Journal of European Public Policy

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/rjpp20

Role of the state and responsibility in governing artificial intelligence: a comparative analysis of AI strategies

Christian Djeffal, Markus B. Siewert & Stefan Wurster

To cite this article: Christian Djeffal, Markus B. Siewert & Stefan Wurster (2022): Role of the state and responsibility in governing artificial intelligence: a comparative analysis of AI strategies, Journal of European Public Policy, DOI: 10.1080/13501763.2022.2094987

To link to this article: https://doi.org/10.1080/13501763.2022.2094987

JOURNAL OF EUROPEAN PUBLIC POLICY
https://doi.org/10.1080/13501763.2022.2094987

Role of the state and responsibility in governing artificial intelligence: a comparative analysis of AI strategies

Christian Djeffal[a]*, Markus B. Siewert [b]* and Stefan Wurster [a,b]*

[a]TUM School of Social Sciences and Technology, Technical University of Munich, Munich, Germany; [b]Munich School of Politics and Public Policy, Munich, Germany

ABSTRACT

Technologies based on artificial intelligence (AI) represent a crucial governance challenge for policymakers. This study contributes to the understanding of how states plan to govern AI with respect to the role they assume and to the way they develop AI in a responsible manner. In different policy instruments across 22 countries plus the European Union, there is considerable variation in how governments approach the governance of AI, both regarding the policy measures proposed and their focus on public responsibility. Analysing a set of policy instruments we find multiple modes of AI governance, with the major difference being between self-regulation-promoting and market-based approaches, and a combination of entrepreneurial and regulatory governance approaches. Our analysis also indicates that the approach to public responsibility is largely independent of the chosen policy mix of AI governance. Therefore, responsibility seems to be a cross-cutting issue that cannot be tied to a specific approach of states towards technology.

KEYWORDS Artificial intelligence; AI governance; policy instruments; state types; responsible research innovation; technology assessment

Introduction

Artificial intelligence (AI) denotes a set of emerging general-purpose technologies.[1] As such, it can have a variety of impacts that depend on future innovations and are, therefore, hard to foresee (e.g., Djeffal, 2020; Taeihagh, 2021; Trajtenberg, 2019). Take the example of social media, which was hardly conceivable 20 years ago and is now widely driven by AI applications (Van Dijck, 2013). Through content moderation systems, AI has become the basis of information for people all around the world, for private communication, news consumption, and advertisement. At the same time, AI applications

CONTACT Markus B. Siewert markus.siewert@hfp.tum.de
*All three authors have contributed equally.
Supplemental data for this article can be accessed online at https://doi.org/10.1080/13501763.2022.2094987.

can be used to hijack social media before an election by spreading fake news to influence voter behaviour, whereas social media algorithms themselves can create filter bubbles which are hotbeds of polarization, conspiracy theories, and hate speech (Chessen, 2017; Djeffal, 2019a; United Nations General Assembly, 2019) and AI is again used to address these problems. With AI fuelling democractic dreams as well as nightmares, proposals about how to govern social media are hotly debated and range from self-regulation to (co-)regulation and even replacement by public agencies (e.g., European Commission, 2020a).

The example above points towards the general quandary for policymakers of governing AI (Taeihagh, 2021; Taeihagh et al., 2021; Ulnicane et al., 2020), which is familiar with other emerging technologies (Kuhlmann et al., 2019; Mandel, 2009). There are many reasons to support the development of AI, as it holds vast promise for future economic growth, societal development, and ecological progress (Bughin et al., 2019; Djeffal, 2019b; Khodabandeh et al., 2020). Yet, AI has also caused concern due to recurring examples of discrimination, opacity, and accountability, as well as its increasing energy consumption, and its reinforcement of existing power structures within society and the economy at large (e.g., Broussard, 2019; Crawford, 2021; Dauvergne, 2021).

In this paper, we explore *how governments navigate the challenges of governing AI to fulfil goals like stimulating innovation, mitigating risks and assuming responsibility for the social impacts of AI, all at the same time*. The main questions in this regard are how governments and policymakers construct the *role* of the state regarding technology governance and how they establish public *responsibility*. Analysing the national AI strategies of 22 countries and the European Union (EU), we examine the policy mix and regulatory regimes, both regarding the types of policy instruments proposed and the approaches taken to public responsibility. Our analysis offers the first systematic picture of the level of activity, the state concepts derived from the policy measures proposed and the stance taken on public responsibility. Moreover, by bringing together the literature on governance regimes (Borrás & Edler, 2020; Erdelyi & Goldsmith, 2018; Mandel, 2009; Marchant et al., 2020) and responsible technology development (European Commission, Directorate-General for Research and Innovation & Stilgoe, 2019; Felt & Wynne, 2007; Sarewitz, 2011; Stilgoe & Guston, 2017;) – which have thus far been at the centre of separate academic debates – we examine whether responsible AI governance can be linked to the general approaches states take.

To address these questions, we first clarify the main theoretical concepts concerning the different roles of states and their view of responsibility in the context of AI governance. The next sections then present the empirical findings concerning the policy mix and resulting state types, the uptake of and different approaches to responsibility, and finally the interrelationship

between the types of regulatory regimes and approaches to responsibility. We conclude with an outlook on the impact of our findings concerning future research on the governance of AI.

The role of the state and public responsibility in AI governance

Questions about the role of the state and responsible development are key questions regarding the governance of any technology (Borrás & Edler, 2020; Kuhlmann et al., 2019; Simonis, 2013; Ulnicane et al., 2020). Building on existing theoretical work, we conceptualize different modes of technology governance and public responsibility which will then guide our further empirical analysis.

Defining state roles in the governance of AI

The role of the state in the governance of AI, or any technology in general, can be envisaged in different ways (see for the following, Borrás & Edler, 2020; Erdelyi & Goldsmith, 2018; Gasser & Almeida, 2017; Kuhlmann et al., 2019; Mandel, 2009; Marchant et al., 2020; Sarewitz, 2011; Ulnicane et al., 2020). Conceptually, we distinguish two analytical dimensions: First, governments can play a proactive role in the development of AI technologies (*strong state intervention*), or they take a more passive stance by stepping back and giving private actors and/or the markets as much leeway as possible in the governance of AI (*weak state intervention*). Second, governments can concentrate on regulating potential risks of AI technologies (*enclosure-and-control approach*), or they can prioritise the deployment of AI and see their role primarily in promoting its development (*stimulation approach*).

As displayed in Figure 1, intersecting these two dimensions creates four ideal-typical regimes of governance which differ in the way that the state assumes a different role towards technology governance: (i) the *entrepreneurial state*, (ii) *the market-oriented state*, (iii) the *regulatory state*, and (iv) the *self-regulation-promoting state*.

Turning to the combination of a strong state which actively stimulates the development and deployment of AI technologies, this state type resembles Mazzucato's (2011) notion of the *entrepreneurial state*. Contrary to the small-state, free-market doctrine that often receives credit for a country's innovation performance, Mazzucato argues that public investments in innovation and technology, based on a proactive state as prime risk taker, can result in high innovation performance. In line with the technology-push approach (Di Stefano et al., 2012), the entrepreneurial state, therefore, stresses the role of public activities in fundamental research by creating

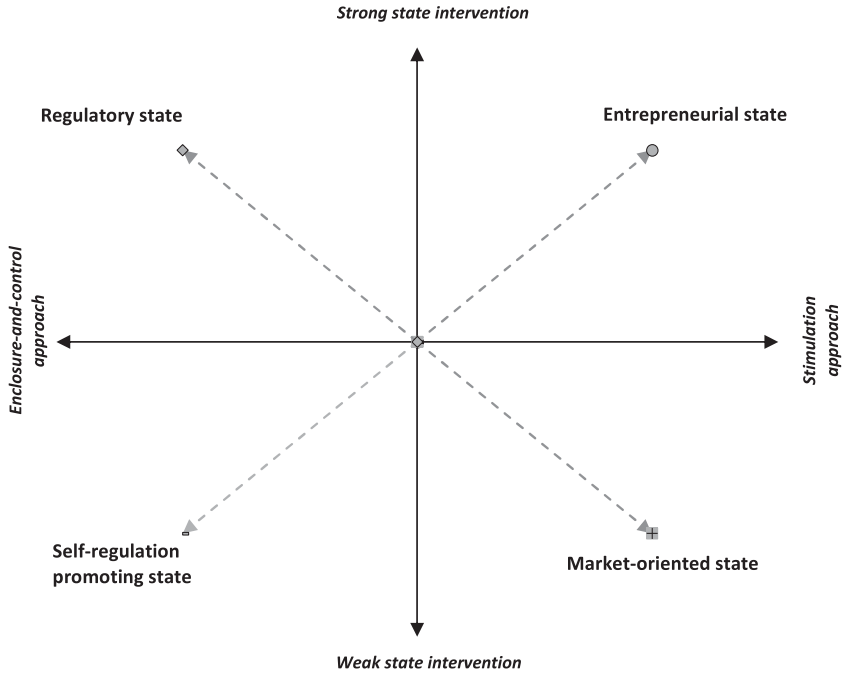


Figure 1. Governance dimensions of AI and resulting state types. Note: Authors' own depiction.

> a highly networked system of actors harnessing the best of the private sector for the national good over a medium to long-term horizon. It is the state as catalyst, and lead investor, sparking the initial reaction in a network that will then cause knowledge to spread. (Mazzucato, 2011, pp. 19–20)

From this perspective, it is the state which actively shapes the development of technologies like AI, e.g., as initiator of research and infrastructure projects, as promoter or co-designer of specific technical solutions, or as main customer and user (Borrás & Edler, 2020).

Juxtaposed regarding the extent of state intervention is the *market-oriented state* which relies on minimal state interference and soft regulation which are usually non-binding or based on self-governance mechanisms (Bisson et al., 2010; Kim, 2007). From this perspective, private actors, companies, and markets are the main drivers of technological innovation. Policymakers and governments, on the other hand, are depicted as being too slow and lacking the expertise to make prudent decisions on highly technical and complex topics like AI (Marchant & Wallach, 2013; Thierer et al., 2017). Moreover, state-led regulations are frequently portrayed as inhibiting innovation through unnecessary red-taping, whereas state interventions – in line with demand-pull approaches – are associated with inefficient processes

and government failure. Applied to AI, the answer is a 'light-touch approach' (Thierer et al., 2017, p. 50) where the state is not completely inactive, but, besides intervention in case of severe market failure, uses soft, market-friendly governance tools to strengthen market processes and to create leeway for private entrepreneurs to promote innovations in AI from the private sector (Marchant et al., 2020; World Bank, 2021, p. 10).

The traditional *regulatory state*, in contrast, is mainly characterized by hard governance instruments and direct interventions by the state (Braithwaite, 2011; Majone, 1997). By regulating innovations and technologies, the main objective is to protect public goods and to ensure citizens' safety. Accordingly, governments and policymakers not only step in once markets fail – i.e., *ex-post* – but design *ex-ante* mechanisms in a proactive manner, and in doing so, resume the regulatory authority of the state (Borrás & Edler, 2020; Mittelstadt, 2019; Nemitz, 2018).

The *self-regulation-promoting state*, finally, pertains to the idea that the industry and other relevant stakeholders within a field undertake collective efforts to regulate themselves and to impose self-restricting mechanisms. The role of the state is thus limited to function as facilitator, observer, or certifier of private initiatives to regulate and govern AI rather than direct engagement (Borrás & Edler, 2020). While the goal is risk prevention, this is not achieved through direct state intervention but rather soft regulatory instruments like 'codes of conduct, organizational and technical industry standards, quality seals and certification bodies, ombudsmen and arbitration/mediation boards and ethic committees / commissions' (Saurwein et al., 2015, p. 40; similar Marchant et al., 2020, pp. 10–16).

Two conceptual notes are important here: First, the described governance regimes are ideal types in the Weberian sense, which means that they are stylized abstractions which cannot be empirically observed in their pure form. We use them as heuristics, and it is an empirical question how closely actual governance arrangements fit these ideal-types – a task which we will take up below. Second and different from the Weberian notion, we conceive of the two dimensions and four state types as not mutually exclusive. For instance, governments can follow a self-regulation-promoting approach regarding some issues of AI governance, while simultaneously showing a strong entrepreneurial approach on others. Or a state's approach might be mainly characterized by a mix of entrepreneurial and market-oriented governance instruments (see section one of the online supplemental appendix for an illustration of this reasoning).

Defining approaches to public responsibility in AI governance

The question of what responsible technology development actually is, has been discussed for decades with responsibility remaining a contested

concept. Questions of public responsibility are not *a priori* tied to the role of the state since the latter does not capture to what extent governments and policymakers assume responsibility for societal impacts. Several aspects have been highlighted, such as expert assessments of technologies and knowledge production for decision makers (Brooks, 1976), participatory approaches based on opinions, concerns, and ideas of ordinary users (Hennen, 2012; Ornetzeder & Kastenhofer, 2012), procedural aspects highlighting anticipation and quickness (Guston & Sarewitz, 2002), or the progressive realization of values through technologies (Felt & Wynne, 2007). These approaches in scholarship are mirrored in developments and shifts in practice (Stilgoe & Guston, 2017): The first shift occurred in the 1970s, when technology assessment (TA; Hasselbalch, 2018) was institutionalized in the United States through the creation of the Office of Technology Assessment in 1972, and subsequently through similar institutions in many OECD member states (Ely et al., 2014; Van Zwanenberg et al., 2009). As its main objective, TA

> […] should forecast, at least on a probabilistic basis, the full spectrum of possible consequences of technological advance, leaving to the political process the actual choice among the alternative policies in the light of the best available knowledge of their likely consequences. (Brooks, 1976, p. 20)

However, the general stance of TA was continuously challenged for different reasons (European Commission, Directorate-General for Research and Innovation & Stilgoe, 2019; Stilgoe & Guston, 2017; Wynne, 1975), such as its static and passive approach, the lack of participation, especially by citizens and those affected by the technologies. One alternative was coined responsible innovation or as responsible research and innovation (RRI). RRI was set out to provide a governance structure that addresses the whole cycle of research, development and use of technologies and influence governance in a more profound manner (European Commission, Directorate-General for Research and Innovation & Stilgoe, 2019). Against this backdrop, we can distinguish three general stances towards responsibility: *deference of responsibility*, an '*orthodox' approach to technology assessment*, and *responsible research and innovation*.

Our research represents an attempt to decouple responsibility indicators from specific technologies and to generalize them in a way that is applicable to government strategies. Table 1 illustrates the three approaches which display an increasing activity level regarding responsibility. While the approach to defer responsibility contains no activity at all and is characterized by the absence of instruments alluding to responsibility, the TA approach, on the other hand, requires measures of states safeguarding social interests and values. The RRI approach, finally, symbolizes the highest activity level and contains all dimensions of responsibility. We further operationalize the three approaches to public responsibility along the following four categories.

Table 1. Forms of responsibility governance.

Form	Description	Responsibility Dimensions			
		Implementation	*Safeguarding*	*Participation*	*Proactiveness*
Responsible Research and Innovation (RRI)	Proactive and continuous all-encompassing efforts for responsibility.	Actual implementation with specific measures	Yes	Yes	Yes
Orthodox Technology Assessment (TA)	Flagging of responsibility issues that are dealt with in the ordinary political process.	Rhetorical implementation without specific measures	Yes	No	No
Deferral of responsibility	No specific mention of responsibility	No	No	No	No

Note: Based on close reading of the literature cited in the main text.

The first category is the degree to which responsibility is actually included in the strategies. There are generally three options: to abstain from action, to flag an issue for further deliberation by the competent authorities, or to implement ideas for a more integrated governance model with specific measures impacting research, development, and implementation of technologies (Guston, 2014; Sarewitz, 2011). The second category is the aim of safeguarding in terms of rights, freedoms, and legitimate interests in society. The third category concerns participation referring to the question of whether those affected by the technology have effectively participated in the context of policy instruments (Wilsdon & Willis, 2004). One of the main arguments here is to challenge the objective and neutral nature of expert statements as opposed to the views of those affected by a technology. Different methods of participation integrate these views interactively and help to gain a more complete picture of what the actual impacts of the technology are (Fisher et al., 2006). The fourth and last category is the focus on the potential for value realization through technology, which we call proactiveness. This dimension is specifically applicable in cases of general-purpose technologies like AI. It speaks to the possibility to guide innovation to realize other values. As an attempt to govern emerging technologies in an all-encompassing manner, RRI has worked towards governing innovation to realize values directly (Felt & Wynne, 2007).

Technology governance through government strategies

The governance of AI is occupying more and more space on the political agenda with governments and policymakers at the national and international level actively devising AI strategies in which they set out their goals regarding the governance of AI and the means to achieve them (Fatima et al., 2020; Radu, 2021; Ulnicane et al., 2020). As a form of

'meta-governance', governmental strategies fulfil several functions, of which two are particularly important for this study (Casado-Asensio & Steurer, 2014; Rayner & Howlett, 2009). First, they can be used to define core objectives regarding the future deployment of AI. In doing so, governments point out issue areas and priorities for which they plan to build up necessary capabilities to be at the vanguard of the AI revolution. Moreover, they lay down the respective measures and mechanisms concerning how they want to achieve these goals. Second, government strategies signal commitment to different (groups of) stakeholders which are part of the AI ecosystem. This signalling pertains to the formulation of and adherence to certain norms, values, and standards, but also the question of responsible development of AI in collaboration with other societal actors. By analysing official AI strategies, we thus gain insights into how governments and policymakers envisage and design the role of the state in steering of AI-based technologies and the responsibility of the state towards society and its citizens.

Our empirical analysis is based on the tracking of policy instruments in national AI strategies from 22 countries plus the EU.[2] We conducted a qualitative content analysis to identify relevant policy instruments defined as any means, techniques, or mechanisms by which the government aims to achieve its AI-related policy objective (Hood & Margetts, 2007; Howlett, 2019). A team of trained research assistants together with the authors screened all documents and manually coded respective policy measures to capture the type of instrument used and its reference to responsible development of AI (for further information on the codebooks including a detailed description of the coding procedure, please see section three, four and five of the online supplemental appendix Section six and seven of the online supplemental appendix display the raw data). In sum, we identified 1.829 individual policy measures across all 23 governmental AI strategies under study, which means, on average, 79.6 instruments per AI strategy. Since each measure can theoretically include more than one instrument, the overall number can be higher if we look at types of policy instruments.

The role of the state in AI governance

Looking at national AI strategies, which policy instruments do governments use and what role of the state is conveyed regarding the governance of AI? We first examine how active a state is concerning the governance of AI and what types of instruments are most prominently employed. Building on this descriptive picture of policy instrument usage, we can then systematise the instrument toolkit by empirically tracing the theoretical state types developed above (revisit Figure 1). Taken together, this offers a systematic

picture on the modes of governance and emerging regulatory regimes as they are portrayed in AI strategies.

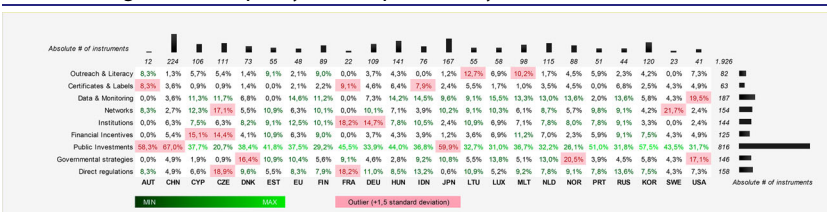
Governance of AI: level of activity and policy instrument use

The number of policy instruments is a measure for the level of activity (see Knill et al., 2012; Schaffrin et al., 2015 in the context of energy policy). On average, we find 83.7 policy instruments per AI strategy, though the level of activity in governing AI varies considerably between countries: China is, by far, the most active state with 219 instruments, followed by Japan (164) and the Republic of Korea (114). All three countries have in common that they are at the vanguard of the digital transformation and early adopters of AI (OECD, 2019). Yet, this does not seem to be a necessary condition since Hungary (135), the Czech Republic (103), or Cyprus (101) show a high level of activity, too; although they are currently ranked at the lower end in digitalization indices (European Commission, 2020b). Austria (10), Sweden and France (both 21) show comparatively low levels of activity, which is surprising for Austria and Sweden, as both countries usually count as prime examples of active states.

Table 2 gives an overview on the toolkit that is used in the governance of AI based on nine instrument subtypes within the four categories of authority, finance, organization, and information (Hood & Margetts, 2007; Howlett, 2019; see section two and three of the online supplemental material for the codebook and a detailed description of the types of instruments, how we defined and coded them).[3]

The largest category, by far, are *financial instruments* with a total of 941 policy measures, including both public investments and financial incentives. Direct public investments (816 policy instruments) into AI infrastructures or government-funded programmes in areas like education, health or public

Table 2. AI governance policy toolkit per country.



Notes: Data based on a qualitative content analysis of 23 governmental AI strategies. See also section four and six in the online supplemental appendix. The percentage values indicate the ratio of policy instrument types per governmental strategy, whereas absolute numbers are given in italics at the upper and right end of the table. For instance, out of the 55 instruments identified in the AI strategy of Lithuania, 12.7% refer to outreach and literacy. Red highlighting of percentages signals outlier cases whose ratio of instruments lies 1.5 standard deviations or more above the mean for the instrument use (row percentages).

administration massively dominate, whereas incentive-based financial instruments (125) like tax breaks to increase research in and the uptake of AI or vouchers for vocational training and further education are only rarely mentioned. While tools of public investments thus play a large role in all AI strategies, they are especially prominent in China, Japan, and Korea, but also Austria and Portugal. The AI strategies of Cyprus, the Czech Republic, Malta, and Estonia, on the other hand, are characterized by a comparatively large number of incentive-based instruments which is (almost) double the average across all strategies.

The category *authority,* comprising direct regulations as well as governmental strategies, includes 304 policy instruments. Here, hard measures such as laws, directives, and other regulatory tools (158) are as common as soft measures including policy guidelines and frameworks (146). Prominent issues addressed via direct regulation are, inter alia, immigration laws to attract skilled labour, threats related to cybersecurity, or issues related to data protection. Moreover, flexible tools of governance like regulatory sandboxes to test and, if necessary, update existing legislation are also frequently mentioned. Looking at the national strategies more closely, we see that countries like France and the Czech Republic, but also Indonesia and Russia predominantly propose direct regulations. Non-binding government arrangements or the commission of strategic planning documents, frameworks, and guidelines, on the other hand, make up a large proportion of the national AI strategies of Norway, Denmark, and the United States.

A total of 298 instruments refer to *organizations,* covering networks and other institutions. These instruments pertain to the establishment of networks (154) to facilitate dialogue and collaboration between relevant stakeholders via hubs, hackathons, or other public-private fora, as well as the creation of institutions (144) like ethic councils, advisory boards, or governmental steering groups on AI. Networking instruments are prominent in Sweden and the Czech Republic, while the AI strategies of Germany, France and the EU include many instruments with the goal of setting up formal institutions.

Finally, we found 332 *informational instruments*, which include the three subcategories data and monitoring, certificates and labels, as well as outreach and literacy measures. Here, the largest subtype refers to data gathering and monitoring (187), a set of tools whose main purpose is to collect, monitor, analyse, and share and analyse data on AI-related trends and developments. Looking at differences between national AI strategies, it becomes apparent that such tools are especially important in the United States. Public outreach measures (82), on the other hand, include AI literacy programmes for the public at large or specific segments of it, as well as public information campaigns to raise awareness for and increase trust in AI-based technologies.

Instruments like these are prominently featured in the AI strategies of the Baltic states like Lithuania and Estonia, but also Malta. Measures aiming at the self-regulation of businesses as well as informational tools for consumers are categorized as certificates, standards, and labels (63), and are frequently mentioned in the AI strategies of France, Austria, and Indonesia.

In sum, this first descriptive account highlights the variance in the level of activity and the policy mix across the analysed AI strategies. For a more systematic analysis of the respective governance regimes, we next turn to the exploration of state types.

State types in the governance of AI

Given that governments adopt a comprehensive AI strategy, what mode of governance and resulting state types can we identify?[4] To address this question, we empirically map the four ideal-typical state types, i.e., the entrepreneurial state, the market-oriented state, the regulatory state, and the self-regulation-promoting state. Table 3 gives an overview how we operationalize each state type (revisit Figure 1) based on the mix of policy instruments we expect it to display (for the following, we draw on Bisson et al., 2010; Borrás & Edler, 2020; Marchant et al., 2020; Saurwein et al., 2015).

AI governance arrangements in line with the entrepreneurial state includes an above average number of instruments related to governmental strategies, direct public investments, network building as well as public outreach and literacy. The market-oriented state, in contrast, largely refrains from proposing direct regulations or making direct public investments. This, however, does not imply a total waiver of state involvement. Instead, we

Table 3. Expected policy instrument mix of the four state types.

Dimensions	Entrepreneurial state	Market-oriented state	Regulatory state	Self-regulation-promoting state
Authority				
Direct regulations	neutral	low	high	low
Governmental strategies	high	neutral	neutral	low
Finance				
Public investments	high	low	neutral	low
Financial incentives	neutral	high	neutral	neutral
Organization				
Institutions	neutral	neutral	high	low
Networks	high	neutral	low	high
Information				
Data & Monitoring	neutral	neutral	high	neutral
Certificates & Labels	neutral	high	neutral	high
Outreach & Literacy	high	neutral	neutral	neutral

Notes: 'High' indicates that we expect a high number of policy instruments to be generic for the respective state concept; 'low' indicates that we expect a low number of policy instruments to be generic for the respective state concept; 'neutral' means that we do not have clear theoretical assumptions, i.e., it can be both high or low numbers of policy instruments.

assume the facilitation of a conducive AI ecosystem via market-friendly instruments including incentive-based financial instruments in combination with soft measures like voluntary standards, certificates, or labels. The regulatory state is characterized by an above average number of laws and other forms of hard regulation as well as the creation of institutions like oversight bodies. Additionally, data collection and monitoring instruments should play an important role, whereas networking tools should be underdeveloped. Finally, the self-regulation-promoting state is constituted by an above average use of instruments related to networks and certificates and labels to make private engagement visible, while stepping back from direct regulations, government strategies and public investment programmes.[5]

Table 4 and Figure 2 present the results of a fuzzy-set ideal type analysis (Fiss, 2011; Kvist, 2007; detailed in section eight of the online supplemental appendix). It provides us with 'a calculus of compatibility' (Kvist, 2007, p. 474), or in other words, the proximity of the theoretically defined state types with the empirically observed patterns in AI governance. The higher the values, the stronger the affinity. Thirteen national AI strategies are connected to the concept of the entrepreneurial state. Japan, China, and the Republic of Korea display a strong affinity to the entrepreneurial state, but

Table 4. State type characteristics of AI strategies.

Country	Entrepreneurial state	Market-oriented state	Regulatory state	Self-regulation-promoting state
AUT	0.03	**0.53**	0.25	**0.68**
CHN	**0.71**	0.50	**0.75**	0.25
CYP	**0.74**	0.37	**0.62**	0.41
CZE	**0.59**	0.43	**0.69**	0.47
DNK	0.48	0.24	**0.57**	0.22
EST	**0.59**	**0.54**	0.25	0.50
EU	0.20	0.50	**0.51**	0.47
FIN	**0.73**	0.49	**0.62**	0.43
FRA	0.04	**0.55**	0.33	**0.65**
DEU	**0.75**	0.34	**0.63**	0.43
HUN	**0.79**	0.40	**0.76**	0.45
IDN	0.40	0.39	**0.90**	0.29
JPN	**0.79**	0.50	0.30	**0.62**
LIT	0.41	**0.54**	0.49	**0.60**
LUX	**0.54**	**0.54**	0.35	**0.55**
MLT	**0.72**	0.31	**0.80**	0.24
NLD	**0.76**	0.47	**0.73**	0.34
NOR	**0.58**	0.45	**0.81**	0.36
PRT	0.33	0.38	0.26	**0.56**
RUS	0.10	**0.61**	0.43	**0.68**
KOR	**0.73**	0.42	**0.54**	0.35
SWE	0.10	**0.53**	0.18	**0.72**
USA	0.31	**0.58**	0.38	**0.57**

Notes: Values of 0.85 or higher signal a very strong affinity with the respective state type; values between 0.84 and 0.70 reflect a strong affinity; values between 0.69 and 0.55 reflect a stable but weak affinity; values between 0.54 and 0.51 display an unstable affinity; values below 0.5 should be interpreted vice versa.

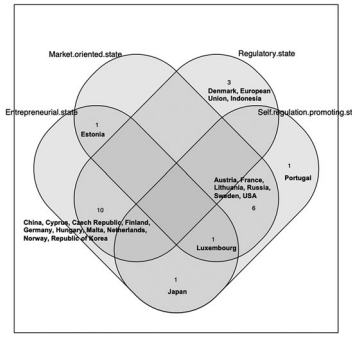


Figure 2. State type characteristics of AI strategies.

also several northern and central European countries like Finland, Netherlands, Germany, or Hungary. Malta and Cyprus also fit well, whereas Estonia, the Czech Republic, Luxembourg, and Norway show only a weak affinity. Eight AI strategies, i.e., Austria, Estonia, France, Lithuania, Luxembourg, Russia, Sweden, and the United States, are linked to the market-oriented state. The proximity, however, is much lower with Russia (value of 0.61) and USA (0.58) showing the highest affinity. From the thirteen states that are linked to the regulatory state, Indonesia (0.90) stands out as having a particularly strong resemblance, trailed by the AI strategies of Norway, Malta, Hungary, and China. Finally, nine states can be characterized as self-regulation-promoting states, with Sweden, but also Austria, Russia, and France showing the closest resemblance to this state type.

Looking at these empirical patterns of AI governance state types, several points are particularly worth pointing out. First, our results indicate that state approaches to AI governance indeed resemble our theoretically defined state concepts since all governmental strategies exhibit a high resemblance with at least one state type. Critical cases are Luxembourg and the European Union which show only weak relations to their respective state types. Second, it strikes the eye that only a few governmental AI strategies belong to one state type: Among the four states with a unique AI approach (Indonesia, Denmark, EU, and Portugal), Indonesia stands out with its clear strong regulatory approach towards AI. Third, and this is the mirror-image, empirical reality is characterized by a mix of governance approaches. In particular, we observe that the market-oriented state frequently goes hand in hand with the self-regulation-promoting state (in six cases), whereas the entrepreneurial state often coincides with the regulatory state (in ten cases). This indicates that the conceptual divide is between the market-oriented and self-regulation-promoting approach, on the one hand, and the regulatory and entrepreneurial approach, on the other.

In sum, a dominant group of countries is signalling in their strategies that a proactive state should do both, promote, but also regulate AI-related aspects. Although dimensions are not balanced since some countries have a closer proximity to the entrepreneurial state, while others have more in common with a regulatory state,[6] we interpret the prominence of this combination as a sign that for a large subset of AI governance arrangements in our sample both aspects of guarding and facilitating are crucial and that the state's role here is seen as both the protector and the producer of collective goods.

Public responsibility in policy instrument use

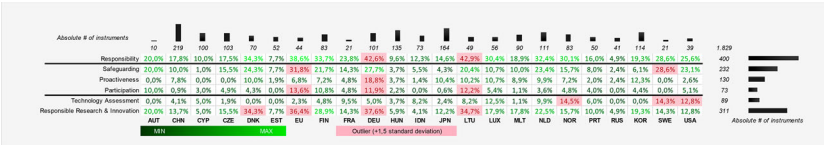
AI strategies also reveal the stance which policymakers take concerning public responsibility. This topic is not necessarily tied to one of the state types. Therefore, we determined the general stance of each state towards responsibility before looking at the relationship between state types and responsibility

Public responsibility in AI governance: level of activity and instrument use

The analysis of policy instruments reveals where the states stand between orthodox TA and RRI. The overall number of RRI instruments is higher, than TA which means that this approach has become more important as opposed to orthodox TA. While there is no complete deference of responsibility, the data clearly shows that in all countries the majority of instruments have no direct link to public responsibility.

Table 5 presents the data on policy instruments that address the responsible governance of AI (see also section five and seven in the online supplemental appendix). All in all, 400 out of the 1.829 instruments relate to responsibility; in other words, more than one-fifth of all policy measures (21.9 per cent) contain a reference to a responsible governance of AI. There is, however, considerable variation regarding the stance of states towards responsibility. In particular, states like Germany and Lithuania followed by the Scandinavian countries Denmark, Finland, and Norway and the EU, Luxembourg and the Netherlands stand out with 30 per cent to 40 per cent of its policy instrument mix showing a reference to responsibility. At the other end of the spectrum, we find Russia, Estonia and Hungary, Cyprus, Indonesia, and Japan which show a below-average activity when it comes to highlighting some form of responsible governance (see row 'responsibility' in Table 5). In sum, states seem to have quite different views to what extent responsibility needs to be ensured in the governance of AI.

Table 5. Policy instruments with responsibility reference.



Notes: Data based on a qualitative content analysis of 23 governmental AI strategies. See also section five and seven in the online supplemental appendix. The percentage values indicate the percentage of measures with different forms of responsibility assignment per governmental strategy, whereas absolute numbers are given in italics at the upper and right end of the table. For instance, out of the 49 instruments identified in the AI strategy of Lithuania, 42.9% contain a reference to public responsibility. Moreover, 34.7% of all instruments in the national AI strategy of Lithuania fall into the category of RRI and 8.2% into the category of TA Red highlighting percentages signal outlier cases whose ratio of instruments lies 1.5 standard deviations or more above the mean for the instrument use (row percentages).

States show very different levels of activism concerning the implementation of responsibility. They also differ regarding whether they lean towards an orthodox TA approach or an RRI approach (see rows 'technology assessment' and 'responsible research & innovation' in Table 5). One interesting observation in that regard is that some states have (almost) no instruments that can be categorized as orthodox TA. For Austria, Denmark, Estonia, Russia, and the Republic of Korea no instrument refers to TA, while China, the Czech Republic, Germany, the EU, Japan, and Malta are all below 5 per cent. The ratio of instruments that are devoted to an RRI approach compared to instruments that warrant a TA approach is 311 to 89. We can thus infer that most states treat AI governance as a task that needs active approaches to ensure responsibility. However, there is a group of states like Norway, Sweden, and the United States having an even balance between TA and RRI instruments.

Looking at the dimensions more granularly (see respective rows in Table 5), it stands out that states view the topic of participation very differently. While there is a large group for which participation plays (almost) no role including China, Estonia, Japan, Russia, and Sweden, most states can be categorized around the mean of 4.7 per cent. Lithuania, Germany, Finland, and the EU display a high ratio of participatory instruments. Participation measures include questions like creating a public debate through organizations (Norway, 2020, p. 60), or to focus on employers and employees (Czech Republic, 2019, p. 30). The most relevant dimension is the goal of safeguarding values with 12.7 per cent of all instruments. These instruments often relate to regulation pointing to 'defined rights and responsibilities' (United Kingdom, 2018, p. 16) or 'innovative regulation in the context of data marketplaces in order to increase legal certainty and transparency' (Luxembourg, 2019, p. 19).

There is a tension between the dimensions of safeguarding and proactiveness, since the former aims to protect values in the use of AI for other purposes, while the latter assumes that AI is used for the realization of ethical, legal, and social values. The dimension of proactiveness comprises 130 measures (7.1 per cent). While this is less than the dimensions of safeguarding, there is a group of states in which proactive measures almost equal safeguarding measures such as China, Luxembourg, and Malta. Goals of proactive realization include the 'quality of urban life and the productivity and safety of the construction industry' (South Korea, 2019, p. 40), or 'early warning programs for energy resource consumption and environmental pollutant discharge' (China, 2017, p. 20).

Linking public responsibility and state types in AI governance

Finally, we are interested in the question whether the different state types and regulatory regimes also imply specific notions of public responsibility? Concerning the expected relationship, we can formulate the following hypotheses (based on European Commission, Directorate-General for Research and Innovation & Stilgoe, 2019; Stilgoe & Guston, 2017; Wynne, 1975): For the marked-oriented state, we assume that its orientation towards weak state intervention will result in a deferral of responsibility. The self-regulation-promoting state fits either with a deferral of responsibility or a TA approach. Conversely, it would be possible that entrepreneurial states take a proactive approach regarding responsibility with a strong RRI approach. However, since they have an active interest in advancing technology and are themselves drivers of technological development, one could also expect that they defer responsibility. Concerning the regulatory state, we can assume that it has a high level of responsibility, however here the question is whether we deal with a TA or an RRI approach. It could be argued in favour of a combination with TA which tends to rely on the ordinary processes in government and, therefore, is only in need of a general assessment of the technology. However, the impetus to regulate might also motivate those states to engage deeper into the governance of technologies and to employ an RRI approach.

To shed light on possible linkages, we correlate the four state types with the different dimensions of responsible AI governance (see section eight in the online supplemental material). In sum, the results show – contrary to our theoretical expectations – rather low (or no) correlations between state types and the uptake of responsibility. This applies in particular when we look at the relative shares and not at the absolute numbers of responsibility instruments in an AI strategy. Regarding the absolute numbers of responsibility instruments, we find significant positive relationships with entrepreneurial and regulatory state types. This is not only true for the total number of

responsibility instruments, but also if we look at the absolute numbers of instruments for the subcategories safeguarding, responsible research & innovation (RRI), proactiveness (only for entrepreneurial states) and technology assessment (TA, only for regulatory states). While participation instruments are in no significant relationship to any of our identified state types, we also find no significant positive correlations of states with a market-oriented or self-regulation-promoting approach and any dimension of responsibility. For the absolute numbers of instruments and the subcategories safeguarding, proactiveness and RRI we find significant negative correlations with a self-regulation-promoting state. If we consider the percentage of responsibility instruments per strategy, we only find significant (positive) correlations for instruments associated with proactiveness and entrepreneurial state types.

Since the correlations are rather weak and apart from proactive instruments and entrepreneurial state orientation not robust, we interpret these findings with caution. One interpretation would be to associate entrepreneurial states with RRI, regulatory states with orthodox TA and market oriented and self-regulation promoting states with deference. The fact that there is no strong relation also shows something else. Other than commonly assumed, we find responsibility to be a crosscutting topic that is not tied to any state type. This narrative would reinforce that all states promote responsibility, but they pursue this aim through different means.

Conclusion

The analysis of AI strategies allowed us to reconstruct how governments and policymakers view their role regarding technology and public responsibility for AI-based technologies. We have shown that there is considerable variation in how states approach the governance of AI concerning the level of activity and the particular policy mix, both regarding the policy instruments proposed and the focus on responsibility.

One of our main findings is that we see a major difference between states which follow a *marked-oriented* and/or *self-regulation-promoting* approach, on the hand, and a group of countries which combine both *entrepreneurial* and *regulatory* governance approaches, on the other hand, with some countries showing a closer affinity to one type than the other. This key finding qualifies dominant narratives in public and political debates which frequently paint a picture according to which policymakers have to choose between the promotion of emerging technologies *or* the regulation of corresponding risks. At least for the field of AI, the majority of governments demonstrate the willingness to do both. Instead, we find a central watershed between active states and those that hold back from state interference. To what extent this pattern also applies for other emerging technologies is of central importance not only from an empirical perspective but also further

theory development. It points out that the dimensions of active versus passive states should be put more at the centre of the governance debate on emerging technologies. In contrast, the distinction between regulating and furthering technologies is less relevant.

We find in our analysis that the topic of public responsibility is not tied to any specific state type. It is therefore an independent topos of politics and technology irrespective of how governments approach technology. The level of engagement of AI strategies in instruments promoting responsibility varies substantially. However, all strategies include some form of responsibility. Building upon the respective discourses especially in STS, we find that RRI is a much more common form of governance towards AI than orthodox TA, highlighting the fact that governments take participatory and proactive actions seriously when thinking about the implementation of responsibility.

Our study has several limitations that indicate important future research avenues. First, with more and more countries passing national AI strategies, it will be crucial to enlarge the sample, particularly including more countries from the Global South. Second, we have only focused on the number of policy instruments mentioned in AI strategies. While this is an important first step to get a more systematic and evidence-based picture on the emerging governance of AI-based technologies, the next generation of research needs to expand this, for instance, by looking more closely at the issue areas that have been addressed, whether some instruments come with more commitment, resources etc. than others, whether certain mixes of policy instruments are used for specific problems. Third, our analysis is limited to the stage of policy formulation, which could be expanded both backward and forward along the policy process. Future studies will need to examine the implementation of the policy measures and whether these have an effect on AI-related policy outcomes. To this end, we also need to assess AI governance structures and ecosystems in a more comprehensive manner by looking beyond AI strategies at implementation and practices. For instance, our analysis puts the USA in the market-oriented and self-regulation-promoting approach. As pointed out before, our analysis might rather show how the USA – and other states – picture themselves in their national AI strategy, namely as the main proponent of a market-oriented state, whereas the implementation of AI governance might look very different (see in line with this also Mazzucato, 2011).

Yet, we have still learned something about the self-ascription of governments and the way that they communicate their goals and plans. To explain why specific approaches are conducted by different countries and how these approaches play out in practice should be the focus of further research. Scholars are just at the beginning of systematically exploring the tools, arrangements, and ecosystems with which governments plan to steer the future trajectory of AI and other digital technologies that shape our present and will shape our future.

Notes

1. Following our previous work (König et al., 2022, p. 11),

   > we understand AI systems as implementations of digital technology that are designed to interact with a given environment and to realize predefined goals by processing data in ways which allow the system to modify its behaviour (Russell & Norvig, 2016). The program that guides the behaviour of the system is informed by data fed into the system and patterns derived from this data.

2. Section one in the online supplemental material outlines the case selection strategy. Out of 51 states that belong to the EU, OECD or G20, 13 countries did not have any official AI strategy at the end of 2020. This changed for some of these countries, such as Great Britain, in the following months; however, data collection ended at the end of 2020 so that strategies passed after this date were not included. From the remaining 38 countries, we excluded those which did not have a standalone AI strategy or had passed only sector-specific ones. Some countries have unofficial AI strategies which are not passed by government bodies and are therefore excluded. We treat the EU as a sovereign entity which is a member of the OECD and part of the G-20; more importantly, it is a key actor that shapes the development, deployment, and governance of AI.
3. Some instruments that did not directly show in the data have been left out. One example are state-owned companies. It is contended that those companies played a big role in strategy making in Russia (Petrella et al., 2021). The actors making the strategy are not identical with the instruments in the strategy. The fact that the Russian strategy opted for self-regulation might benefit the interests of state-owned companies as they have been part of self-regulatory efforts in the past.
4. Our analysis only includes comprehensive national strategies, i.e., a government strategy which cover the full spectrum and not only focuses on a specific policy sector like industry or research and development. For more information which countries drop from our sample, visit section two in the online supplemental appendix. Why certain states have not passed such a strategy is beyond the scope of our analysis. One possible interpretation is that countries doing so take an intentional decision by following implicitly or explicitly a laissez-faire state approach of non-intervention. Other explanations could be low state capacity, a low level of AI development in a country, being a general laggard concerning the digital transformation, among others. For the interpretation of national AI strategies including only a few policy instruments, the fallback option of having no strategy at all should, however, be kept in mind.
5. We focus on key components of the expected instrument mix for each state type. This does not mean that, for example, a regulatory state might not also provide financial incentives – e.g., via regulated tariffs or allowed return to investment – or that a market-oriented state would not build institutions to establish or protect specific market mechanisms. However, we would argue that a high usage of these instruments is not an essential component of the respective state type.
6. The same applies to the pair of market-oriented and self-regulation-promoting states.

Acknowledgements

The authors thank the two anonymous reviewers for the time and efforts spend on commenting previous versions of this manuscript. The final paper profited immensely from their feedback and suggestions. We would also like to extend our thanks to Jeremy Richardson, Berthold Rittberger and Tim Büthe for their constructive comments and for steering the manuscript through the publication process. Finally, we gratefully acknowledge the support by Sarah Kluge, Margareta Maier, Bernadette Spitzer and Lukas Wörlein during the process of data collection.

Notes on contributor

Christian Djeffal is Professor for Law, Science and Technology at Technical University of Munich (Germany). *Markus B. Siewert* is Senior Lecturer and Managing Director of the TUM Think Tank at the Munich School of Politics and Public Policy, Technical University of Munich (Germany). *Stefan Wurster* is Professor for Policy Analysis at the Munich School of Politics and Public Policy, Technical University of Munich (Germany).

ORCID

Markus B. Siewert http://orcid.org/0000-0002-8710-531X
Stefan Wurster http://orcid.org/0000-0003-2532-0249

References

Bisson, P., Kirkland, R., & Stephenson, E. (2010). *The market state*. McKinsey & Company.

Borrás, S., & Edler, J. (2020). The roles of the state in the governance of socio-technical systems' transformation. *Research Policy*, *49*(5), 103971. https://doi.org/10.1016/j.respol.2020.103971

Braithwaite, J. (2011). The regulatory state? In R. E. Goodin (Ed.), *The Oxford Handbook of political science* (pp. 1–23). University Press.

Brooks, H. (1976). II. Technology assessment in retrospect. *Newsletter on Science, Technology & Human Values*, *1*(4), 17–29. https://doi.org/10.1177/016224397600100403

Broussard, M. (2019). *Artificial unintelligence: How computers misunderstand the world*. MIT Press.

Bughin, J., Hazan, E., Allas, T., Hjartar, K., Manyika, J., Sjatil, P. E., & Shigina, I. (2019). *Tech for good. Smoothing disruption, improving well-being*. McKinsey Global Institute.

Casado-Asensio, J., & Steurer, R. (2014). Integrated strategies on sustainable development, climate change mitigation and adaptation in western Europe: communication rather than coordination. *Journal of Public Policy*, *34*(3), 437–473. https://doi.org/10.1017/S0143814X13000287

Chessen, M. (2017). *The Madcom future: How artificial intelligence will enhance computational propaganda, reprogram human culture, and threaten democracy … and What can be done about it*, Atlantic Council.

Crawford, K. (2021). *Atlas of AI. Power, politics, and the planetary costs of artificial intelligence*. Yale University Press.

Dauvergne, P. (2021). The globalization of artificial intelligence: Consequences for the politics of environmentalism. *Globalizations*, *18*(2), 285–299. https://doi.org/10.1080/14747731.2020.1785670

Di Stefano, G., Gambardella, A., & Verona, G. (2012). Technology push and demand-pull perspectives in innovation studies: Current findings and future research directions. *Research Policy*, *41*(8), 1283–1295. https://doi.org/10.1016/j.respol.2012.03.021

Djeffal, C. (2019a). AI, democracy, and the law. In A. Sudmann (Ed.), *The democratization of artificial intelligence: Net politics in the era of learning algorithms* (pp. 255–284). transcript Verlag.

Djeffal, C. (2019b). Sustainable development of Artificial Intelligence (SAID). *Global Solutions Journal*, *4*, 186–192.

Djeffal, C. (2020). Artificial intelligence and public governance: Normative guidelines for artificial intelligence in government and public administration. In T. Wischmeyer & T. Rademacher (Eds.), *Regulating artificial intelligence* (pp. 227–293). Springer Verlag.

Ely, A., van Zwanenberg, P., & Stirling, A. (2014). Broadening out and opening up technology assessment: Approaches to enhance international development, coordination and democratisation. *Research Policy*, *43*(3), 505–518. https://doi.org/10.1016/j.respol.2013.09.004

Erdelyi, O. J., & Goldsmith, J. (2018). Regulating artificial intelligence proposal for a global solution. In J. Furman, G. Marchant, H. Price, & F. Rossi (Eds.), *Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society* (pp. 95–101). Association for Computing Machinery.

European Commission. (2020a). *Proposal for a Regulation of the European Parliament and of the Council on a Single Market For Digital Services (Digital Services Act) and Amending Directive* 2000/31/EC (COM(2020) 825 final.

European Commission (2020b) *Digital Economy and Society Index (DESI) 2020*. Brussels: European Commission. https://digital-strategy.ec.europa.eu/en/library/digital-economy-and-society-index-desi-2020

Fatima, S., Desouza, K. C., & Dawson, G. S. (2020). National strategic artificial intelligence plans: A multi-dimensional analysis. *Economic Analysis and Policy*, *67*, 178–194. https://doi.org/10.1016/j.eap.2020.07.008

Felt, U., & Wynne, B. (2007). *Taking European knowledge society seriously: Report of the expert group on science and governance to the science, economy and society directorate, directorate-general for research, European commission (EUR: Vol. 22700)*. Office for Official Publications of the European Communities.

Fisher, E., Mahajan, R. L., & Mitcham, C. (2006). Midstream modulation of technology: Governance from within. *Bulletin of Science, Technology & Society*, *26*(6), 485–496. https://doi.org/10.1177/0270467606295402

Fiss, P. C. (2011). Building better causal theories: A Fuzzy set approach to typologies in organization research. *Academy of Management Journal*, *54*(2), 393–420. https://doi.org/10.5465/amj.2011.60263120

Gasser, U., & Almeida, V. A. F. (2017). A layered model for AI governance. *IEEE Internet Computing*, *21*(6), 58–62. https://doi.org/10.1109/MIC.2017.4180835

Guston, D. H. (2014). Understanding 'Anticipatory Governance'. *Social Studies of Science*, *44*(2), 218–242. https://doi.org/10.1177/0306312713508669

Guston, D. H., & Sarewitz, D. (2002). Real-time technology assessment. *Technology in Society*, *24*(1-2), 93–109. https://doi.org/10.1016/S0160-791X(01)00047-1

Hasselbalch, J. A. (2018). Innovation assessment: governing through periods of disruptive technological change. *Journal of European Public Policy*, *25*(12), 1855–1873. https://doi.org/10.1080/13501763.2017.1363805

Hennen, L. (2012). Why do we still need participatory technology assessment?. *Poiesis & Praxis*, *9*(1-2), 27–41. https://doi.org/10.1007/s10202-012-0122-5
Hood, C., & Margetts, H. (2007). *The tools of government in the digital age*. Macmillan.
Howlett, M. (2019). *The policy design primer*. Routledge.
Khodabandeh, S., Ransbotham, S., Kiron, D., Candelon, F., Chu, M., & LaFountain, B. (2020). *Expanding AI's impact with organizational learning*. MIT Sloan Management Review and Boston Consulting Group. https://sloanreview.mit.edu/projects/expanding-ais-impact-with-organizational-learning/.
Kim, Y. T. (2007). The transformation of the East Asian States: From the developmental state to the market-oriented state. *Korean Social Science Journal*, *59.2*, 31–60.
Knill, C., Schulze, K., & Tosun, J. (2012). Regulatory policy outputs and impacts: Exploring a complex relationship. *Regulation & Governance*, *6*(4), 427–444. https://doi.org/10.1111/j.1748-5991.2012.01150.x
König, P., Wurster, S., & Siewert, M. B. (2022). Consumers are willing to pay a price for explainable, but not for green AI. Evidence from a choice-based conjoint analysis. *Big Data & Society*, *9*(1), 1–13. https://doi.org/10.1177/20539517211069632
Kuhlmann, S., Stegmaier, P., & Konrad, K. (2019). The tentative governance of emerging science and technology: A conceptual introduction. *Research Policy*, *48*(5), 1091–1097. https://doi.org/10.1016/j.respol.2019.01.006
Kvist, J. (2007). Fuzzy set ideal type analysis. *Journal of Business Research*, *60*(5), 474–481. https://doi.org/10.1016/j.jbusres.2007.01.005
Majone, G. (1997). From the positive to the regulatory state: Causes and consequences of changes in the mode of governance. *Journal of Public Policy*, *17*(2), 139–167. https://doi.org/10.1017/S0143814X00003524
Mandel, G. N. (2009). Regulating emerging technologies. *Law, Innovation and Technology*, *1*(1), 75–92. https://doi.org/10.1080/17579961.2009.11428365
Marchant, G., Tournas, L., & Gutierrez, C. I. (2020). Governing emerging technologies through soft law: Lessons for artificial intelligence—an introduction. *JURIMETRICS*, *61.1*, 1–18.
Marchant, G. E., & Wallach, W. (2013). Governing the governance of emerging technologies. In G. E. Marchant, K. W. Abbott, & B. Allenby (Eds.), *Innovative governance models for emerging technologies*, (pp. 136–152). Edward Elgar Publishing.
Mazzucato, M. (2011). *The entrepreneurial state*. Demos.
Mittelstadt, B. (2019). Principles alone cannot guarantee ethical AI. *Nature Machine Intelligence*, *1*(11), 501–507. https://doi.org/10.1038/s42256-019-0114-4
Nemitz, P. (2018). Constitutional democracy and technology in the age of artificial intelligence. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *376*(2133), 20180089. https://doi.org/10.1098/rsta.2018.0089
OECD. (2019). *Artificial intelligence in society*. OECD Publishing.
Ornetzeder, M., & Kastenhofer, K. (2012). Old problems, new directions and upcoming requirements in participatory technology assessment. *Poiesis & Praxis*, *9*(1-2), 1–5. https://doi.org/10.1007/s10202-012-0116-3
Petrella, S., Miller, C., & Cooper, B. (2021). Russia's Artificial Intelligence strategy: The role of state-owned firms. *Orbis*, *65*(1), 75–100. https://doi.org/10.1016/j.orbis.2020.11.004
Radu, R. (2021). Steering the governance of Artificial Intelligence: National strategies in perspective. *Policy and Society*, *40*(2), 178–193. https://doi.org/10.1080/14494035.2021.1929728

Rayner, J., & Howlett, M. (2009). Introduction: Understanding integrated policy strategies and their evolution. *Policy and Society*, *28*(2), 99–109. https://doi.org/10.1016/j.polsoc.2009.05.001

Russell, S. J., & Norvig, P. (2016). *Artificial intelligence: A modern approach* (3rd ed). Pearson.

Sarewitz, D. (2011). Anticipatory governance of emerging technologies. In G. E. Marchant, B. R. Allenby, & J. R. Herkert (Eds.), *The growing gap between emerging technologies and legal-ethical oversight: The pacing problem* (pp. 95–105). Springer Netherlands.

Saurwein, F., Just, N., & Latzer, M. (2015). Governance of algorithms: Options and limitations. *info*, *17*(6), 35–49. https://doi.org/10.1108/info-05-2015-0025

Schaffrin, A., Sewerin, S., & Seubert, S. (2015). Toward a comparative measure of climate policy output. *Policy Studies Journal*, *43*(2), 257–282. https://doi.org/10.1111/psj.12095

Simonis, G. (2013). Technology governance. In G. Simonis (Ed.), *Konzepte und Verfahren der Technikfolgenabschätzung* (pp. 161–186). Springer.

European Commission, Directorate-General for Research and Innovation & Stilgoe, J. (2019). *Monitoring the evolution and benefits of responsible Research and Innovation*. Publications Office. https://data.europa.eu/doi/10.2777/285467.

Stilgoe, J., & Guston, D. H. (2017). Responsible research and innovation. In U. Felt, R. Fouché, C. A. Miller, & L. Smith-Doerr (Eds.), *The Handbook of Science and Technology studies* (pp. 853–880). MIT Press.

Taeihagh, A. (2021). Governance of Artificial Intelligence. *Policy and Society*, *40*(2), 137–157. https://doi.org/10.1080/14494035.2021.1928377

Taeihagh, A., Ramesh, M., & Howlett, M. (2021). Assessing the regulatory challenges of emerging disruptive technologies. *Regulation & Governance*, *15*(4), 1009–1019. https://doi.org/10.1111/rego.12392

Thierer, A., Castillo O'Sullivan, A., & Russell, R. (2017). *Artificial intelligence and public policy*. Mercatus Center, George Mason University.

Trajtenberg, M. (2019). Artificial Intelligence as the next GPT: A political-economy perspective. In A. Agrawal, J. Gans, & A. Goldfarb (Eds.), *The Economics of Artificial Intelligence. An agenda* (pp. 175–186). University of Chicago Press.

Ulnicane, I., Knight, W., Leach, T., Stahl, B. C., & Wanjiku, W.-G. (2020). Framing governance for a contested emerging technology: Insights from AI Policy. *Policy and Society*, *40.2*, 158–177. https://doi.org/10.1080/14494035.2020.1855800

United Nations General Assembly. (2019). *Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression*. A/74/486. https://www.undocs.org/A/74/486.

Van Dijck, J. (2013). *The culture of connectivity - a critical History of Social media*. Oxford University Press.

Van Zwanenberg, P., Ely, A., & Stirling, A. (2009). *Emerging technologies and opportunities for international science and technology foresight STEPS Working Paper* 30. Brighton: STEPS Centre, 4–13.

Wilsdon, J., & Willis, R. (2004). *See-through science: Why public engagement needs to move upstream*. Demos.

World Bank. (2021). *Harnessing Artificial Intelligence for development on the post-COVID-19 era: A review of national AI strategies and policies*. World Bank.

Wynne, B. (1975). The Rhetoric of consensus politics: A critical review of technology assessment. *Research Policy*, *4*(2), 108–158. https://doi.org/10.1016/0048-7333(75)90028-1