection Regulation. [27] There is awareness of harms from data mining and profiling, [28] and discussion over liability and risks in decision-making systems. [29] There is even discussion of copyright interests in AI generated works [30] but this discussion, as with discussion of risks and harms of online content and materials in the 1990s, remains piecemeal and rooted in specific harms or risks rather than the systemic risk of AI and ML.

The wider discourse that is taking place is drawing us away from law, or even traditional models of command and control or co-regulation and governance, towards soft self-regulation and codes of practice. This ethical model, discussed further below, has seen the adoption of codes of practice for general AI [31] and for data-driven health and care technology, [32] among others. However, as we shall discuss, ethical standards for such systemic risks are insufficient, particularly in so far as they assume that risks are individualised and that the key to their management is the choices an individual consumer makes within the market place. Based upon our experience of internet regulation and governance from 1995 to today, such an approach will lead to future regulatory failures.

If one were to predict the outcome of this based upon our experience of the internet regulation case study, it does not make for happy reading. There is a nascent debate on the regulation of AI with a number of proposals, beyond ethics, put forward. The first, from Matthew Scherer, is the creation of a state regulator who would certify AI following safety testing. [33] Variations on this theme come from Andrew Tutt, [34] and Olivia Erdélyi and Judy Goldsmith who recommend a new international artificial intelligence organisation which might bring binding commitments from states. [35] However as we might predict from the internet governance debate they are being met by a number of arguments. Firstly, there is AI Libertarianism: (1) the market will regulate; [36] (2) there is no one government or regulator who has authority, or even the legitimacy to regulate; [37] (3) the community is the source of legitimate authority to regulate. [38] Then in time there will be AI realism: (1) the market cannot control this; (2) key players will set the agenda and should be the focus of regulation; (3) regulation should focus on discrete risks and harms rather than processes or structures. Much later may come the realisation that as governments stood by a few large corporations have stolen a march and have become self-governing within the sphere through contractualisation. Now it may be that AI and ML will not follow the same path as internet regulation as the risks and harms are more clearly defined than with the internet and as a result, governments will move more quickly this time. Pessimistically though the early indicators are this is not the case.

One clear systemic risk of AI and ML is the "black box" issue. This is the problem that arises when an algorithmic system makes decisions which prove extremely difficult to explain in a way that the average

person can understand. In essence while it is possible to observe incoming data (input) and outgoing data (output) in algorithmic systems, but their internal operations are not very well understood. The problem of the "black box" has been much discussed in academic (and wider) circles. One of the best-known recent discussants from a regulatory standpoint is Frank Pasquale. His 2015 book *The Black Box Society [39]* is, or was, most people's introduction to the problem from a regulatory standpoint. However, in the four years since publication there has been little development of the question of how best to regulate black boxes. In his 2018 paper The Artificial Intelligence Black Box and the Failure of Intent and Causation , [40] Yavar Bathaee was forced to conclude that the current legal-regulatory approach to the black box problem, a right to receive an explanation in cases of automated decision-making, "poses an immediate threat to intent and causation tests that appear in virtually every field of law". [41]

A further concern is reliance on ethics and soft governance. In a retreat from regulation many have proposed the answer to AI governance may be found in ethics. [42] This should be resisted on two grounds: one empirical, one normative. The normative objection is that this trend, to produce lists of desiderata of good behaviour, causes the law, including regulation, to be marginalised in the debate in favour of a focus on these soft forms of governance. This "ethics washing" leads to significant problems. Where companies have a voluntary, ethical commitment in tension with a legal, commercial duty, it is not difficult to see why compliance with the legal duty wins out. Furthermore, the data and AI profession lacks key characteristics of professions in which a soft governance approach works - there are no longstanding norms of good behaviour, no well-established methods for translating principles into practice, and no licensing body. The empirical objection may be found in the now defunct, and similar, debate around internet and data ethics in the 1980s and 1990s. Then as now there were arguments over whether ethics were to be preferred to stricter forms of governance. Indeed, if one reads James Moor's classic 1985 paper, What Is Computer Ethics? clear parallels with the current debate on AI ethics emerge as Moor observes the ethical risks: "Computers are logically malleable in that they can be shaped and molded to do any activity that can be characterized in terms of inputs, outputs and connecting logical operations ... . Because logic applies everywhere, the potential applications of computer technology appear limitless. The computer is the nearest thing we have to a universal tool. Indeed, the limits of computers are largely the limits of our own creativity." [43]

This ethical debate remained vibrant through the 1980s and 1990s but generally died out in the early 2000s as it became clear that regulation and governance was required and ethics were too soft to control a sphere of increasing sophistication and commercial value. Again, if we assume the we are roughly at 1991/92 on the internet regulation timeline the discussion of ethics is to be expected. Our experience though is that the lure of soft regulation through ethical codes of practice were a crutch for governments who did not want to set hard standards. Eventually though with contractualised regulation replacing ethics the folly of that error would become apparent.

4. The Regulation We Need

The debates on regulation of the internet and now of AI and ML swirl around issues of authority and legitimacy of different types of bodies or groups to regulate; the efficacy of different types of intervention, such as licensing, code, contracts, governmental rules; organisational structures, in particular the territorial mismatch of national governments and transnational operators; conflicts between systems, such as ethics and legal duties; motivations of different actors: corporates, legislatures, governments; and whose normative values should dominate: those of individuals making individual choices, or those of the 'market', in practice the dominant platform providers, or those of the state - but which state?

Debates also tend to be focused on the technology and its associated actors; it is relatively rare to draw on analyses of regulatory systems in other domains, or to abstract back even further to consider what a regulatory system consists of. It should be remembered that 'regulation' or indeed 'regulatory governance' are terms which carry a range of connotations in both public and academic debate.

So some preliminary definitional work is required to avoid misunderstandings. [44] First, the terms ''state'' and ''non-state'' are used throughout this section to distinguish in broad terms those regulators which have a legal mandate and those which do not - while recognizing that in practice the two are interrelated in a myriad of different types of relationship, and indeed state actors may be regulated by non-state actors. A hierarchy of state-non-state cannot be assumed. By regulation (and regulatory governance) is meant sustained and focused attempts to change the behaviour of others in order to address a collective problem or attain an identified end or ends, usually but not always through a combination of rules or norms and some means for their implementation and enforcement, which can be legal or non-legal. [45] The regulatory functions can be exercised primarily by one actor or dispersed between a number of actors within a system. The greater the dispersal and fragmentation of actors in the performance of regulation, including the definition of the problem/goals, the greater the polycentricity of the regime. A regulatory regime, system or network is a set of interrelated actors who are jointly attempting to address a particular set of problems to achieve a set of goals, its boundaries are defined by the definition of the problem being addressed, and it has some continuity over time.

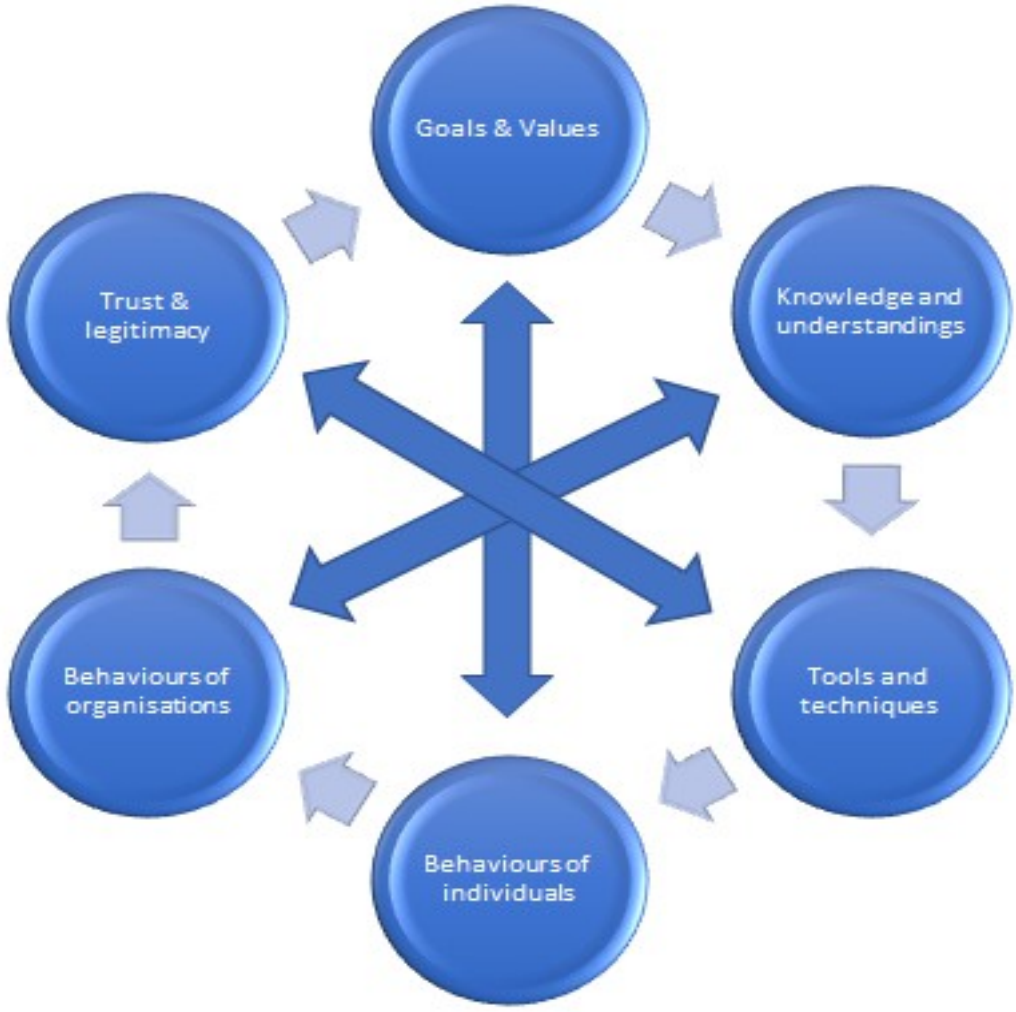
The arguments developed here draw on the decentering or polycentric analysis of regulatory systems. [46] At its conceptual core are five central notions: complexity, both conceptually and in terms of the actors and organisations involved; the fragmentation of power, capacities and responsibilities; the inevitable interdependencies between all actors within the system or network, not least regulators and regulatees; the inherent ungovernability of actors due to their ability to exercise agency and choice; and the rejection of a clear distinction between public and private in the performance of regulation. The decentred analysis thus draws attention away from individual regulatory bodies, be they at the national or global level, and emphasizes instead the multitude of actors which constitute a regulatory regime or regulatory network in a particular domain, and the interactions within and between them. [47] Moreover, it emphasises that regulatory strategies can be hybrid - combining governmental, private and other non-governmental actors, multifaceted - using a number of different strategies simultaneously or sequentially, and often indirect - including using the positional power of intermediaries or gatekeepers such as insurers, auditors, advisors and others. In the context of accountability, this aspect of the decentering analysis requires recognition of the multiple points of accountability within a regulatory regime and of the form that accountability mechanisms may have to take to be effective.

Once we understand regulation or regulatory governance conceptually as a complex, polycentric system composed of several elements, we can quickly see that in practice, regulation is usually complex, messy and highly imperfect; that addressing difficult problems involves complicated interactions of multiple people and organisations with conflicting, or at least differing, interests, understandings and values; and often requires the creation, adaptation and implementation of myriads of techniques which may or may not cut across one another. As such, it is not surprising that regulation often fails. What is surprising is that it ever succeeds at all.

So how can we stand back to analyse a regulatory system, and indeed to create a framework which can help us to design a system whilst also thinking about how different configurations of a system might work in practice?

Building out from the decentred or polycentric analysis, we should think of regulation as a particular form of social system with six key elements, all of which constantly interact to produce a dynamic system. [48]

Figure 2: Regulatory systems - an analytical framework



<

A logical starting point in theory, though not necessarily in practice, is with goals and values - what is the regulatory system trying to achieve, and which values is it trying to uphold? The standard economic justification for regulation prevalent since the 1980s is that regulation is there to correct market failures. But as we have seen, regulation has always been about more than that, or indeed not about that at all: as the histories above illustrate, it has been focused on coordination or the management of scarce resources, and / or managing risks. Regulation can also be aimed at controlling power; and/or introduced to uphold fundamental values of a particular group such as, in some societies at least, principles of equality, non-discrimination, the rule of law and the administration of justice, and very prominently in current debates, privacy.

Identifying what it is that a regulatory system is meant to be achieving can be harder than it seems, as goals and values are often poorly articulated, or inchoate, or simply conflicting, as current debates are demonstrating. It may be possible to get agreement on some of the more technical desiderata of AI and ML, such as the agreement on technical standards to enable coordination or inter-operability, or on scientific assessments of the scientific quality and robustness of different algorithms and their reliability and appropriateness to the different tasks they are being used to perform, though even here assessments

can be scientifically contested [49] even before we enter into the discussions on social interpretations of science. And different groups (both within and between countries) will and do differ on clearly value laden questions of the ethical principles relating to the sourcing, use, protection and ownership, as well as the reliability of the data in the datasets on which algorithms learn, [50] and the uses to which different modes of AI are put in different contexts. They are also highly likely to differ on far more fundamental values and in particular on the appropriate trade-offs between them, such as privacy and security, or individual rights vs those of the collective, trade-offs which become more acute when deciding how to manage a technology which provides societal benefits but also the potential for significant harm. We also know from long histories of regulating technologies (and of using technologies as instruments of regulation) that apparently technical questions cannot and should not be separated from questions relating to subjectivity, ethics or values. We may be comfortable with the outcome that the algorithms of Amazon and Netflix can produce highly differentiated results from similar data sets, [51] but once AI moves from the 'discretionary' parts of our lives, such as online purchasing of consumer goods, to the core elements of our governance, judicial decision-making, healthcare, educational and welfare systems, then the values become increasingly contested, the trade-offs more acute, and the stakes considerably higher.

To date, as noted, much of the discussion on AI has related to the role of goals and values, discussed in the language of ethics. But identifying, and agreeing, goals and values is only one element of a regulatory system: essential, but by no means sufficient. Regulation also requires people and organisations to change their behaviours - so understanding these second and third elements of any regulatory system is critical to understanding its dynamics and enhancing its effectiveness. Regulation may be directed at getting individuals to change their behaviours, frequently individuals as consumers. We need to have a highly sophisticated understanding of how and why consumers behave as they do if they are to attempt to change their behaviours, and we are only beginning to understand behaviours online, their relationship with offline behaviours, and how platform providers deliberately manipulate behaviours in the way in which advertisers have tried to do for centuries, but in ways which data and AI now enable to be more sophisticated by orders of magnitude. In contrast, in policy making, particularly where dominated by economists, the limitations of the rational actor model of consumer behaviour have been assumed for far too long, and although inroads are being made through increased use of psychology and behavioural sciences, [52] that model remains remarkably tenacious in some literatures and policy discussions. Moreover, it is not only the behaviour of consumers which is of relevance, but of all involved in regulatory systems. That would include in this case the designers of AI and, to the extent it exhibits agency, of AI itself.

Further, regulatory systems are comprised of a series of organisational actors. These may be arranged into more or less formal multi-level structures operating at the global, regional, national and/or sub-national level, or into looser multi-lateral configurations, and/or they may compete, collaborate or simply co-exist. [53] Financial regulation provides an interesting model. Following the financial crisis, the G20 created the Financial Stability Board, which sets principles of regulation which both G20 and non-G20 countries agree to implement through a cascading system of rules down to the regional and national level. There is a system to monitor implementation and to review the impacts and effectiveness. As noted, various aspects of the debate on internet regulation and now that of and ML rightly focus on organisational design. But designing and creating organisational structures is not enough. Those undertaking any regulatory functions require the necessary capacities and associated resources to undertake those functions, both material and human (funding, expertise, organisational systems and processes, the ability to learn), and social and political (power, authority, legitimacy), depending on the functions they are exercising. [54] Those who are auditing compliance require different capacities to those who are setting the rules or imposing sanctions for their breach, for example. Organisational actors

also need the motivation to use those capacities to further the orientated goals of the regulatory system, which may not necessarily align with their own interests. The 'well-intentioned, ill-informed' and 'ill-intentioned, well-informed' regulatee are familiar characters in the compliance literature, for example. [55] Furthermore, any regulatory system for AI will have to extend beyond national jurisdictions or beyond anyone company, no matter how large. So international cooperation will be critical. But it is likely that goals will be contested and interests and capacities misaligned - there is nothing unusual about that. So understanding and anticipating the dynamics of the interactions of regulators is essential both to analysing and to building and dynamically maintaining regulatory systems or networks.

Moreover, interactions between organisational actors are critical to understanding the dynamics of a regulatory regime or network. Regulation is often a process in which one set of organisations (regulators) act on another set of organisations (regulatees, or those they are regulating). The interactions between regulators and regulatees, for example the debates about how to enforce rules, how to gain compliance, are well researched. [56] But regulators, and others seeking to develop regulatory systems, also need to focus on the context of those organisations they are regulating: what is the market structure, who are the dominant players; what are their incentives; how does the market function? This includes but goes well beyond the dominant platforms and into the far reaches of those developing and deploying AI, including those in universities. We know that the internal governance and operation of those being regulated is critical to the success or failure of regulation. [57] But we also need to focus on an area which is often neglected, which is on the internal organisational dynamics of regulators themselves. [58] Productive interactions and unproductive dysfunctionalities can arise in all cases, as the long history of regulatory failures, and regulatory successes, tells us. [59] And importantly, hybridity draws our attention to the fact that large organisations can be at once regulatees, implementing standards imposed from elsewhere, and regulators themselves - developing systems to motivate and ensure compliance with others' rules as well as their own.

The fourth element of any regulatory system are the knowledge and understandings that regulators, and others, have of what it is they are regulating. [60] This comprises not only technical knowledge, based on particular epistemologies including what is seen to constitute valid knowledge, but system knowledge of the context regulation is operating in. This element is particularly important where the focus of a regulatory regime is on managing risks. For example, it was largely due to failures to understand the actual operation of the financial markets which led to the financial crisis. [61] We have argued above that it has been the dominance of a particular way of 'seeing' and understanding the internet which has led to 'not seeing' the structural and systemic role which it plays, and therefore the risks that it poses and impacts that it can have. In the context of AI, if we are not to make the same mistakes, then deep engagement with those who are developing AI is essential. AI is, in significant part, a socially created technical system of calculative models or devices [62] - understanding the key concepts which are being deployed, the decision rules and selection criteria for what is included and what is excluded, and the validation criteria are all critical to developing a sophisticated understanding of the technology. [63] But we also need to understand the market and other contexts in which AI is being used and deployed. As the history of telecoms and internet regulation above illustrates, it is the cognitive framing of the technology and of the nature of the 'problem' it presents, combined with deep-rooted political philosophies on the legitimate role of the state which can lead to failures to both see and accept the need for government-led structural interventions.

How regulators perceive the world they operate in and the problems they have to address (and the acceptability of any solutions they may devise) are thus key to the fifth element, which is the design and operation of regulatory tools and techniques . It is this element which is usually where debates are most focussed. When should regulation be applied: at the stage of entry of a technology (approval); and/or to the manner of its use; and/or to providing compensation if it causes harm? What role can economic

techniques (price controls, taxes) play in changing behaviours? When should regulation focus on market structures and when is it sufficient to focus on firms' or consumers' behaviour? To what extent can we rely on the parity of contracting power within a market to address the problems and when might we need to regulate those contracts to ensure collective values are upheld and/ or power asymmetries addressed? When is it appropriate and possible to use 'nudge' techniques to change behaviour, or when are rules required? If rules are required, what should be their legal status? What should be their form: should regulators use standards, rules and/or principles and in what combination and in what contexts? Can and should we use regulation by technology, including AI? What are the most effective ways of gaining compliance? What sanctions should be imposed for breach and by whom? How and when should regulation be evaluated, and, again, by whom?

All these are very familiar questions to anyone used to thinking about regulation. [64] In considering how to answer them in the case of AI, and more particularly the use of AI in different contexts, we can learn from regulatory systems which have preceded it, both in from closely related areas such telecommunications and the internet, but in particular from other areas of regulation of risk and technologies, including calculative technologies such as financial models, and ethically contested areas such as genetic engineering in plants and humans. [65] In general, where risks fall on individuals and can be compensated for, then regulation is comprised of ex post liability regimes (e.g. negligence, contract, statutory regimes of product liability or food safety) which may or may not be supplemented by state or non-state based oversight and enforcement. This is the 'develop, deploy, regulate' model noted above. Where the risks of individual harm are such that it is considered a more precautionary approach is required, then those making such products or delivering such services may be required to have specific authorisation (financial services, for example) (develop, regulate, deploy). Where the risks fall on individual but are non-compensatable or 'deep regret', such as threat to life, then there more onerous ex ante requirements are imposed through licensing and ongoing monitoring and enforcement, and standards of consent to be exposed to the risk are higher (such as licensing of pharmaceuticals and consent to medical treatment). On the other hand, where risks are systemic even though still compensatable, then systems which rely on individual consent to the risk are inadequate, and again ex ante regulation is required as well as ex post remediation. Payment systems are a good example, and indeed the argument above is that the internet should have been seen as posing such systemic risks. Finally, where the risks are systemic and the harms non-compensatable or non-remediatable, then regulatory regimes are ex ante and usually highly restrictive, with regulation of development imposed as well as deployment, requiring extensive trialling and close regulation: such as the use of genetically modified organisms, stem cell treatments, or in the case of aviation or nuclear power. We return to this element below.

Finally, but most importantly, is the element of trust and legitimacy, and thus accountability. All regulators need a political and a social licence to act, whoever they are. The need for trust and legitimacy is as critical for a company operating under a self-imposed self-regulatory regime as it is for a national regulator or a transnational organisation imposing regulatory norms on others, including those which are unlinked to governments such as IETF and W3C, and those acting as regulators need to work proactively to create that legitimacy. regulatory system needs to be trusted and perceived as legitimate by a critical number of legitimacy communities in order for it to function, even if it is not universally seen as legitimate. These include those who are relying on it to protect or support them, as citizens or consumers, and those it is seeking to regulate. There are four core legitimacy and accountability demands which are usually made by such legitimacy communities, in different combinations, and which we can see echoed in the debates on internet regulation and now of AI: claims based on constitutional values (rule of law, procedural fairness, accountability); claims based on normative values (attainment of justice, ethics, sustainability and so forth); claims based on democratic values (dialogue, participation, representation,

accountability (again); and claims based on functional performance (such as effectiveness, expertise, efficiency). But the demands of each group or legitimacy community can pull in different directions, so maintaining trust and legitimacy is an ongoing task requiring transparency and continual engagement [66] and is particularly difficult in the context of managing risks.

This systems framework will not be attractive to those seeking set menus of solutions. It deliberately breaks away from adopting a 'toolbox' approach to designing regulation and associated accountability mechanisms which has been so prevalent for so long. Instead it provides a framework for enabling us to think systematically about each part of any regulatory system. It is important to understand also that any one system does not exist in isolation but frequently operates in interaction with other systems, in important and complex ways. Nevertheless it is a framework for designing regulatory systems, for understanding their dynamics, for analysing deep rooted causes of failures, for thinking through the potential impacts of changes in any part of the system, and for helping us understand how each element would need to operate and be accountable if regulation is to be both effective and trusted.

5. The Regulatory Action

It is too late for us to put AI and ML back into a box. It may be that in areas which are already heavily regulated, such as medical products and applications, then the use of AI or ML will require prior regulatory approvals. But even if they are caught in an existing regulatory net, there is little evidence that regulators have the necessary capacity properly to evaluate all the actual and potential uses of AI in their regulatory domains. Asymmetries of knowledge and skills are amplified in the highly technical area of AI. And we can see from current debates in multiple areas that existing regulatory systems simply do not capture the use of AI and ML, allowing them to operate on the edges of existing regulatory perimeters or escape them entirely. The current domination by corporate players means that AI is likely to be developed and marketed in a similar fashion to internet products and online services. There will be both a consumer market and a commercial market for products and services and in all likelihood they will be regulated, if at all, in piecemeal fashion. But as noted, AI is also being rapidly used by governments themselves to deliver welfare provision (education, healthcare) [67] and exercise core functions of government (policing, justice) and indeed in the function of regulation itself. [68] Furthermore, we know from the long histories of regulation in other areas that companies, government bodies, NGOs and others will seek to reassure governments and consumers that formal regulation is not required; that they can and will act ethically and adopt such devices as codes and ethics boards to demonstrate that commitment. However, we also know from history that a commitment to ethics is important, indeed essential, for effective regulation, but is rarely sufficient on its own in the absence of very specific conditions which rarely exist in a highly competitive market.

However, the current debate around AI ethics, fuelled by academics [69] has become the focus of both governmental and intergovernmental discourse. The UK government has responded to nascent AI and ML challenges by issuing general guidance on Understanding *Artificial Intelligence Ethics and Safety [70]* which requires anyone in the public sector involved in the design, production, and deployment of an AI project to consider ethical considerations which arise at every stage of their project. There is also sector specific guidance such as the Code of conduct for data-driven health and care technology issued at the same time [71] and which also employs an ethical framework. The focus on ethics is so strong that the new advisory body for AI in the UK, has ethics in its title. The Centre for Data Ethics and Innovation was set up to "identify how we can enjoy to the full the potential benefits of data-driven technology within the ethical and social constraints of a liberal democratic society." [72] At a European level the High-Level Expert Group on Artificial Intelligence, which as an aside had four lawyers and seven philosophers/ethicists, also focused on ethical standards over legal/regulatory

ones. [73] Although their trustworthy AI framework requires that AI should be lawful, this is simply a requirement that it "complies with all applicable laws and regulations". [74] Thus lawful AI means AI which meets the general requirements of law and regulation, there is no indication nor intent to suggest specific regulation for AI, or indeed to amend those laws and regulations to accommodate the very particular challenges AI poses.

If we are to seek to control the way corporates and governments use AI and ML, then ethics cannot substitute for law or other forms of formal regulation. Unlike academic proposals, new regulatory regimes rarely land newly minted, in perfect form and onto a blank canvas: they are always situated in an existing context often thick with existing norms and rules, with existing organisational structures, and amongst actors with particular behaviours, cognitive frameworks, capacities and motivations. This paper is at a minimum a call for lawyers, and for regulators more generally, to get involved in the debate and to drive the discussion on from ethical frameworks to legal/regulatory frameworks and how they might be designed, but it is also a call to recognise the dynamics and composition of any regulatory governance system, even when introducing relatively minor changes, let alone seeking to design more radical approaches. It is also a call to adopt a more differentiated approach to the different types of risk that different modalities or technologies of AI and ML can pose when developed and used by different actors in different contexts, and to continually test our understandings of the risk/benefit trade-offs involved. As the financial crisis demonstrated, if we build our regulatory system on the basis of a fundamental misunderstanding of the dynamics of the system we are seeking to regulate, including its technologies, the result can be disastrous.

It is not within the scope of this paper to design a new framework for regulating AI. However, we suggest that whilst it is important that the overall regime for AI regulation is coherent, it does not need to, and indeed should not, operate in isolation from existing regulatory regimes. Where an activity is already regulated under a specific regulatory regime, then the use of AI in the development or deployment of that activity, for example in the development of medical treatments or devices, is captured within the perimeter of an existing regulatory regime. Those regulators need to develop norms for the use of AI, and quickly, but the mechanism is there. In areas where AI is being used where there is currently no regulation or it falls at the edges of existing regimes, then we will have to rely on existing legal principles. Reed, for example, argues the application of general legal principles, in particular human rights, can provide an interim framework for the general regulation of AI. But there are limits to the degree to which general legal frameworks, such as the law of negligence, may adequately be used to manage risks or attribute liability in ways which achieve overall societal goals. [75]

There are also risks that if we leave it to existing regimes to respond then we will end up not with a coherent system but with patchwork regulation in which there are overlaps and underlaps, with conflicting goals and logics. Moreover, enforcement systems which rely on individuals to bring cases to court can be less effective than public enforcement systems for very well documented reasons. [76] Coordination both in design and operation is required. However, we do not need to do nothing whilst new integrated systems are being developed. Furthermore, using the quite familiar framework of risk regulation outlined above to analyse what types of risk particular uses of AI is posing in which contexts could be a highly productive way to begin to develop regulatory regimes which are appropriately tailored to its use.

In his recent article, Reed in effect takes this risk-based approach to explore how liability for decisions made using AI could or should be attributed, at least as an interim measure whilst more tailored regimes are developed. He proposes that liability should be attributed using principles of transparency about the reasoning method being used. However, as he notes, we need to distinguish between ex ante transparency, where the decision-making process can be explained in advance of the AI being used, and ex post transparency, 'where the decision-making process is not known in advance but can be

discovered retrospectively by testing the AI's performance in the same circumstances. Any law mandating transparency needs to make it clear which kind of transparency is required. [77] Importantly, only some algorithmic methods lend themselves to ex ante transparency, notably those relying on decision trees. Here the reasoning can be set out in advance. However, in the case of other algorithmic technologies, such as neural networks, the machine is learning as it processes the data and it is not possible to set out the reasoning in advance. It is also not possible, or at least not easy, to explain the reasoning ex post. [78] Requiring ex ante transparency would in effect prohibit the use of that particular technology, even where it may produce a superior result. Nonetheless it is possible to test neural network technologies, for example, for reliability and replicability, and to examine the process by which the particular algorithm was developed including the data sets, training methods and testing processes. Either no transparency or e x post transparency should be sufficient, he argues where the harm falls on individuals and is compensatable (e.g. autonomous vehicles). However, where there is no clear societal benefit and harms are systemic and non-compensatable, such as breaches of human rights, then only systems in which ex ante transparency is possible should be permitted. [79] Furthermore, simply demanding transparency is likely to be effective without considering who the information is being conveyed to, and their ability to understand it. [80] Numerous examples exist of disclosure requirements which end up baffling consumers as they prioritise comprehensiveness over comprehensibility.

This structured way of using a risk/benefit calculus for analysing what form of transparency should be required takes us into the very familiar territory, for regulationists, of the risk-based regulation of new technologies, outlined above. It thus helps take us towards a more systematic development of a regulatory regime for AI and ML. We can see steps being taken in this direction within the EU with respect to data. Already under the General Data Protection Regulation [81] data controllers are required to process data in a transparent manner and to give an explanation of processes used in data profiling. [82] With respect to AI, as a first step annual transparency reports from AI developers could prove this interim solution until more formal regulation is developed. This is in line with the recommendation from the Communications and Digital Committee of the House of Lords that "data controllers and data processors should be required to publish an annual data transparency statement" [83] and to the recent consultation of the UK Information Commissioner on guidance on explaining AI based decisions. [84]

But transparency can only go part of the way; we need a more robust, holistic and coherent system for regulating the development and use of AI and ML. How we design, create, and operate those regulatory systems will be critical. If we allow the regulation and governance models for AI to drift for the next 5-10 years, as happened with the internet between 1995-2010, we will find ourselves in the same position in 20 years' time with respect to AI as we do now with regard to online content and activity: at the mercy of a small number of companies who regulate the market and activity through private contractual ordering and (mostly) outside the direct control and influence of public regulators including states.

* Professor of Law and Strategic Director of Innovation, LSE.

**Professor of Law, LSE.

[1] U. Beck, Risk Society, Towards a New Modernity (Sage, 1992); A. Giddens, Consequences of Modernity (Polity, 1990); A. Giddens, Modernity and Self-Identity: Self and Society in the Late Modern Age (CUP, 1991).

[2] However, it should be noted that only basic safety-specific forms of regulation applied such as speed limits, warnings, registration details and weight limits (steam powered locomotives were damaging roads). There were no rules on other harmful aspects of motorised vehicles such as noise, lighting etc.

[3] The Broadcasting Fairness Doctrine, Congressional Digest : 227,228,256. October 1987.

[4] A. Thierer, 'Unnatural Monopoly: Critical Moments in the Development of the Bell System Monopoly,' The Cato Journal , Fall 1994.

[5] Thierer, n.4.

[6] The United States it must be noted was not the only jurisdiction to subscribe to the natural monopoly theory which presumed that redundant telephone infrastructure was economically inefficient and that monopoly power could simply be tempered through regulation. Most European states created a state utility (monopoly) supplier of telecommunications services.

[7] See the Centre for Connected and Autonomous Vehicles and the Law Commission project Automated Vehicles .

[8] A vibrant literature developed in the 1990s around disruption and governance. On disruption see N. Negroponte, Being Digital (Knopf, 1995); M. Castells, The Rise of the Network Society (Wiley-Blackwell, 1996); C. Sunstein Republic.com (Princeton UP, 2001). On governance see H. H. Perritt Jr., 'Cyberspace Self-Government: Town-Hall Democracy or Rediscovered Royalism?' 12 Berkeley Technology Law Journal 413 (1997); J.R. Reidenberg, 'Lex Informatica: The Formulation of Information Policy Rules through Technology' 76 Texas Law Review 553 (1997-1998); F.H. Easterbrook, 'Cyberspace and the Law of the Horse' (1996) University of Chicago Legal Forum 207; L. Lessig, 'The Law of the Horse: What Cyberlaw Might Teach' 113 Harvard Law Review 501 (1999).

[9] Local loop unbundling began in the 1990s and gathered regulatory pace in the UK in 2000 when, in anticipation of the Regulation on Local Loop Unbundling (EC/2887/2000) a new condition 83 was inserted into BTs Telecommunication Act Licence which set out the co-location products BT were required to offer, the conditions which applied to the supply of these products and unbundled loops. At the same time the Commission opened investigations into a number of "bundled" products offered by incumbent telecoms providers including France Telecom/Wanadoo. See also speech of Pierre-Andre Buigues, 'European Policy on Local Loop Unbundling: Competition Law Background and Problems of Implementation' available from: https://ec.europa.eu/competition/speeches/text/sp2001_043_en.pdf .

[10] J.P. Barlow, A Declaration of Independence for Cyberspace , < https://www.eff.org/cyberspace-independence >.

[11] D.R Johnson & D.G Post 'Law and Borders - The Rise of Law in Cyberspace' (1996) 48 Stanford Law Review 1367.

[12] See further C. Reed & A. Murray, Rethinking the Jurisprudence of Cyberspace (Edward Elgar, 2018), 6-7.

[13] Perritt Jr., n.8 above.

[14] Basic Books, 1999.

[15] 'Against Cyberanarchy' 65 University of Chicago Law Review 1199 (1998).

[16] 'Network Neutrality, Broadband Discrimination' 2 Journal of Telecommunications and High Technology Law 141 (2003).

[17] A. Murray, 'Symbiotic Regulation' 26 John Marshall Journal of Computer & Information Law 207 (2008). See also A. Murray, The Regulation of Cyberspace: Control in the Online Environment (Routledge, 2006).

[18] I. Brown & C. Marsden, Regulating Code: Good Governance and Better Regulation in the Information Age (MIT Press, 2013).

[19] C. Reed, Making Laws for Cyberspace (OUP, 2012).

[20] Reed & Murray, n.12 above.

[21] Such as the Copyright and Related Rights in the Information Society (InfoSoc) Directive, Dir.2001/29/EC.

[22] See the Framework Decision 2008/913/JHA of 28 November 2008 on combating certain forms and expressions of racism and xenophobia by means of criminal law.

[23] Intermediary or platform responsibility/liability focuses on network gatekeepers. These are, in the words of Karine Barzilai-Nahon who is credited with creating this theory, "an entity (people, organizations, or governments) that has the discretion to exercise gatekeeping through a gatekeeping mechanism in networks and can choose the extent to which to exercise it contingent upon the gated standing." From K. Barzilai-Nahon, 'Toward a theory of network gatekeeping: A framework for exploring information control' 59 Journal of the American Society for Information Science and Technology 1493 (2008), 1497. In intermediary or platform responsibility/liability the regulator directs their interventions to these gatekeepers using their ability to control information flows to control the wider populace.

[24] For a discussion of the particular role of what she calls Internet Information Gatekeepers see E. Laidlaw, Regulating Speech in Cyberspace: Gatekeepers, Human Rights and Corporate Responsibility (CUP, 2015). See also Reed & Murray, n.12 above at 5.1.

[25] See eg the German Network Enforcement Act or NetzDG Law or the UK Online Harms White Paper. For further discussion on the relationship between regulators and platforms see Laidlaw, n.18 above, at ch.6 and J. Van Dijck, T. Poell & M. De Waal, The Platform Society (OUP, 2018), ch.7.

[26] L. Lessig, Code and Other Laws of Cyberspace , n.14 above, 53-4.

[27] S. Wachter and B.Mittelstad, "A Right to Reasonable Inferences: Re-thinking Data Protection Law in the Age of Big Data and AI", (2019) Columbia Business Law Review 494.

[28] House of Commons Digital, Culture, Media and Sport Committee, "Disinformation and 'fake news': Final Report" HC 1791 (2019)

[29] B. Casey, "Amoral Machines, or: How Roboticists Can Learn to Stop Worrying and Love the Law", 111 Northwestern University Law Review 231 (2017)

[30] J. Grimmelmann, "There's No Such Thing as a Computer-authored Work" (2016) 39 Columbia Journal of Law & Arts 403; M.E. Kaminski, 'Authorship, Disrupted: AI Authors in Copyright and First Amendment Law' 51 UC Davis Law Review 589 (2017-2018).

[31] High-level Expert Group on Artificial Intelligence, Ethics Guidelines for Trustworthy AI: https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai .

[32] HM Government, Code of conduct for data-driven health and care technology , July 2019: https://www.gov.uk/government/publications/code-of-conduct-for-data-driven-health-and-care-technology/initial-code-of-conduct-for-data-driven-health-and-care-technology .

[33] M.U. Scherer, 'Regulating Artificial Intelligence Systems: Risks, Challenges, Competencies, and Strategies' 29 Harvard Journal of Law & Technology 353 (2016).

[34] A. Tutt, 'An FDA for Algorithms' 69 Administrative Law Review 83 (2017)

[35] O. Erdelyi & J. Goldsmith, 'Regulating Artificial Intelligence Proposal for a Global Solution', in AAAI/ACM Conference on Artificial Intelligence , Ethics and Society (2018).

[36] See G. Gurkaynak, I. Yilmaz & G. Haksever, 'Stifling artificial intelligence: Human perils' [2016] 32 Computer Law and Security Review 749.

[37] Discussed by P. Nemitz, 'Constitutional democracy and technology in the age of artificial intelligence' Philosophical Transactions of the Royal Society A 376 20180089 (2018).

[38] See C. Cath et al, 'Artificial Intelligence and the "Good Society": the US, EU, and UK approach', 24 Science and Engineering Ethics 505 (2018).

[39] Harvard UP 2015.

[40] 31 Harvard Journal of Law & Technology 889 (2018).

[41] Bathaee, n.40, 938.

[42] See e.g. M. Taddeo and L. Floridi, "How AI can be a force for good", Science 361, 751 (2018); C. Cath, "Governing artificial intelligence: ethical, legal and technical opportunities and challenges", Philosophical Transactions of the Royal Society A 376: 20180080 (2018).

[43] J. Moor, "What Is Computer Ethics?" Metaphilosophy 16.4 (1985): 266

[44] J. Black, 'Constructing and contesting legitimacy and accountability in polycentric regulatory regimes' (2008) 2 Regulation & Governance 137.

[45] For review of differing definitions see C. Koop and M. Lodge, 'What is Regulation? An Inter-Disciplinary Concept Analysis' (2017) 11 Regulation and Governance 95; see also K. Yeung, 'Algorithmic Regulation: A Critical Interrogation' (2018) 12 Regulation and Governance 505.

[46] J. Black, 'Decentring Regulation: Understanding Regulation and Self-Regulation in a 'Post-Regulatory' World (2002) Current Legal Problems 102.

[47] B. Eberlein, K.W. Abbot, J.Black, E. Meidinger, & S. Wood, 'Transnational Business Governance Interactions: Conceptualizations and Framework for Analysis' (2014) 8(1) Regulation and Governance 1.

[48] This framework draws substantially on, and is a development of, a succession of previous work including J. Black, 'Learning from Regulatory Disasters' (2014) 10(3) Policy Quarterly 3; J. Black, 'Reconceiving Financial Markets - From the Economic to the Social' (2013) 14(2) Corporate Law Studies 401; J. Black, J, 'Paradoxes and Failures: 'New Governance' Techniques and the Financial Crisis' (2012) 75(6) Modern Law Review 1038; and R. Baldwin and J. Black, 'Really Responsive Regulation' (2008) 71(1) Modern Law Review 59.

[49] P. Domingos, The Master Algorithm: How the Quest for the Ultimate Learning Machine Will Remake Our World (Basic Books, 2015).

[50] On which see eg L. Gitelman (ed), 'Raw Data' is an Oxymoron (MIT Press, 2013).

[51] Domingos, n. 49 above.

[52] See e.g. C. Sunstein, How Change Happens (MIT Press, 2019); M. Leiser, 'The problem with 'dots': questioning the role of rationality in the online environment' (2016) 30 International Review of Law, Computers & Technology 191.

[53] See Eberlein et al, n.47 above.

[54] J. Black, 'Enrolling Actors in Regulatory Systems' (2003) Public Law 63; for variants see eg C. Hood and H. Margetts, The Tools of Government in the Digital Age (Palgrave Macmillan, 2007).

[55] For review see R. Baldwin, M. Lodge and M. Cave, Understanding Regulation (OUP, 2013).

[56] Baldwin, Lodge and Cave, n.55.

[57] E.g. in the context of the financial crisis see Senior Supervisors Group, Risk Management Lessons from the Global Banking Crisis of 2008 (October 2009); OECD, Corporate Governance and the Financial Crisis: Key Findings and Main Messages' (June 2009).

[58] See for example, National Commission on the BP Deepwater Horizon Oil Spill and Offshore Drilling, Deepwater - the Gulf Oil Disaster and the Future of Offshore Drilling, Report to the President (2011).

[59] On the considerable literature on 'enforced self-regulation' or 'meta-regulation' see C. Coglianese and E. Mendelson. 'Meta-Regulation and Self-Regulation', in R. Baldwin, M. Cave and M. Lodge (eds), Oxford Handbook of Regulation (OUP, 2010).

[60] See for example, Scott, JC, Seeing Like a State: How Certain Schemes to Improve the Human Condition Have Failed (Yale UP, 1999).

[61] See for example Financial Services Authority, "The Turner Review: A Regulatory Response to the Global Banking Crisis" (FSA, 2009).

[62] For the development of this notion in a different setting see M. Callon and F. Muniesa, "Economic Markets as Calculative Collective Devices" (2005) 26(8) Organization Studies 1229.

[63] For comparisons between different types of methodologies see Domingos, n.49, above.

[64] See for example Baldwin, Lodge and Cave, n.55 above; R. Brownsword, Law, Society and Technology: Reimagining the Regulatory Environment (Routledge, 2019).

[65] In the UK debates, for example, the UK Human Genetics and Embryology Authority is often referred to as an example of a regulator which embeds ethical principles in its licensing decisions.

[66] See Black, n.44 above.

[67] See Reform, Thinking on its own: AI in the NHS (2018): https://reform.uk/research/thinking-its-own-ai-nhs ; L. Rouhiainen, ' How AI and Data Could Personalize Higher Education' Harvard Business Review 14 October 2019: https://hbr.org/2019/10/how-ai-and-data-could-personalize-higher-education .

[68] See Yeung, n.45 above; M. Hildebrandt, 'Algorithmic regulation and the rule of law', Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences , 376 (2128):20170355

[69] D.J. Gunkel, Critical Perspectives on AI, Robots, and Ethics (MIT Press, 2012); C. Allen, W. Wallach & I. Smit, 'Why Machine Ethics?' (2006) Intelligent Systems, IEEE 21(4); Mittelstadt, Daniel, Allo, Taddeo, Wachter & Floridi, 'The ethics of algorithms: Mapping the debate' (2016) (2) Big Data & Society 1.

[70] Government Digital Service and Office For Artificial Intelligence, June 2019: https://www.gov.uk/guidance/understanding-artificial-intelligence-ethics-and-safety

[71] Above, n.32.

[72] Centre for Data Ethics and Innovation, Introduction to the Centre for Data Ethics and Innovation : https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/787205/CDEI_Introduction-booklet.pdf .

[73] Above n.31.

[74] Above n.31, 5.

[75] C. Reed, 'How should we regulate artificial intelligence?' Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences , 376 (2128) 20170360, 2018.

[76] See Baldwin, Lodge and Cave n 55 above for a full discussion.

[77] Baldwin, Lodge and Cave n 55.

[78] For a fuller discussion see Domingos, n.49 above.

[79] Reed, n.75, 6-8.

[80] Reed, n.75, 7.

[81] Regulation (EU) 2016/679.

[82] See Recital 71. There is much controversy about this provision. See Wachter, Mittlestadt & Floridi, 'Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation' 7 International Data Privacy Law , Volume, 76 (2017); Edwards & Veale, 'Slave to the Algorithm? Why a "Right to an Explanation" Is Probably Not the Remedy You Are Looking For' 16 Duke Law & Technology Review 18 (2017). It should be noted the Wachter, Mittlestadt & Floridi paper is itself controversial see O. Williams, 'How Big Tech funds the debate on AI ethics' New Statesman 6 June 2019: https://www.newstatesman.com/science-tech/technology/2019/06/how-big-tech-funds-debate-ai-ethics .

[83] House of Lords Select Committee on Communications, Regulating in a Digital World , HL Paper 299 (2019), [81].

[84] https://ico.org.uk/about-the-ico/ico-and-stakeholder-consultations/ico-and-the-turing-consultation-on-explaining-ai-decisions-guidance/ .