Global competition and convergence of AI law

Mireille Hildebrandt

I. Introduction

Providing a comparative account of AI law is like nailing a pudding to the wall. AI law is very much a moving target and by the time this chapter is published, the horizon itself may have shifted. In a sense, it is not possible to provide a comparative overview, because no dedicated 'AI law' is in force today, unless we include data protection law, the regulation of algorithmic decision making or algorithmic behaviours (for instance chatbots, recruitment support or social security fraud detection) or dedicated safety laws that apply to embedded systems (for instance medical devices or aircraft). Although omnibus legislation targeting AI is not in force (yet), a race seems to be going on between China, the US and the EU to get AI legislation right (Smuha 2021). This, in itself, is an interesting observation, as the more usual approach is to highlight the race to develop, deploy and monetise AI, framing its regulation as an obstruction or a distraction.

Obviously, 'getting AI legislation right' means different things to different people, institutions and jurisdictions. A caricature sketch would probably frame this as the US creating as much free space for innovation as possible, largely relying on the competitive incentives of economic markets and funding in the context of the military. As always, actual policies are far more nuanced, see section III.2 below. To continue the caricature, one could frame China as planning for innovation that is in line with the political objectives of the Communist party, see III.3 below for a more complex and rich reality. Finally, Europe can be framed as wanting to have their cake and eat it to, meaning that the development of innovative AI 'solutions' should be both enabled and constrained by way of a legislative framework that shapes economic markets to make them conducive to efficiency, effectiveness, safety, health and fundamental rights, see III.4 below for the details.

A comparative account of AI law will thus inevitably be an account of upcoming AI law, hampered by the opacity that is inherent in the process of law-making. I do not have access to the relevant negotiations and would not claim to have read all the white papers, legislative proposals, assessment reports and advisory expert opinions to produce a coherent inventory of what is in the making on a global scale. In point of fact, taking AI systems seriously, it may have been better to follow the advice of Wolff (2019), deploying a dedicated AI system to compare relevant developments in the making of AI law, based on AI's assumed ability to aggregate and process incredible amounts of data in a neutral and disinterested manner. The only reason for me to spend time on this chapter, according to Wolff, would be that such a system has not yet been developed.¹ Once it becomes available, however, the chapter can be written by such an AI system or largely based on its output, allegedly offering a far better version of the current chapter, unhindered by my personal affinities and limitations. Wolff's naïve enthusiasm reads like a parody of all that comparative law stands for, replacing the rich complexity and diversity of its goals and methods by proxies that are more conducive to computation, while ignoring the fact that a proxy is not what it stands for and requires keen attention to what is lost in translation. It would be interesting to know what the winner of the 2000 Grand Prize of the International Academy of Comparative Law, Patrick H. Glenn (2014), would have thought of computational comparative law and what legal tradition he would find to be more conducive to such an approach. This is not to say that I will attempt to measure approaches on the scale of 'a good AI society' as others have done (Cath et al. 2018), because I believe that AI law should enable a diversity of 'good societies', based on different

¹ Perhaps this is an attempt: (Cognilytica 2021). It contains a chapter 'Measuring Laws and Regulations', which I could not access; the price of the report is 995 USD.

understandings of the public good, which implies a particular understanding of the relationship between technology, law and ethics (Hildebrandt 2020a; 2015).

The example of computational comparative law raises a number of salient questions that will be the focus of this contribution. To begin with, we need to come to terms with the question of what is meant with AI, avoiding both over- and underinclusive definitions. In the next section (II), I first explore the definitional conundrum within the research domain of AI, followed by an account of how various public bodies that prepare AI legislation aim to define the scope of their object. In section III, I provide a brief overview of the different approaches to dedicated AI legislation by the OECD, the US, China and the EU. Finally, in section IV, I discuss the Brussels effect of the EU's AI Act that may diffuse the EU approach across a variety of jurisdictions whose corporations wish to compete on the internal market of the EU.

II. Defining AI

II.1 Working definitions in the domain of computer science

In the context of computer science, we can distinguish two types of AI. One based on symbolic logic, resulting in knowledge-based systems building on prior knowledge that is formalised in the form of decision trees. These systems are deterministic by definition whether very simple or highly complex. The other type is based on various kinds of machine learning, resulting in patterns, concepts or other output that build on statistical inferences from historical or streaming training data whose distribution is assumed (and hoped) to be replicated in future data. The first type is often called 'good old fashioned AI' (GOFAI) (cf. Aronson 2003), allegedly replicating explicit human reasoning, the second 'AI the modern approach' (AIMA) (cf. Russell and Norvig 2020), allegedly replicating the unconscious workings of the human brain (Hinton 1990). Current frontline research in AI often combines both approaches in so-called 'hybridAI' (Marcus and Davis 2020). Within the research domain of artificial intelligence, AI has often been defined as the ability of computing systems to perform tasks or to develop capacities typical for human intelligence, even though this type of definition is deemed controversial (Wang 2019; Monett et al. 2020). As a proxy for AI, this definition feeds into the narrative of replacement, suggesting that replacing humans by machines will improve both efficiency and effectiveness, because machines are supposedly cheaper, faster and unimpeded by emotions, affinities or fatigue.

Another relevant distinction is that between narrow and general AI, or weak and strong AI, where narrow or weak AI is only capable of performing dedicated tasks whereas general or strong AI would be capable of performing any human or other task with ever increasing efficiency and effectiveness, for instance based on its ability to crunch unprecedented amounts of data based on interconnected layers of complex computations (artificial neural networks, deep learning). This distinction is sometimes topped by something called 'superintelligence' or 'singularity', which has generated fears of humanity being overcome and/or overruled by its own creations (Bostrom and Ryan 2015). Both general and super AI have been depicted as flawed concepts (Fjelland 2020), raising flawed expectations, perhaps distracting attention from the threats posed by the actual state of the art in AI.²

In 2019 a special issue of the Journal of General Artificial Intelligence was devoted to a working definition proposed by Pei Wang, who first quotes some of the typical definitions coined by the

² This is why advocacy for responsible AI often focuses on AI as it is now, see e.g. <https://ainowinstitute.org>.

founding fathers of AI, for instance the definition of Newell and Simon (1976, cited in Wang 2019), who took their cue from human action:

By 'general intelligent action' we wish to indicate the same scope of intelligence as we see in human action: that in any real situation behavior appropriate to the ends of the system and adaptive to the demands of the environment can occur, within some limits of speed and complexity.

Wang – in the end – proposes the following definition of intelligence, whether human or not:

Intelligence is the capacity of an information-processing system to adapt to its environment while operating with insufficient knowledge and resources.

He thereby targets AI systems with a high level of autonomy (basically artificial general intelligence), capable of successfully adapting to their environment based on uncertainties and various types of limitations. This seems to be an excellent definition, emphasizing the ability to respond to the unexpected under conditions of bounded rationality (which is mostly a bug not a feature, (cf. Gigerenzer and Selten 2001). Legislative recommendations do not and should not, however, restrict themselves to such ambitious levels of AI, as this would unnecessarily restrict the scope of application and thereby the scope of legal protection.

II.2 Legislative definitions

Surveying the distinctions, definitional camps and ongoing science wars about what AI stands for, it seems wise for legislatures to take some distance from claims made about AI systems that have not yet been developed, instead focusing on the need to both enable and constrain the development, provision and deployment of AI systems that inform current and foreseeable critical infrastructure, social media, private and public transport, housing, fraud detection, finance, smart policing, healthcare and medical research. The goal of AI legislation is not to decide on scientific controversies or to otherwise have the last word on metaphysical issues around AI. Instead, the goals seem twofold across all jurisdictions: first, to stimulate investment in trustworthy AI, while second, protecting against undesirable consequences, whether in terms of safety, health or human rights.

In this light, the proposed legislative definitions are more pragmatic than those used in the domain of AI as a research domain. Proposed AI legislation is focused on potential impact rather than committing itself to a subset of underlying methodologies. Nevertheless, it is not meant to cover just any software system; whereas it should *include* embedded systems (robotics and Internet of Things, connected cars and smart grids), it should *exclude* physical systems not embedded with software (unenhanced machines). This requires keen attention to an appropriate demarcation of the scope of the legislation. I will discuss two types of legislative definition, highlighting either impact or underlying methodologies, noting these are often combined.

As an example of the first, we can look into OECD's 2019 Recommendation AI that defines AI systems as follows:³

³ OECD Recommendation of the Council on Artificial Intelligence, adopted on 22 May 2019 (OECD Legal/0449): <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449>.

An AI system is a machine-based system that can, for a given set of human-defined objectives, make predictions, recommendations, or decisions influencing real or virtual environments. AI systems are designed to operate with varying levels of autonomy.

This clarifies a number of issues: (1) legislation should not target 'AI' in general, the research domain of AI, or inflated models of supposedly forthcoming systems, but identifiable real world systems, (2) it should only target systems that are machine-based, (3) AI systems should be understood in the context of their intended purpose as defined by their human creators, (4) they should be defined in terms of their output, notably predictions, recommendations or decisions, and (5) their impact on real or virtual environments, while (6) taking into account they can have very different levels of autonomy.

An example of the second type of definition, with more focus on underlying methodologies, would be art. (3) of the US National AI Initiative Act of 2020:⁴

The term "artificial intelligence" means a machine-based system that can, for a given set of human-defined objectives, make predictions, recommendations or decisions influencing real or virtual environments. Artificial intelligence systems use machine and human-based inputs to- (A) perceive real and virtual environments; (B) abstract such perceptions into models through analysis in an automated manner; and (C) use model inference to formulate options for information or action.

Even this definition starts with impact ('influencing real or virtual environments'), which is also a prominent component of the definition in art. 3(1) of the proposed EU AI Act:

'artificial intelligence system' (AI system) means software that is developed with one or more of the techniques and approaches listed in Annex I and can, for a given set of human-defined objectives, generate outputs such as content, predictions, recommendations, or decisions influencing the environments they interact with;

This part of the definition is inspired by the OECD 2019 Recommendations, focusing on impact ('influencing the environments they interact with'). However, in Annex I of the EU's proposed AI Act the focus shifts to underlying methodologies:

- a. Machine learning approaches, including supervised, unsupervised and reinforcement learning, using a wide variety of methods including deep learning;
- b. Logic- and knowledge-based approaches, including knowledge representation, inductive (logic) programming, knowledge bases, inference and deductive engines, (symbolic) reasoning and expert systems;
- c. Statistical approaches, Bayesian estimation, search and optimization methods.

In section IV we will return to the AI Act in more detail, including the recently updated draft proposal of the EU Council, which pays even more attention to the underlying methodologies, thus aiming to restrict the scope of application to a smaller subset of systems.

⁴ 'National Artificial Intelligence Initiative Act of 2020'. "TITLE 15 (Commerce and Trade) / CHAPTER 119/ § 9401.

III. Approaches to AI legislation

AI legislation entails more than defining what counts as AI. In this section I will discuss the explicit and implicit aims of ‘regulating’ AI, distinguishing the OECD, US, Chinese and EU approaches. The OECD approach is pivotal here, as it may have a similar effect as the OECD 1980 Recommendation concerning guidelines governing the protection of privacy and transborder flows of personal data, usually referred to as the Guidelines for Fair Information Principles (FIPs). These guidelines have inspired data protection legislation around the world, notably the EU Data Protection Directive and its successor the General Data Protection Regulation (GDPR).

Meanwhile, the Council of Europe (CoE) has instituted the 2019-2021 Ad Hoc Committee on Artificial Intelligence (CAHA), which has voted to recommend a CoE Convention on AI, based on a set of principles that will be published in 2022. This could result in a Convention signed by the CoE (48 European contracting parties, including Russia and Turkey) and e.g. the US, Japan, Australia and China, similar to the CoE Cybercrime Convention. As we have no details at the moment of writing, I will not further discuss this initiative, apart from signalling its potential high impact.

III.1 OECD

The OECD Recommendation tops the definition of AI systems with an additional set of relevant definitions that demonstrate the approach taken:

–AI system lifecycle: AI system lifecycle phases involve: i) ‘design, data and models’; which is a context-dependent sequence encompassing planning and design, data collection and processing, as well as model building; ii) ‘verification and validation’; iii) ‘deployment’; and iv) ‘operation and monitoring’. These phases often take place in an iterative manner and are not necessarily sequential. The decision to retire an AI system from operation may occur at any point during the operation and monitoring phase.

–AI knowledge: AI knowledge refers to the skills and resources, such as data, code, algorithms, models, research, know-how, training programmes, governance, processes and best practices, required to understand and participate in the AI system lifecycle.

–AI actors: AI actors are those who play an active role in the AI system lifecycle, including organisations and individuals that deploy or operate AI.

–Stakeholders: Stakeholders encompass all organisations and individuals involved in, or affected by, AI systems, directly or indirectly. AI actors are a subset of stakeholders.

The distinction between actors (those who participate in design, verification, deployment, operation and monitoring of AI systems) and the wider group of stakeholders (all those involved or affected) is pivotal, and the need for a rich type of knowledge skills and resources for the actors whose design decisions will affect stakeholders nails it when it comes to legislation focused on addressing impact (whether ex ante or ex post). The definition of the AI system lifecycle highlights the crucial role of design, verification, deployment and monitoring, preventing shortcuts that limit attention to the final product.

The Guidelines proceed to spell out a small set of principles. First, that of ‘Inclusive growth, sustainable development and well-being’, requiring a proactive approach of all stakeholders. Second, that of ‘Human-centred values and fairness’, demanding respect for the rule of law, human rights and democratic values, to be implemented by appropriate mechanisms and safeguards. Third,

that of ‘Transparency and explainability’, calling for ‘a general understanding of AI systems’, the need to make stakeholders aware of their interactions with AI systems, asking that those affected by AI systems are enabled to understand the outcome, and to challenge it insofar as it adversely affects them, based on ‘easy-to-understand information on the factors, and the logic that served as the basis for the prediction, recommendation or decision’. Fourth, that of ‘Robustness, security and safety’, insisting that the actors ensure appropriate functioning in the case of ‘normal use, foreseeable use or misuse, or other adverse conditions’, preventing unreasonable safety risk.

These are high level principles, integrating three generations of human rights for the era of AI: first, individual liberty & political rights, to be claimed against governments, and second & third, social & cultural rights that rather function as instruction norms for governments and institutions of international law. We should note, however, that the actors who are indirectly addressed are ‘those who play an active role in the AI system lifecycle, including organisations and individuals that deploy or operate AI’, suggesting that these principles intend to inspire treaties and legislation with direct legal effect for those actors (not just regarding governments), just like the GDPR which builds on the 1980 OECD Guidelines. This implies something like direct horizontal effect, in human rights parlance. We should also note that the stakeholders include those who are adversely affected by AI systems, while several principles articulate obligations that require concomitant rights, such as ‘being enabled to challenge the outcome’. Due to the high level of abstraction and the fact that the principles are part of a Recommendation that is not binding upon OECD member states, the principles are aspirational, conveying intent rather than attributing rights. As the OECD member states have agreed to thus articulate their joint aspirations, we can nevertheless expect these principles to form the backbone of upcoming legislation, treaties and regulation (OECD 2021). The fact that the Recommendation takes an ‘impact’ approach and does not restrict itself to particular methods or techniques implies a generally broad scope as to what *counts as* AI.

III.2 US approach

As with all current AI law, the US approach is a moving target (OECD 2021; Harris 2021). A number of dedicated initiatives have been taken. At the federal level, for instance, aiming to regulate the integration of unmanned aircraft systems into the national airspace system, and AI-powered medical diagnostic systems, and at state or city level, for instance, New York City enacted ‘A Local Law to amend the administrative code of the city of New York, in relation to automated employment decision tools’,⁵ applicable from January 2023, purportedly the first of its kind in the US. Similarly, both at state and federal level, regulation of the use of facial recognition systems by government agencies and law enforcement is a big issue. The US Office of Management and Budget has issued guidance to federal agencies on the regulation of private sector use of AI, with a preference for non-regulatory approaches. Also, the 2020 Executive Order on AI guides federal agencies in the uptake of trustworthy AI in federal government, including a set of Principles for the Use of AI in Government that apply to ‘designing, developing, acquiring, and using AI in the Federal Government’, which will be done in a way that is: ‘(a) Lawful and respectful of our Nation's values. (b) Purposeful and performance-driven. (c) Accurate, reliable, and effective. (d) Safe, secure, and resilient. (e) Understandable. (f) Responsible and traceable. (g) Regularly monitored. (h) Transparent. (i) Accountable.’⁶

⁵ Local Law Int. No. 1894-A, <https://legistar.council.nyc.gov/LegislationDetail.aspx?ID=4344524&GUID=B051915D-A9AC-451E-81F8-6596032FA3F9&Options=Advanced&Search>.

⁶ Executive Order 13960 of December 3, 2020, <https://www.federalregister.gov/documents/2020/12/08/2020-27065/promoting-the-use-of-trustworthy-artificial-intelligence-in-the-federal-government>.

Many of these initiatives are part of the 2020 US National Artificial Intelligence Initiative,⁷ which basically aims to promote investment in and deployment of AI in both the private and the public sector, taking into account the kind of principles summed up above. So far, there is no indication that the US will introduce omnibus legislation that covers the development, sale and deployment of AI across the board in both the private and the public sector.

For this reason, it is important to keep an eye on the policy and enforcement interventions of the Federal Trade Commission (FTC), a regulatory body tasked with the enforcement of consumer law. In a series of decisions based on existing law, the FTC has tried to implement legal constraints on the deployment of AI systems in the private sector, based on, amongst others, the FTC Act, the Fair Credit Reporting Act and the Equal Credit Opportunity Act. This mostly concerns the use of systems that are unfairly biased, while the FTC has qualified the sale of racially biased algorithms as a deceptive practice based on the FRC Act. In 2020 the Director of the FTC Bureau of Consumer Protection published Guidelines for the 'Use of Artificial Intelligence and Algorithms', highlighting that decisions affecting consumers based on AI or algorithms must be transparent, explained, fair, based on robust and empirically sound models and include appropriate accountability mechanisms (Smith 2020; Jillson 2021). This refers to many of the principles named in the 2020 Executive Order, which largely overlap with the principles of the OECD Guidelines.

Finally, the sale and deployment of AI systems squarely falls within the scope of private law liability if damage is caused and the relevant legal conditions of tort law apply, notably concerning negligence (Barfield 2020; Hildebrandt 2020b; Selbst 2020). The latter is a complicated affair because it is not obvious what duty of care applies regarding systems that continue learning after being put on the market, how such duties are distributed between the vendor and those who deploy the system, to what extent harm is foreseeable by either of them and which of them is liable for taking what appropriate precautionary measures. The same goes for systems based on machine learning outputs, if the behaviour or decision-making of such systems cannot be explained or understood. Strict liability based on a product's defect may not apply because it is unclear whether an AI system counts as a product or because the nature of AI systems implies that the defect is not knowable at the moment of sale. This may require legislative intervention or relevant developments at the level of case law. See also section III.4 on the EU approach.

III.3 Chinese approach

China is a fierce regulator in the broader context of AI law. It has enacted the 2021 'Personal Information Protection Law' which provides broad protections similar to the EU's GDPR (though not in relation to government processing of personal data). From 1 March 2022, the 'Provisions on the Management of Algorithmic Recommendations in Internet Information Services'⁸ will be applicable, enacting a whole range of rights and obligations in respect of the use of recommender systems in the private sector, overlapping with many of the OECD AI principles, notably imposing transparency, explanatory and auditing duties, while also giving those targeted a right to configure and even disable recommender systems (RSs) that affect them (Huld 2021). Interestingly, the Provisions require service providers to embed 'mainstream values', to 'actively transmit positive energy' and to 'promote uplifting use of algorithms', while simultaneously making platforms liable for any illegal content they allow, which rings some bells as to China's 2021 Cybersecurity Law that holds platforms

⁷ See the official website of US National Artificial Intelligence Initiative <https://www.ai.gov>.

⁸ Provisions on the Management of Algorithmic Recommendations in Internet Information Services, applicable from 1 March 2022, see http://www.cac.gov.cn/2022-01/04/c_1642894606364259.htm, translation <https://www.chinalawtranslate.com/en/algorithms/>

accountable for hosting illegal or undesirable content. China has also submitted a (Ministry of Foreign Affairs of the people's Republic of China 2021) to the 2021 conference of the United Nations (UN) Convention on Certain Conventional Weapons in Geneva, demonstrating its concern about the global race to develop and deploy AI in the context of military operations. This should not, however, distract attention from a long-standing military strategy of investing in so-called 'trump card' technology in the context of military competition, aiming to 'win' by way of breakthrough technology rather than by having more of existing technology. It may, nevertheless, indicate that China's concerns are the welfare, safety and health of its own population rather than world domination (Roberts et al. 2021).

In 2017 China's State Council issued 'A Next Generation Artificial Intelligence Development Plan' (AIDP), setting out a national AI strategy for the period up to 2030 (Webster et al. 2017).⁹ The plan includes a series of intermediate steps: (1) by 2020, it aims to create a 21 billion dollar AI industry, and establish an initial set of ethical principles, (2) by 2025 it plans to have achieved a major breakthrough in fundamental AI theory, expanding its AI industry to one of 58 billion dollar, and to further develop and codify the 2020 ethical principles for AI and (3), by 2030, it aspires to be the number one global AI innovation hub with a 147 billion dollar industry, while further adapting and extending its AI law to face emerging challenges. The AIDP is managed by the Ministry of Science and Technology (MOST) and aims to steer private companies in the direction of what it considers the ethical development and deployment of AI, by rewarding those who align with the values endorsed by the Chinese Communist Party (CCP). In that sense the Chinese government takes an approach that is both centralised (strategic goals are defined by the CCP) and decentralised (rewarding AI development initiated by private enterprise, if aligned with CCP values and goals). As far as I can see, the AIDP does not define AI, raising a number of thorny issues, e.g. concerning overlap with the regulations on recommender systems, referred to above (noting that recommender systems are defined in the relevant Provisions, meaning the definition will probably be added once legislation is enacted).

In 2021, the MOST published 'The New Generation Artificial Intelligence Ethics' (NGAIE), as scheduled within the first step of the AIDP. Though not binding, the ethical principles it contains should be transformed into binding legislation by 2025. This seems a procedure similar to that of the EU, which first had a High Level Expert Group on AI (HLEGAI) develop Ethical Principles for Trustworthy AI,¹⁰ followed by their further extension and codification in the proposed EU AI Act. The NGAIE has 6 chapters. In chapter 1 (General Provisions) art. 3 summarises six ethical principles: first 'Enhancing the well-being of humankind', second 'Promoting fairness and justice', third 'Protecting privacy and security', fourth 'Ensuring controllability and trustworthiness', fifth 'Strengthening accountability' and sixth 'Improving ethical literacy'. (National Governance Committee for the New Generation Artificial Intelligence, China 2021) These principles largely overlap with those of the OECD Guidelines, discussed above, noting that China is an OECD member state. In chapter 2 (Norms of Management) we find the following norms: art. 5. 'Promotion of agile governance', art. 6. 'Active practice', art. 7. 'Exercise and use power correctly', art. 8. 'Strengthen risk preventions', art. 9. 'Promote inclusivity and openness'. In chapter 3 (Norms of Research & Development): art. 10. 'Strengthen the awareness of self-discipline', art. 11. 'Improve data quality', art. 12. 'Enhance safety, security and transparency' and art. 13. 'Avoid bias and discrimination'. In chapter 4 (Norms of Supply): art. 14. 'Respect market rules', art. 15. 'Strengthen quality control', art. 16. 'Protect the

⁹ An unofficial translation can be found here: <https://www.newamerica.org/cybersecurity-initiative/digichina/blog/full-translation-chinas-new-generation-artificial-intelligence-development-plan-2017/>.

¹⁰ European Commission, Directorate-General for Communications Networks, Content and Technology, Ethics guidelines for trustworthy AI, Publications Office, 2019, <https://data.europa.eu/doi/10.2759/177365>.

rights and interests of users', art. 17. 'Strengthen emergency protection'. In chapter 5 (Norms of Use): Art. 18. 'Promote good use', art. 19. 'Avoid misuse and abuse', art. 20. 'Forbid malicious use', art. 21. 'Timely and Proactive feedback' and art. 22. 'Improve the ability to use'. Chapter 6 is worth quoting in full:

Art. 23. This set of norms is issued by the National Governance Committee of New Generation Artificial Intelligence, and it is responsible for explaining and guiding its implementation;

Art. 24. With actual requirements and needs, management departments at all levels, enterprises, universities, research institutes, associations and other related organizations may formulate more specific ethical norms and related measures based on this set of norms.

Art. 25. This set of norms shall be carried out starting from the date of its publication, and shall be revised in due course according to the needs of economic and social development and the development state of AI.

Throughout the NGAIE we find reference to 'orderly market competition', 'self-discipline', asking all those concerned to promote 'the healthy and sustainable AI in an orderly manner', emphasising that ethical duties are part of a plan, aiming to serve the goals set by the MOST, which issued, explains and guides the implementation of the norms (cf. art. 23).

The emphasis on ethical duties and orderly central planning may not bode well for the many references to individual human rights, such as those enumerated in art. 7 that requires to 'fully respect and respect privacy, freedom, dignity, safety and other rights of relevant stakeholders and other legal rights and interests, and prohibit improper use of power to infringe the legal rights of natural persons, legal persons and other organizations'. The problem is that such human rights may not apply when the government itself is supplier or user of AI. One could say that the Chinese government is acting as an enlightened absolute ruler, who decides what is best for its people, while having the best intentions for their welfare in terms of economic sustainability, education, health and safety. This is best illustrated by art. article 51 of the Constitution: 'Citizens of the People's Republic of China, in exercising their freedoms and rights, may not infringe upon the interests of the State, of society or of the collective, or upon the lawful freedoms and rights of other citizens.' The crucial question is who decides upon an infringement of the State's, society's or the collective' interests, and what such decisional power means for the rights and freedoms attributed in the context of either the Constitution or the promulgated 'New Generation Artificial Intelligence Ethics'. To answer those questions, we must note that individuals have no effective remedies against the state when they find their constitutional rights infringed. In the same vein, the difference between non-binding ethical norms and binding legal norms may be less relevant in this context as in either case an emphasis on ethical *duties* rather than *individual rights* prevails, notably the duties of a government to decide on what is right and wrong and the duties of citizens to prioritise their duties towards 'the collective' as represented by the state. These are thorny issues concerning the checks and balances of the rule of law that should be faced when discussing the scope, purpose and effectiveness of AI law in different jurisdictions.(Zhao 2015)

III.4 The EU approach

The general assumption that underlies the EU approach is that major investment in the development and deployment of AI systems is by default a good thing, provided they are reliable and built and deployed in a way that protects against risks to safety, health and fundamental rights. There is no legal duty to 'embed positive energy' when building AI systems (as in China), but the proposed AI Act's preliminary wordings expound on the Union's values, fundamental rights and

principles, and on the EU's commitment to democracy and the rule of law.¹¹ There are many ways to interpret or criticize the Act (Ebers 2021; Veale and Borgesius 2021); I will restrict myself to a comparative perspective, highlighting the Act's core direction.

The proposed AI Act is part of a larger legal framework meant to create fair competitive markets for the provision of AI systems (Digital Markets Act, see Bertuzzi 2021b), to safeguard fair and accountable digital services and appropriate data governance (Data Governance Act, see Bertuzzi 2021c, and Digital Services Act, see Bertuzzi 2021d) and Data Act (see Bertuzzi 2021a), all geared towards a full-fledged digital transition and based on an ambitious 'European Data Strategy'.¹² In this section I will briefly sum up the main structure of the AI Act as proposed by the EU Commission, and briefly discuss some of the amendments made by the EU Council.

The AI Act deploys a broad, pragmatic definition of AI systems, as we have seen above. The Council has, however, proposed some amendments 'to distinguish artificial intelligence from more classic software and programming' (recital 6). The Council has therefor redefined an AI system as referring to a system (not necessarily software) that

- i. receives machine and/or human-based data and inputs,
- ii. infers how to achieve a given set of human-defined objectives using learning, reasoning or modelling implemented with the techniques and approaches listed in Annex I, and
- iii. generates outputs in the form of content (generative AI systems), predictions, recommendations or decisions, which influence the environments it interacts with;

The techniques and approaches listed in Annex I have not been changed, meaning that knowledge-based or logic-based systems squarely fall within the scope of the Act, provided the conditions above apply. The major change in those conditions is that a system must achieve its human-defined objectives by way of 'learning, reasoning or modelling' to qualify as an AI system. This raises a number of questions, e.g. of whether an excel sheet used to *decide on or make recommendations concerning* individual social security benefits does or does not qualify as an AI system, depending on whether the implementation of the excel sheet's decision tree is considered to be a form of 'reasoning'. Many of such systems, used for recruitment or promotion, social benefits or medical diagnosis are simple, deterministic decision trees. If these systems are not considered AI, automated decision making (ADM) and decision support (ADS) systems will be out of scope, while they may also not be covered by art. 22 of the GDPR (which contains a default prohibition of automated decisions with major impact). They may fall within the scope of the DSA, which however seems to offer less protection than the AI Act would provide for high risk systems. This is especially problematic whenever the decision trees of ADM or ADS systems are based on the output of data-driven systems that learn and/or model, but the main point is that legal protection should depend on potential impact of automation rather than what techniques have been deployed and/or whether the system is simple or complex.

Even if the scope of what counts as AI systems were to be broad, the Act's far-reaching obligations target a small subset of these systems, i.e. those considered high risk. The Act also defines three prohibited AI practices and imposes specific transparency rules for a further subset of systems, even if they do not qualify as high risk. I will focus on the main target of the Act, which concerns 'high risk

¹¹ Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative Acts, Brussels, 21.4.2021 COM(2021) 206 final 2021/0106 (COD).

¹² All relevant documents can be found via https://ec.europa.eu/info/strategy/priorities-2019-2024/europe-fit-digital-age/european-data-strategy_en.

AI systems'. Other than in the GDPR, high risk AI systems are categorically qualified as such in Annex II (mainly embedded systems with high risk impact on health and safety, such as medical devices, connected cars and aircraft) and Annex III (mainly but not only software systems deployed in contexts that suggest a high risk impact on fundamental rights, such as biometric systems, critical infrastructure, education and vocational training, law enforcement and migration, asylum and border control). Whereas the GDPR takes a high risk approach to the regulation of processing of personal data if they are *highly likely to result in risks to fundamental rights*, thus taking a flexible and more casuistic approach, the AI Act qualifies *certain technologies and certain contexts as high risk by definition*. This obviously has advantages (legal certainty) and disadvantages (depending on what systems and contexts are qualified as high risks).

The qualification as high risk is pivotal, because these are the systems that must be *in conformity with* a whole range of detailed legal requirements stipulated in Chapter 2, Title III of the Act, such as a dedicated risk management system, data and data governance, automated logging, documentation and auditing, transparency for and effective human oversight by those who deploy these systems, accuracy, robustness and cybersecurity. Most of the legal obligations enacted by the AI Act address the providers of AI systems (those who develop an AI system or have it developed and first place them on the market or put them into service under their own name or trademark). As most of the requirements are both demanding and salient, the AI Act will no doubt reconfigure the market, creating an incentive structure that favours the provision of reliable, resilient, robust and responsible AI systems.

Providers must draw up and keep at the disposal of competent authorities a *conformity declaration*, based on self-assessment or third party auditing and the Act makes them liable to steep fines if found not to be *in conformity* with the requirements of the Act. This is a matter of supervision by national supervisory bodies and much will depend on them being properly funded and having sufficiently trained technical staff to monitor, investigate and enforce conformity.

Next to such administrative enforcement, the EU has scheduled an update of the private law liability regime, in particular of the 1985 Product Liability Directive. Following the 2019 advice of the Expert Group on Liability and New Technologies on *Liability for Artificial Intelligence*, the European Commission (EC) published its 2020 *Report on safety and liability implications of Artificial Intelligence, the Internet of Things and Robotics*. We can safely assume that the upcoming liability regime for AI will be informed by the requirements underlying the conformity declaration and by some of issues identified in the EC's 2020 report. The report notes that AI, IoT and Robotics have specific characteristics that do not fit with the current private law liability framework, notably *connectivity, autonomy, data dependency, openness* to updates and upgraded after being placed on the market, *opacity* of their decision making and their exposure to *cyberthreats*. The Commission therefore sees the need to update the private law liability regime across the internal market of the EU, in tune with the updated safety regime. This probably involves a new balance between harmonised EU law and national law, making sure that, (1) differences between national liability regimes do not disturb cross-border development, provision, deployment and sale of AI systems, IoT and robotics and (2) that victims are compensated for damage in a way that ensures their trust, noting that liability is an incentive to develop and sell trustworthy systems, in tandem with proper insurance. The second point concerns the definitions of 'product' and 'defect' in the current Product Liability Directive and relates directly to the reversal of the burden of proof as to causality in the Directive, relaxing the more stringent requirements of fault or negligence when damage is brought about in the context of AI, IoT or robotics, combined with compulsory insurance for those liable for the damage.

An interesting question is whether a conformity declaration for AI systems that – upon inspection - do not conform to the requirements of chapter 2, title III of the Act will generate strict liability for providers when damage has been caused. This would provide a practical and effective enforcement mechanism, especially if combined with a right to sue based on collective action (either mandating one's claims or dedicated non-profit organisations suing on their own behalf).

IV. The Brussels effect of the AI Act

Different states may compete to attract corporate investments, by way of their corporate law and tax law, hoping to make such investments more profitable. Within the context of the US this has been called out as causing the so-called Delaware effect, or a 'race to the bottom', forcing states within the US to lower the standards for corporate behaviour under pressures of other states doing the same. The state of Delaware was the prime example here, with half of the nation's incorporations based in Delaware. A similar argument has been put forward as to AI law, claiming that the competition between China, the US and the EU will lead to a geopolitical race to the bottom in an effort to attract the best AI companies to their jurisdiction. The Delaware effect has been denounced, both on empirical grounds (there is no evidence, see Subramanian 2004) and on normative grounds (rejecting economic determinism while highlighting that democracy not economy decides on corporate law, see Greenwood 2005). More interestingly for our case is the California effect, referring to the fact that corporations tend to adopt business rules that comply with the most stringent consumer and environmental laws, assuming this will also make them compliant within less stringent jurisdictions. This effect only occurs if the stringent rules are enacted in a jurisdiction with a large and profitable consumer market, such as California (Vogel 1997).

Anu Bradford (2020) has convincingly argued that similar reasoning applies to the EU's regulatory impact, based on corporations' wish to compete on the internal market of the EU, due to the large size of that market, the regulatory capacity of the EU, the stringency of the enacted rules (assuming that adopting the most stringent rules provides access to all other – less stringently regulated - markets), inelasticity of the regulated targets (assuming that products and services cannot be provided without EU jurisdiction) and the non-divisibility of the global market (which would allow corporations to adapt their products and services to different economic regions). The claimed effect was termed the Brussels effect, and it would be interesting to see whether it also applies to AI law. Bradford addresses the main arguments against the Brussels effect, arguing that the effect is not bad for innovation and that it is neither a form of protectionism nor one of regulatory imperialism. Though we cannot discuss her argumentation in detail, it is clear that her assessment of a potential decline of the Brussels' effect must be followed-up with regard to the regulation of AI: the emergence of China as a global economic power with a massive consumer market may disrupt the power of the EU market, protectionism as a force of deglobalisation may render competition on the EU market irrelevant and new technologies that destabilise the non-divisibility of the global market (giving the example of 3-printing) may enable corporations to differentiate their business rules depending on the target region.

Time will tell whether we are moving towards competition between regulatory regimes with a potential race to the bottom, or towards a race to the top. We may actually move towards convergence, especially with regard to trustworthiness in terms of safety and health, while nevertheless diverging when it comes to human rights, democracy and the rule of law. It is crucial to acknowledge that whatever 'time will tell' does depend on legislative choices made in the coming years, and should not be attributed to technological or economic determinism; both the design of AI systems and the behaviour of economic markets is enabled and constraint by the legal frameworks of private, public, criminal and international law. As AI is moving into all other legal domains, they

will all have to deal with the emergent uncertainties and the potential impact of AI systems, based on the question what world ‘we’ want to live in.

V. References

- Aronson, Jay E. 2003. ‘Expert Systems’. In *Encyclopedia of Information Systems*, edited by Hossein Bidgoli, 277–89. New York: Elsevier. <https://doi.org/10.1016/B0-12-227240-4/00067-8>.
- Barfield, Woodrow, ed. 2020. *The Cambridge Handbook of the Law of Algorithms*. Cambridge Law Handbooks. Cambridge: Cambridge University Press. <https://doi.org/10.1017/9781108680844>.
- Bertuzzi, Luca. 2021a. ‘LEAK: Draft Impact Assessment Sheds Some Light on Upcoming Data Act’. *Www.Euractiv.Com*. 1 November 2021. <https://www.euractiv.com/section/data-protection/news/leak-draft-impact-assessment-sheds-some-light-on-upcoming-data-act/>.
- . 2021b. ‘EU Parliament’s Key Committee Adopts Digital Markets Act’. *Www.Euractiv.Com*. 23 November 2021. <https://www.euractiv.com/section/digital/news/eu-parliaments-key-committee-adopts-digital-markets-act/>.
- . 2021c. ‘Data Governance: New EU Law for Data-Sharing Adopted’. *Www.Euractiv.Com*. 1 December 2021. <https://www.euractiv.com/section/digital/news/data-governance-new-eu-law-for-data-sharing-adopted/>.
- . 2021d. ‘MEPs Adopt Digital Services Act in a Key Committee Vote’. *Www.Euractiv.Com*. 14 December 2021. <https://www.euractiv.com/section/digital/news/meps-adopt-digital-services-act-in-a-key-committee-vote/>.
- Bostrom, Nick, and Napoleon Ryan. 2015. *Superintelligence: Paths, Dangers, Strategies*. MP3 Una edition. Audible Studios on Brilliance Audio.
- Bradford, Anu. 2020. *The Brussels Effect: How the European Union Rules the World*. New York: Oxford University Press.
- Cath, Corinne, Sandra Wachter, Brent Mittelstadt, Mariarosaria Taddeo, and Luciano Floridi. 2018. ‘Artificial Intelligence and the “Good Society”: The US, EU, and UK Approach’. *Science and Engineering Ethics* 24 (2): 505–28. <https://doi.org/10.1007/s11948-017-9901-7>.
- Cognilytica. 2021. ‘Worldwide AI Laws and Regulations 2021’. Cognilytica. <https://www.cognilytica.com/document/worldwide-ai-laws-and-regulations-2021/>.
- Ebers, Martin. 2021. ‘Standardizing AI - The Case of the European Commission’s Proposal for an Artificial Intelligence Act’. SSRN Scholarly Paper ID 3900378. Rochester, NY: Social Science Research Network. <https://doi.org/10.2139/ssrn.3900378>.
- Fjelland, Ragnar. 2020. ‘Why General Artificial Intelligence Will Not Be Realized’. *Humanities and Social Sciences Communications* 7 (1): 1–9. <https://doi.org/10.1057/s41599-020-0494-4>.
- Gigerenzer, Gerd, and Reinhard Selten. 2001. *Bounded Rationality : The Adaptive Toolbox*. Cambridge, Mass.: MIT Press.
- Glenn, Patrick. 2014. *Legal Traditions of the World: Sustainable Diversity in Law*. Oxford, United Kingdom ; New York, NY.
- Greenwood, Daniel J. H. 2005. ‘Democracy and Delaware: The Mysterious Race to the Bottom/Top’. *Yale Law & Policy Review* 23 (2): 381–454. <https://www.jstor.org/stable/40239643>.
- Harris, Laurie A. 2021. ‘Artificial Intelligence: Background, Selected Issues, and Policy Considerations’. R46795. US Congressional Research Service. <https://crsreports.congress.gov/product/pdf/R/R46795>.
- Hildebrandt, Mireille. 2015. *Smart Technologies and the End(s) of Law. Novel Entanglements of Law and Technology*. Cheltenham: Edward Elgar.
- . 2020a. ‘11. Closure: On Ethics, Code and Law’. In *Law for Computer Scientists*. Oxford University Press. <https://lawforcomputerscientists.pubpub.org/pub/nx5zv2ux>.
- . 2020b. ‘Private Law Liability for Faulty ICT’. In *Law for Computer Scientists and Other Folk*. Oxford: Oxford University Press. <https://doi.org/10.1093/oso/9780198860877.003.0008>.

- Hinton, Geoffrey E. 1990. '20 - Connectionist Learning Procedures'. In *Machine Learning*, edited by Yves Kodratoff and Ryszard S. Michalski, 555–610. San Francisco (CA): Morgan Kaufmann. <https://doi.org/10.1016/B978-0-08-051055-2.50029-8>.
- Huld, Arendse. 2021. 'Recommendation Algorithm Regulations - China's Cybersecurity Regime'. Business Intelligence on legal, tax, and operational issues in China from a practical perspective. *China Briefing News* (blog). 7 September 2021. <https://www.china-briefing.com/news/recommendation-algorithm-regulations-china-cybersecurity-regime/>.
- Jillson, Elisa. 2021. 'Aiming for Truth, Fairness, and Equity in Your Company's Use of AI'. Federal Trade Commission. *Business Blog FTC* (blog). 19 April 2021. <https://www.ftc.gov/news-events/blogs/business-blog/2021/04/aiming-truth-fairness-equity-your-companys-use-ai>.
- Marcus, Gary, and Ernest Davis. 2020. 'Insights for AI from the Human Mind'. *Communications of the ACM* 64 (1): 38–41. <https://doi.org/10.1145/3392663>.
- Ministry of Foreign Affairs of the people's Republic of China. 2021. 'Position Paper of the People's Republic of China on Regulating Military Applications of Artificial Intelligence (AI)'. https://www.fmprc.gov.cn/mfa_eng/wjdt_665385/wjzcs/202112/t20211214_10469512.html.
- Monett, Dagmar, Colin W. P. Lewis, Kristinn R. Thórisson, Joscha Bach, Gianluca Baldassarre, Giovanni Granato, Istvan S. N. Berkeley, et al. 2020. 'Special Issue "On Defining Artificial Intelligence"—Commentaries and Author's Response'. *Journal of Artificial General Intelligence* 11 (2): 1–100. <https://doi.org/10.2478/jagi-2020-0003>.
- National Governance Committee for the New Generation Artificial Intelligence, China. 2021. 'The Ethical Norms for the New Generation Artificial Intelligence, China'. *China-UK Research Centre for AI Ethics and Governance* (blog). 27 September 2021. <https://ai-ethics-and-governance.institute/>.
- Newell, Allen, and Herbert A. Simon. 1976. 'Computer Science as Empirical Inquiry: Symbols and Search'. *Communications of the ACM* 19 (3): 113–26. <https://doi.org/10.1145/360018.360022>.
- OECD. 2021. 'State of Implementation of the OECD AI Principles: Insights from National AI Policies'. Paris: OECD. <https://doi.org/10.1787/1cd40c44-en>.
- Roberts, Huw, Josh Cows, Jessica Morley, Mariarosaria Taddeo, Vincent Wang, and Luciano Floridi. 2021. 'The Chinese Approach to Artificial Intelligence: An Analysis of Policy, Ethics, and Regulation'. *AI & SOCIETY* 36 (1): 59–77. <https://doi.org/10.1007/s00146-020-00992-2>.
- Russell, Stuart, and Peter Norvig. 2020. *Artificial Intelligence: A Modern Approach*. 4th edition. Hoboken: Pearson.
- Selbst, Andrew D. 2020. 'Negligence and AI's Human Users'. *Boston University Law Review* 100 (4): 1315. <https://www.bu.edu/bulawreview/files/2020/09/SELBST.pdf>.
- Smith, Andrew. 2020. 'Using Artificial Intelligence and Algorithms'. *Federal Trade Commission* (blog). 8 April 2020. <https://www.ftc.gov/news-events/blogs/business-blog/2020/04/using-artificial-intelligence-algorithms>.
- Smuha, Nathalie A. 2021. 'From a "Race to AI" to a "Race to AI Regulation": Regulatory Competition for Artificial Intelligence'. *Law, Innovation and Technology* 13 (1): 57–84. <https://doi.org/10.1080/17579961.2021.1898300>.
- Subramanian, Guhan. 2004. 'The Disappearing Delaware Effect'. *Journal of Law, Economics, & Organization* 20 (1): 32–59. <https://www.jstor.org/stable/3554918>.
- Veale, Michael, and Frederik Zuiderveen Borgesius. 2021. 'Demystifying the Draft EU Artificial Intelligence Act'. SocArXiv. <https://doi.org/10.31235/osf.io/38p5f>.
- Vogel, David. 1997. *Trading Up: Consumer and Environmental Regulation in a Global Economy*. Cambridge, Mass.: Harvard University Press.
- Wang, Pei. 2019. 'On Defining Artificial Intelligence'. *Journal of Artificial General Intelligence* 10 (2): 1–37. <https://doi.org/10.2478/jagi-2019-0002>.

- Webster, Graham, Rogier Creemers, Paul Triolo, and Elsa Kania. 2017. 'China's Plan to "Lead" in AI: Purpose, Prospects, and Problems'. *New America* (blog). 1 August 2017. <http://newamerica.org/cybersecurity-initiative/blog/chinas-plan-lead-ai-purpose-prospects-and-problems/>.
- . 2017. 'Full Translation: China's "New Generation Artificial Intelligence Development Plan" (2017)'. *New America* (blog). 1 August 2017. <http://newamerica.org/cybersecurity-initiative/digichina/blog/full-translation-chinas-new-generation-artificial-intelligence-development-plan-2017/>.
- Wolff, Lutz-Christian. 2019. 'Artificial Intelligence Ante Portas: The End of Comparative Law?' *The Chinese Journal of Comparative Law* 7 (3): 484–504. <https://doi.org/10.1093/cjcl/cxz020>.
- Zhao, Jun. 2015. 'China and the Uneasy Case for Universal Human Rights'. *Human Rights Quarterly* 37 (1): 29–52. <http://www.jstor.org/stable/24518273>.