Available online at www.sciencedirect.com

ScienceDirect

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

Computer Law & Security Review

A European Agency for Artificial Intelligence: Protecting fundamental rights and ethical values

Bernd Carsten Stahl[a,*], *Rowena Rodrigues*[b], *Nicole Santiago*[b], *Kevin Macnish*[c]

[a] *De Montfort University, Leicester, UK*
[b] *Trilateral Research, London, UK*
[c] *Sopra Steria, Hemel Hempstead, UK*

ARTICLE INFO

Keywords:
AI ethics
European artificial intelligence board
European agency
Artificial intelligence
Human rights
AI Act
European regulation on AI

ABSTRACT

Following years of intensive international debate of the ethical and human rights implications of artificial intelligence (AI)-related technologies, there are numerous proposals to legislate and regulate these technologies. One aspect of possible legislative frameworks for AI is the creation of a new regulator or other body with the remit to provide oversight of AI. This article reviews the ethical and human rights challenges as well as proposed mitigation strategies, in order to the discuss how a regulatory body might be designed to address these challenges. It focuses on a particular form that a new body might take, more specifically on a potential European Agency for AI. Based on a multi-step methodology of stakeholder interaction, the article proposes a terms of reference for such an Agency and discusses the characteristics it would need to display to ensure that it could adequately engage with current and future ethical and human rights challenges arising from the development, deployment and use of AI. This proposal is then contrasted with the proposed European Artificial Intelligence Board included in the draft European Regulation on AI (the AI Act).

1. Introduction

Artificial intelligence (AI) brings significant economic and social benefits. Many companies and governments are positioning themselves to harness the benefits. Most industrialised countries are developing policies to foster the uptake of AI in different sectors. The flipside of AI's expected benefits is equally widely discussed. AI raises several ethical and human rights concerns. It can lead to biases and discrimination, raise privacy and security concerns, exacerbate economic and other inequalities, lead to political and power imbalances and restructure human interaction, thoughts and lives. There is an extensive debate about how societies can retain the benefits of AI while identifying, addressing, and mitigating its disadvantages. Here, an overarching approach is called for that promotes the use of AI for human flourishing. Such an approach requires attention to various aspects, particularly respect for human rights and ethical values. It calls for education and awareness raising which translate into responsible activities at the individual, professional and corporate level. It furthermore calls for the creation of appropriate governance struc-

* Corresponding author at: De Montfort University, The Gateway, Leicester, LE19BH, UK
E-mail addresses: bstahl@dmu.ac.uk (B.C. Stahl), rowena.rodrigues@trilateralresearch.com (R. Rodrigues), nicole.santiago@trilateralresearch.com (N. Santiago), kevin.macnish@soprasteria.com (K. Macnish).

https://doi.org/10.1016/j.clsr.2022.105661

tures to regulate, provide guidance, promote incentives for responsible behaviour, and collect good practice. Legislation, while certainly not being a panacea, plays a key role in establishing an ecosystem for AI that promotes human flourishing.

Here, we focus on one specific regulatory and governance aspect, i.e., an Agency that sits at the heart of a legislative framework for AI. More specifically, we investigate what such an Agency could look like in the context of the European Union (EU). We choose the EU as an example because it exemplifies many of the challenges and benefits that such an Agency could face or generate. AI technologies are not confined to a geographical area or jurisdiction and many AI applications have a global character. At the same time, legislation is typically linked to jurisdictional boundaries. The EU as a regional body has legislative power across its Member States which raises interesting questions concerning the principle of subsidiarity, i.e., the question which issues should be dealt with at the Member State level and which ones should be addressed on the European level. Further, there have been many calls for a new regulator for AI, big data, and robotics, including calls from the EU Parliament and some Member States for the creation of an EU Agency for AI. This article answers the research question: *how should a European Agency for AI be designed, to promote human flourishing?* This question is clearly of practical and policy interest, e.g., as evident in the proposal of an AI Board in the legislative proposal for AI (see (European Commission, 2021a, b)) may well mandate the creation of such an agency, but the details of its design are likely to require further discussion, to which this article aims to contribute. The question is furthermore of academic interest from a variety of perspectives. The article also contributes to the discussions on governance of digital technologies more broadly by showing how various governance structures might be aligned and optimised for addressing AI concerns with the help of a new body.

The article is based on legal and conceptual analyses of AI, technology law and ethical and human rights issues. It takes into account the current institutional environment of the EU. It is furthermore informed by desktop research, a focus group and consultation with experts undertaken to explore the feasibility of an EU Agency for AI. Based on these insights, we present several features that such an Agency could have. We critically discuss and evaluate these features. We do so in the context of a discussion of legislative interventions into the AI innovation ecosystem (Rodrigues, Santiago, et al., 2020, p. 6).

The article proceeds as follows. The first section discusses the ethical and human rights challenges that AI raises. This is followed by a review of proposals for the creation of new regulators, agencies or other institutions for AI. We then outline the methodology underpinning the research. This leads to a discussion of findings which contribute to our proposals of the features of a European AI Agency. The art article then analyses the proposal for European Artificial Intelligence Board outlined in the European AI Regulation/AI Act (European Commission, 2021a) and makes recommendations for its improvements.The conclusion reviews the findings, highlights fields of further research and confirms the article's contribution to knowledge.

2. AI and its consequences

The scope of a statutory body overseeing AI will to a large extent be determined by the definition of the term AI. Despite its broad use, there is no clear and universally accepted definition. Some see AI as a branch of computer science, the branch that "studies the properties of intelligence by synthesizing intelligence" (Stone et al., 2016, p. 13). This raises the question of what constitutes or comes within the scope of intelligence. Raj et al. propose that AI requires a machine to perform cognitive tasks associated with human minds, including "perceiving, reasoning, learning, interacting with the environment, problem solving, decision-making, and even demonstrating creativity" (Rai et al., 2019, p. iii). The OECD suggests that an AI system can "make predictions, recommendations, or decisions influencing real or virtual environments" (OECD, 2019, p. 7).

The difficulty of determining in detail what falls under these definitions, what counts as intelligence and at what point a system should be considered as AI has led to the embrace of rather broad definitions, in particular in policy contexts Hall and Pesenti (2017, p. 8)., for example, in a report for the UK government suggest that AI describes "a set of advanced general purpose digital technologies that enable machines to do highly complex tasks effectively." The European Commission (2020a, p. 2), in a White Paper underpinning the EU legislation on AI proposed that "AI is a collection of technologies that combine data, algorithms and computing power." This definition may be too inclusive, but it is caused by the difficulty of clearly delineating the term. This difficulty is not confined to policymakers. An extensive study of the academic and other AI-related discourses undertaken by the publisher Elsevier (2018) shows that AI may be better understood as a set of interrelated discourses rather than an easily defined topic.

For the purpose of this article, we have used the definition of AI from the European Commission's High-level expert group on AI (AI HLEG):

> Artificial intelligence (AI) systems are software (and possibly also hardware) systems designed by humans that, given a complex goal, act in the physical or digital dimension by perceiving their environment through data acquisition, interpreting the collected structured or unstructured data, reasoning on the knowledge, or processing the information, derived from this data and deciding the best action(s) to take to achieve the given goal. AI systems can either use symbolic rules or learn a numeric model, and they can also adapt their behaviour by analysing how the environment is affected by their previous actions. As a scientific discipline, AI includes several approaches and techniques, such as machine learning (of which deep learning and reinforcement learning are specific examples), machine reasoning (which includes planning, scheduling, knowledge representation and reasoning, search, and optimization), and robotics (which includes control, perception, sensors and actuators, as well as the integration of all other techniques into cyber-physical systems) (AI HLEG, 2019a).

However, we do not offer or propose this as an authoritative definition of AI because we agree that AI can be seen as a "catchphrase for a cluster of technologies embedded in social systems" (Latonero, 2018, p. 8). This cluster of technology notably includes machine learning and those techniques that support it, such as machine learning, deep learning, artificial neural networks etc. (LeCun et al., 2015). We expect the definition will be clarified in the anticipated European AI regulatory framework and this should be followed by the Agency. Critiques of the definition put forward in the draft EU AI regulation (European Commission, 2021a) illustrate the ongoing challenge of constructing a clear definition in the policy context (Smuha et al., 2021). Furthermore, the definition will likely require periodic updating to ensure relevancy, taking into account the dynamic nature and applications of AI. We propose the Agency itself should be instrumental in further assessing and clarifying the definition for usefulness and applicability.

2.1. *Ethical and human rights challenges of AI*

This article explores one regulatory option for AI, namely the creation of a new regulatory agency. A starting point of the discussion of ethical and human rights issues of AI can be the benefits that AI promises. These are significant and affect many aspects of modern life. The most frequently cited benefits arise from AI's potential to automate, rationalise and improve economic processes and products by making better use of data. The economic benefits are difficult to quantify but are broadly believed to be enormous, which drives much of the policy debate (European Commission, 2020a House of Commons Science and Technology Committee, 2016; House of Lords, 2018). However, AI is not just about wealth creation, but can also be targeted to address specific issues. A frequently cited example is the use of AI in healthcare that has the potential to improve diagnoses and improve availability (Haque et al., 2020 Topol, 2019). Similar examples can be found in many other areas, such as AI for security (Babuta et al., 2020 Richards et al., 2020), education, (Stone et al., 2016), transport (Horizon 2020 Commission Expert Group to advise on specific ethical issues raised by driverless mobility (E03659), 2020) and many others. AI can be used to promote sustainability, both in the immediate sense of preserving the environment but also in the broader sense of building economically and socially sustainable structures as outlined by the UN's sustainable development goals (Coeckelbergh, 2020a Tjoa and Tjoa, 2016). These broader aims to use AI for desirable purposes are often called "AI for good" (Berendt, 2019 International Telecommunication Union, 2017; Taddeo and Floridi, 2018). AI can enable new ways of thinking that can help design positive futures with numerous and typically unforeseen potential ethical and other benefits.

This vista of benefits drives much AI research and policy. It is contrasted by concerns that AI can raise ethical issues and lead to infringements of human rights. Many of these concerns are anchored in specific characteristics of AI. Most current AI systems require large datasets for training and validation purposes. Where these datasets contain personal data, questions of privacy and data protection arise (EDPS, 2016, 2020). New technical capabilities also raise new data protection challenges (Veale et al., 2018). These are often closely linked to concerns about reliability (AI HLEG, 2019a), safety (BmVI, 2017) and security (AIEI Group, 2020 Brundage et al., 2018).

AI can affect individuals, for example when biases, often previously hidden in training data, get incorporated into AI models (CDEI, 2019) and lead to unfair discrimination (Access Now Policy Team, 2018). The black box nature of many AI systems contributes to this concern due to a lack of transparency and accountability (Hagendorff, 2019 Spiegelhalter, 2020). AI systems can act autonomously (Shneiderman, 2020) in the sense that they make or structure decisions about humans in ways that the recipients of these decisions do not understand and thus cannot query.

In addition to these concerns that are focused on the impact of AI on the individual, the wide-spread use of AI in modern society can have problematic social consequences, particularly for minorities and vulnerable groups (Rodrigues, 2020). The economic benefits promised by AI may exacerbate existing inequalities, raising questions of justice and distribution (Muller, 2020). There are concerns about the impact of AI on employment (Kaplan and Haenlein, 2019 Willcocks, 2020), the possibility of worker surveillance and the generalised exploitation of the majority of individuals by organisations able to harness the capabilities of AI (Zuboff, 2019). Partly facilitated by its economic impact, AI can influence and at the same time undermine many societal processes. There are significant concerns about its impact on politics in democratic societies, which are open to surveillance and manipulation (Coeckelbergh, 2020b). Further concerns are raised by AI use for military purposes (Boden, 2018 Richards et al., 2020) and more broadly about how AI can structure and shape the spaces for action that humans perceive as being open to them (Coeckelbergh, 2019a).

In addition to these immediate concerns, there are worries about future possibilities of AI systems moving beyond their current confines, becoming conscious and acquiring the ability to self-improve, leading to the possibility of a singularity (Kurzweil, 2006) where machines surpass humans' cognitive potential. While this sort of development may be impossible with current level of technological development and remains highly contested, it is important in a policy context because of a long history of these questions (LaGrandeur, 2020) and their high level of public and media interest.

So far, we have not commented on what constitutes an ethical issue. Ethics as the branch of philosophy that deals with moral questions, with questions of what is perceived to be right and wrong and how such a distinction can be made offers millennia of discourse to draw on. Different ethical theories focusing on concepts such as duty, consequences or care provide different insights into what could count as an ethical issue and why. We propose, drawing on ideas stemming from the ancient ideas of virtue ethics (Aristotle, 2007) that humans have an intrinsic desire to flourish and that a limitation of flourishing constitutes an ethical issue. These figures of thought are well-established in the philosophy of technology and ethics of computing (Bynum, 2006 Vallor, 2016). We use the term 'flourishing' here as a way to address the full breadth of obstacles that impede individuals' and groups' opportunities to fulfil their potential (Figs. 1 and 2).

The use of the term flourishing is already well-established to denote ways of dealing with ethical and social concerns related to AI The UK House of Lord report (2018, p. 125)., for example, posits that "All citizens have the right to be educated to enable them to flourish mentally, emotionally and economically alongside artificial intelligence." Responsible AI, according to Dignum (2019, p. 119) is concerned with "human responsibility for the development of intelligent systems along fundamental human principles and values, to ensure human flourishing and well-being in a sustainable world." This is seen before the background that AI technologies do not necessarily contribute to flourishing (UNESCO, 2020). Flourishing is difficult to pin down in substantive terms, as it is linked to personal needs and preferences. The content of flourishing is therefore typically defined in reference to existing social norms, cultural beliefs and "humanity's best interests" (Fjeld et al., 2020, p. 61) or the public good (EDPS, 2020). There are, however, some more specific criteria that allow determining the progress of human flourishing with regards to well-defined and recognised aims. One of these is the promotion of the UN's Sustainable development goals (European Parliament, 2020 Griggs et al., 2013). The other measure of flourishing is linked to the degree to which humans can enjoy their human rights (Latonero, 2018). A discussion of the relationship between ethical concepts such as human flourishing and the legal tradition of human rights is beyond the scope of this article. Suffice it to say that we agree with the World Economic Forum's (2019) position that ethics and human rights are complementary and mutually reinforcing. We therefore do not speculate what such flourishing would look like in detail (Wolbert et al., 2019) but assume that ethical issues such as those listed above are key amongst obstacles human flourishing and that addressing them and strengthening human rights promotes human flourishing (Table 1).

2.2. *Responses to the challenges*

There is an intensive discussion of how these challenges and other issues related to AI can be addressed. There are various policy initiatives from international organisations (Council of the European Union, 2020 OECD, 2019; UNESCO, 2020) and national and local governments. There are also numerous initiatives driven by industry and attempts to shape corporate governance in ways to be more receptive to challenges of AI (R Clarke, 2019a). Beyond corporate structures, there are attempts to develop or strengthen standardisation (IEEE, 2017) and computing professionalism (ACM, 2017). Numerous suggestions have been put forward for ethically sensitive development or testing methodologies (AIEI Group, 2020 Berendt, 2019; Dignum, 2019) and there is a quickly growing array of tools to support such efforts (Morley et al., 2019).

One dominant response to ethical issues of AI that is interlinked with many of the above options is the creation of AI ethics guidelines based on a set of principles (Fjeld et al., 2020 Jobin et al., 2019). While such principles are very useful, can have high-level political support and relevance (AI HLEG, 2019b), they have been criticised for various reasons, including their susceptibility to manipulation (Rességuier and Rodrigues, 2020) and lack of enforceability (Mittelstadt, 2019) Nemitz (2018, p. 12). makes the point strongly when he says that "The works on ethic rules for technology can be precursors of the law; they can give orientation on the possible content of legal rules. But, they cannot replace the law, as they lack democratic legitimacy and the binding nature which allows enforcement with the power of government and the judiciary." Similar sentiments can be found elsewhere, often based on the suspicion that ethical approaches can be used for ethics washing (Wagner, 2018), as an attempt by powerful players to manipulate and/or prevent regulation (Hagendorff, 2019 Ochigame, 2019). A further argument in favour of legal regulation is that many of the ethical issues of AI either have human rights implications or can be described as human rights infringements, which would suggest that legal remedies are appropriate.

We do not see ethics and legislative interventions as contradictory but as complementary ways of achieving social goals and promoting human flourishing. This is not a new insight and the need for mutually supportive interaction between ethics and the law has been remarked upon (R Clarke, 2019b. Coeckelbergh, 2019b; Vesnic-Alujevic et al., 2020; World Economic Forum, 2019). The question is thus not so much *whether* legal measures should be used to address ethical and human rights issues in AI, but *what shape* they should take to achieve this aim.

The field of potential legislation for AI is large and goes beyond the limitations of this article (Rodrigues, Panagiotopoulos, et al., 2020). Relevant legislation also touches on other areas such as big data (Mantelero, 2017). Here we focus on one aspect, which is the option of creating a new regulatory agency or body with mainly soft law powers. There have been calls for a new regulator for AI, big data and robotics, including from researchers (Erdélyi and Goldsmith, 2018; W Wallach and Marchant, 2019), civil society organisations (Miller and Ohrvik-Stott, 2018), policy advisory bodies (Datenethikkommission, 2019) and political bodies (Council of Europe, 2019 European Parliament, 2017;, 2020). In this article, we take up these calls for the creation of such a new body and explore its possible design and tasks. We focus on a European Agency for AI, but in discussing it, take into account international developments to ensure we provide a broad picture of roles and tasks of such an agency.

3. Feasibility of a new AI agency

In this section, we explore the feasibility of a new agency which we see as a regulatory body. Regulatory bodies, such as the one discussed here have different objectives and scopes. They encompass, e.g., bodies appointed (by the government) to establish standards, and/or policies for compliance, bodies or public authorities that regulate an entire sector, independent authorities that uphold rights in public interest and carry out enforcement action, and independent advisory bodies (with statutory footing) tasked by government to investigate and advise on technology regulation matters.

There are many forms such a regulator can take, based on established practice (e.g., a Commissioner, digital authority, fundamental rights protection agency, independent watchdog, inspectorate, licensing body or authority, network of regulators, professional conduct authority, professional regulator,

public sector regulator, standards agency, statutory registration board, supervisory agency or task force). For a discussion of possible regulators for civil drone practice, see Finn and Wright (2016). In some cases, private organisations can also take the role of regulators, for example such smartphone platforms that may effectively regulate privacy (van Hoboken and Fathaigh, 2021) or social media companies that have a role in regulating digital disinformation (Marsden et al., 2020). These are examples based on established practice. Some of these, such as the Digital Authority (UK) and the network of regulators (European Commission, 2020a) have been researched and/or considered as potential options for AI regulation, while others listed have not been deeply explored in policy or academia. We start with a summary of views of potential AI regulators.

3.1. *Summary of views and positions on new AI regulators*

This section summarises views and positions within the EU on new AI regulators within the EU. At the time of writing (June 2021), there were no AI- or big data-specific regulatory bodies in the EU at either transnational or Member State level (Rodrigues et al., 2019). Despite calls from the EU Parliament and some Member States for the creation of a regulatory body, the EU has not yet set up such as body on this (though a European Artificial Intelligence Board has been proposed in the EU AI Regulation). Section 5 of the article already coves the proposed European Artificial Intelligence Board and compares the two proposals and makes recommendations to improve the proposed AI Board. Some EU Member States have broadly assigned regulatory concerns regarding AI and big data to pre-existing regulatory agencies/bodies.

3.1.1. *International-level views and positions*

At the international (e.g., UN) level, no new regulators have been established. Some calls have been made for new regulatory bodies, such as the proposal for an International Artificial Intelligence Organisation (IAIO) (Erdélyi and Goldsmith, 2018) but these have as yet received little traction in terms of public policy. On 11 September 2019, the Committee of Ministers of the Council of Europe set up an Ad-hoc Committee on Artificial Intelligence (CAHAI) (Council of Europe, 2020) to examine the feasibility and potential elements of a legal framework for the development, design and application of artificial intelligence, on the basis of broad multi-stakeholder consultations, based on the Council of Europe's standards on human rights, democracy and the rule of law.

3.1.2. *EU institutions views and positions*

Different positions on new regulators are evident. In 2017, the EU Parliament called on the European Commission (the Commission) to, amongst other things, create an EU Agency for Robotics and Artificial Intelligence (European Parliament, 2017 Villaronga and Golia, 2019). The European Commission did not consider the creation of a new agency necessary, but instead proposed the creation of "a high-level advisory body on robotics and artificial intelligence which could provide knowledge and expertise to the Commission". This became the High-Level Expert Group on Artificial Intelligence ("AI HLEG"). Some Member States (e.g., France) have called for a new regulator at the EU-level.

The European Commission (2020a) *White Paper On Artificial Intelligence - A European approach to excellence and trust* discusses a European governance structure on AI in the form of a framework for cooperation of national competent authorities which could have a variety of tasks: as a forum for a regular exchange of information and best practice, identifying emerging trends, and advising on standardisation activity as well as on certification. It could also play a key role in facilitating the implementation of the legal framework, such as through issuing guidance, opinions and expertise. To that effect, it should rely on a network of national authorities, as well as sectorial networks and regulatory authorities, at national and EU level. This proposal by the Commission therefore sees the role of a European governance structure and regulatory body as similar to that developed in the General Data Protection Regulation (GDPR), which institutes a forum for exchange of information and best practice at the EU-level through the European Data Protection Board (EDPB) while relying on a network of national authorities (i.e., the Data Protection Authorities) for enforcement of the Regulation. As such, it envisages the creation/use of national-level regulators akin to the DPAs. However, this approach has not as yet been embraced by Member States.

In April 2021, the Commission released its first draft version of a European regulation for AI (European Commission, 2021a). The draft includes a proposal for the creation of a European AI Board, discussed in Section 5.

3.1.3. *EU member states' views and positions*

The Commission's (2020a) White Paper on AI has noted that, "member States are pointing at the current absence of a common European framework. … If the EU fails to provide an EU-wide approach, there is a real risk of fragmentation in the internal market, which would undermine the objectives of trust, legal certainty and market uptake." A recent report confirms that within Europe, there are currently no new regulators dedicated to AI use or development (Rodrigues et al., 2019). Where AI is a deep concern in specific sectors, most EU Member States use existing sectoral regulators. This has occurred primarily in the areas of finance and competition law, where there are strong, well-established existing regulatory frameworks. New regulation of AI does not appear to have occurred in other regulation-heavy areas, such as healthcare or security, beyond existing laws on export restrictions in the case of security. An overview of all European Member States with regards to their approach is available in the EU's digital government factsheets (European Commission, 2019).

3.1.4. *Non-EU countries views and positions*

The UK government's Committee on Standards in Public Life (2020) explicitly stated that it does not see the need for an AI regulator. The UK's Centre for Data Ethics and Innovation (CDEI) functions in the UK as a government advisory body to investigate and advise on how the benefits of data-enabled technologies, including AI, can be maximised, and to identify measures needed to strengthen and improve the way data and AI are used; promote best practice and advice on how Government should address potential gaps in the regulatory landscape. This body is not a regulator in the traditional sense but

a regulatory advisory body. There is also the Office for Artificial Intelligence, a joint BEIS-DCMS unit that oversees the implementation of the AI and Data Grand Challenge and seeks to "drive responsible and innovative uptake of AI technologies for the benefit of everyone in the UK" (Office for Artificial Intelligence, 2020).

The US Food and Drug Administration (US FDA, 2019) published a discussion paper on a "Proposed Regulatory Framework for Modifications to Artificial Intelligence/Machine Learning (AI/ML)-Based Software as a Medical Device (SaMD)" which, if it were to be enacted in law, would see the FDA take on additional regulatory responsibilities of AI-enabled devices in healthcare in the US. In January 2021, the FDA followed-up with an action plan, which includes further developing the regulatory framework (US FDA, 2021). This suggests that other agencies overseeing areas in which regulation is well-established may also see an increase in responsibilities to include AI and big data regulation.

Other proposed laws in the US have also sought to provide additional/new powers to existing bodies for the regulation of aspects of (or aspects closely linked to) AI and big data. For example, the 2019 Federal Trade Commission (FTC) in the Algorithmic Accountability Act (HR2231) (Clarke, 2019c) would have mandated the Federal Trade Commission (FTC) to require (US FDA, 2019)entities that use, store or share personal information to conduct impact assessments for any high-risk automated decision system that makes a decision (or facilitates a human decision) that impacts consumers, and to reasonably address the results of the impact assessments in a timely manner. The bill provided basic protection for personal data used in automated decision systems on a national level (R Clarke, 2009). However, the bill did not require covered entities to make the results of the algorithmic assessments public, leading some commentators to suggest that there would be insufficient transparency (New, 2019). Likewise, while the proposed act would have required covered entities to conduct required assessments in consultation with external third parties (e.g., independent auditors and independent technology experts) if reasonably possible, some commentators suggested that the bill did not go far enough to require neutrality in the assessment. Finally, the proposed text did not specify how often algorithmic assessments must be updated (using instead the phrase, "as frequently as the Commission determines is necessary"). Some commentators have suggested this could be unduly burdensome given the iterative nature of software development (Rodrigues, Panagiotopoulos, et al., 2020).

Another example is the proposed "Deep Fakes" Accountability Act (HR 3230) 2019 (Clarke, 2019d), which would have obliged the US Attorney General to submit a quinquennial report to Congress containing a description of the impact of intimate and sexual deep fakes on women and marginalized communities, and providing official guidance to Federal prosecutors regarding any potential legal concerns that may impede such prosecutions absent clarification (Rodrigues, Panagiotopoulos, et al., 2020). Some concerns were raised regarding this proposed Act and its potential to introduce a chilling effect on citizens regarding the production of satirical or parodic political videos. This proposed law responds to very specific concerns regarding one element of AI, however, and is a long way from providing an "AI and big data" regulatory framework.

3.2. *Pros and cons of creating an agency*

Several proposals have been put forward since 2015 for new regulators and/or regulatory bodies for new technologies and their assessment. Annex A presents proposals made by EU-level and/or national policymakers and academics which have gained significant traction in academic and popular press. The seven proposals covered include the following:

- An EU Agency for Robotics and Artificial Intelligence;
- An EU Taskforce of field specific regulators for AI/big data;
- A Network of national authorities, as well as sectoral networks and regulatory authorities;
- the proposal for a Digital Authority (UK);
- An independent regulator under the new statutory duty of care for online harms (UK);
- A FDA for Algorithms (US); and
- A Federal Trade Commission to regulate robotics (US).

The analysis of these proposals allows for the extraction of key arguments in favour and against the creation of a regulator. This section contains a summary of the more detailed analysis provided in (Rodrigues et al., 2019 Rodrigues, Panagiotopoulos, et al., 2020).

A key argument in favour of the creation of a new regulator is that the management and regulation of AI involves a complex understanding of technical, legal and ethical considerations. It could therefore be helpful to bring these together in one body. A regulator could furthermore demand access to the technical details of algorithms if it were genuinely independent, leading to the institution of state-backed guarantees, similar to kite marks. A regulator might boost human rights, by:

- Requiring assessments to identify and reduce risks of high-impact automated decisions (e.g., Algorithmic Accountability Act of 2019);
- Protecting democracy and privacy by promoting healthy competition (e.g., anti-trust regulations);
- Safeguarding and enhancing human rights more comprehensively, such as the right to privacy and freedom from discrimination, in the context of predictive algorithms (e.g., register of algorithms used in government); and
- Introducing compliance mechanisms to monitor, prevent and manage risks to human rights (e.g., the legislative framework for independent and effective oversight; legal frameworks for human rights impact assessments).

Furthermore, existing regulations may not be adequate to regulate AI issues; questions are already being raised, e.g., about whether data protection regulation is sufficient to protect privacy. At the EU level, a new regulator will need the ability to respond to reflect the legal specificities of each Member State and respond to the policy and technological needs and priorities. A bespoke system tailored to individual states may not be effective if it is not part of a uniform regulatory approach to AI at a supra-national level.

These arguments in favour of a regulator are counterbalanced by a number of challenges. To start with, it is not clear whether a new regulatory body would be necessary. It may duplicate the work of existing regional or national agencies. As one report highlights, existing regulatory bodies, with the help of new regulations, can and do cover many new challenges posed by AI and/or robotics (House of Lords, 2018).

The creation of a new body may be burdensome and raise new concerns. A lack of transparency and clarity on the operation, powers, scope and relationships of a regulator with regards to other regulatory authorities (e.g., CDEI and the proposed IAIO) could be deeply problematic. A similar limitation relates to the lack of management stability, weak collective control and oversight. There are concerns about mission creep and an over-regulation of the market, leading to worries about the ability to comply, in particular by SMEs.

There is also the risk of regulatory capture ("special interests' manipulation of government agencies regulating network industries" (Wren-Lewis, 2011, p.)). This could be seriously damaging unless measures are taken to ensure that structure and governance of such a regulatory body is appropriately modelled and, in a way, as to promote human flourishing and not protect singular interests (e.g., the AI industry).

Positioning a new body in the existing regulatory environment can prove to be distracting. There is a risk that a new regulator could be viewed as a panacea or replacement for existing frameworks, rather than a means of enforcing those frameworks. A new regulator may excessively focus on the risks of AI while neglecting the relevant benefits and advantages. This might extend to a disproportionate focus on bias and discrimination, at the loss of examination of other fundamental rights and freedoms. Alternately, the regulator may give insufficient consideration to issues of racial and gender bias or privacy. When creating a new body, it will be crucial for it to have relevant competence and understanding of the features and capabilities of AI. There is a risk of providing false assurances of fair, trustworthy and/or ethical AI, a risk that may be exacerbated by a lack of an accountability framework providing for sanctions in the case of failure to apply an option.

Finally, a new body may introduce conflict. Where a new regulator might require the publication of AI-sensitive information, there may be conflicts with intellectual property rights and prohibitions of releasing sensitive information to the public. Resistance may arise from actors to share sensitive information from impact assessments. Furthermore, a lack of consistency and consideration of the legal and political specificities can arise where a new regulator has a supranational effect.

The following proposal for a European Agency for AI is informed by this discussion. Its features are designed in a way to promote possible benefits and proactively address or avoid some of the identified downsides.

4. Features of a European Agency for AI

The above discussion has demonstrated that there are good reasons for the creation of a new body devoted to AI but that the case is not straightforward and needs to take into consideration possible challenges. The exact benefits and problems of a new body will depend, to a large extent, on the shape it takes in practice. We, therefore, set out a detailed account of a possible shape of such a body. This required a decision on some of the basic parameters of the potential body. The first choice was to focus on an EU-level body. This was motivated by the high level of policy-activity around AI in the EU that includes several high-level proposals for a regulatory body. An EU-level body is furthermore of interest because it must bridge the gap between national-level regulation and international policy, something that is arguably of key importance for a global technology such as AI. In order to develop a convincing account of the detail of an EU-level body we used 'terms of reference' (ToR) to describe the characteristics of this body. ToRs are used in governance where they define basic features of project, committee etc., typically including aspects such as purpose, roles and functions, structure and membership. The ToR provide a common understanding of scope, objectives and operational processes. We used the concept of a ToR as part of our methodological approach to determine the characteristics that an EU-level body could display.

4.1. *Methodology*

Based on the theoretical insights concerning AI in general and the feasibility of an AI agency, we drafted an initial ToR (Rodrigues, Santiago, et al., 2020, p. 6). We then undertook several steps to reach out to a broad range of stakeholders and develop the ToR further. In order to achieve this and reach broad coverage, we undertook three steps in the SHERPA project: a) a stakeholder board meeting, b) a focus group, c) a revised draft ToR, d) a set of 10 expert interviews and e) a final round of open feedback. This work was guided by three core parameters: 1. Explore the feasibility of a bespoke/new regulator for AI and big data at the EU and/or Member State levels, 2. Assess factors indicating the need for a regulator and how it would fit or interact with existing regulatory bodies and structures and 3. Outline its ToR. We recognise there is a rich body of regulation literature which one might engage from the theoretical point of view, but we have limited this article in its expression of the ToR to that of a practical policy implementation perspective.

a) A stakeholder board meeting took place on 23 March 2020. It was organised as a face-to-face meeting but, due to Covid, was then held as an online event. The board was assembled as a standing resource of the project underlying this research. It consists of about 30 members representing industry, civil society, policy and research. For this event, 13 members attended (7 male, 6 female). Participants were provided with a draft ToR based on the above overview in advance of the meeting and were then asked to discuss the following questions:
   1. Do we need a new regulator/body for AI and big data at the EU or national level?
   2. What international, EU or national policy directions are relevant to consider in the creation of such a new regulator?
   3. If no new regulator is deemed necessary, what other regulatory options are the most desirable and feasible?

The stakeholder board discussion was captured in a protocol of the proceedings which provided the data for the analysis underlying this article.

b) On 26 June 2020, a focus group with five experts was held to discuss a new AI regulator. A Google Form questionnaire (see Annex 2) was circulated to the participants in advance of the meeting. The questionnaire was completed by nine invited participants. Detail on the analysis is provided in (Iordanou et al., 2020).
c) Building on findings of the above activities, we revised the ToR. In doing so, we took into consideration the Framework for cooperation of national competent authorities outlined in the European Commission White Paper (2020a), and the European Parliament Draft Report (2020). We looked at the role of various EU bodies, especially their remits and functions (e.g., Agency for the Cooperation of Energy Regulators (ACER), European Data Protection Board (EDPB), European Data Protection Supervisor (EDPS), EU Taskforces, European Union Agency for Cybersecurity (ENISA), Fundamental Rights Agency (FRA)). This research fed into the drafting of the ToR.
d) We then carried out a consultation in August-September 2020 with 10 experts using interviews. The experts were given a draft ToR document and asked specific questions about the proposed terms. Experts came from a range of European and national institutions, and included policy makers, legal advisors, a Member of Parliament, industry experts, civil society and policy analysts. Using their feedback, and consulting some founding regulations, e.g., the ACER Regulation, the General Data Protection Regulation (GDPR), European Union Agency for Fundamental Rights Regulation, and operational documents for existing EU agencies, the ToR was revised and finalised.
e) The final steps included feedback on the penultimate draft of the ToR which was shared with the Stakeholder Board (see point a), interested stakeholders, project partners and colleagues.

4.2. *EU agency for AI – proposed terms of reference*

Based on the steps described above, we propose the ToR outlined below. While a ToR also covers operational principles, reporting and auditing, evaluation, review and, of course, funding and sustainability, for the purposes of this article, we focus on the (a) purposes, role and functions and (b) structure of the Agency. We also assess its ability to address ethical and human rights concerns and promote human flourishing (Rodrigues et al., 2020).

4.2.1. *Nature*

The European Agency for AI should be an **independent European Union agency with legal personality.** The advantages of a separate and independent legal entity are well-recognised - i.e., it would meet the need for special expertise and make the 'credible commitments' required in relation to its objectives. It is essential that its work is of high quality and not influenced by political or contingent considerations. Its independence should be both in relation to its functioning and of the persons managing it. The Agency should operate as a point of reference establishing trust and confidence by virtue of its independence and its work. I Its legal basis should lie in the European Union AI legislation and/or a new founding regulation that would set out its mandate and operational procedures.

Additionally, in its nature and scope, the transversal and cross-sector nature of AI must be considered. As AI has many diverse forms and applications, it is subject to different current regulations and is therefore relevant to the mandate and function of many existing EU agencies and institutions. The Agency should complement the work of sector-specific approaches and not duplicate the roles of existing agencies/authorities. To determine precisely the gaps and potential for merging roles and functions, the Commission must conduct a full regulatory impact assessment on the creation of an Agency to clarify the remit and function vis-a-vis other regulatory bodies.

4.2.2. *Purposes, role and functions*

The complexity of the ethical and human rights issues of AI and the fact that the new Agency would enter into a dynamic environment will require it to fulfil a number of purposes. The Agency should foster effective application and enforcement of existing EU and national legislation, including the new legal regulation on AI (if/when implemented). It should assist, and coordinate with EU institutions, bodies, agencies and competent authorities to promote and protect Union values and fundamental rights, including those enshrined in the EU Charter of Fundamental Rights, from adverse impacts of AI in all phases of the AI lifecycle and particularly safeguard the rule of law. Furthermore, its purpose should include to promote and strengthen common governance, consistency and harmonised approach across the European Union, and work to reduce fragmentation of responsibilities. It should provide the relevant institutions, bodies, agencies and authorities of the Community and its Member States, when implementing Community law, with information, assistance, expertise and recommendations on the regulation of AI to support them when they take measures or formulate courses of action within their respective spheres of competence to fully respect EU values and fundamental rights and their enforcement. In addition, it should collaborate, deliver advice to sectoral and national bodies upon request and improve regulatory capabilities in the regulation of AI, and deliberate on and discuss AI developments (e.g., as identified by the European Commission's AI Watch (2020b)), as they happen to determine which need binding rules, or other regulatory measures.

In order to fulfil its purpose, the Agency would need to take over a number of roles and functions. The following list highlights capability of the new body. The Agency should:

1. Make **Recommendations addressed to the European Parliament, the Council, or the Commission for legislative amendments and adjustments**, after carrying out fitness-for-purpose checks, to boost implementation and enforcement of legislation at the EU-level, related to AI;
2. **Identify potential red lines or restrictions,** i.e., thresholds, boundaries, limits which should not be crossed for

AI development, deployment and use that violates human rights and/or has significant negative societal impacts addressed to the EU institutions and to feed into European Commission Decisions;
3. **Develop and promulgate general guidance** on legal concepts and regulatory issues of AI -based on its discussions and deliberations of AI developments;
4. **Set benchmarks for enforcement** and present its position via Opinions or Intervention/Enforcement Advisories for enforcement authorities;
5. **Support, work with and advise EU-level institutions, bodies and agencies,** e.g., European Commission, European Parliament and the Council, EDPB, EDPS, FRA **and national competent authorities in Member States** to fulfil their ethical and human rights obligations and to protect the rule of law where AI is researched, commissioned, developed, deployed and used;
6. **Maintain an AI risk alert system** (notifications of direct or indirect risks to human life and health) to competent authorities from the development, deployment and use of AI systems, services and products via its Network;
7. **Assist in coordinating the mandates and actions** of the national competent authorities of Member States;
8. **Develop harmonised and objective criteria for risk assessment** and/or conformity assessment including certification of ethical and human rights requirements of AI and related technologies in the EU and issue guidance on their application;
9. **Monitor and/or coordinate** the evaluation of the operation of conformity assessment and/or certification schemes established for such purposes;
10. **Cooperate, liaise, exchange information, promote public dialogue**, best practices and training activities with international, EU, national AI regulatory bodies and/or supervisory authorities think-tanks, civil society, technology community, and other underrepresented stakeholders when proposing new regulations;
11. **Ensure complementarity and synergy** between its activities and other Community programmes and initiatives;
12. **Promote the adoption of regulatory sandboxes** to allow live testing of the distribution and use of AI innovations by stakeholders, including the public, in a controlled environment under regulators' supervision;
13. **Promote the Union's AI approach through international cooperation** with relevant bodies such as the UN, OECD, G20 and regional legal orders by participating/inviting participation in common interest regulation-related activities. It should also play a greater coordinative role in shaping the development of international standards for AI, through the International Organisation for Standardisation (ISO) and International Electrotechnical Commission (IEC) to harmonise the technical requirements of AI.

These roles and functions are critical based on our research and stakeholder consultation. These are minimum requirements for the Agency. The Agency should have the possibility of modifying its roles and functions in line with needs that cannot be foreseen at the present time.

The Agency should operate based on the following principles: respect for human rights/human-centric approach; independence, and impartiality; fairness; transparency; proactivity; good governance, integrity and good administrative behaviour; collegiality, inclusiveness and diversity; cooperation; efficiency and modernisation.

4.2.3. *Structure*

The Agency structure and composition should reflect its independent nature. Additionally, the Agency should be more than a gathering of national regulators. The following indicative structure is in line with the features of existing decentralised European Union agencies, though it should be adapted to meet the specific need of a new regulation on AI.

In order to fulfil maintain independence and effectively fulfil its mandate, the Agency could be comprised of:

A **Management Board** to ensure the Agency carries out its tasks. It would have authority over the Agency's work programme, budget and annual report. The Management Board would also make final decisions on the adoption of Decisions, Opinions, Recommendations, Guidelines, Advisories and other documents (prepared by the Executive Board). The Management Board would nominate an Executive Director. The members of the Management Board would be appointed by the European Parliament, the Commission (in equal number) and the Council (same number as the other two plus one extra) in equal number. The European Commission would be represented in the Board without the right to vote. The term of office of the members of the Board will be four years, renewable once. Directly reporting to the Management Board would be:

- An **Executive Board** responsible for preparing Decisions, Opinions, Recommendations, Guidelines, Advisories and other documents (coordinated by Rapporteurs) to be adopted by the Management Board. The Executive Board would consult and/or delegate specific tasks to the scientific/technical committee, the advisory committee and its working groups and Rapporteurs.
- An **Executive Director** responsible for managing the Agency and performing duties independently. The Director would be the legal representative of the Agency and would be in charge of its day-to-day management; prepare and participate in the work of the Boards and have the overall responsibility for implementing the decisions adopted by the Board, draft, consult upon, adopt and publish opinions, recommendations and decisions. The Executive Director would also be responsible for implementing the Agency's annual work programme under the guidance of the scientific/technical committee and under the administrative control of the Management Board.
- Ad hoc **Working groups** of experts convened to address specific technical and scientific matters not addressed by other committees in the Agency. Experts could be drawn from e.g., Confederation of Laboratories for Artificial Intelligence Research in Europe (CLAIRE), European AI Private-Public Partnership (AI, Data and Robotics Partnership) and other EU-funded AI projects. The Working Groups would be established by the Executive Director, in consultation with the Advisory Committee.

- A **Secretariat** responsible for providing analytical, administrative and logistical support to the Management Board and the Agency (e.g., preparation of documents, and organising meetings and communication). The head of the Secretariat would be responsible for the due and timely performance of its tasks.

A **network of EU and Member State national competent authorities** to facilitate the exchange of information between the Agency and EU Member States for the consistent and effective application of the EU AI regulatory framework and the AI risk alert system. Relevant agencies that could be represented include: the EDPB, EDPS, Frontex, eu-LISA, EASO, EIGE, EMCDDA, CEPOL, Europol, Eurojust, EURODAC, FRA, EQUINET, EIOPA, EBA, ESMA, field-specific agencies and national competent authorities (or designated AI bodies) from the EFTA, EEA states, etc. The Network will closely cooperate with the Agency and help monitor the impacts of the AI Regulation. The Agency will be open to the participation of third countries which have concluded agreements with the Union and which have adopted and are applying the relevant rules of Union law in the field of AI. In addition to Member State representation at the Agency, the Agency should also have dedicated presence in each Member State (via notification of one national competent authority as its main liaison).

A **Scientific and Technical Committee** that ensures the high scientific quality of the work of the Agency and guides its work by means of scientific objectivity.

An A**dvisory Committee** that focuses on regulatory issues relevant to stakeholders and brings them to the attention of the Agency. The committee should have mixed representation of diverse backgrounds, including industry (small and large companies) and civil society organisations, particularly those representing those vulnerable groups.

A **Conformity and Risk Assessment Committee** to coordinate the mandates and actions of the national competent authorities of Member States. The committee would also develop harmonised and objective criteria for risk assessment and/or conformity assessment, including certification of ethical and human rights requirements of AI and related technologies in the EU.

4.3. *Discussion*

The proposal for the ToR is based on a strong understanding of the regulatory environment and current EU legislative activities, and taking into account stakeholder feedback and concerns. We realise, however, that it is not authoritative, but only one possible proposal amongst many that could be used to govern AI for human flourishing. Its purpose is to contribute to the policy debate and help to shape the contours and influence the eventual regulatory body, which we believe is likely to come into existence early in this decade. In this section we discuss why we think that an Agency displaying the characteristics proposed above would be well-placed to address the ethical and human rights issues of AI and promote human flourishing.

The proposed Agency or any other regulatory or other body to be created will form part of a broader regulatory environment. Just like AI can be seen as an ecosystem of many inter-related technologies, institutions and stakeholders (OECD, 2019), the same applies for the regulatory environment. This ecosystem includes legislators (international, regional and national), regulators, alongside executive bodies, companies, professional bodies, research groups and other organisations as well as numerous individuals acting in various roles, promoting a broad array of interests.

For the regulatory ecosystem to be successful, it will need to be able to address the challenges that arise from AI, as discussed earlier. amongst these challenges there is the question of the definition of AI. This is a crucial question because the definition will determine the scope and reach of the governance ecosystem. At the same time, it seems unlikely that providing a comprehensive definition will be possible, as the concept of AI encompasses different families of technologies (Samoili et al., 2020). In addition, the ethical and human rights consequences of AI can be determined more or less narrowly. A regulatory body such as the European Agency for AI will therefore provide the flexible approach to these questions, providing a workable definition, scoping and be able to update these definitions in accordance with technical and social developments. It can work with legislatures and executives to provide expert input and guidance (see section 0, recommendation 1, 5)

The ability to flexibly develop definitions and refine them is not only important for the concept of AI itself, but for other aspects of the AI ecosystem as well. All governance structures need to be based on the best available knowledge which is likely to develop quickly. In order to collect this knowledge and maintain it, there will need to be knowledge repositories that make knowledge available and that also act as reliable and transparent gatekeepers and centres of scientific excellence that provide knowledge for the development of policy and regulations. The European Agency for AI as described in the ToR can cover this role (see section 0, recommendations 1, 3, 6,).

One challenge of the AI debate is its global nature based on global AI technologies combined with the local nature of regulatory responses Jobin et al.'s (2019). review of ethics guidelines has shown that there are shared core positions but also much variance. A similar picture is likely to apply to policies and governance. In order to avoid a patchwork of regulations that can be easily side-stepped and come with their own challenges, there need to be mechanisms that allow for the exchange of knowledge and mutual learning between and across jurisdictions. Our proposed ToR gives the Agency this role (see section 0, recommendations 5, 8, 9, 11, 12)

The complex and fast-shifting nature of AI furthermore means that in addition to scientific and procedural knowledge, there are open questions about the way in which non-experts understand the technologies and how they interpret possible ethical and human rights issues. While legitimate regulation must be driven by democratically appointed bodies of representative democracy, it is therefore likely that a full understanding of the phenomena in question requires further engagement with other stakeholders. The relationship between stakeholder engagement and established democratic governance mechanisms can be difficult, but the sort of sensitive governance structure required for AI, also called tentative governance (Kuhlmann et al., 2019), will most likely require mech-

anisms to understand stakeholder views (see section 0, recommendation 10).

5. The proposed European artificial intelligence board

In April 2021, the European Commission set out its proposal for an EU regulation for the governance of AI (the Regulation). To "facilitate a smooth, effective and harmonized implementation" (European Commission, 2021a) of the regulatory framework, the Commission calls for establishing a new European Artificial Intelligence Board (the Board) (European Commission, 2021a). At this stage, the proposal only includes high-level information on the mandate, structure, and tasks of the Board; a full Terms of Reference has not yet been developed. The proposed Regulation must be adopted by the Parliament and Member States before coming into effect. In summer 2021, the draft Regulation text was opened for public consultation via the Commission's 'Have your say' portal, and will be revised in response to feedback from the public, European Parliament, and EU Member States.

This section summarises the Commission's proposal for the Board and contrasts it with our proposal for an Agency, concluding with recommendations.

5.1. *Commission proposal for a European artificial intelligence board*

The Board would be mandated to "advice and assistance to the European Commission" for three purposes: (1) contribute to cooperation with Member States; (2) coordinate and contribute to guidance and analysis on emerging issues in AI, and (3) assist in ensuring consistent application of the law (European Commission, 2021a). The specific Board tasks are laid out in Article 58 of the draft Regulation:

> *Article 58 Tasks of the Board*
> When providing advice and assistance to the Commission in the context of Article 56(2), the Board shall in particular:
>
> a) collect and share expertise and best practices amongst Member States;
> b) contribute to uniform administrative practices in the Member States, including for the functioning of regulatory sandboxes referred to in Article 53;
> c) issue opinions, recommendations or written contributions on matters related to the implementation of this Regulation, in particular
> i. on technical specifications or existing standards regarding the requirements set out in Title III, Chapter 2,
> ii. on the use of harmonised standards or common specifications referred to in Articles 40 and 41,
> iii. on the preparation of guidance documents, including the guidelines concerning the setting of administrative fines referred to in Article 71.

Along with the Commission, the Board is also expected to "encourage and facilitate the drawing up of codes of conduct intended to foster the voluntary application to AI systems of requirements related for example to environmental sustainability, accessibility for persons with a disability, stakeholders participation in the design and development of the AI systems and diversity of development teams on the basis of clear objectives and key performance indicators to measure the achievement of those objectives" (European Commission, 2021a).

The proposed Regulation itself does not lay out the Board's operational principles, nor does it specify any details on executing these tasks, but directs the Board to adopt its own rules of procedure (European Commission, 2021a).

The proposed Board would be chaired by the Commission, with representation from the EU Member States and the EDPS. Chairing duties include convening meetings, preparing agenda, and providing administrative and analytical support (European Commission, 2021a). The Board may establish subgroups or invite external experts and observers to meetings

Table 1 – Comparison of functions of the AI Agency and AI Board.

Comparison of functions of the proposed AI Agency and AI Board	
Proposed European AI Agency	Proposed AI Board (based on tasks specified in draft Regulation Article 58)
1. **Make recommendations** addressed to the European Parliament, the European Council, or the Commission for legislative amendments	Falls within Article 58(c)
2. Identify potential **red lines or restrictions for AI development, deployment and use** that violates human rights and/or has significant negative societal impacts	Partially falls within Article 58(c)
3. **Develop and promulgate general guidance** on legal concepts and regulatory issues of AI	Falls within Article 58(c)
4. **Set benchmarks for enforcement**	Missing
5. **Support and advise EU-level institutions, bodies and agencies and national competent authorities in Member States** to fulfil their ethical and human rights obligations and to protect the rule of law	Falls within Article 58(c)
6. **Maintain an AI risk alert system**	Missing
7. **Assist in coordinating the mandates and actions** of the national competent authorities of Member States	Could falls within Article 58(c)
8. **Develop harmonised and objective criteria** for risk assessment and/or conformity assessment	Could fall within Article 58(c)
9. **Monitor and/or coordinate** the evaluation of the operation of conformity assessment and/or certification schemes	Partially falls within Article 58(a)
10. **Cooperate, liaise, exchange information, promote** public dialogue, best practices and training activities	Falls within Articles 58(a) and 69(2).
11. **Ensure complementarity and synergy** between its activities and other Community programmes and initiatives	Falls within Article 58(b)
12. **Promote the adoption of regulatory sandboxes**	Falls within Article 58(b)
13. **Promote the European Union's AI approach through international cooperation**	Could fall within Article 58(b)

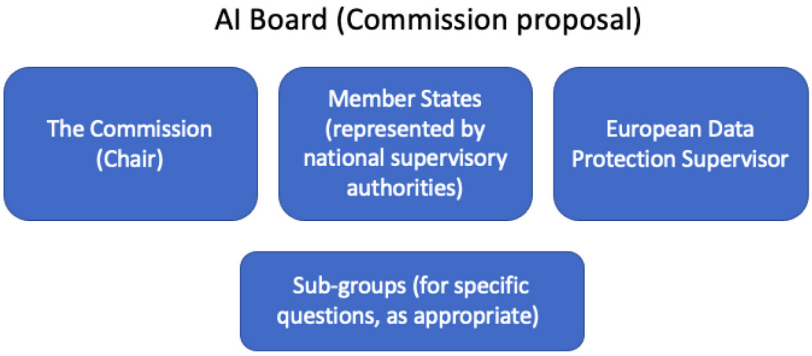


Fig. 1 – Structure of proposed AI Board.

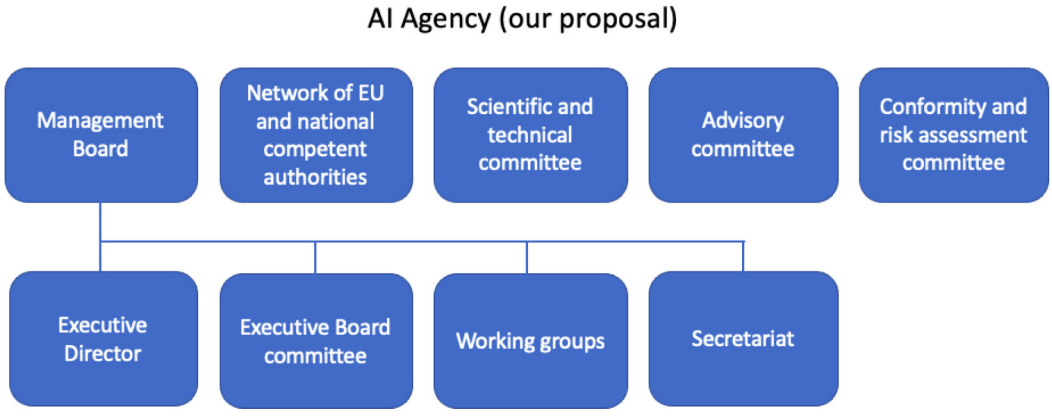


Fig. 2 – Structure of proposed AI Agency.

as needed. The Board may also facilitate exchange with 'interested their parties' and other EU bodies, offices, agencies and advisory groups where appropriate (European Commission, 2021a). The Board is anticipated to be staffed by five full-time equivalent (FTE) (European Commission, 2021a).

5.2. *Comparison of the European AI agency and the AI board*

The two proposals are aligned in some ways but differ in others. Both proposals envisage a mandate to support and facilitate implementation of the EU regulatory framework for AI to help ensure cooperation, coordination and consistent application of EU law. Both bodies, as proposed, play an advisory role and do not have rule-making authority. The proposals, however, differ in terms of structure, with our proposal for an Agency having a more independent structure that includes permanent representation from diverse bodies, stakeholder groups and technical advisory support.

5.2.1. *Comparison of purposes, role and functions*

There are some similarities in the purpose, role and functions (e.g., make recommendations, guidance/written contributions, supporting harmonisation, collection and sharing of best practice including in relation to sandboxes). However, some tasks included in our proposal for the European Agency are missing from the proposal for a Board. Other tasks of the Agency are included in the proposed Regulation, but not under the proposed Board's direct remit. For example, the proposed Regulation already includes an exhaustive list of prohibited AI systems (i.e., red lines), which means that the responsibility to identify red lines would not lie with the proposed Board.

5.2.2. *Comparison of the structure*

The proposed Board would be limited to permanent representation from Member States, the Commission, and the EDPS. This differs from our recommendation for the Agency, which is more independent from the Commission and includes permanent representation from diverse stakeholder groups on standing committees, including a scientific and technical committee and an advisory committee. Furthermore, the proposed Board does not have any management or secretariat structures outlined yet.

The following figures illustrate the differences between our proposal and the Commission's.

5.3. *Recommendations to improve the proposed AI board*

For the reasons discussed above, we recommend the following issues are are addressed as the draft AI Regulation is re-

vised and finalised, and decisions are taken in regard to the proposed AI Board.

Increase independence by decreasing alignment with the Commission. The Board should be independent, providing support across the EU to the Commission, Parliament and Council. The Board should not be limited to providing opinions and recommendations to the Commission. Independence would also be increased with broader representation (see below).

Enhance diversity with permanent representation from broader stakeholders. The Board should include representation from more EU bodies than just the Commission and EDPS. Additionally, the Board should establish permanent mechanisms to draw on the expertise of diverse stakeholder groups on permanent standing committees (e.g., ethics, scientific and technical, human rights).

Expand the task remit to include more roles and functions. The Board should have a broader mandate to enable it to better support effective application and enforcement of the law. This should include, for example, developing and promulgating general guidance on legal concepts and regulatory issues associated with AI for Parliament, Council, Member States, AI industry and the general public.

Adopt operational principles. In addition to rules of procedure, the Board should adopt operational principles to guide its work. These principles should include: respect for human rights/human-centric approach; independence, and impartiality; fairness; transparency; proactivity; good governance, integrity and good administrative behaviour; collegiality, inclusiveness and diversity; cooperation; efficiency and modernisation.

6. Conclusion

In this article, we have provided an answer to the research question: *how should a European Agency for AI be designed, to promote human flourishing?* We reviewed the discussion of the ethical and human rights issues of AI and possible mitigation strategies, which suggest that regulatory bodies, such as the proposed European Agency for AI will need for form part of the governance ecosystem required to address these issues. A review of the feasibility of such an agency explored the current state of debate, indicating that the European-level is an interesting and well-advanced area where such a body may come into play. We then developed a ToR for such an agency, based consultations with stakeholders. We finally discussed why and in what way an Agency developed according to our ToR would form part of the regulatory AI ecosystem.

The article contributes to the academic discussion of AI governance and AI ethics which form part of the broader field of technology governance, which is of interest to a number of disciplines ranging from law, information systems and science and technology studies to the philosophy of technology. The combination of conceptual review and empirical engagement should ensure that the ToR will prove to be of interest to scholars from these various fields. More importantly, we see this article as a step outside of the academic ivory tower and an attempt to make a practical contribution to a current policy debate. In light of the current policy discourse, it seems likely that new regulatory bodies for AI will come into existence in the near future. Their exact purposes, roles and tasks will form part of political debates. We believe that these policy discussions can benefit from the insights offered here and we thus see the article as a contribution to the ideal of research-based policy development.

There is of course much work to be done. We have focused on the European level for reasons outlined earlier, but also for the practical reason that the authors work at the European level and are familiar with it. Similar considerations would need to be explored at the national level, where regulatory bodies may encounter different requirements and take different forms. Different cultural settings and political traditions on other regions may call for a different shape of a possible regulator. In addition, there will need to be bodies that allow for global coordination, for which some initial recommendations already exist (Wendell Wallach and Marchant, 2018).

The final shape of any AI regulatory body or bodies will be the outcome of political negotiations. Their precise nature is thus impossible to predict. One aspect worthy of further anticipatory reflection, however, will be the question of success criteria. How would we be able to tell whether the creation of a particular regulatory body achieves its purposes? This clearly depends on the final definition of the body's mission and objectives. But there is a larger question of whether and how progress in ethics, human rights and human flourishing can be measured and attributed to particular interventions, such as the creation of a regulatory Agency. This needs further research. However, we believe that our proposal has theoretical and practical impact. The interaction between scientific/technical and regulatory expertise is required if we want our research to provide a positive answer to Walsham's (2012) question whether we are making a better world with ICTs.

Declaration of Competing Interest

None.

Data Availability

Data will be made available on request.

Acknowledgement

This This article draws from research undertaken by the EU-funded projects SHERPA (www.project-sherpa.eu) and the Human Brain Project SGA3 (www.humanbrainproject) that have received funding from the European Union's Horizon 2020 Framework Programme for Research and Innovation under GA no 786641 and GA no 945539.

Annexes

Annex A: Summary of proposals for new regulatory bodies and their assessment

This annex presents a summary of proposals for regulatory bodies at the EU, Member State or non-EU country level. It presents the bodies, proposer, year of the proposal, the level they have been proposed, what they regulate (AI, big data), gaps filled, identified pros and cons, and presents an assessment, i.e., are they feasible and sustainable (e.g., supported by policy and market incentives) and future-proof? Or are they be adversely affected by future developments e.g., technological, policy changes, social demands?[1]

Name of proposed body and proposer	**EU Agency for Robotics and Artificial Intelligence[1]** **Proposer: EU Parliament** **Year: 2017**
Level /country	EU-level
What does it regulate	Robotics and Artificial Intelligence
Gap filled/need met	Currently there is no systematic public tracking system for autonomous systems that have been deployed. Suppliers of autonomous systems may keep records (such as system serial numbers) for deployed systems, but there is no public visibility into these proprietary records.[2]
Identified pros	A uniform registration system would facilitate safety notifications and recalls, similar to the vehicle identification number system for cars today.[3]
Identified cons	Rejected by European Commission, which founded the AI HLEG instead to investigate further.[4]
Final assessment	Insufficient detail provided to establish feasibility, sustainability and future-proofing.[5]

[1]European Parliament, "Civil Law Rules on Robotics, European Parliament Resolution of 16 February 2017 with Recommendations to the Commission on Civil Law Rules on Robotics", 2017.
[2]Rodrigues et al., SHERPA D3.3, op.cit., 2019, p.143.
[3]Ibid.
[4]European Parliament, "Follow up to the European Parliament Resolution of 16 February 2017 on Civil Law Rules on Robotics", February 2017.
[5]Rodrigues et al., SHERPA D3.3, op.cit., 2019, p.143.

Name of proposed body and proposer	**The EU Taskforce of field specific regulators for AI/big data[1]** **Proposer: SHERPA Stakeholder Board** **Year: 2019**
Level /country	EU-level
What does it regulate	AI and big data
Gap filled/need met	It might help address shortcomings in the areas of cooperation, coordination, consistent application of Union law related to AI/big data, also, e.g., cross-border risks from AI and big data applications.[2]
Identified pros	It will promote cooperation on AI/big data legal issues and provide clarity at the EU-level. The task force could create a good collaborative environment for EU AI policy and regulation and promote the adoption of a unified message on AI/big data regulation to the extent possible/required.[3]
Identified cons	Task forces are limited often by the capacity of their members. If established without a clear mandate, it might duplicate the work of existing EU agencies. It might cause further frustrations amongst stakeholders. Changing regulatory culture of the EU. Managing conflicts, limited resource. Funding issues and personnel turnover.[4]
Final assessment	Feasibility and sustainability depend on internal and external buy-in and EU political will to create and keep it going. It might also be affected by competing priorities of the different bodies that might be expected to house and/or form it. The task force would also need to allay the concern of participating bodies that such participation may conflict with their primary mission.[5]

[1]Rodrigues et al., SHERPA D3.3, op.cit., 2019, p.53, 159.
[2]Rodrigues et al., SHERPA D3.3, op.cit., 2019, p.159.
[3]Ibid.
[4]Rodrigues et al., SHERPA D3.3, op.cit., 2019, p.159.
[5]Ibid.

[1] Pros, cons and assessments of each proposal are drawn, as referenced, from SHERPA report D3.3 (Rodrigues et al., op. cit., 2019) and SIENNA report D4.2 (Rodrigues et al, op. cit., 2019) There is no novel analysis in this summary.

Name of proposed body and proposer	**Network of national authorities, as well as sectoral networks and regulatory authorities, at national and EU level.**[1] **Proposer: European Commission** **Year: 2020**
Level/country	EU-level
What does it regulate	AI
Gap filled/need met	Current fragmentation of responsibilities and standards, increase capacity in Member States, and make sure that Europe equips itself progressively with the capacity needed for testing and certification of AI-enabled products and services. "Some specific features of AI (e.g. opacity) can make the application and enforcement of existing legislation more difficult. For this reason, there is a need to examine whether current legislation is able to address the risks of AI and can be effectively enforced, whether adaptations of the legislation are needed, or whether new legislation is needed."[2] Current absence of a common European framework.[3] Mandatory conformity assessment for high-risk AI applications.[4] A forum for a regular exchange of information and best practice, identifying emerging trends, advising on standardisation activity as well as on certification. It should also play a key role in facilitating the implementation of the legal framework, such as through issuing guidance, opinions and expertise.[5]
Identified pros	Would support a European governance structure on AI in the form of a framework for cooperation of national competent authorities. A clear European regulatory framework would build trust amongst consumers and businesses in AI.[6] A solid European regulatory framework for trustworthy AI will protect all European citizens and help create a frictionless internal market for the further development and uptake of AI as well as strengthening Europe's industrial basis in AI.[7]
Identified cons	Conformity assessment could be burdensome on SMEs.[8]
Final assessment	Feasible: Given existing European frameworks for oversight of related issues, and calls for increased regulatory oversight by large member states (France and Germany), this could be a feasible option. Sustainable: "Such a regulatory framework should be consistent with other actions to promote Europe's innovation capacity and competitiveness in this field."[9] "It should establish close links with other EU and national competent authorities in the various sectors to complement existing expertise and help existing authorities in monitoring and the oversight of the activities of economic operators involving AI systems and AI-enabled products and services."[10] "Must ensure socially, environmentally and economically optimal outcomes and compliance with EU legislation, principles and values."[11] Future-proof: "Given how fast AI is evolving, the regulatory framework must leave room to cater for further developments. Any changes should be limited to clearly identified problems for which feasible solutions exist."[12] "Particular account should be taken of the possibility that certain AI systems evolve and learn from experience, which may require repeated assessments over the life-time of the AI systems in question."[13]

[1]European Commission, White Paper, op. cit., 2020.
[2]European Commission, White Paper, op. cit., 2020, p.10.
[3]Ibid.
[4]European Commission, White Paper, op. cit., 2020, p.23.
[5]European Commission, White Paper, op. cit., 2020, p.24.
[6]European Commission, White Paper, op. cit., 2020, p.10.
[7]Ibid.
[8]European Commission, White Paper, op. cit., 2020, p.23.
[9]European Commission, White Paper, op. cit., 2020, p.10.
[10]European Commission, White Paper, op. cit., 2020, p.24.
[11]European Commission, White Paper, op. cit., 2020, p.10.
[12]Ibid.
[13]European Commission, White Paper, op. cit., 2020, p.23.

Name of proposed body and proposer	**Digital Authority**[1] **Proposer: UK House of Lords** **Year: 2019**
Level/country	UK
What does it regulate	AI - does not have independent monitoring, oversight or enforcement mechanisms.
Gap filled/need met	Assessment of regulation in the digital world and recommendations where additional powers are necessary to fill gaps; Establish an internal centre of expertise on digital trends which helps to scan the horizon for emerging risks and gaps in regulation; Help regulators to implement the law effectively and in the public interest; Inform Parliament, the Government and public bodies of technological developments; Provide a pool of expert investigators to be consulted by regulators for specific investigations; Survey the public to identify how their attitudes to technology change over time, and to ensure that the concerns of the public are taken into account by regulators and policy-makers; Raise awareness of issues connected to the digital world amongst the public; Ensure that human rights and children's rights are upheld in the digital world; Liaise with European and international bodies responsible for internet regulation, addressing shortcomings in cooperation, coordination, consistent application of EU law related to AI/big data.[2]
Identified pros	It is expected to help regulators to implement the law effectively and in the public interest and bring a new consistency and urgency to regulation. It could help eliminate overlaps in legislation.[3] One of its key proposed functions includes ensuring that human rights and children's rights are upheld in the digital world.
Identified cons	May become over-prescriptive.[4] The expense involved in setting up a new body.[5] Overregulation of the digital world.[6] Mission creep.[7]
Final assessment	Proposal details unclear (with respect to its functions, instruction remit, relationships with other bodies).[8] No independent monitoring, oversight and enforcement mechanisms.[9] Effectiveness will depend on proper funding, ability to coordinate and instruct different regulators, ability to remain politically impartial and independent of the Government, and democratic scrutiny.[10] Its sustainability will depend on the policy and funding model adopted and its usefulness in regulating the digital world.[11]

[1]House of Lords, Regulating in a Digital World: 2nd Report of Session 2017–19. Report, 2019. https://publications.parliament.uk/pa/ld201719/ldselect/ldcomuni/299/299.pdf.

[2]House of Lords, op. cit., 2019.

[3]Hill, Rebecca, "UK Peers Suggest One Big "Digital Authority" to Watch the Tech Watchers, Tighten up Regulation", *The Register*, 9 March 2019. https://www.theregister.co.uk/2019/03/09/lords_communications_committee_internet_regulation/.

[4]Rodrigues et al., SHERPA D3.3, op.cit., 2019, p.55.

[5]Rodrigues et al., SHERPA D3.3, op.cit., 2019, p.218.

[6]Ibid.

[7]Rodrigues et al., SHERPA D3.3, op.cit., 2019, p.218.

[8]Rodrigues et al., SHERPA D3.3, op.cit., 2019, p.43.

[9]Rodrigues et al., SHERPA D3.3, op.cit., 2019, p.46.

[10]Rodrigues et al., SHERPA D3.3, op.cit., 2019, p.61.

[11]Rodrigues et al., SHERPA D3.3, op.cit., 2019, p.218.

Name of proposed body and proposer	**New statutory duty of care for online harms**[1] **Proposer: UK Government - White Paper** **Year: 2019**
Level/country	UK
What does it regulate	AI - A new regulatory scheme regarding online harmful user-generated content ("UGC"), to be managed by a new independent regulator.[2]
Gap filled/need met	The new regulatory scheme would apply to entities that offer services or tools that allow users to: • share or discover UGC, or • interact with each other online. Currently most of the covered entities are only required to remove or restrict harmful content when they are notified or become aware of it, but not to take proactive steps to monitor for harmful content.[3]
Identified pros	Designed to reduce illegal, dangerous and otherwise harmful UGC and user interactions. Achieving this goal supports the right to life, freedom from slavery, the right not to be discriminated against, and the right to participate in free elections. The regulatory scheme is intended to reduce deception, intimidation, and other content or activities that promote harm to individuals or groups. Implementation will increase market demand for (and spur further development of) tools to analyse content and online behaviour.
Identified cons	This regulatory scheme does not directly regulate smart information systems. (AI is one tool that covered entities may use to monitor for harmful UGC.)[4] Excessive monitoring and control of UGC can jeopardize free expression, freedom of assembly, and the right to privacy.[5] Places too much faith in technology to help covered entities comply with the oversight requirements.[6] Lacks a clear delineation of legal but "harmful" content to be regulated.[7] Does not identify responsibility for oversight of the regulator.[8]
Final assessment	Feasible: Insufficiently defined to assess. Sustainable: Insufficiently defined to assess. Future-proof: The regulator would have the ability to update codes of practice to reflect changes in technology, but rapidly changing technology may make it challenging for the regulator and covered entities to keep up.[9]

[1]Department for Digital, Culture, Media and Sport and Home Office, "Online Harms White Paper", White Paper (House of Commons, 8 February 2019). https://www.gov.uk/government/consultations/online-harms-white-paper.

[2]Ibid.

[3]Department for Digital, Culture, Media and Sport and Home Office, op. cit., 2019.

[4]Digital Action, "Online Harms White Paper: Seven Expert Perspectives", April 2019.

[5]Digital Action, op. cit., 2019.

[6]Ibid.

[7]Digital Action, op. cit., 2019.

[8]Ibid; Smith, Graham, "Cyberleagle: Users Behaving Badly – the Online Harms White Paper", 18 April 2019, https://www.cyberleagle.com/2019/04/users-behaving-badly-online-harms-white.html.

[9]Rodrigues et al., SHERPA D3.3, op.cit., p.199.

Name of proposed body and proposer	**FDA for Algorithms**[1] **Proposer: Andrew Tutt** **Year: 2017**
Level/country	USA
What does it regulate	AI - A new specialist federal-level regulatory agency to be created to regulate algorithmic safety. (1) to organize and classify algorithms into regulatory categories by their design, complexity, and potential for harm (in both ordinary use and through misuse). (2) to prevent the introduction of algorithms into the market until their safety and efficacy has been proven through evidence-based pre-market trials. (3) broad authority to impose disclosure requirements and usage restrictions to prevent algorithms' harmful misuse.[2]
Gap filled/need met	Gaps in current remedies offered by tort/civil and criminal law.[3]
Identified pros	Could develop comprehensive policy.[4] Could quickly respond to new products and practices.[5] Could also ensure that consumers are adequately protected.[6] Could add significant value in the centralized expertise.[7]
Identified cons	As covered entities are not required to make assessments public or to report assessments to authorities, it's unclear how the FTC or state AGs would become aware of the need for assessments.[8] Negative impact on innovation.[9] Resource constraints.[10] Too soft or too tough a mandate.[11] Challenge in determining what is excessive and/or insufficient regulation.[12] Excessive regulatory authority.[13] Addressing any internal knowledge gaps.[14]
Final assessment	The US FDA is part-funded by federal budget authorization and the other part is paid for by industry user fees. Sustainability of the proposed (new) FDA for algorithms will have to be similarly ensured and guaranteed.[15] It is susceptible to policy changes (e.g., deregulation) and the restriction of its powers by changes to policy/legislation.[16] An FDA for algorithms would need a depth of technical know-how, and a rich diversity of expertise to grasp the breadth of society; it would also need distinct trigger points on when to review and at what level of scrutiny.[17]

[1] Tutt, Andrew, "An FDA for Algorithms", *Administrative Law Review* 69, no. 1, 2017, pp. 83–124, https://doi.org/10.2139/ssrn.2747994 ; Groth, Olaf J., Mark J. Nitzberg, and Stuart J. Russell, "AI Algorithms Need FDA-Style Drug Trials", *Wired*, 15 August 2019. https://www.wired.com/story/ai-algorithms-need-drug-trials/.

[2] Tutt, op. cit., 2017.

[3] Rodrigues et al., SHERPA D3.3, op. cit., 2019, p.230.

[4] Tutt, op. cit., 2017.

[5] Tutt, op. cit., 2017.

[6] Ibid.

[7] Tutt, op. cit., 2017.

[8] Ibid.

[9] Tutt, op. cit., 2017.

[10] Rodrigues et al., SHERPA D3.3, op. cit., 2019, p.230.

[11] Ibid.

[12] Rodrigues et al., SHERPA D3.3, op. cit., 2019, p.230.

[13] Ibid.

[14] Rodrigues et al., SHERPA D3.3, op. cit., 2019, p.230.

[15] Ibid.

[16] Rodrigues et al., SHERPA D3.3, op. cit., 2019, p.230.

[17] Ibid.

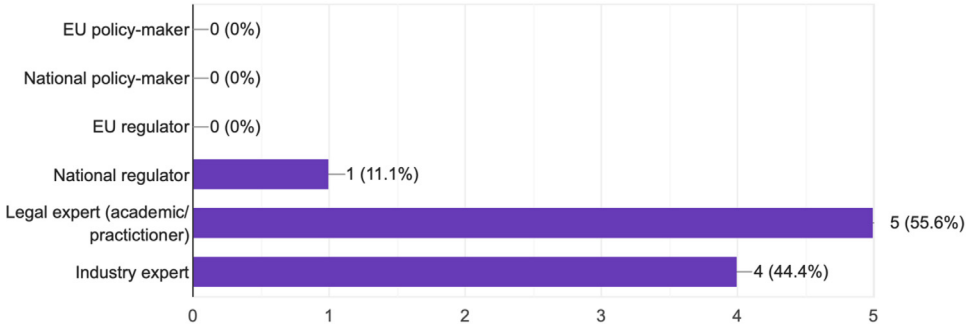
Name of proposed body and proposer	**US Federal Trade Commission to regulate robotics[1]** **Proposer: Various, Woodrow Hartzog** **Year: 2015**
Level/country	USA
What does it regulate	AI - Federal Trade Commission ("FTC") to be given primary responsibility for overseeing regulation of autonomous systems.[2]
Gap filled/need met	Autonomous systems are currently regulated (inconsistently) by multiple federal agencies based on their function.[3]
Identified pros	Could build a rich cross-industry knowledge base and experience base for regulation of a wide spectrum of autonomous systems, avoiding knowledge "silos".[4]
Identified cons	Regulation of autonomous systems in highly specialised environments (such as medical uses) or in environments presenting risk of injury or death to bystanders (drones and autonomous vehicles) may require specialized knowledge that is already in place in other agencies.[5] The FTC's jurisdiction does not extend to federally regulated financial institutions, common carriers, or non-profit organisations.[6] The FTC does not have the power to approve or certify medical devices, passenger vehicles or aircraft.[7]
Final assessment	Feasible: Yes (the FTC already regulates a wide variety of businesses).[8] Sustainable: Yes, supported by policy.[9] Future-proof: Possibly. Since its establishment in 1914, the FTC has shown the ability to adapt its regulatory approach to new technologies and new issues. However, this is primarily a political choice and will require political will to be adopted and implemented.[10]

[1]Hartzog, Woodrow, "Unfair and Deceptive Robots", *Maryland Law Review* 74, no. 4, 2 June 2015, p.785.
[2]Hartzog, op. cit., 2015.
[3]Rodrigues et al., SHERPA D3.3, op. cit., 2019, p.235.
[4]Rodrigues et al., SHERPA D3.3, op. cit., 2019, p.235.
[5]Ibid.
[6]Rodrigues et al., SHERPA D3.3, op. cit., 2019, p.235.
[7]Rodrigues et al., SHERPA D3.3, op. cit., 2019, p.235.
[8]Rodrigues et al., SHERPA D3.3, op. cit., 2019, p.52.
[9]Rodrigues et al., SHERPA D3.3, op. cit., 2019, p.235.
[10]Rodrigues et al., SHERPA D3.3, op. cit., 2019, p.62.

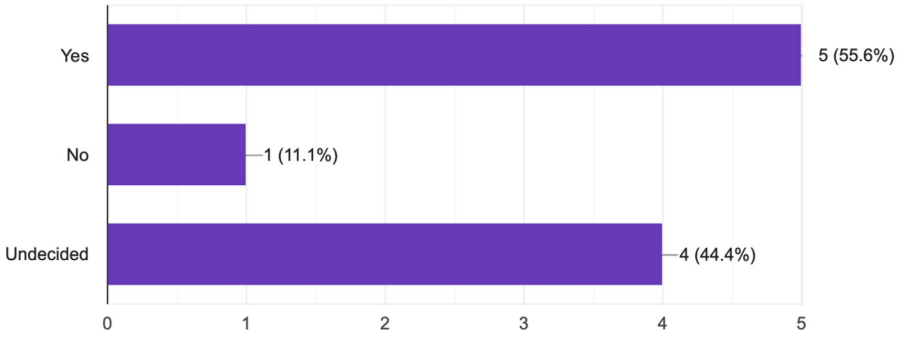
Annex B: Focus Group: New AI regulator: questionnaire

The objective of the focus group questionnaire was to gather thoughts on the feasibility of a bespoke new regulator for AI and big data at the EU-level and determine ToR. It was administered using Google Forms in advance of the focus group meeting held on 26 June 2020. The results and discussion fed into this report, particularly the development of the ToR. Nine participants completed the questionnaire and the raw results are presented here.

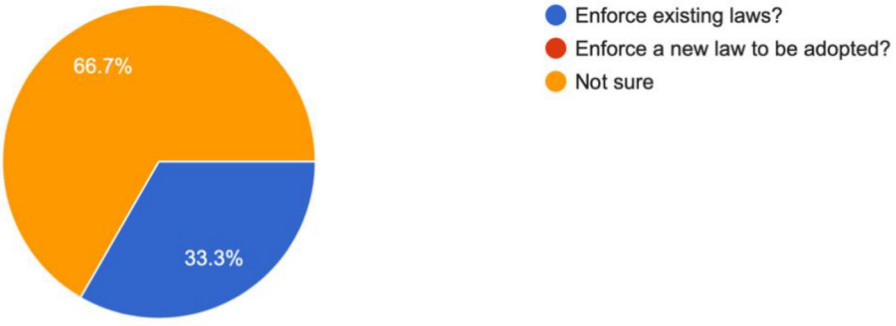
You are:
9 responses



Question 1: Do we need a new or bespoke regulator for AI and big data at the EU level?
9 responses



Question 2: Should the new regulator:
9 responses



Question 3: What should its mandate be? E.g., should it protect specific human rights? Or regulate particular applications? Or particular industries? Regulate use and implementation? 9 responses

All of the above

Protecting human rights, operating as a regulatory advisor, and collaborating in creating best practices for the industry ("light standardization")

Differentiate between more/less consequential applications. And regulate more consequential applications, e.g., in health, judiciary, security, finance, employment. Build all rights in the Charter of Fundamental Rights of the EU.

At one hand it should cover the possibility for developing models. I don't believe there is shortcoming in privacy-legislation.

Assess a new or existing technology in view of its potential (or actual) impact on society and individuals with the aim to regulate its distribution and use

Protect all kinds of rights but also the emergence of a strong AI & big data european industry

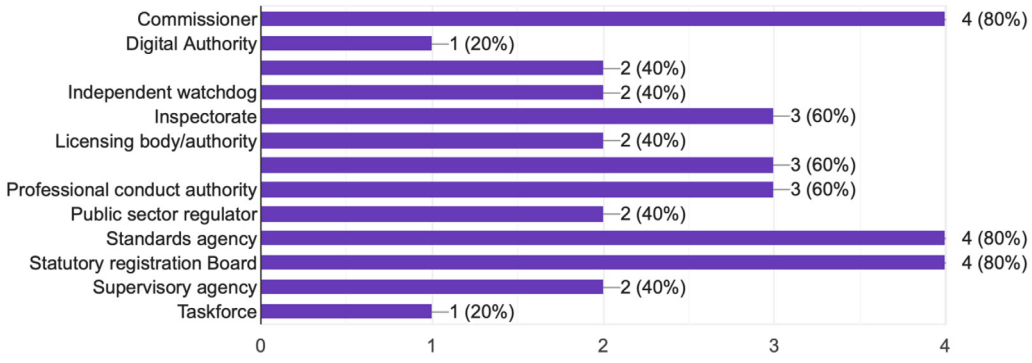
A new regulator is undesirable given the fuzzyness of the target. Three years ago we worried a lot about Big Data, now it is AI and Big Data, tomorrow it will be called differently. What it is that needs monitoring/regulating is unclear. Focus on technology is wrong. I am by far convinced a new regulator is required, hence the questions below are premature

Regulate implementation of AI, and conduct audits (notably) on the basis of complaints from citizens/end users/customers/etc.

The main problem is not Privacy, Security, Artificial soldiers, Anthropomorphism, Fake news, Ethics etc. Really important problems but used for covering the most important and underlying ones; and for misleading us: The AI revolution (economic, social, cultural, ..) is empowering whom?? Ai is not scientifically orientated but servant of business; it's creating unbelievable powers over society and new capitalistic power concentration.

Question 4: Which of the following types of bodies would NOT be a good model for a new AI regulator at the EU-level? Please tick all that should not be considered.

5 responses



Question 5: How should the new regulatory body be constituted? Who should its members be?

7 responses

policy makers, legal experts, AI academicians, industry experts

Academics, policymakers/regulators, industry experts and civil society organizations should for an advisory board for the regulator. Data Protection Agency model could be working, if those agencies would stay efficient and agile but well-resourced (not the case with the DP agencies, atm)

Not sure at the moment.

Specific AI regulators/experts, technology developers (the industry), end users (multiple and with different background), lawyers (human-rights, customer rights, and any other relevant legal-counsellor), any potential stakeholder which could be impacted by the technology at stake

It should be integrated inside an already existing EU institution legal experts and data scientists, alongside with a group of sector experts specific to each case at stake.

Not only business men, technologists, scientists, politicians but also social stakeholders (social movements (like women movement), trade unions, philosophers, …)

Question 6: What would its role and functions be? 7 responses

1) To identify the criteria to be taken into account during the design, evaluation, and use of AI systems, both general, and application/domain/industry specific ones, and 2) to inspect (and report) the use of AI systems in public and private sector.

To enforce and supervise the selected application areas/themes. For example, face recognition could be one area/theme. The amount of areas/themes should be limited to a few to keep it efficient, and the themes could be updated over time.

Not sure at the moment.

Ensure all outcomes or uses of a technology are accounted for or predicted. This should be part of a risk-assessment process. Based upon multiple analyses conducted by the regulator (or body of regulators), it should be ensure that the technology is compliant with and does not breach any of the EU/UN human rights. The use and implementation of this technology should be regulated according to the results of these (and other) analyses in order to guarantee the beneficial impact of the technology on society.

Monitor industry behaviour/practices

Regulate implementation of AI tools (as much in the private as in the public sector), making audits and investigations, raising awareness and teaching citizens/end users/customers etc. about the risks, benfits, and main principles of AI.

Identify the emergent problems (ethical, political, and social); analyse them, discuss possible causes and solutions; proposals of interventions

Question 7: How should it be governed? To whom would it report? 6 responses

-

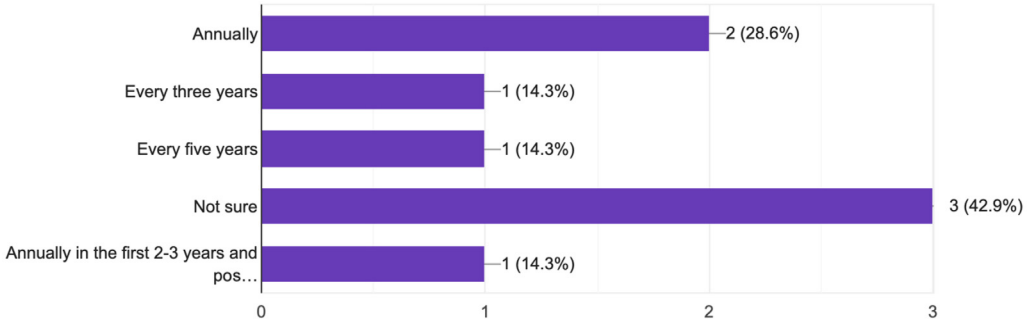
Not sure at the moment.

Members of states and European parliament.

EC it should be an independant authority, reporting (accountability) to the public at large, and responsible before the Commission.

To scientifico associations and universities; government; public opinion

Question 8: How often should its terms of reference be reviewed?
7 responses



Question 9: Are there existing regulators for other issues (e.g., data protection, health) that are good models? 7 responses

GDPR and Health related ones may be good examples

Data protection agencies are operating with a relatively good model, but might need a bit more extra resources.

Don't know.

Health/pharma regulatory models

No information to answer

Member states' data protection authorities are a good example, but such a regulator should be primarily constituted at EU level with subsidiaries in Member States (i.e. more decisionnal power than EDPB, for instance).

I don't know

Question 10: What best practices could/should be carried over to a bespoke regulator for AI?
5 responses

I would emphasize collaboration with the universities and civil society. Instead of having straightforward regulatory agency, the best practices should be created and maintained with the close collaboration with other organizations.

Don't know.

Supervison and regulation

N/A

Discussion on these issues in AI conference and meeting; in the AI courses for students. Meeting of AI people with social scientists, humanists, philosophers,…

Question 11: What would be some challenges and barriers to its success? 6 responses

To grow too bureautic. The agency should stay relatively small, compared to the size of each country. It should emphasize open collaboration and discussion with different parties, instead of being buried in the "regulatory machine".

How can the EU regulate technologies that are largely developed in the US and China? difficulties in fully understanding and predicting the use of a technology risk assessment (understanding the side effects of technology) communication amongst stakeholders could be hampered by the different walks of life and vocabulary Potential lack of transparency from stakeholders or misunderstandings regarding the use and potential application

Asymmetric regulation, enforcement, other regions no regulation

Access to, and understanding of, relevant information from undertakings under investigation; granting such powers at EU level for Member States; having enough staff to deal with sufficient number of cases/year to be efficient.

To reduce the concentration of economic and political powers over society; to increase people awareness; to readdress research towards knowledge as priority, not technology.

Question 12: Are there any other policy or other considerations that need to be taken into account? 6 responses

-

How can the EU promote the development of AI by EU member states? See, e.g. CLAIRE (https://claire-ai.org/). potential influence on other (unforeseen) fields (for instance, an AI application for the healthcare sector might also impact or have a spill-over effect on other sector of society)

European industry and jobs

A clear legal framework on AI is needed (new laws or clear interpretation of existing ones in relation to AI) for such a regulator to be able to manage its tasks.

AI is an anthropological radical revolution of human mind, sociality, knowledge, powers.

REFERENCES

Access Now Policy Team. The Toronto declaration: protecting the right to equality and non-discrimination in machine learning systems; 2018 Toronto: Access No. Retrieved from https://www.accessnow.org/cms/assets/uploads/2018/08/The-Toronto-Declaration_ENG_08-2018.pdf.

ACM. Statement on algorithmic transparency and accountability; 2017.

AI HLEG. Ethics guidelines for trustworthy AI. Brussels: European Commission - Directorate-General for Communication; 2019a Retrieved from https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai.

AI HLEG. Policy and investment recommendations for trustworthy Artificial Intelligence. Brussels: European Commission - Directorate-General for Communication; 2019b Retrieved from https://ec.europa.eu/digital-single-market/en/news/policy-and-investment-recommendations-trustworthy-artificial-intelligence.

AIEI Group. (2020). From Principles to Practice - An Interdisciplinary framework to operationalise AI ethics, VDE /Bertelsmann Stiftung, p. 56. Retrieved from https://www.ai-ethics-impact.org/resource/blob/1961130/c6db9894ee73aefa489d6249f5ee2b9f/aieig—report—download-hb-data.pdf

Aristotle. The Nicomachean ethics. Filiquarian Publishing, LLC; 2007.

Babuta A, Oswald M, Janjeva A. Artificial Intelligence and UK National Security - Policy Considerations. Royal United Services Institute for Defence and Security Studies; 2020 (Occasional Paper)Retrieved from https://rusi.org/sites/default/files/ai_national_security_final_web_version.pdf.

Berendt B. AI for the Common Good?! Pitfalls, challenges, and ethics pen-testing, *Paladyn*. J Behav Robot 2019;10(1):44–65.

BmVI. Ethik-Kommission: Automatisiertes and Vernetztes Fahren. Bundesministerium für Verkehr und digitale Infrastruktur; 2017 Retrieved from https://www.bmvi.de/SharedDocs/DE/Anlage/Presse/084-dobrindt-bericht-der-ethik-kommission.pdf?__blob=publicationFile.

Boden MA. Artificial intelligence: a very short introduction. Oxford, United Kingdom: OUP Oxford; 2018 Reprint edition.

Brundage, M., Avin, S., Clark, J., Toner, H., Eckersley, P., Garfinkel, B. et al.. (2018). The malicious use of artificial intelligence: forecasting, prevention, and mitigation, *arXiv:1802.07228 [cs]*. Retrieved from http://arxiv.org/abs/1802.07228

Bynum TW. Flourishing Ethics. Ethics Inf Technol 2006;8(4):157–73.

CDEI. *Interim report: review into bias in algorithmic decision-making*. Centre for Data Ethics and Innovation; 2019 Retrieved from https://www.gov.uk/government/publications/interim-reports-from-the-centre-for-data-ethics-and-innovation/interim-report-review-into-bias-in-algorithmic-decision-making.

Clarke R. Privacy impact assessment: its origins and development. Comput Law Secur Rev 2009;25(2):123–35.

Clarke R. Principles and business processes for responsible AI. Comput Law Secur Rev 2019;35(4):410–22.

Clarke R. Regulatory alternatives for AI. Comput Law Secur Rev 2019;35(4):398–409.

Clarke YD. All Info - H.R.2231 - 116th Congress (2019-2020): Algorithmic Accountability Act of 2019; 2019c Retrieved August 28, 2021, from https://www.congress.gov/bill/116th-congress/house-bill/2231/all-info.

Clarke YD. H.R.3230 - 116th Congress (2019-2020): defending each and every person from false appearances by keeping exploitation subject to accountability Act of 2019; 2019d Retrieved August 28, 2021, from https://www.congress.gov/bill/116th-congress/house-bill/3230.

Coeckelbergh M. Technology, narrative and performance in the social theatre. In: Kreps D, editor. Understanding digital events: bergson, whitehead, and the experience of the digital. New York: Routledge; 2019a. p. 13–27.

Coeckelbergh M. Artificial Intelligence: some ethical issues and regulatory challenges. Technol Regul 2019b:31–4.

Coeckelbergh M. AI for climate: freedom, justice, and other ethical and political challenges. AI Ethics 2020a Retrieved from. doi:10.1007/s43681-020-00007-2.

Coeckelbergh M. AI Ethics. The MIT Press; 2020b.

Council of Europe. Unboxing Artificial intelligence: 10 steps to protect human rights; 2019 Retrieved from https://www.coe.int/en/web/commissioner/view/-/asset_publisher/ugj3i6qSEkhZ/content/unboxing-artificial-intelligence-10-steps-to-protect-human-rights.

Council of Europe. CAHAI - Ad hoc Committee on Artificial Intelligence. Artif Intell 2020. Retrieved January 22, 2021, from https://www.coe.int/en/web/artificial-intelligence/cahai.

Council of the European Union. *Presidency conclusions - the charter of fundamental rights in the context of artificial intelligence and digital change* (No. 11481/20). Brussels: Council of the European Union; 2020 Retrieved from https://www.consilium.europa.eu/media/46496/st11481-en20.pdf.

Datenethikkommission. Gutachten Der Datenethikkommission - Kurzfassung. Berlin: BMI; 2019 Retrieved from https://www.bmi.bund.de/SharedDocs/downloads/DE/publikationen/themen/it-digitalpolitik/gutachten-datenethikkommission-kurzfassung.pdf;jsessionid=965B44F9D6794C002FA97D8D6CBF8E8D.1_cid295?__blob=publicationFile&v=4.

Dignum V. Responsible artificial intelligence: how to develop and use ai in a responsible way. 1st ed. Springer; 2019 2019 edition.

EDPS. Artificial intelligence, robotics, privacy and data protection (Room document for the 38th international conference of data protection and privacy commissioners). Marrakech: EDPS; 2016 Retrieved from https://edps.europa.eu/sites/edp/files/publication/16-10-19_marrakesh_ai_paper_en.pdf.

EDPS. EDPS Opinion on the European Commission's White Paper on Artificial Intelligence – A European approach to excellence and trust (Opinion 4/2020) (Opinion No. 4/2020), EDPS; 2020 Retrieved from https://edps.europa.eu/sites/edp/files/publication/20-06-19_opinion_ai_white_paper_en.pdf.

Elsevier. Artificial intelligence: how knowledge is created, transferred, and used - trends in China, Europe, and the United States. Amsterdam: Elsevier; 2018 Retrieved from https://www.elsevier.com/?a=827872.

Erdélyi OJ, Goldsmith J. Regulating artificial intelligence: proposal for a global solution. Proceedings of the 2018 AAAI/ACM conference on AI, ethics, and society. New York, NY, USA: Association for Computing Machinery; 2018. p. 95–101 Retrieved from https://doi.org/10.1145/3278721.3278731.

European Commission. Digital Government Factsheets - 2019 | Joinup; 2019 Retrieved January 22, 2021, from https://joinup.ec.europa.eu/collection/nifo-national-interoperability-framework-observatory/digital-government-factsheets-2019.

European Commission. White Paper on Artificial Intelligence: a European approach to excellence and trust (White paper No. COM(2020) 65 final), Brussels; 2020a Retrieved from https://ec.europa.eu/info/files/white-paper-artificial-intelligence-european-approach-excellence-and-trust_en.

European Commission. AI Watch | Knowledge For Policy. European Commission; 2020b Retrieved January 23, 2021, from https://knowledge4policy.ec.europa.eu/ai-watch_en.

European Commission. *Proposal for a regulation on a European approach for artificial intelligence* (No. COM(2021) 206 Final). Brussels: European Commission; 2021a Retrieved from https://digital-strategy.ec.europa.eu/en/library/proposal-regulation-european-approach-artificial-intelligence.

European Commission. Artificial Intelligence January 8. Shaping Europe's Digital Future - European Commission; 2021b Retrieved January 30, 2021, from https://ec.europa.eu/digital-single-market/en/artificial-intelligence.

European Parliament. Civil Law Rules on Robotics - European Parliament resolution of 16 February 2017 with recommendations to the Commission on Civil Law Rules on Robotics (2015/2103(INL)), , Pub. L. No. P8_TA(2017)0051; 2017 Retrieved from http://www.europarl.europa.eu/sides/getDoc.do?pubRef=-//EP//NONSGML+TA+P8-TA-2017-0051+0+DOC+PDF+V0//EN.

European Parliament. *DRAFT REPORT with recommendations to the commission on a framework of ethical aspects of artificial intelligence, robotics and related technologies* (No. 2020/2012(INL)). European Parliament, Committee on Legal Affairs; 2020 Retrieved from https://www.europarl.europa.eu/meetdocs/2014_2019/plmrep/COMMITTEES/JURI/PR/2020/05-12/1203395EN.pdf.

Finn RL, Wright D. Privacy, data protection and ethics for civil drone practice: a survey of industry, regulators and civil society organisations. Comput Law Secur Rev 2016;32(4):577–86.

Fjeld J, Achten N, Hilligoss H, Nagy A, Srikumar M. Principled artificial intelligence: mapping consensus in ethical and rights-based approaches to principles for AI; 2020 Retrieved from https://dash.harvard.edu/handle/1/42160420.

Griggs D, Stafford-Smith M, Gaffney O, Rockström J, Öhman MC, Shyamsundar P, et al. Sustainable development goals for people and planet. Nature 2013;495(7441):305–7.

Hagendorff T. The ethics of AI ethics – an evaluation of guidelines; 2019 arXiv:1903.03425 [cs, stat]. Retrieved from http://arxiv.org/abs/1903.03425.

Hall W, Pesenti J. Growing the artificial intelligence industry in the UK. London: Department for Digital, Culture, Media & Sport and Department for Business, Energy & Industrial Strategy; 2017 Retrieved from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/652097/Growing_the_artificial_intelligence_industry_in_the_UK.pdf.

Haque A, Milstein A, Fei-Fei L. Illuminating the dark spaces of healthcare with ambient intelligence. Nature 2020;585(7824):193–202.

Horizon 2020 Commission Expert Group to advise on specific ethical issues raised by driverless mobility (E03659). Ethics of connected and automated vehicles : recommendations on road safety, privacy, fairness, explainability and responsibility; 2020 Retrieved from http://urn.kb.se/resolve?urn=urn:nbn:se:hb:diva-23829.

House of Commons Science and Technology Committee. Robotics and artificial intelligence; 2016 Retrieved from http://www.publications.parliament.uk/pa/cm201617/cmselect/cmsctech/145/145.pdf.

House of Lords, H. of L. AI in the UK: ready, willing and able?. London: Select Committee on Artificial Intelligence; 2018 Retrieved from https://publications.parliament.uk/pa/ld201719/ldselect/ldai/100/100.pdf.

IEEE. The IEEE Global initiative on ethics of autonomous and intelligent systems; 2017 Retrieved February 10, 2018, from https://standards.ieee.org/develop/indconn/ec/autonomous_systems.html.

International Telecommunication Union. AI for good global summit report 2017. Geneva: International Telecommunication Union; 2017 Retrieved from https://www.itu.int/en/ITU-T/AI/Documents/Report/AI_for_Good_Global_Summit_Report_2017.pdf.

Iordanou K, Christodoulou E, Antoniou J. *SHERPA Deliverable 4.2: evaluation Report* (SHERPA project deliverable No. 4.2); 2020 Retrieved from.

Jobin A, Ienca M, Vayena E. The global landscape of AI ethics guidelines. Nat Mach Intell 2019;1(9):389–99.

Kaplan A, Haenlein M. Siri, Siri, in my hand: who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence. Bus Horiz 2019;62(1):15–25.

Kuhlmann S, Stegmaier P, Konrad K. The tentative governance of emerging science and technology—a conceptual introduction. Res Policy 2019;48(5):1091–7.

Kurzweil R. The singularity is near. London: Gerald Duckworth & Co Ltd; 2006.

LaGrandeur K. How safe is our reliance on AI, and should we regulate it? AI Ethics 2020 Retrieved from. doi:10.1007/s43681-020-00010-7.

Latonero M. Governing artificial intelligence: upholding human rights & dignity. Data&Society; 2018 Retrieved from https://datasociety.net/wp-content/uploads/2018/10/DataSociety_Governing_Artificial_Intelligence_Upholding_Human_Rights.pdf.

LeCun Y, Bengio Y, Hinton G. Deep learning. Nature 2015;521(7553):436–44.

Mantelero A. Regulating big data. The guidelines of the Council of Europe in the context of the European data protection framework. Comput Law Secur Rev 2017;33(5):584–602.

Marsden C, Meyer T, Brown I. Platform values and democratic elections: how can the law regulate digital disinformation? Comput Law Secur Rev 2020;36 :105373.

Miller C, Ohrvik-Stott J. Regulating for responsible technology - capacity, evidence and redress: a new system for a fairer future. London: Doteveryone; 2018 Retrieved from https://www.doteveryone.org.uk/wp-content/uploads/2018/10/Doteveryone-Regulating-for-Responsible-Tech-Report.pdf.

Mittelstadt B. Principles alone cannot guarantee ethical AI. Nat Mach Intell 2019. Retrieved from https://papers.ssrn.com/abstract=3391293.

Morley, J., Floridi, L., Kinsey, L. and Elhalal, A. (2019). From what to how- an overview of AI ethics tools, methods and research to translate principles into practices, *arXive*. Retrieved from https://arxiv.org/abs/1905.06876

Muller C. The impact of artificial intelligence on human rights, democracy and the rule of law (No. CAHAI(2020)06-fin). Strasbourg: Council of Europe, Ad Hoc Committee on Artificial Intelligence (CAHAI); 2020 Retrieved from https://rm.coe.int/cahai-2020-06-fin-c-muller-the-impact-of-ai-on-human-rights-democracy-/16809ed6da.

Nemitz P. Constitutional democracy and technology in the age of artificial intelligence. Philos Trans R Soc A 2018;376(2133).

New J. How to fix the algorithmic accountability act. Center for Data Innovation; 2019 Retrieved from https://datainnovation.org/2019/09/how-to-fix-the-algorithmic-accountability-act/.

Ochigame R. The invention of "Ethical AI": how big tech manipulates academia to avoid regulation. Intercept 2019. Retrieved from https://theintercept.com/2019/12/20/mit-ethical-ai-artificial-intelligence/.

OECD. *Recommendation of the council on artificial intelligence* (OECD Legal Instruments). OECD; 2019 Retrieved from https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449.

Office for Artificial Intelligence. About us, GOV.UK; 2020 Retrieved August 28, 2021, from https://www.gov.uk/government/organisations/office-for-artificial-intelligence/about.

Rai A, Constantinides P, Sarker S. Next-generation digital platforms: toward human–AI hybrids. Mis Quart 2019;43(1):iii–x.

Rességuier A, Rodrigues R. AI ethics should not remain toothless! A call to bring back the teeth of ethics. Big Data Soc 2020;7(2) :2053951720942541.

Richards L, Brockmann K, Boulanini V. Responsible artificial intelligence research and innovation for international peace and security. Stockholm: Stockholm International Peace Research Institute. Retrieved from; 2020 https://reliefweb.int/sites/reliefweb.int/files/resources/sipri_report_responsible_artificial_intelligence_research_and_innovation_for_international_peace_and_security_2011.pdf.

Rodrigues R. Legal and human rights issues of AI: gaps, challenges and vulnerabilities. J Respons Technol 2020;4 :100005.

Rodrigues R, Panagiotopoulos A, Wright D, Hatzakis T, Laulhé Shaelou S, Grant A, et al. *SHERPA Deliverable 3.3 Report on regulatory options* (online resource No. Project deliverable). SHERPA project 2020a Retrieved from https://doi.org/10.21253/DMU.11618211.v2.

Rodrigues R, Santiago N, Macnish K, Antinou J, Wright D. *SHERPA Deliverable 3.6: Feasibility of a new regulator and proposal for a European Agency for AI* (Project Deliverable No. D3.6), SHERPA project; 2020b Retrieved from.

Rodrigues R, Siemaszko K, Warso Z. Analysis of the legal and human rights requirements for AI and robotics in and outside the EU (Deliverable of the SIENNA project No. D4.2), Zenodo; 2019 Retrieved from https://zenodo.org/record/4066812#.YAqrFej7SUk.

Samoili S, Lopez Cobo M, Gombez Gutierrez E, De Prato G, Martinez-Plumed F, Delipetrev B. AI WATCH. defining artificial intelligence. Luxembourg: Publications Office of the European Union; 2020 Retrieved from http://dx.doi.org/10.2760/382730.

Shneiderman B. Design lessons from AI's two grand goals: human emulation and useful applications. IEEE Trans Technol Soc 2020;1(2):73–82.

Smuha, N.A., Ahmed-Rengers, E., Harkens, A., Li, W., MacLaren, J., Piselli, R. et al. (2021). How the EU can achieve legally trustworthy AI: a response to the European Commission's proposal for an artificial intelligence act, *Available at SSRN*, ⟨13:italic ⟩https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3899991⟨/13:italic⟩

Spiegelhalter D. Should we trust algorithms? Harvard Data Sci Rev 2020;2(1). Retrieved from https://hdsr.mitpress.mit.edu/pub/56lnenzj.

Stone P, Brooks R, Brynjolfsson E, Calo R, Etzioni O, Hager G, et al. Artificial Intelligence and Life in 2030. One hundred year study on artificial intelligence: Report of the 2015-2016 Study Panel. Stanford, CA: Stanford University; 2016 http://ai100.stanford. edu/2016-report. Accessed: September 6: 2016.

Taddeo M, Floridi L. How AI can be a force for good. Science 2018;361(6404):751–2.

Tjoa AM, Tjoa S. The role of ICT to achieve the UN sustainable development goals (SDG). In: Mata FJ, Pont A, editors. ICT for promoting human development and protecting the environment. Springer International Publishing; 2016. p. 3–13.

Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. Nat. Med. 2019;25(1):44–56.

UK Committee on Standards in Public Life. Artificial intelligence and public standards; 2020 LondonRetrieved from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/868284/Web_Version_AI_and_Public_Standards.PDF.

UNESCO. *First version of a draft text of a recommendation on the ethics of artificial intelligence* (No. SHS /BIO / AHEG-AI / 2020/4 REV). Paris: UNESCO; 2020 Retrieved from https://unesdoc.unesco.org/ark:/48223/pf0000373434.

US FDA. Proposed regulatory framework for modifications to artificial intelligence/machine learning (AI/ML)-based software as a medical device (SaMD); 2019 Retrieved from https://www.fda.gov/media/122535/download.

US FDA. Artificial intelligence and machine learning in software as a medical device. FDA 2021. Retrieved from https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-software-medical-device.

Vallor S. Technology and the virtues: a philosophical guide to a future worth wanting. Oxford University Press; 2016.

van Hoboken J, Fathaigh RÓ. Smartphone platforms as privacy regulators. Comput Law Secur Rev 2021;41 :105557.

Veale M, Binns R, Edwards L. Algorithms that remember: model inversion attacks and data protection law. Philos Trans R Soc A 2018;376(2133) :20180083.

Vesnic-Alujevic L, Nascimento S, Pólvora A. Societal and ethical impacts of artificial intelligence: critical notes on European policy frameworks. Telecomm Policy 2020 :101961.

Villaronga EF, Golia AJ. Robots, standards and the law: rivalries between private standards and public policymaking for robot governance. COMPUT Law Secur Rev 2019;35(2):129–44.

Wagner B. Ethics as an escape from regulation: from ethics-washing to ethics-shopping. In: Bayamlioglu E, Baraliuc I, Janssens LAW, Hildebrandt M, editors. Being profiled: cogitas ergo sum. Amsterdam: Amsterdam University Press; 2018. p. 84–90.

Wallach W, Marchant G. Toward the agile and comprehensive international governance of AI and robotics [point of view]. Proc. IEEE 2019;107(3):505–8.

Wallach Wendell, Marchant GE. An agile ethical/legal model for the international and national governance of AI and robotics. In: Wallach Wendell, editor. Control and responsible innovation in the development of AI and robot. The Hastings Center; 2018. p. 45–59 Retrieved from https://www.thehastingscenter.org/wp-content/uploads/Control-and-Responsible-Innovation-FINAL-REPORT.pdf.

Walsham G. Are we making a better world with ICTs? Reflections on a future agenda for the IS field. J Inf Technol 2012;27(2):87–93.

Willcocks L. Robo-Apocalypse cancelled? Reframing the automation and future of work debate. J Inf Technol 2020;35(4):286–302.

Wolbert LS, Ruyter DJD, Schinkel A. What kind of theory should theory on education for human flourishing be? Br J Educ Stud 2019;67(1):25–39.

World Economic Forum. *Responsible use of technology* (White paper). Geneva: WEB; 2019 Retrieved from http://www3.weforum.org/docs/WEF_Responsible_Use_of_Technology.pdf.

Wren-Lewis L. Regulatory capture: risks and solutions. In: Estache A, editor. Emerging issues in competition, collusion, and regulation of network industries. CEPR; 2011. p. 147–69.

Zuboff PS. The age of surveillance capitalism: the fight for a human future at the new frontier of power. 01 ed. Profile Books; 2019.