Industrial Policy for Emerging Technologies: The Case of Narrow AI and the Manufacturing Value Chain as Blueprint for the Industrial Metaverse

Dietlmeier, Simon Frederic

University of Cambridge, Department of Engineering, Institute for Manufacturing, Sidney Sussex College

12 June 2024

Industrial Policy for Emerging Technologies: The Case of Narrow AI and the Manufacturing Value Chain as Blueprint for the Industrial Metaverse

Simon Frederic Dietlmeier – Corresponding Author
Email sfd37@cam.ac.uk
Phone +49 151 19417611

Supervisor:

Dr Eoin O’Sullivan

University of Cambridge
Centre for Science, Technology and Innovation Policy (CSTIP)
Institute for Manufacturing
Department of Engineering
17 Charles Babbage Rd
Cambridge, CB3 0FS

Industrial Policy for Emerging Technologies:
The Case of Narrow AI and the Manufacturing Value Chain

Abstract

In this paper, a qualitative model is inductively developed describing a dynamic "policy mix" -system of innovation enabling and outbalancing dimensions for the deployment of narrow artificial intelligence (AI) in the manufacturing value chain. A literature review first identifies and summarizes general policy recommendations on AI as an emerging technology presented by authors prior to this research. In the empirical part, policy dimensions and suggestions of policy remedies with a focus on the manufacturing value chain were taxonomized based on exploratory interviews with 37 international elite experts on AI across several stakeholder groups. The findings were refined in a survey with participants of the workshop "AI in Manufacturing" organized by the European Commission. The dimensions build the foundation for an industrial policy in the form of a "four-wing industrial policy system model" that can unleash the value of narrow AI in the manufacturing value chain and addresses barriers to scale-up. It represents a qualitative modelling approach and confirms previous views in the literature that innovation policies need to be thought as "policy mix" and systems. A case study of the European Union's policy mix for AI validates the model empirically based on additional interviews with ten European civil servants.

Keywords: Artificial Intelligence • Emerging Technologies • Manufacturing • Value Chain • System • Policy Mix

1 Introduction

Manufacturing firms have discovered the potential of narrow artificial intelligence (AI) as emerging technology for an application in their own manufacturing value chain. The desire is to make processes more efficient and to capitalize on various value propositions the technology promises for the fourth industrial revolution (Schwab, 2017, p. 12). Consulting firms have attributed AI a great economic potential and value-add to the manufacturing sector when the technology is deployed along the manufacturing value chain (McKinsey & Company, 2018; Accenture, 2019). The current state, however, seems to showcase a rather low adoption rate, like for many other technologies within the spectrum of "Industry 4.0" (Atkinson, 2019, p. 15). The CEO of a global strategy consultancy indeed stressed in one of the background interviews for this paper on AI: "It is overhyped, where it is also underhyped. It is overhyped on the magical part, it is underhyped on the practical part."

Policymaking nevertheless considers the technology to be a major driver for future prosperity. Many national AI strategies have been adopted, and many policy initiatives are being developed worldwide (see FTI, 2018; Deloitte, 2019). All of those strategies have, however, started with a non-sectoral view of AI development or adoption and comprehended it rather as a transversal topic without focusing on the specific context of its application. A leading British politician asserted in another interview: "I don't think AI at this present moment, certainly not narrow AI, is completely exceptional. It needs to be dealt with in that context in a wholly different way. What I do think AI does though, it makes us think more than previously about how technologies are applied. So, I do think it's a kind of wake-up call a little bit for us when applying new technology."

Justified by the importance of the technology for the economy and policymaking, the present paper therefore poses the following research question: What policy remedies are needed to unleash the value AI can create in the manufacturing value chain and to address the barriers for scaling AI applications in the manufacturing value chain? Given the limited adoption rate of AI in the manufacturing value chain at the time when this research was conducted but high future value promises, it could be hypothesised that industrial policy has to play a major role in steering further adoption. As Oqubay (2015, p. 18) writes, industrial policy is "a strategy that includes a range of implicit or explicit policy instruments selectively focused on specific industrial sectors for the purpose of structural change in line with a broader national vision and strategy". Industrial policy was therefore assumed as being well suited to support the adoption of AI systems in the manufacturing value chain. However, it cannot be stressed enough that any such policy remedies are to be designed purely for a case of narrow AI, compared to other existing approaches that develop policy recommendations for artificial general intelligence (cf. Baum, 2018; Bostrom et al., 2018).

The results of this research indicate that an industrial policy which supports the deployment of AI systems in the manufacturing value chain should be comprehended as a "policy mix" -system of innovation enabling and counterbalancing dimensions that ensure human-centric, safe, and ethical deployment of the technology. This system needs to be dynamically adaptable over time and can react in anticipation of future automation and autonomy. The study also shows that there are aspects of the technology that require the due care and special attention by policymakers. A generalizability for emerging technologies and transfer of the policy model to cases of other newly emerging technologies might be possible.

2 Theoretical Background

Several concepts need to be introduced for an ontology clarification in the context of this study. The terminology combines ideas from the policy, economic and legal academic literature, which reflects the interdisciplinary character of the research. Policy modelling has been frequently applied in research on policy systems.

2.1 Terminology

The manifold definitions of AI could fill books, but in this paper a focus is laid on its applied characteristics. To account for the complexity of elements that must have the right format, like data and model, an ontological discussion is avoided by using the phrase "AI system" in the findings and discussion sections.

Calo (2017) argues for the use of the term "AI policy" other than "AI governance" or "AI ethics" to describe policy recommendations for AI due to "a degree of finality once promulgated" (p. 408). He suggests a few advantages through the usage of this term. It "admits of the possibility of new laws but does not require them" (p. 409), and it entails planning for later effects which are not yet foreseeable. Moreover, in the sense of a public policy, it can be influenced by industry, "but it is not its role ultimately to set it" (p. 410). The term policy is used in line with this definition.

Newly introduced are policy dimensions. They are defined as fields of action that require policymaking remedies for the purpose of creating an industrial policy for artificial intelligence deployment in the manufacturing value chain. A remedy is defined as a measure that either unleashes value of AI systems usage for the manufacturer or helps to bypass one of his barriers to adoption.

Governance is considered as an equivalent to the tangible policy structure that administers the policy from a procedural view. Regulation defines approaches to set formalized rules that should be adhered to by certain targeted stakeholders under the umbrella of policymaking. Ethical considerations form the moral boundaries of what should be facilitated by the policy, especially in the context of platforms (cf Gawer, 2014).

The research aims at formulating a framework that illustrates an industrial policy for the manufacturing sector with the strategic goal of facilitating the deployment of AI technologies in the manufacturing value chain. An industrial policy for AI is then a sub-field of the more generalized theoretical concept of an "AI policy". Industrial policy is defined in line with Cimoli et al. (2009, p. 2):

> "The notion of 'industrial policy' is understood here in a quite expansive manner. […] Industrial policies, in this broad sense, come together with processes of 'institutional engineering' shaping the very nature of the economic actors, the market mechanisms and rules under which they operate, and the boundaries between what is governed by market interactions, and what is not."

An industrial policy for AI can be understood as a governmental approach to form a sector-specific comprehensive industrial policy (cf. Graham, 1994), an approach which is modelled in the present paper for the development of the manufacturing sector as part of a polity's industrial portfolio.

Conceptual basis for the deployment of AI in the manufacturing sector forms Porter's (1985) idea of a value chain, which can provide a standardized and generic overview about the different functions of a manufacturing firm. A value chain consists of two types of firm activities in the firm (Figure 1).

Primary activities describe the process by which firms add value to their products and ultimately achieve their competitive advantage. It is characterized by the steps inbound logistics, operations and production, outbound logistics, the customer interface based on marketing and sales, as well as the potential after-sales service for the product.

Supporting activities serve as supporting mechanism for the firm's primary activities. This includes the establishment and maintenance of the firm's infrastructure, the relations with current and potential employees, technology development based on research and development (R&D), as well as the procurement of input material and equipment.

AI systems could potentially be deployed in all these functions, and an industrial policy would aim to support the uptake of the technology in primary and support activities, whilst at the same time regulating the modalities of this deployment such as on boundary sovereignty.

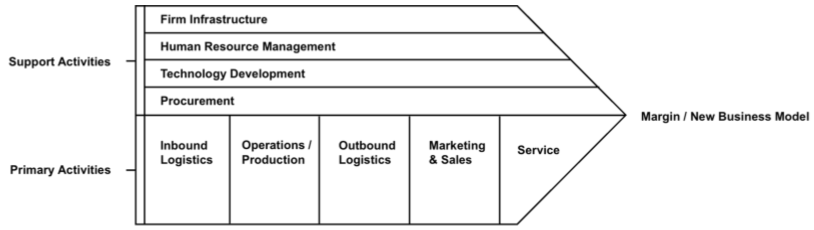


Figure 1 The Manufacturing Value Chain Functions following Porter (1985)

2.2 Modelling of Policy Systems

Foundational element of deriving policy recommendations for AI systems is the capability of modelling dimensional categories that require the attention and management of stakeholders (cf Freeman, 1984). The academic literature has developed various options of how such policy models – often comprising a mix of policy instruments – have been influencing the thinking in relevant circles (cf Magro & Wilson, 2013; Li & Garnsey, 2014). This adaptability might be due to the inter-temporal complexity of innovations, as approximated via S-Curves (cf Ghanbarnejad, 2014).

Within the frameworks provided, there can be logical and (co-)evolutionary approaches differentiated (cf Moore, 1993; Sotarauta & Srinivas, 2006). These comprise a perspective from a wider lens (macro – meso – micro) under which a model is constructed (cf Niderman & March, 2019), as well as the developmental approach differentiating deductive from inductive logic, or most complex likely abduction (cf Staat, 1993). Finally, the unit of analysis might require consideration, in a business context to distinguish between organizations like enterprises and networks, as well as depending on spatial viewpoints like regional or boundary-related value exchange (cf Padmore et al., 1998; Padmore & Gibson, 1998).

3 Literature Review

A systematic literature review on the policy issues and recommendations presented by other authors with regard to the technology "artificial intelligence" was performed prior to the empirical research. This was taken as theoretical input for the questionnaire development and to identify the state-of-the-art in the academic debate on AI policy.

3.1 Approach of the Systematic Literature Review

The systematic approach was chosen to capture the different policy angles and only academic papers were included to avoid researcher bias towards "interest group"-driven recommendations by private and public actors. It follows a nine steps approach adapted from Boland et al. (2013). *Scopus*™ was selected as academic search engine platform, and three review questions were proposed:

- Which different research perspectives exist in the interface of AI and policy?
- What policy problematics with regard to AI are identified?
- What policy recommendations do the authors give?

For the search string, it was decided to choose the key word "artificial intelligence" over further refinements to sub-methods like "machine learning" or "neural networks" due to the unclear ontology in the field, and the number and types of relevant papers had not changed after several tests of alternative keywords. Five key words were chosen in combination to the primary phrase: "regulation", "governance", "policy", "politics", and "innovation". A limitation to title and abstract was applied due to the inflationary use of "artificial intelligence" as keyword. Subject areas were limited to "social sciences", "business", "economics" and "arts", as a search in the subject areas engineering, computer science or mathematics had revealed predominantly technical papers. The search was conducted in Q2/2019 and the presented articles in the literature review include only articles that were published before. The search resulted 1.037 hits, which was refined to 191 articles after screening title and abstract.

It was found that many papers were uncited and non-peer reviewed. Therefore, conference papers, working papers, book chapters and books were included into the literature review in line with Heinonen et al. (2013, p. 343). Of the selected papers, a forward and backward search based on the reference lists was conducted (cf. Webster and Watson, 2002).

Quality criteria for the selected articles were journal origination, reputation of the author in the field, but also originality of the policy perspective presented. Most of the articles originated only after 2012 and with a steep increase from 2016. A total of 79 papers was selected into the final paper set after a full-text review.

The identified articles can be separated into three streams of research. A first stream thematizes only AI sub-policy aspects on governance and regulatory issues. The second stream considers AI policies as a bundle of parametrized policy remedies in the form of more comprehensive policy models. The third stream focuses on sectoral aspects. The present paper describes a comprehensive policy modelling approach with a sectoral focus, which represented a research gap by embedding the study into a combination of the second and third stream of literature (Figure 2).

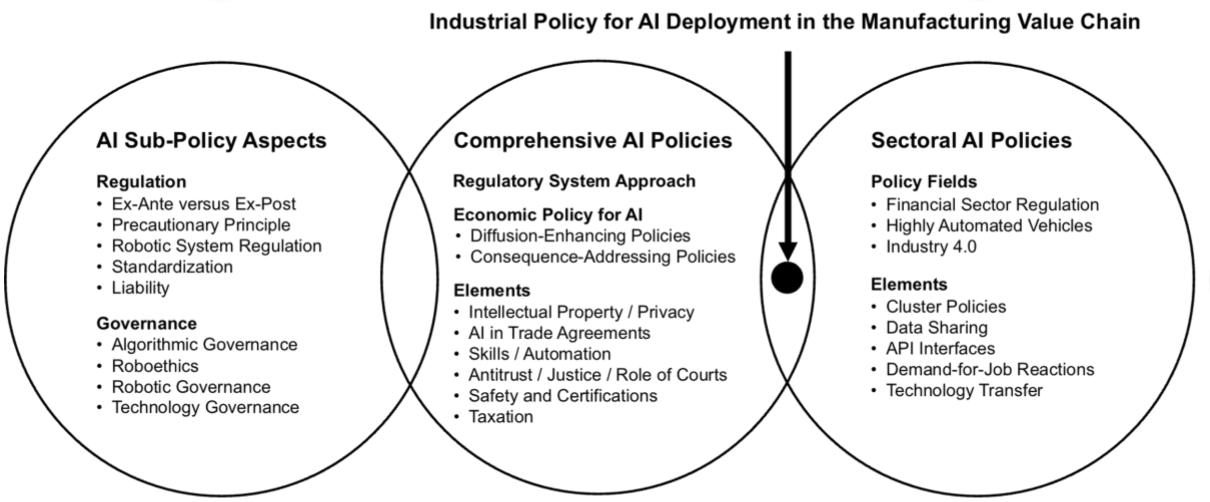


Figure 2 Summary of Existing Research Prior to this Study

3.2 AI Sub-Policy Aspects

Most articles in the first stream on AI sub-policies are related to ethical discussions on the characteristics of the technology, which call for ex-ante attention by policymakers. Baum (2017) suggested to steer the development of "beneficial AI", which is aligned to societal goals; not only with regulation, but also by utilizing intrinsic accelerating factors like social norms or "allies" of the technology. Cath et al. (2018) urge governments to develop a "Good AI Society" based on the concept of a "mature information society" by Floridi (2016), a policy goal she shows is at least partially included in three countries' strategy and legal frameworks. A number of papers further contribute to the domains of AI governance and AI regulation.

3.2.1 Governance

The most fundamental problems for AI have been researched in the adjunct field of algorithmic governance.

Ananny and Crawford (2018) discuss the theoretical concept of transparency and whether this creates accountability for a "black box" system, arguing in a typology of ten transparency limitations that this "ideal state" can hardly be realized. This however imposes liability concerns for malfunctioning AI systems (Russell and Norvig, 2016, p. 1036). Burrell (2016) introduced the concept of "opacity" of machine learning (ML) algorithms as a desired secrecy mechanism of state or private actors. Other issues quoted are potential inequality and discrimination arising through algorithms, which would require regulatory solutions (Pasquale, 2015; Barocas and Selbst, 2016; Goodman and Flaxman, 2017). Kroll et al. (2016) synthesize a "technological toolkit" that can help policymakers to ensure a fair and accountable decision-making process. Gandy (2010) called for a social movement to raise awareness of such issues. Rouvroy (2008) assessed legal implications arising out of privacy concerns in a human-virtual system and conclude that law would be needed to create a normative framework for such issues. Others have concurred (Agre and Rotenberg, 1998; Austin, 2003; Friedewald et al., 2007). Institutionalists have repeatedly argued for dedicated agencies to regulate such problems (Tutt, 2016; Scherer, 2015). And Wallach (2011) proposed a mechanism that can monitor and manage emerging technologies from a policy-side with the means of expert workshops and foresight.

Out of this research string, the fields of AI ethics and "roboethics" have emerged. Iphofen and Kritikos (2019) make the normative argument that any such AI system must be developed through an "ethics-by-design" approach or otherwise not being used as an "autonomous moral agent". A perspective that is deepened in relation to intentional malicious AI (Brundage et al., 2018). It is also discussed whether robots should be granted human rights and a legal personality (Wurah, 2017; Mcnally and Inayatullah, 1988; Solaiman, 2017; Sparrow, 2016). Boesl and Bode (2018) proposed a robotic governance framework as addition to roboethics to assemble relevant stakeholders and to engage in a discourse embedded in a code of conduct.

Boesl and Bode (2016) propose a holistic concept of "technology governance" applicable to "robotics, automation and artificial intelligence", which combines ethical, moral, and social aspects with political and economic realities (p. 421). They rely in the case of AI on Naisbitt's (1982) notion of a megatrend, suggest a "dynamical framework of guidelines", introduction of a stakeholder dialogue and the development of a guiding framework. They also see the need to take market pull and push factors into account that frame innovation. Moreover, they could imagine a differentiation of different technologies, with AI being steered by a dedicated AI governance. They further suggest an implementation cycle that starts with soft laws and standards, which would be embedded into corporate governance principles.

3.2.2 Regulation

In terms of regulation, the perspectives diverge between proponents for ex-post versus ex-ante regulation. Some argue that regulation should depend on the maturity of the technology, as much of the debate around this technology would be caused by the imagination of "anthropomorphizing AIs" (cf. Gurkaynak et al., 2016). Others take the opposite view to call for an early ex-ante regulation of narrow AI, for instance by ethical impact assessments that allow regulators to validate AI systems (Benrimoh et al., 2018).

Clarke (2005) assessed the impact of a precautionary principle in the development of future technologies, under which the development of AI could be banned by policymakers (see also Bodansky, 1991; Majone, 2002; Sunstein, 2002). And Maas (2018) argues that opaque narrow AI applications are already so interconnected, that a "normal" accident could lead to a cascade effect of further failures and needs tentative precautionary policymaking approaches

This is proposed to be merged with regulatory approaches for robotic systems (Palmerini et al., 2016), which potentially needs to differentiate between systems and would revise existing regulations. Bertolini (2013) make the liability case dependent on whether the autonomous character changes the robot from being an object or becoming a cognitive subject, highlighting the role of case law (see also Asaro, 2016, and Čerka et al., 2015, on robotic liability). Villaronga and Golia (2019) discuss the relative advantages of standard setting versus law-making for robot regulation. They argue that the lack of social legitimacy of standards is problematic but suggest the integration of impact assessments into policies and a multiple-stakeholder process to set the standards. By introducing a systems theory framework, they conclude that a combination of private standards and non-political considerations is necessary.

More generalized for future emerging technologies, Székely et al. (2011) develop "speculative" findings and "empirical" findings for law making approaches to regulate the future. They argue that transparency of the technologies and a sense for responsibilities of stakeholders might decrease, auditing might become more complicated, and inequalities increase. Timing of interventions, self-regulation approaches and stakeholder communication with lawmakers, also in order to establish governance rules, are recommended by them. Wright (2011) develops a general policy framework for an ethical impact assessment of future emerging technologies.

3.3 Comprehensive AI Policies

More and more authors have moved away from discussing isolated characteristics of the technology but take a holistic view of the policy issues and possible reactions.

A comprehensive approach has been undertaken by Agrawal et al. (2019) who try to construct an "Economic Policy for AI". Based on working papers from leading economists introduced during two *NBER* Conferences, they provide a consolidated analysis to describe how a price-drop in predictions induced by ML algorithms impacts society. The rational lies on the character of AI as a "general purpose technology", which is considered by Bresnahan and Trajtenberg (1995) as productivity-enhancing. Agrawal et al. (2019) conclude that policymaking on AI can be organized into diffusion-enhancing and consequence-addressing policies.

Agrawal et al. (2019) consider intellectual property assurance as a main driver of economic growth induced by AI, to which privacy regulations would be a contradicting obstacle (see also Goldfarb and Tucker, 2012). The inclusion of AI-related topics into future trade agreements is seen as rent-enabling mechanism. Finally, they argue that a liability risk would hamper the diffusion of AI.

Moreover, Agrawal et al. (2019) identify the impact on jobs, on inequality and antitrust issues as policy fields that might need to be addressed by policymakers. Acemoglu and Restrepo (2018) had argued that both, an increase in productivity, but also decrease in labor demand would be possible and show a negative effect of excessive automation. The rise in inequality is attributed to a skills-bias due to disproportional worsening of wages for workers and an increase in capital share. Counterbalancing policies would have to adjust the "social safety net". Agrawal et al. (2019) conclude that antitrust issues, through the accumulation of market power by information and communication technology (ICT) firms and the important role of data, could play a future role in policymaking initiatives (cf. Goolsbee (2018) on pricing issues due to power accumulation on the seller side).

Though most working papers were not presented above, the consolidated summary by Agrawal et al. (2019) gives an interesting overview about the identified economic policy scenarios possibly induced by AI. The methodological basis of the research approach, however, is a meta-level synthesis of essays from leading economists without much empirical corroboration. A similar list of economic policy factors can be found in Furman and Seamans (2019). They discuss the development of a universal basic income as public policy reaction to automation caused by AI, with the speed of AI adoption as alternating variable.

Other authors have tried to construct policy frameworks either through taxonomizing or normative approaches. Calo (2017) defines key questions that an AI policy could tackle, but also states the problematic of defining inclusion criteria as to why a certain question and not the other should be presented and resolved in a taxonomic approach. Altogether, he describes a roadmap approach that policymakers could take when trying to develop some acquaintance with the field. He identifies five areas that require policy attention: Justice and Equity, Use of Force, Safety and Certification, Privacy and Power, Taxation and Displacement of Labor. In this topical grouping, he provides an interesting overview about AI policy-related issues and touches on a variety of fundamental problems presented above.

The regulatory system proposed by Scherer (2015) builds on the democratic concept of separation of power. He foresees the role of legislatures in the creation of a legal act phrased "The Artificial Intelligence Development Act", which introduces the principles of an AI regulation. Due to the lack of expertise and need for delegation in a complex technology like AI, the creation of an agency is suggested as a mean to ensure expertise and political independence in a regulatory oversight. Finally, he stresses the role of courts as ex-post ruling instance to solve liability cases arising out of the use of AI.

For advanced AI, Bostrom et al. (2018) provide several desiderata to be considered by policymakers. It could be best compared to a taxonomy of AI issues that might inform farfuture policymaking. A proceeding that he calls “vector field approach”. Equally hypothetical approaches are the frameworks by 6 (2001a; 2001b) and Stahl and Wright (2018).

3.4 Sectoral AI Policies

Very few papers consider policy-related issues of the technology “artificial intelligence” in a specific sector’s or industry’s context.

Wall (2018) discusses some financial regulatory implications. He concludes that the key problem for AI value creation in the financial environment is the availability of underlying data, to which only some firms have access and therefore a competitive advantage. Two possible solutions to circumvent this data accumulation would be the enforcement of data sharing and the pooling of data by several smaller banks. Other authors have also begun to publish perspectives on “RegTech” issues. Aziz and Dowling (2019) show practical implications of ML use with regard to regulatory issues. Van Liebergen (2017) concludes that the use of ML even within the financial sector, and resulting problematics, are context-dependent and require auditing reactions in a regulated environment.

For the manufacturing context, the literature is very limited. The focus is clearly on the regulation of highly automated vehicles (HAV) and surrounding regulation, as well as ethical considerations. Schuelke-Leech et al. (2019) use a “big data analytics technique” of public and private documents to establish a narrative between the discussions on policy-level and amongst technical experts on HAVs. They show that a time lag between the development and policy-reaction to the development exists; principal-agent problems and many of the fundamental problems, like responsibility and liability, apply. Crane et al. (2016) conducted a survey of regulatory issues regarding autonomous connected vehicles and summarized it in a framework. And a whole stream of philosophical and ethical considerations around the issue has emerged, as well (see Goodall, 2014; Himma, 2009; Hevelke and Nida-Rümelin, 2015). This context is somewhat different to the considerations presented in the present paper, since the AI is then not applied by the manufacturer but built into the manufacturer's product with entirely different implications. In the present case, however, the manufacturer can be seen as the consumer of the technology, with different problematics compared to a corporate consumer who doesn't underly consumer protection.

The application of AI in the manufacturing context is typically described as a factor that enables the connectivity of machines within a factory as part of the fourth industrial revolution, subsumed under the politicized term "industry 4.0" (Schwab, 2017). Kirchberger (2017) is one of the very few academic authors who frames the topic of AI under the broader advancements of industry 4.0 policies, so is Park (2018) who suggests that industry 4.0 and with-it AI requires new and innovative cluster policies. Hristov (2017) suggests supporting the AI industry as part of approaches to further stimulate growth in the Internet of Things. Tabandeh (1994) identifies AI's characteristics "complexity, localization, uncertainty, capital intensiveness and awareness" as the major hurdles for international technology transfer. Borgogno and Colangelo (2019) analyse the efforts of the *European Commission* in building a common European data market. They discuss the necessity of data to facilitate innovation in AI in particular and stress the importance of application programming interfaces (APIs) as enablers of data sharing. Further, they argue that a sector-specific regulation with standardized APIs as core concept might be better suited than horizontal approaches to ensure the success of the free-flow for non-personal data. Overall, they recommend regulatory intervention in data sharing problematics.

A first step towards a sectorial view of AI deployment implications has been undertaken by Bessen (2018). He developed an economic model to predict the demand-for-jobs reactions in three industries – textile, steel and automotive – that is likely caused by the deployment of AI technologies in these industries. The analysis shows that under an elastic demand environment, if AI would not entirely automate human tasks, an increase of jobs can be expected that outgrows job losses. This view is shared by Brynjolfsson et al. (2018) and Felten et al. (2018) who expect a different type of jobs to be affected by ML as compared to earlier automation . They stress the transformative character rather than replacement rational of AI. On the opposite, Frey and Osborne (2017) had found earlier that a high number of job roles is potentially at risk of automation (see also Autor, 2015). Although Heinen et al. (2017) welcome the efforts of these types of research studies, they also highlight some of the methodological weaknesses such studies have and argue for a less alarmistic view. They propose an ordo-liberal economic policy framework for AI that is open to different scenarios of the technology's further development and adaptable to it. Boyd and Holton (2018) concur from a sociological point of view and argue for a "normative openness".

4 Methodology

The policy proposals presented in the literature review laid the foundation for the empirical part. In this study, a relativist world view was assumed, as it relies on personal opinions from experts and stakeholders. The result is therefore arguably socially constructed, and the epistemology underlies an interpretivist assumption.

4.1 Research Approach

An inductive qualitative research approach follows Edmondson and McManus' (2007, p. 1160) recommendations for methodological fit in a nascent field. The research question was formulated to allow for an "open-ended inquiry about a phenomenon of interest", applicable to the new technology "artificial intelligence".

A first step was the collection of primary interview data for the qualitative development of a taxonomy. In line with Ketokivi and Mantere (2010, pp. 320-322), the strategic response to the reasoning dilemma arising out of this qualitative induction approach is the idealization of the findings, "used to model the phenomenon under study": policy remedies for AI deployment in the manufacturing value chain. Although this allows for an abstraction of findings, which is one of the strengths of idealization, the "intersubjective idealization of induction in a particular way 'mystify[s]' the entire process of theoretical explanation" (Ketokivi and Mantere, 2010, p. 322). In response to this inductive reasoning dilemma, a standardized survey counterbalances in a second step the problem. It allows for a valid interpretation of the open-ended data for meaning (Edmondson and McManus, 2007, p. 1160).

The research strategy follows the grounded theory school of thought introduced by Glaser and Strauss (1967) to transform the data into a new theory. Data collection and analysis alternate and iterate (Edmondson and McManus, 2007, p. 1163). This was achieved in a multi-stage data collection approach for research activity on four simultaneous levels (Figure 3).

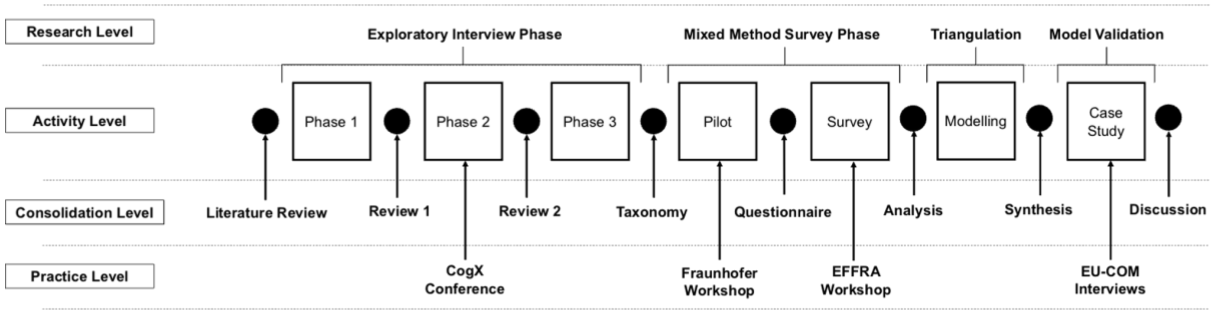


Figure 3 Multi-Stage Research Approach

4.2 Interview Data Collection

The exploratory interview process was conducted with international elite experts on AI from policy, industry, and academic professions, and designed to facilitate their creative thinking (Osborn, 1953; Isaksen et al., 2010). Interviews were individualized to the expected knowledge and expertise areas of the respective interview partner. The questionnaire was therefore mostly unstructured to capture the manifold facets and aspects of the phenomenon in question sought to be modelled. Thereby, different leads gained in the course of the discussions could be further explored with additional interview partners until theoretical saturation could be assumed. A question pool was prepared and adjusted in both review phases. Brinkmann and Kvale (2015) was followed for advice on interview best practice. The interviews were recorded and transcribed. After every phase, themes and responses were reviewed, and ideas were taken to the next phase. This theoretical sampling and permanent comparison adhered to the grounded theory strategy as suggested by Strauss and Corbin (1990).

Overall, 37 interviews were scheduled during three phases. The first group of eight interviews were selected to explore the scope of the topic in May 2019, and interviewees were of previous acquaintance with the author to allow for an honest assessment of the field and the research project. The second group comprised ten interviews, which were conducted at a conference venue in London in June 2019. Having the motto "The Festival of AI and Emerging Technology", the *CogX*™ conference assembles annually many of the most influential experts on AI in London, and face-to-face interviews were held to capture a visionary perspective on the research topic. Interviewees were selected and approached based on their appearance in the conference programme if they had speaking slots with relevance for the research topic. Finally, nineteen more interviews were conducted during the third phase predominantly on the phone, with some interviewees previously invited and followed-up during the conference or identified as experts on *LinkedIn™*. The sampling process therefore relied on purposive sampling (Easterby-Smith et al., 2012, pp. 228-229). Interviews lasted ~32 minutes on average.

It was aimed for a balanced mix of views, and a mix of different hierarchy levels of the interviewees was realized. Thirteen different nationalities were included, with a majority of respondents from Germany (13) and the United Kingdom (10). Twelve experts were involved in the development process of the earliest national AI strategies; three of them in leading roles in the USA, the United Kingdom, and Estonia, all countries with the most advanced and much-noticed strategies at the time. Interviewees included senior level government officials and a politician responsible for AI, two former U.S. Presidential Innovation Fellows, a former Special Assistant to the U.S. President for Economic Policy, three corporate CIO/CDOs, senior consultants, industry association representatives, as well as leading academics and data scientists. Figure 4 shows the distribution of stakeholder groups and interview partners.

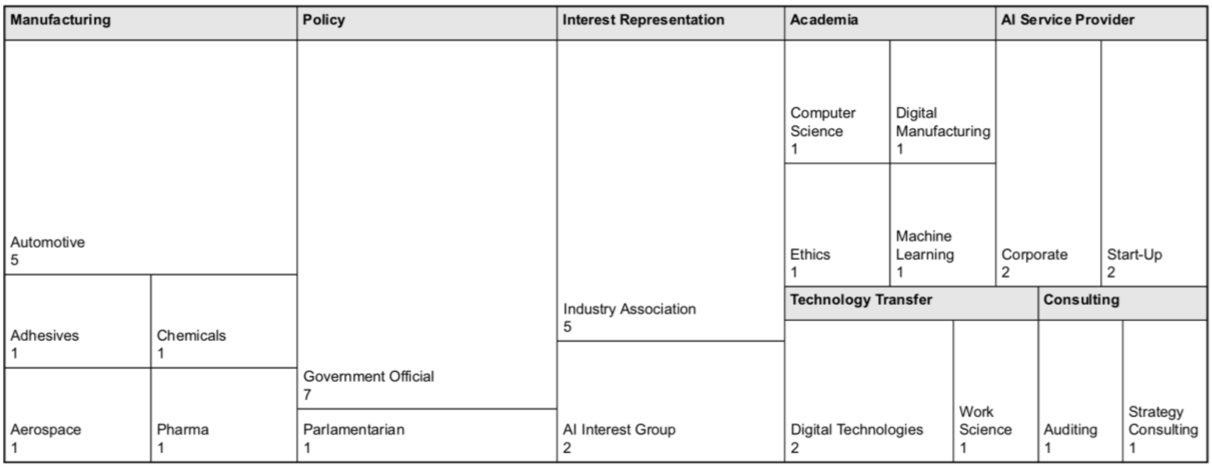


Figure 4 Overview about the Interviewees' Professional Background

4.3 Survey Data Collection

Outcomes of the interview phase allowed to answer the research question qualitatively and created the input for a taxonomy of policy remedies. The findings should then be validated and refined based on survey data. For this purpose, the most tangible policy propositions were selected and transformed into singular sentences as Likert-style items of a new questionnaire.

This questionnaire was designed according to principles presented by Easterby-Smith et al. (2012, p. 239) and formed by the bipolar Likert-scale items. Each respondent could express his agreement or disagreement to a policy remedy on an ordinal scale in a five-column ranking from positive (agree) through neutral (not sure) to negative (disagree). It was administered in two phases.

A first version was developed and tested in a small-scale pilot investigation with four participants of the workshop “Unternehmensdialog Künstliche Intelligenz” organized by *Fraunhofer IAO – Institute for Industrial Engineering* in June 2019 in Stuttgart. Participants were hand selected and representative of the variety of individuals for the main survey (Sapsford and Jupp, 2006, p. 103). *SurveyMonkey*™ was used as platform to program the questionnaire and emails were sent to administer the survey (follows Scherb et al. 2012). It underwent substantive remodelling after the test. The final questionnaire consisted of 15 matrix-questions with Likert-style item batteries that reflected the main categories of the taxonomy. It was split into three sub-questionnaires to reduce the scope for a respondent.

The pre-sampled respondents of the main survey were participants of the workshop “AI in Manufacturing” held in July 2019 in Brussels, which was organized jointly by the *European Commission*, the *Big Data Value Association (BDVA)* and the two European public private partnerships (PPP) *European Factories of the Future Research Association (EFFRA)* and *euRobotics*. Sampling followed the same logic of purposive sampling as did the interview process and the selection for the pilot study by comprising similar stakeholder groups. These were predominantly advanced manufacturing experts and data scientists from industry, academia, and policy professions. The sample was split in three sub-groups based on the surname’s alphabetical order, and the three different survey parts were sent out by email to the respective recipient group. This approach was necessary to reduce the response time of each individual respondent based on observations during the pilot study. The result is a good overall conversion rate of more than 40% (88 responses out of 212 invitations to participate). Overall, the demographic statistic of the survey implies a reasonable ratio of private sector (~ 48%), academia (~ 27%), public sector (~ 15%) and other professions, with respondents who were in majority very familiar with AI applications owed to the pre-selected character of the sample.

4.4 Data Analysis and Synthesis

For the analysis and taxonomy building as input for the questionnaire and system modelling, the qualitative research software *NVivo™* was used (see also Yearworth, 2010). This allowed for a bottom-up and inductive theme generation as nodes based on open coding of interview data. The survey results are evaluated and displayed in stacked bar charts. The free software environment *R* was used to program the results under adaptation of the package *HH* following Heiberger and Robbins (2014). The weighted average of every Likert-style item was calculated and forms the basis of the interpretation of results. Policy remedies with an average result of finding at least "rather agreement" by the survey participants were chosen as element for the policy model. Even though the survey questionnaire is per definition a quantitative instrument, its results are analysed and interpreted qualitatively.

Perspectives captured in this survey are therefore linked with the qualitative interview findings and "subjective views and interpretive patterns becoming visible in interviews are transformed into items of the subsequent standardized survey" (Flick et al., 2007, pp. 40-43). This provides a data and methodological triangulation based on the definition provided by Flick et al. (2007, p. 41), as the study develops "perspectives [that] can be substantiated by using several methods and/or in several theoretical approaches. Both are or should be linked." Thereby, the mixed-methods design follows the third proposition by Miles and Huberman (1994, p. 41), where qualitative research is a precondition for the quantitative method and could not be skipped (see also Barton and Lazarsfeld, 1955).

5 Findings

The manufacturing sector was claimed by most interviewees to have a high use case potential for the application of AI systems. It was, however, recognized that the challenges of installing such AI systems in the value chain are often underestimated despite advanced manufacturing techniques prevailing in many industrialized countries. Indeed, there was an overarching consensus that the manufacturing sector is a late adopter or even laggard of this technology's adoption. Targeted policy remedies were generally agreed to potentially constitute a decisive factor in accelerating or preventing the adoption of an emerging technology like narrow AI in the manufacturing sector.

Main result of the interview process is a taxonomizing approach describing policy remedies for narrow AI deployment in the manufacturing value chain. This allowed for a qualitative modelling of a comprehensive industrial policy. In addition to this taxonomy and model, the interviews also yielded insights into the potential use cases for AI systems in the different manufacturing value chain functions (cf. Porter, 1985) and the barriers to AI adoption in manufacturing. These findings are summarized in Appendices A and B.

5.1 Taxonomy of AI Policy Remedies

The taxonomy (Appendix C) was developed based on the qualitative interviews and lays the basis for this analysis. It consists of five policy dimensions: an employment dimension, innovation dimension, regulation dimension, data dimension, and governance dimension. Most policy remedies within each dimension address the barriers to the scalability of the technology's adoption or describe instruments to steer the deployment, and many interviewees stressed the importance of a human-centric application. A major difference to prior works on AI policy is the emphasis of the technology's deployment rather than development.

All findings are presented after the survey results were triangulated with the insights gained in the taxonomizing approach. The survey thereby served as "filter" to refine the interviewees' suggestions in the five policy dimensions. In the stacked bar charts below, those items are highlighted that show an average value of "rather agreement" to the item, which was then taken as basis for the final policy model.

5.1.1 Innovation Dimension

The innovation dimension describes a category of policy remedies that are dedicated to foster the development of novel use cases for the AI deployment and the necessary innovative environment that supports this objective. Figure 5 shows the results of the survey for the innovation dimension. Overall, a clear mandate was given to policymakers for accelerating the deployment of AI with policy remedies, a result that is not surprising in view of the participants' composition.

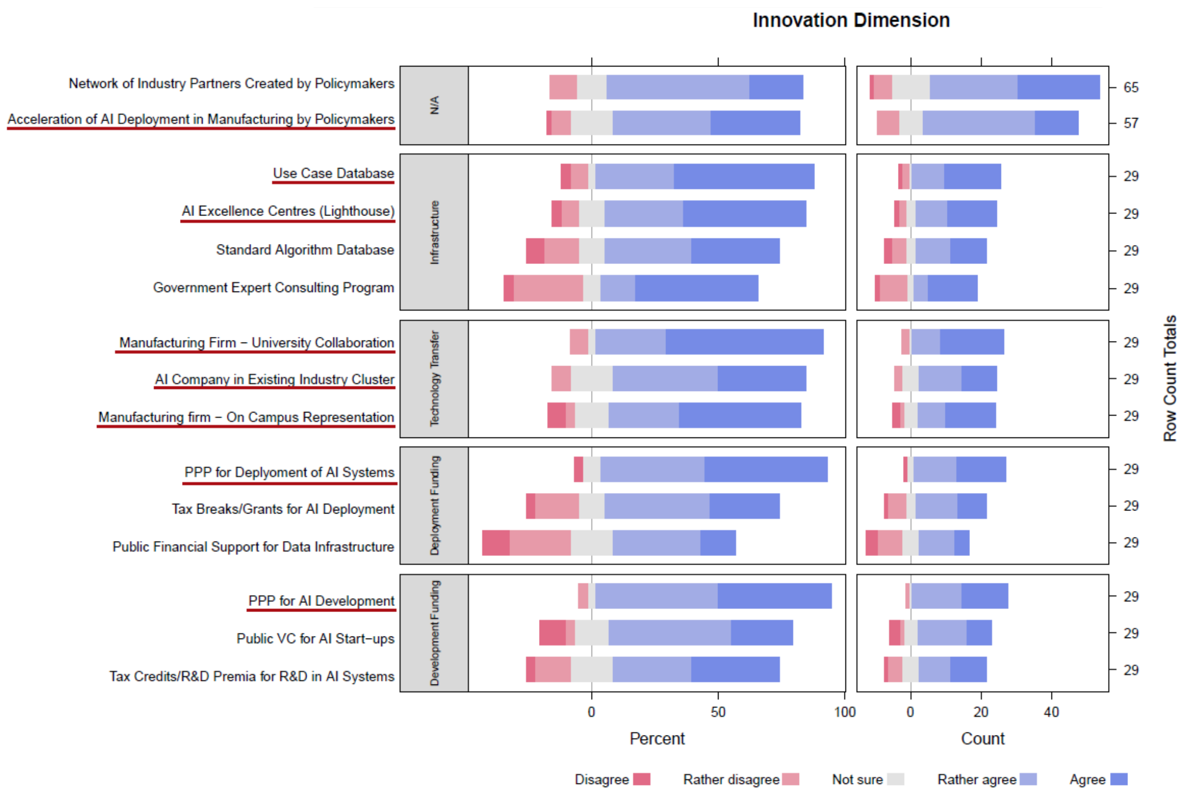


Figure 5 The Innovation Dimension

All policy remedies that can be attributed to a soft, non-interventionist policymaking approach are preferred by the respondents. The emphasis is on the creation of an ecosystem, in which innovation for the deployment of AI is enabled and fostered. To create this ecosystem, the importance of technology transfer and industry-university collaborations was found to be most significant, alongside the development of an infrastructure where industrial firms can familiarize themselves with best practice examples of industrial AI applications and applied research. This could be supported by a use case database. On the funding side, the joint undertaking of such innovation in the form of public-private-partnerships was highly supported by the respondents, whilst direct financial incentives for firms like tax breaks did not find the overall support. This reflects observations during the interviews, where the reactions on this policy remedy were equally divided.

Figure 6 further corroborates this result, as it confirms interview findings that the dominant rational for firms to introduce an AI system remains the value proposition of a business case and is not justified purely based on an experimental budget.

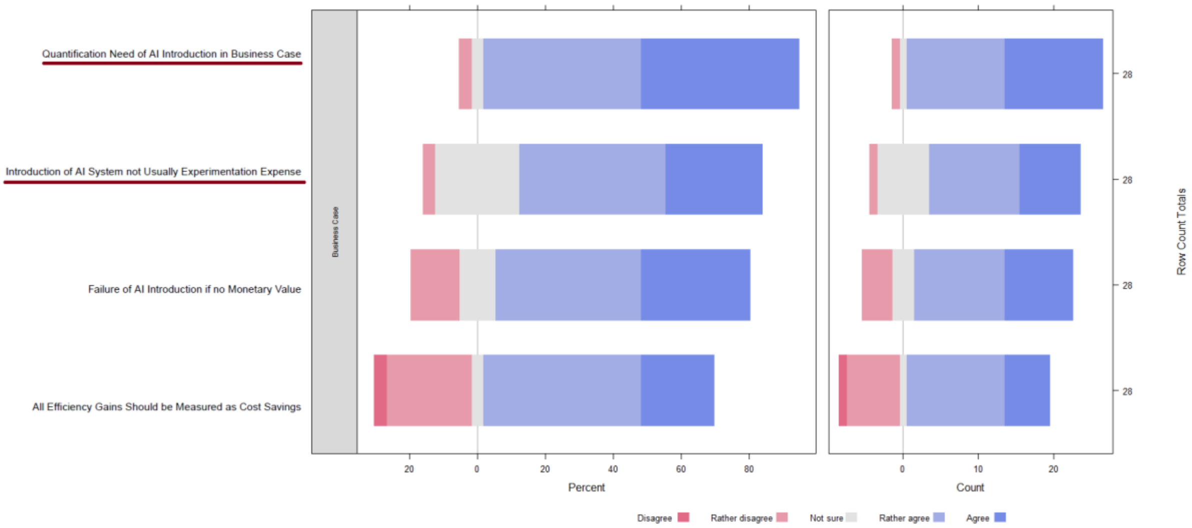


Figure 6 The Need for a Business Case

At the same time, this also confirms that value is not necessarily only financially provided but can be measured with other key performance indicators (KPIs). Firms have many potential value-driven use cases options to choose from, as one respondent described:

> "Industrial companies are just starting to really be in a position where they find these new data driven tools. It's almost like during the Garden of Eden and there's a plentiful bounty of things you could work on certainly. Usually we're just looking for that kind of juiciest fruit."

Appendix A shows that value could be generated with the application of AI systems in all functions of the value chain. The value propositions are manifold and could partially even be externalized, for instance with higher energy efficiency. It can also be inferred that some applications are likely to lead to a reduction in headcount, but many others not. Moreover, there will be a difference between sectors and industries in the measurability of value and practicability of use cases, as for example pharma firms can hardly validate non-deterministic systems for drug production with regulators.

The usefulness of an algorithm database and expert consulting program initiated by policymakers were doubted in the survey. The latter is somewhat in contradiction to the opinions during the interviews, where an educational mechanism for implementation of AI systems for firms was appreciated. Further research on the experiences with such programs could be interesting. The formation of an industry-wide network with partners bound together by the target to further accelerate the deployment of the technology receives overall consent; an ecosystem approach would support this approach, but whether this needs to be initiated by policymakers and not the manufacturing industry was doubted by respondents.

Finally, three strategic factors (Figure 7) were identified to which policymakers should tailor their innovation remedies for AI.

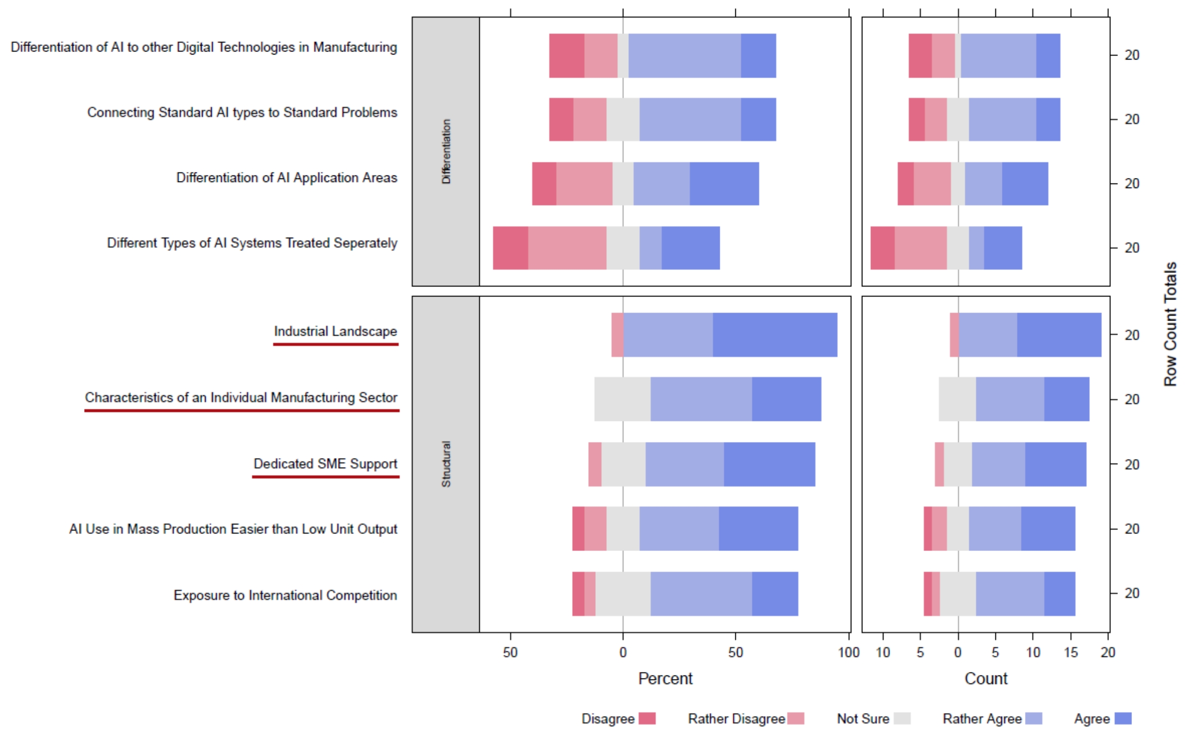


Figure 7 The Three Strategic Factors

The characteristics of the industrial landscape are an important prerequisite to judge the overall state of digitization for a certain region, which could vary considerably also within a country and requires different types of support. Moreover, the properties of the individual manufacturing sector were also found to make a difference during the interviews, which was confirmed in the survey. And finally, the need to provide dedicated support for SMEs was unanimously highlighted by the study's participants. Other possible structural factors like the type of production system or competition exposition of the firm were rather not confirmed, a result that is intuitive given the high granularity any innovation program would have to display. This might also be the reason why any further differentiation into type of AI algorithms or application areas was not supported by the respondents.

An interesting result is the overall scepticism to treat AI methods different to other digital technologies that are deployed in the manufacturing value chain. At least in the innovation dimension, this leads to the conclusion that there is a need for predominantly technology-agnostic programs and funding.

5.1.2 Employment Dimension

Skills and work-related considerations play an important role in the employment dimension. In this policy dimension (Figure 8), an interesting result has emerged. A majority of respondents rejected the idea that AI systems should only be deployed if a reduction in headcount is avoided. In case of an "efficiency gain", deployment was still considered a valid usage scenario, even if no employment gap exists – a likely scenario for some value chain functions according to Appendix A. However, it was emphasised that a human-centric approach of the deployment is necessary, and therefore respondents stressed the desire for social cohesion. This was perceived as a major risk and problematic of the technology, as one policymaker admitted anonymously: "It will be very tricky, and this is something that actually keeps me up at night".

AI systems could also serve as a means to improve work conditions, and factory workers might need to increasingly learn as teachers for such systems with increasing maturity of the technology. A grave intervention into market forces by policymakers was nevertheless rejected, which would be thinkable, for instance, by compensating of factory workers for their value-added work as AI teachers. Though it was emphasised that clear guidelines for the use of employee data in AI systems is important.

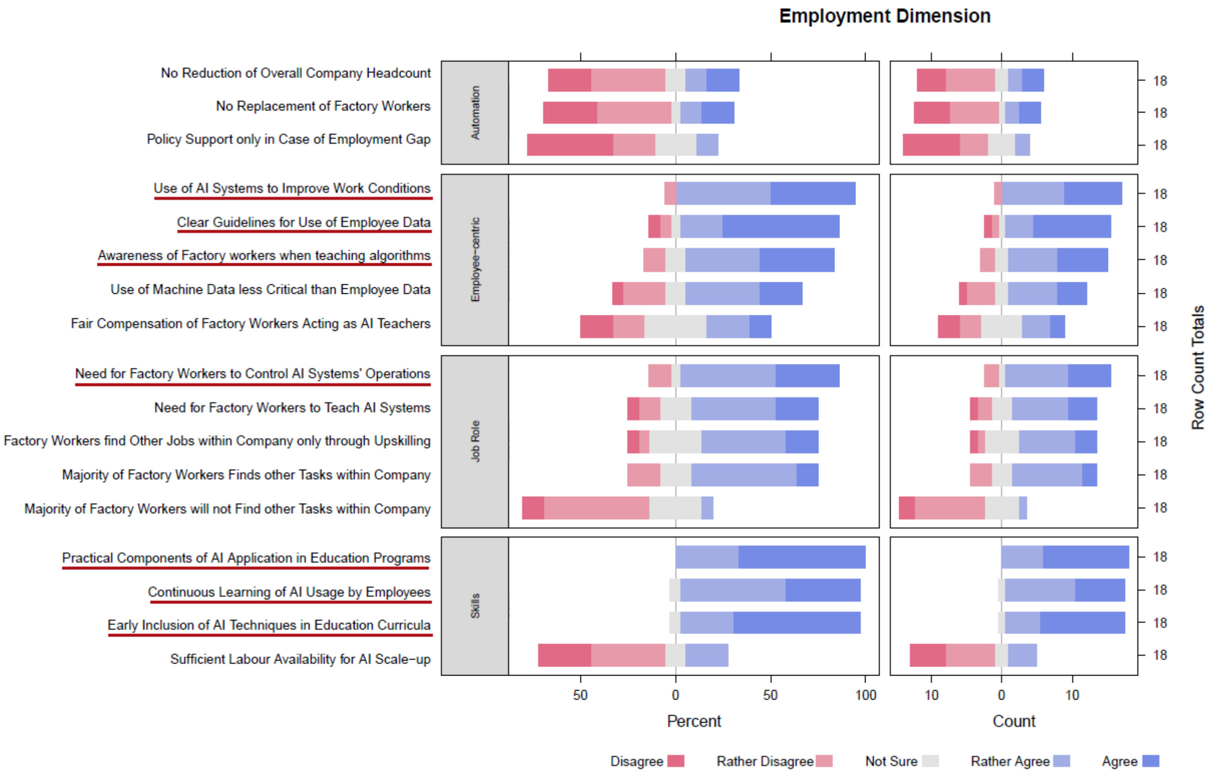


Figure 8 The Employment Dimension

Finally, the experts were optimistic that factory workers will be able to find other tasks if their original operation would be overtaken by an AI system. It was expected that novel job roles for factory workers might involve the usage of an AI system and the control of its performance and operations. The teaching of such systems was expected to be a role for engineers or inspectors.

Table 1 provides a further ranked indication as to what exact operational problems manufacturing firms might face in the course of an accelerated introduction of AI in the value chain. This can serve as an indication for the exact skill sets that would be expected from employees with increasing maturity of the systems. The operational problems comprise the data capturing and formatting, system control and change management, maintenance, performance measurement and teaching of the AI.

Table 1 Ranked Operational Problems

Rank	Operational problem
1	Gathering the correct data in the right format at the right time
2	Control for error in the AI predictions
3	Accommodate for changes in the system space
4	Training of factory workers on how to teach AI systems
5	Building-up a data gathering infrastructure
6	Analysing collected data and AI predictions
7	Cleaning the datasets
8	Maintenance of AI systems
9	Ensuring health and safety in robotics
10	Measuring the performance of AI systems
11	Training of employees on how to use AI systems

The overall strongest consensus amongst respondents in any policy dimension was found to be the need to create a sufficient skills base. In the case of AI, this would entail programming education early on in a curriculum, also from a practice-oriented perspective. Moreover, the importance of engaging in continuous learning activities as suggested by many studies was confirmed. Finally, a large majority of experts believed that there is not enough skilled labour available for an AI scale-up, which is a call for immediate action by policymakers should an increased usage of such systems be envisaged by society. As some interviewees mentioned, however, this will rather require dedicated vocational training, apprenticeship, and upskilling programs, and not the creation of a large amount of PhD positions. Programs like *Robotics First* in the USA or *appliedAI* in Germany were mentioned as an example, as for most future tasks the application and control of the algorithms will be important.

5.1.3 Regulation Dimension

Regulatory approaches are often discussed when aiming to steer the deployment of an emerging technology like AI, and the regulation dimension summarizes important policy remedies that try to counterbalance implications of the technology that are unfavourable for society. This dimension (Figure 9) confirms in nearly all respects the expectations that were derived from the interviews, as most experts had favoured light-touch regulation in early maturity stages.

In the survey, two different scales were created in order to test the required strength and rigidity of the regulatory approach, one with softer regulatory mechanisms and one with stronger mechanisms. Company-internal measures that provide transparency and accountability for decision-makers when trying to introduce AI systems in a manufacturing firm were clearly favoured by the respondents The approval rate for internal audits was very high, which suggests this measure would be a helpful initial regulation to create oversight about AI deployment: "Have you truly got a plan of the AI systems that you want to build? Have you started a data collection methodology which allows you to collect training data for their systems and testing data with those systems? Then let's look at the controls for that." Therefore, such an audit would entail elements that check and clean the gathering of data and datasets for errors and bias, control for model creep, and develop mitigation strategies for technological rather than organisational risk.

Figure 9 The Regulation Dimension

Ethical guidelines for AI were endorsed in addition to corporate governance guidelines. It was deemed to be a task for policymakers to examine existing proposals of company-internal guidelines and principles that were developed simultaneously and heterogeneously, which could then be harmonized and in collaboration with industry and standardization bodies.

However, the option to create more intrusive guidelines was rather rejected, for example as prescriptive project guidelines for AI systems' deployment. It was tentatively surprising that the creation of standards for the use of selected AI applications did not find overall support. One explanation would be the difficulty to set the optimal standards for a low maturity emerging technology, as an early standardization could potentially hinder further and better technology development of superior solutions. A surprising result was a rather critical view of regulatory sandboxing approaches by the survey's respondents, which had found strong support during the interviews; it is therefore advisable not to break-up regulation without due consideration, when it comes to the testing of new AI systems in real application scenarios.

And a clear mandate for the development of guidelines and principles was given to policymakers regardless of whether or not a human would be in the loop of an AI system's supervision. Many other regulatory issues need to be considered context dependent though, for instance subject to the value chain function that would apply the AI system. An AI used in customer analytics clearly poses other ethical problematics and implications compared to an algorithm for tool-path optimization in a machine.

More formalized regulation was deemed important in two regards: a clear guidance in health and safety issues caused by embodied AI, as well as a resolution on liability issues arising in the case of an AI malfunction. This was highlighted not only for the use of AI in a product, but also when the technology is applied in the value chain. Policymakers were thus called to at least review existing regulation in this regard and to further observe the implications of malfunctioning AI. Whether concrete action by administrative bodies should be taken was seen controversially nevertheless, as one respondent describes the dilemma:

> "At the end of the day, there are industrial accidents all the time. If anything, I would imagine that the usage of AI would reduce the number of accidents overall rather than increase them. Because if you programmed the software to make sure that safety is its priority, we will actually look for ways of making

> production safer rather than less safe. If anything, the correct usage of the technology should result in a safer workplace rather than a less safe."

Survey respondents were also undecided what role the judiciary should play in resolving liability cases based on precedent ruling, which might be due to the mix of respondents from common law and civil law countries.

Overall, the more intrusive regulatory approaches become, from risk classification systems for algorithms, to certification of the AI systems' usage in the value chain, to external auditing requirements, the more they were rejected by survey respondents. This triangulates with the findings from the interviewees, where it was perceived undesirably in the technology's early maturity stage to apply rigid regulatory approaches before more powerful AI application scenarios were also seen in manufacturing use cases. Therefore, it is advisable to flank the technology's maturity process with a light-touch regulatory approach, that lays a flexible the foundation for possible future regulatory action, as the development of AI systems could – like many other ICT technologies in the past – advance very rapidly. For this scenario, study participants expected policymakers to be prepared for providing more guidance with an increase of autonomy of AI systems.

5.1.4 Data Dimension

The data dimension (Figure 10) describes a bundle of policy remedies that aim to improve data availability and sharing as basis for the application of AI systems. It was a dimension that resulted in mixed views by respondents of the survey, even though interviewees repeatedly stressed the necessity of creating a better access to industrial and manufacturing data for AI model development and training.

Even though policymakers were endorsed to develop clear guidelines for the sharing of non-personal data, companies should not be legally forced to use non-disclosure agreements (NDAs) and private contractual clauses for safeguarding this process. However, NDAs were generally perceived as useful means to agree on data sharing principles for industrial process-type data.

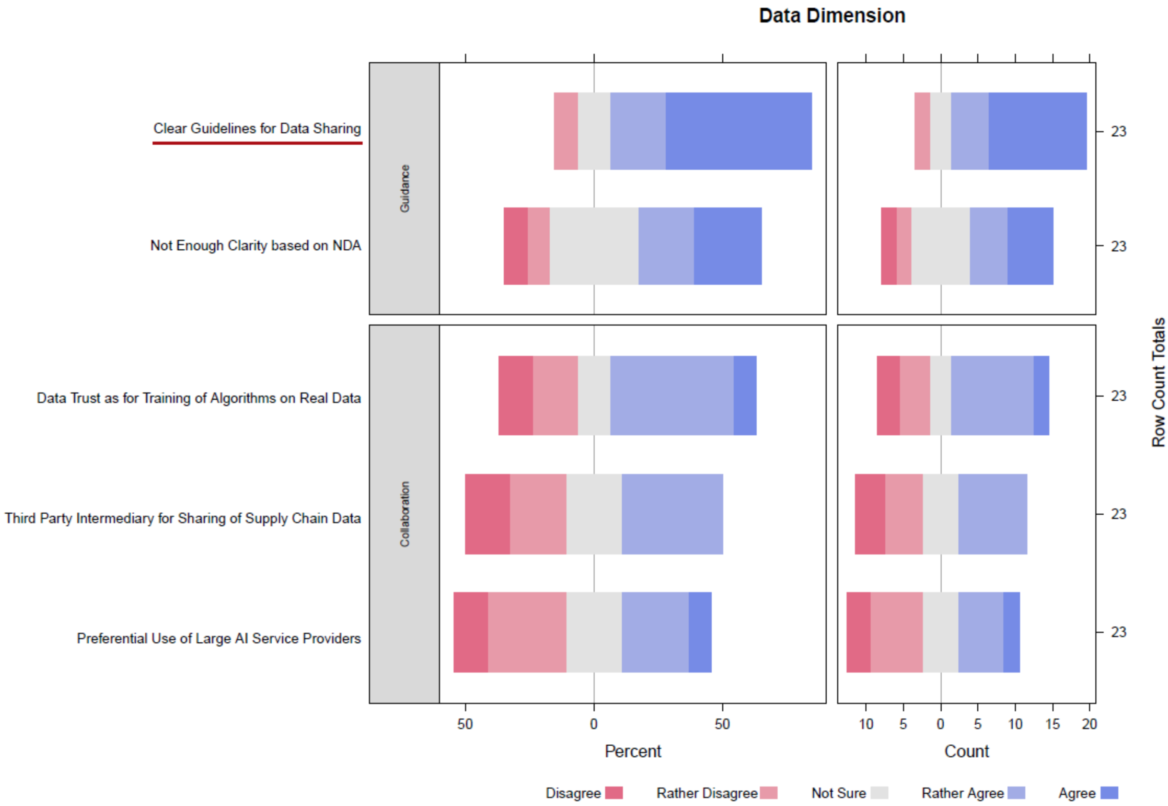


Figure 10 The Data Dimension

The approach of institutionalizing data sharing with the means of third-party intermediaries, for example enabled by a government institution like the data trust in the United Kingdom, led to a split decision picture. This was supported by some interviewees, but the technical problematics of sharing non-standardized datasets was acknowledged. The repeatability of an AI model for manufacturing tasks was deemed difficult due to non-identical information structures of data generated in different firms. Therefore, the benefits of an institution like a data trust were expected to be limited in an early maturity stage of the technology but could become a helpful policy remedy with further maturity.

Sharing of data would be further motivated by rewards for the providers of the datasets, for instance when better AI services are guaranteed afterwards. Small service providers, such as start-ups, were perceived to rather not be disadvantaged when offering their services to industrial firms compared to large corporate service providers, if they provide a clear value proposition. All in all, the data dimension was found to be very much dependent on further standardization efforts in the AI space.

5.1.5 Governance Dimension

Policymakers will have to find modes of collaboration that allow for the further coordination of policy remedies in the innovation, employment, regulation, and data dimensions. In a governance dimension (Figure 11), they are called for including representatives from industry into the further AI deployment process. Study participants favoured the creation of advisory bodies early on, which should allow industry to provide their expertise in the policymaking efforts, and in creating a network across different industries to engage in discussions about the deployment of AI in manufacturing. This should create synergies between industries.

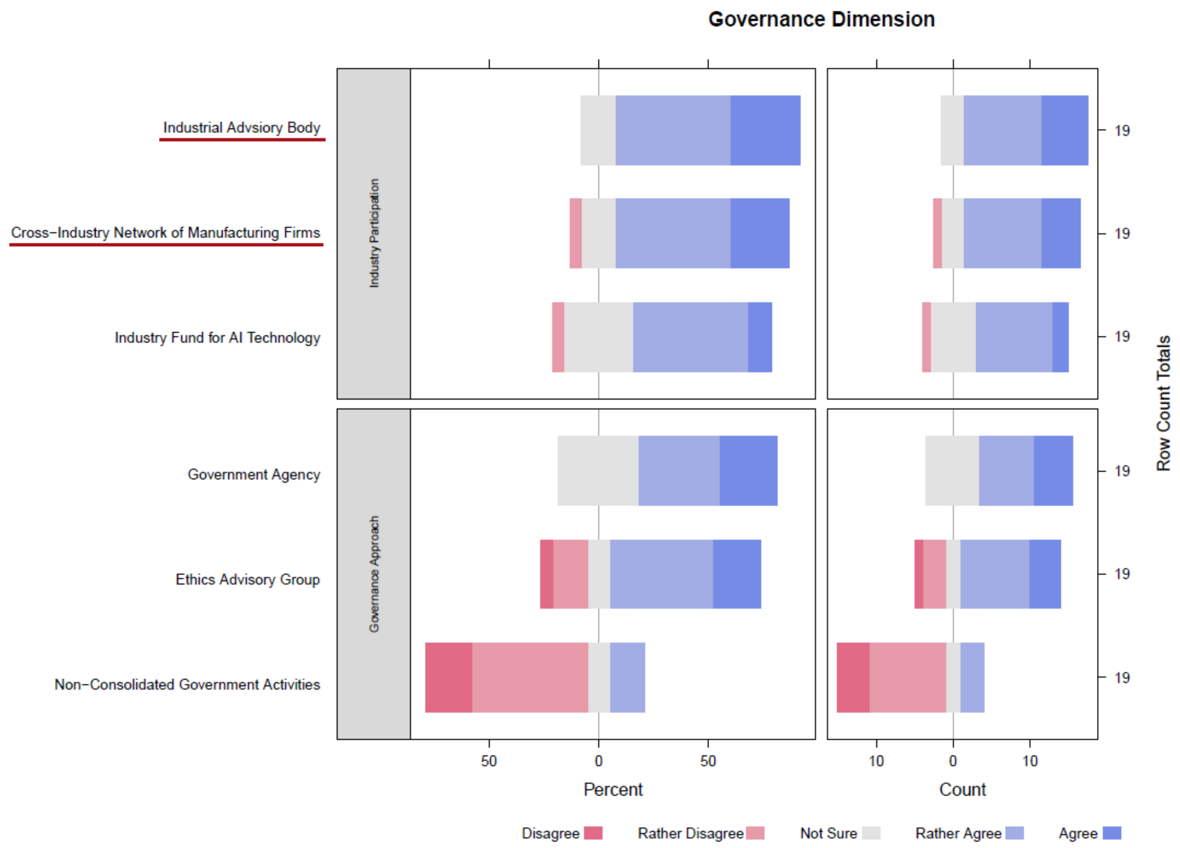


Figure 11 The Governance Dimension

The proposal of a dedicated industry fund for AI scale-up rather than pioneering development was only reluctantly supported by the respondents of the survey, although experts mentioned the lack of scale-up funding in the AI industry:

> "What I'm looking at is an industrial technology fund, that is led by a consortium of decision makers of the industry and will invest in AI type of technologies, applications, startups, entrepreneurs to work with collectively and then try out these solutions across the industry in order to be able to find what works at scale."

An optimal governance approach might involve some form of administrative structure and coordination, which could be steered by a specialized agency:

> "What we think that could be useful is to have an agency, a supporting agency which has AI expertise and supports other bodies in Germany, like the 'Gesundheitsamt' or 'BaFin'. To support those bodies to implement standards and processes to control algorithms."

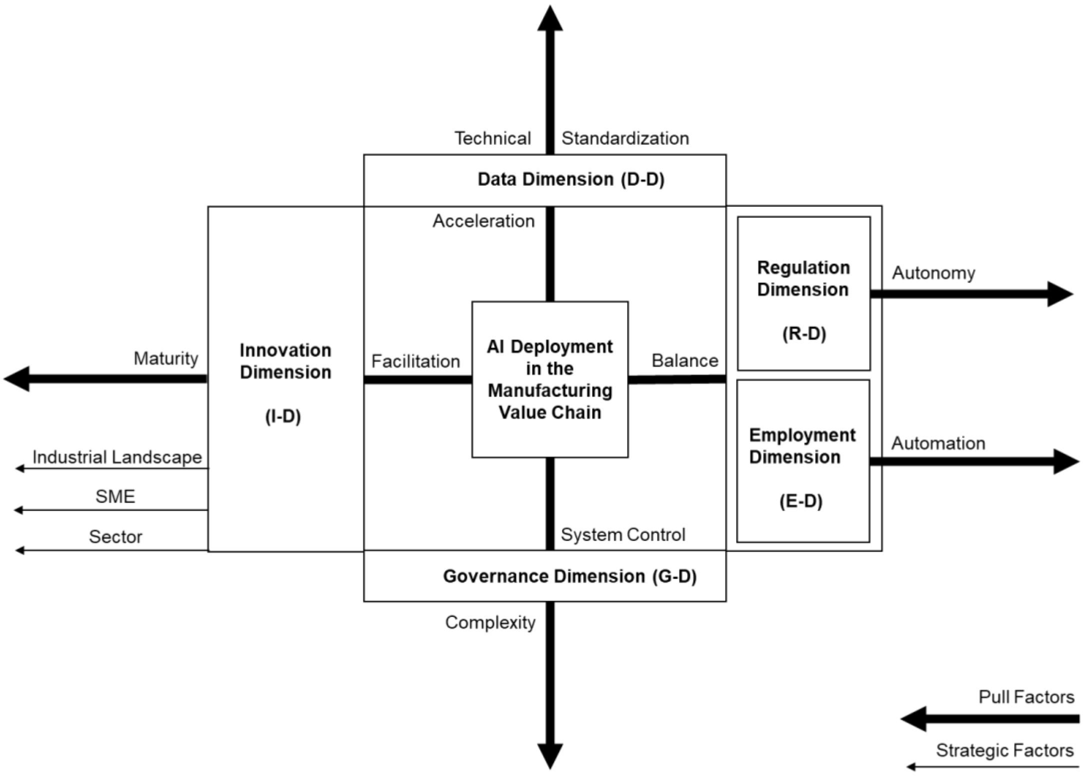
Although the respondents were rather undecided whether this should be a universally recommended governance approach, the necessity to work interdisciplinary on AI policymaking was reinforced. Siloed or non-consolidated government activities were rejected. Exemplary forms of governance could be the creation of a dedicated team between departments like in the United Kingdom's *Office for Artificial Intelligence*, or at least regular interdepartmental meetings on the topic. The creation of an ethics advisory group was clearly not recommended as single governance mechanisms that could deal with the implications and policymaking necessary for the deployment of AI systems in the manufacturing value chain.

5.2 Development of the "Four-Wing Industrial Policy System Model"

The previously developed taxonomy of policy remedies forms the qualitative basis for a model of the variable "industrial policy for the deployment of AI systems in the manufacturing value chain" as displayed in Figure 12. This is best described as a systemic model of a policy mix.

On the horizontal axis, three policy dimensions form the main wings of the model: the innovation dimension, and opposite the regulation and employment dimensions. Whilst the innovation dimension can be seen as the wing that facilitates AI deployment, the latter dimensions outbalance the system to an equilibrium. This reactive wing of checks and balances is needed to create a human-centric, responsible, and safe application of AI systems in the manufacturing value chain.

A vertical axis is formed by the data dimension, opposite the governance dimension. The governance dimension controls and stabilises the system. This becomes especially important if the data dimension's full potential is unleashed by technical standardization and increasing data sharing willingness of manufacturing firms. The data dimension functions then as accelerating wing of the entire policy system.



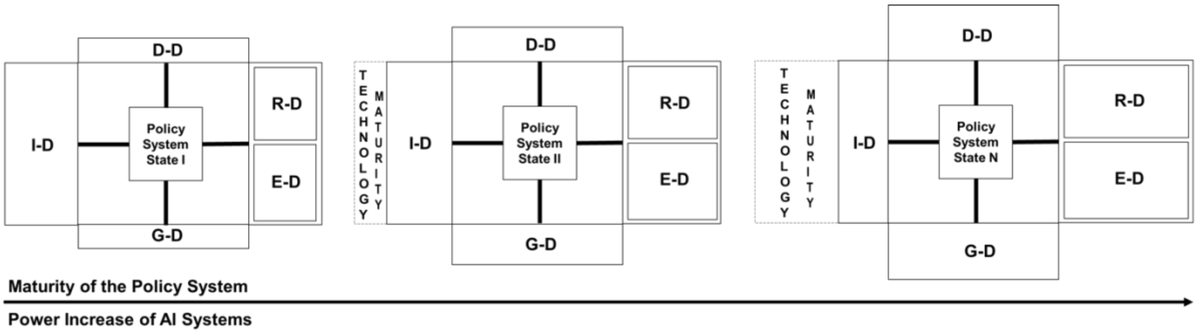


Figure 12 The "Four-Wing Industrial Policy System Model"

In the Innovation dimension, three factors provide policymakers a strategic focus: they reflect the need to adjust the policy system with additional support for SME companies, to tailor policy remedies to specific industries and to adjust policy remedies depending on the characteristics of an industrial landscape and its digitization state.

The pull factors lead to an overall increase of policymaking activity in the respective dimension, but the innovation dimension mainly expands with an increasing maturity of the technology itself. This in turn allows for a reduction of traditional demand- and supply-side innovation remedies. On the opposite wing, an increase in autonomy of AI systems increases the need for regulatory reaction, for example with compulsory risk classifications of algorithms or certifications. An increase in the usage of AI systems as automation tool will trigger reaction in the employment dimension with skills- and work-related measures like job guarantees. Setting of technical standards to facilitate data sharing and AI model training on pooled data accelerates the policy system and requires additional policy reaction, for instance based on successful institution building and data playgrounds. The governance dimension adjusts the system back into an equilibrium and expands with increasing complexity due to more sophisticated and targeted governance approaches.

As displayed in the second part of Figure 12, an increase in power of AI systems is then directly correlated with an increase in maturity of the policy system, caused by the pull factors. In consequence, the policy system model forms an expanding equilibrium over time. Adjustments of the policymaking activities in the five dimensions let the policy system mature dynamically in parallel to the power increase of the technology. This approach takes the spectrum of potentially different maturity states of AI systems in the manufacturing value chain into account and might require several simultaneous states of the policy system, as one respondent describes the situation:

> "The lower end of AI, if we could call it like this, I think has reached a certain maturity; especially around maintenance, diagnostics, risk analysis – we have different applications with that. I think this is reasonable mature. They are scaling reasonably well. […] When you move on the other end of the spectrum, I mean in industrial applications, when you really go to something that is cognitive, this is for me at the very pilot phase, at the very high end of the spectrum. Which is more like autonomous processes, like bringing intelligence to a SCADA system, having the ability to react almost in real-time to a pattern of events."

As the analysis has further shown, the five dimensions can be concretised for initial policymaking recommendations, as long as the technology is classified as "emerging" and in early maturity state. This is modelled in Figure 13.

The innovation policy thereby focuses on the creation of an ecosystem. A regulatory approach should be light-touch in the beginning, mainly comprising AI principles, guidance (such as for internal audits), health and safety guideline revisions, and a harmonization of guidelines. The employment dimension needs to emphasise skills creation, but also has to think ahead on issues of social policy, such as the impact of automation on society and privacy problematics due to the inclusion of employee data. Finally, this is controlled by a multidisciplinary and coordinated governance approach, as well as with policy support for standardization efforts and data sharing.

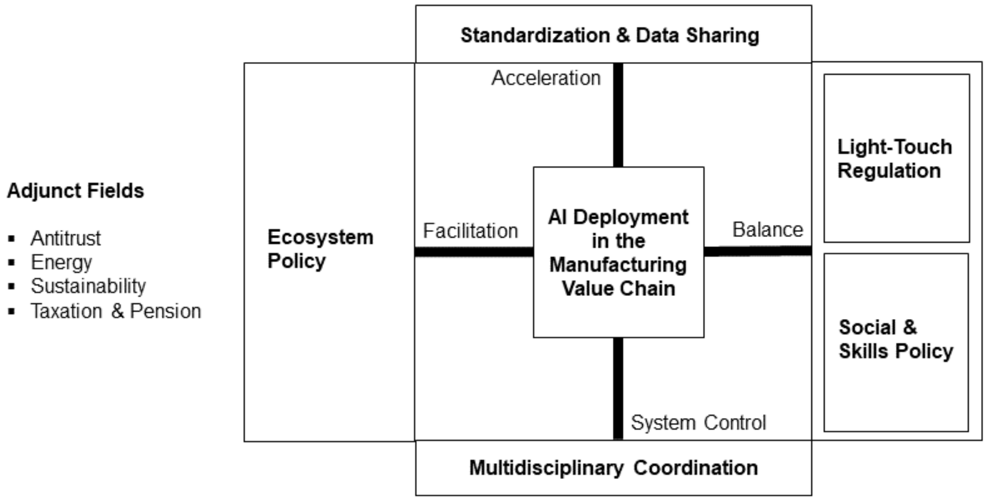


Figure 13 Concretized Policy Mix in Early Maturity State

Four adjunct fields of policy action were identified during the interviews, but it was avoided to triangulate them in the survey as they gain in relevance mainly in later maturity states of the technology. The application of more powerful AI systems might require comprehensive energy policy and sustainability strategies that can create the necessary power generation infrastructure to facilitate the use of these AI systems. An increase in automation could require in some countries additional social policy remedies like a universal basic income, which potentially leads to new fiscal or taxation measures. This could be financed with increased deficit spending or by introducing innovative taxation approaches like robot taxation. Value-add facilitated by AI systems would then not entirely be retained by the private sector should the overall number of employees be reduced. Finally, antitrust issues could occur due to an increase in market power of firms that can steer access to data or the maintenance of AI systems, which would require interventions based on competition law.

6 Empirical Validation of the Findings

To validate the findings of the empirical part, a fourth phase of unstructured interviews for a short explanatory case study was conducted with ten civil servants in four directorate-generals of the *European Commission* (*DG GROW*, *DG CONNECT*, *DG COMP*, *DG RTD*). These interviews allowed to summarize the *European Commission*'s policy system on AI in its early phase mid-2019 based on the taxonomy of policy remedies and the derived model. The case study describes the policy system in early maturity state prior to the proposal of the *Artificial Intelligence Act (AI Act)* in 2021, which was the European Union's (EU) first regulation draft on AI (European Commission, 2021). It also confirms that the policy system underwent the dynamic maturity process as proposed by the model. In the following, the EU's policy implementation for each of the taxonomy's dimensions is highlighted.

1) **Innovation dimension**. The Commission's *Forum on Important Projects of Common European Interest* defined three *Key Strategic Value Chains*: the Industrial Internet of Things, Clean Connected and Autonomous Vehicles, and Smart Health. This approach allowed member states to closely collaborate and to establish projects and funding vehicles for projects that were defined within these key strategic value chains. They were exempt from state aid rules and could also entail projects on AI. In such a case, projects could be considered as potential target for "good state aid" and were allowed to being promoted on an equal level by member states. General conditions for the existence of a market failure for the provision of state aid were given: it should be considered if creditors object to providing funding for AI (especially speculative R&D), or when industry – in the case of cleantech – would not be able to conduct projects due to a lack of ROI. Early research funding vehicles were also set out. Moreover, the EU's innovation focus was found to follow an ecosystem approach similar to the model's proposition, which should support the technology transfer potential. *Digital Innovation Hubs* provide digital expertise for industrial SMEs, also on AI, supported by *AI Excellence Research Centres*. This reflects

another strategic factor of the model. Under *Horizon 2020*, the PPPs *Big Data Value Association (BDVA)* and *euRobotics* on the technology push-side are complemented by *Factories of the Future* and *Spire* from the deployment side to provide best practice examples in the sphere of Industry 4.0 and AI in manufacturing. A new PPP dedicated on AI was planned under the new programming period.

2) **Employment dimension.** Job creation is one of the priorities of the Commission and social cohesion was addressed by Commission President Ursula von der Leyen with the proposal of a European minimum wage. Moreover, the industrial strategy sets out a strong social element. And the EU industrial days and *High-Level Skills for Industry* meetings thematized the skills needed for the future of work in 2030. The earliest approach of steering activities for AI in the EU were ethical principles developed by the *High-Level Expert Group on AI,* which should ensure a human-centric development and deployment of AI. This was also comprehended as a major differentiation to AI strategies in other, predominantly autocratic countries with geopolitical ambitions. Especially China neglected this aspect at the time and focused mainly on technological progress, and the EU was perceived as setting the precedent for such an ethical approach.

3) **Regulation dimension**. The *Communication on Artificial Intelligence* (European Commission, 2018a), the *Coordinated Plan on Artificial Intelligence* (European Commission, 2018b), and the *Communication on Building Trust in Human-Centric Artificial Intelligence* (European Commission, 2019) formed the foundation of the European Union's AI strategy at the time of this case study. With these documents, the Commission and the member states set the policy goals of a coordinated and ethical development and deployment of AI prior to a dedicated regulation draft with the *AI Act*. Each of the member states was asked to develop its own AI strategy, which was supposed to take country-specific properties into account. This reflects one of the strategic factors

found in this study. Furthermore, a mapping exercise was conducted by the *European Commission* that also identified the fit of the existing regulatory framework with the deployment of AI. As part of this exercise, the machinery directive that regulates the safety standard of equipment was reviewed and in a consultation process, the novel requirements of AI were identified for the directive's revision. The *High-Level Expert Group on AI* also developed an assessment list for the development and use of AI in companies. This could be seen as the first step towards a harmonized non-compulsory audit and risk analysis for the use of AI systems. This constitutes a light-touch regulatory approach as demanded by the model, before the *AI Act* was drafted to reflect the increase in power of AI systems and enhanced regulatory need. Moreover, *DG Comp* has begun to observe a potential antitrust risk caused by the servicing of AI models, for example in autonomous cars where car manufacturers could exclude external competition by AI service providers.

4) **Data dimension**. A commission target is a European data economy for the digital single market, which acknowledges the need to enable data sharing across country borders based on first principles for B2B data sharing activities. The initiative *AI4EU* created an AI on-demand-platform that should make AI algorithms widely available. And an objective of *DG Connect* is the creation of a digital industrial platform that brings the digital actors and industry together, to overcome the defragmentation of the industrial landscape and to enable interoperability between service providers. Aside from that, GDPR has already regulated the data sharing of customer and employee data, and the regulation on the free flow of non-personal data across EU member states' borders should facilitate data sharing of industrial and cloud data.

5) **Governance dimension**. An ethical approach on AI was especially manifested by the formation of the *High-Level Expert Group on AI* in the EU. With the *Ethics Guidelines for Trustworthy Artificial Intelligence* and the *Policy and Investment Recommendations for*

> *Trustworthy AI*, the group has provided a comprehensive set of policy recommendations on AI development and deployment; many of those are similar to observations made during this research. The interdisciplinary character was acknowledged by regular inter-service meetings between the DGs, where also manufacturing-related issues are discussed. *DG Research & Innovation* has even formed a unit coined Industry 5.0, that acknowledges the need to think the industrial development socially and human-centric. Moreover, the Industry 2030 Group, consisting of industry representatives, has published a report that also sets out policy recommendations for AI. *DG Grow* has looked at the deployment of AI industrial applications in the project *Critical Industrial Applications of Artificial Intelligence*. This project aimed at assessing the uptake of critical AI applications at the industry level.

This case study demonstrated the usefulness of the developed taxonomy and model to comprehend policymaking on AI as a system. It is now possible to understand the interconnections between the different policy fields. It can be inferred that the *European Commission* implicitly converges its policy instruments towards such a system and develops dedicated policy activities in all dimensions.

7 Discussion

It could be shown in this research that it is necessary to think the deployment of AI in five policy dimensions that facilitate, accelerate, outbalance and control the deployment of AI technologies in the manufacturing value chain. The following central thesis is derived:

> **An industrial policy for the deployment of narrow AI systems in the manufacturing value chain requires the creation of a "policy mix" -system of innovation enabling and counterbalancing dimensions that ensure a human-centric, safe and ethical deployment of the technology.**

This empirical demonstration has provided a model of a comprehensive AI industrial policy for the manufacturing sector. Like Agrawal et al. (2019), who propose the need for diffusion-enhancing and consequence-addressing policies based on the theoretical work by other scholars, the model proposes the need for innovation enabling and counterbalancing dimensions. Implicitly, it has thereby confirmed the public policy theory that innovation policy requires a "policy mix" of different policy elements. The term "policy mix" has originally been used in economic debates about the relative importance and design of fiscal and monetary policy (Mundell, 1962). It is nowadays more and more used in an innovation context (Flanagan et al., 2011, p. 702). Basis is the idea that policymakers might need to establish a mix of policy instruments, as well as objectives when supporting new technologies (Branscomb and Keller, 1999). The sole application of funding and R&D instruments alone is deemed unsuccessful:

> "A general recognition that innovation-driven economic success depends on more than traditionally conceived S&T policies - exemplified by the emergence of 'systemic' rationales and new typologies of innovation policies that emphasize the role of 'indirect' as well as traditional 'direct' measures, 'demand-side' as well as 'supply-side' instruments […] This implies that instruments intended to achieve other policy goals (such as procurement, regulation, education, tax measures, etc.) can or should be 'co-opted' to achieve the goals of innovation policy." (Flanagan et al., 2011, p. 703)

This clearly envisages the use of balancing elements such as regulation and skills improvement as potential elements of such a policy mix, which had been suggested by Gunningham and Sinclair (1999) in the environmental context before. Flanagan et al. (2011, p. 703) further attest:

> "We believe that the uptake of the term reflects the realisation that modern states are increasingly characterised by the dispersal of power, not merely upwards and downwards from the national level to supra- and sub-national actors, but also outwards to quasi-state and non-state actors."

This applies in the present case. The European Commission's "Communication on Artificial Intelligence" had insisted top-down that member states are to develop their own national AI strategies, and much of the policy activity has been transferred to the High-Level Expert Group of Artificial Intelligence, which is more or less a non-state advisory body. This however increases the complexity of policymaking:

> "As Hay (1999, p. 322, original emphasis) notes, the state as a 'complex and institutionally fragmented system (of systems) has no innate propensity to proactive and reflexive transformation as a system'. In order to engage in a process of reflexive self-transformation, the state must not only co-ordinate its multiple activities but also 'the process by which these are reconstituted and re-coordinated'." (Flanagan et al., 2012, p. 704)

They also suggest the use of policy dimensions to showcase that within and between the dimensions interactions would occur when conceptualizing the interactions in the policy mix (p. 710). The empirical findings of this paper are also in line with propositions made by other authors before, who suggested that the interactions forming the policy can be viewed like an overall "system", that preferably rests in an equilibrium (Aghion et al., 2009, p. 682-683; Edquist, 2013). It was also suggested that such a system is to be dynamic over time (Kay, 2006). Finally, evolutionary economics gives an indication that normative claims and normative policymaking can be the legitimizing part of any such system (Witt, 2003; Stilgoe et al., 2013), as it was found to be necessary in the present case through its ethical elements.

The research has provided new insights, as the ideas of a comprehensive AI policy were merged with a sectoral view and constituted an industrial policy for narrow AI. The emerging, uncertain and dynamic character of the technology makes regulation difficult and bears the risk of regulatory misfit. Through its transversal character, policymakers might easily lose focus in designing their instruments and forget to include contextual stakeholders (cf Freeman, 1984). The systemic character through subcomponents like the data and supervisory system make policymaking more complex. As second-tier optimization, it requires in many cases a pre-investment into infrastructure for unlocking value, which could incentivize an innovation policy inaction on AI.

A comparably novel approach to policymaking in the emerging technology sphere promises the idea of co-evolutionary networked models, which are derived from biological if not ecological determinants of inter-node behaviour and connectors (cf Moore, 1993; Sotarauta & Srinivas, 2006). In the long-run, this allows key stakeholders an ex-post validation of anticipated network evolvement, which might look considerably adapted from ex-ante expected modelling (cf Padmore et al., 1998; Padmore & Gibson, 1998). The derived theoretical implications from the presented model in this research extend prior desiderations by an ecosystem-policy view, in whose realm interconnected policy actors engage through flowing information without rigid reporting systems (cf Li & Garnsey, 2014; Bostrom et al., 2018).

The analysis of existing literature on artificial intelligence has further shown that papers, which provide policy recommendations or policy aspects of the technology, have emerged very recently. Many are essay-style ethical or legal publications and entirely of a non-empirical basis, although many recommendations seem valid in the light of the present study. Problematic is that some policy recommendations are thought for advanced AI, whilst others are for narrow AI, caused by the divide in “presentists” and “futurists” as explained by Baum (2018). Policymakers will hence need to tailor their instruments according to application case (cf Magro & Wilson, 2013).

This paper addresses the presentist-type of research and tries to bridge the divide between comprehensive and sectoral policy developments. It uses as context the manufacturing sector and approaches the development of a comprehensive industrial policy for the deployment of the technology in the manufacturing sector. A generalizability for different technologies with altering specifications might hypothetically be postulated by modularity, a theoretical assumption also displayed by platforms embedded in multi-stakeholder fora ecosystems (cf Gawer, 2014). These specifications could depend on a multitude and magnitude of stimuli within an AI system’s reach depending on the platform type (cf Tiwana, 2014).

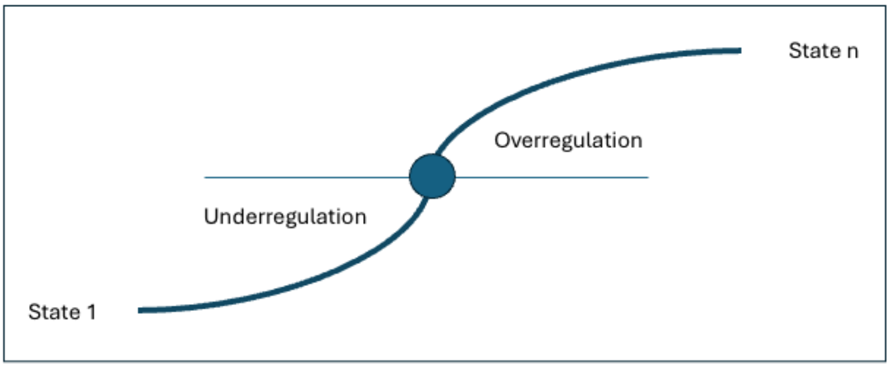


Figure 14 Gliding Maturity of the Policy System to AI Technology Maturity Differential

The reaction by policy stakeholders might cause a response that varies from its impact gravity. In an emerging technologies early stages, this could be an environment with spare regulatory exposure, whereas in late maturity stages, especially after S-curve jumps, the policy requirements might be well-advanced to innovation-hampering (cf Ghanbarnejad, 2014). Finding the balance may be difficult, but achievable by design – albeit potentially in constantly self-renewing steady-states.

8 Limitations

Throughout the paper, the idea of an industrial policy has sometimes been used interchangeably with aspects of innovation policy. Critics might argue however that an industrial policy and an innovation policy are not congruent, as the first will try to stimulate a certain industry or sector in growth with a strategic objective, the latter wants to stimulate further advancement in a technology with the means of policy instruments. In the present case, however, both policy types can be seen as target-congruent. Artificial Intelligence technologies have been available since the 1960s, and the recent increase in computing power led to a massive increase in available technology, especially through basic research. What is lacking for further scale-up and practical innovation is arguably the innovative step of applying the technology in practice. This needs an innovation policy for application, or just viewed from the opposite perspective, an industrial policy. Magro and Wilson (2013) treat industrial policy as policy domain of an innovation policy system anyways.

The most apparent limitation is the lack of a universally accepted definition of AI. This bears the risk that interviewees applied different ontologies that were blended by the researcher. The reliability of the findings is therefore somewhat uncertain, but it matches with many perspectives brought forward by consultancies or government documents. The validity of this research from a public policy perspective is considerably higher than previous non-empirical papers, as any normative claims provided in the development of the industrial policy system model are based on mixed-method primary data. A generalizability of the result could be subject to further research, as it might potentially be applicable to other sectors, as well. An application of the framework to administer other emerging technologies seems possible if they are identified as general-purpose technologies like AI.

In the methodological approach, there are some limitations, as well. The literature review is based on the limited academic literature only. It would make sense to include the large amount of grey literature and government studies in future research or to survey other fields of research, as well, like studies on innovation and digitization of industry. Although interviewees were selected based on expected competence, this happened only partially following specific criteria and sometimes also due to coincidences or by chance. Nevertheless, it was taken care that every single interview adds research value from a pre-defined perspective to cover a vast parameter space. Few of the “phase 3”-interviews were scheduled after the survey was already administered.

The survey’s overall response rate is good but divided into three sub-parts due to the extensive scope. Moreover, the participation was higher in the first group, which led to more valid results for this part of the taxonomy. The split into three groups also led to an inhomogeneous ratio of professions between the survey parts. This effect is weaker due to an equal split between taxonomy dimensions. Using the weighted average of an ordinal scale as decision criterion can be viewed critically from a statistical perspective, but it serves well as an indicator for a tendency of overall agreement or disagreement, which was the purpose of triangulation. The small N survey made the use of a correlation/reliability analysis and calculation of Cronbach’s alpha unrealistic.

9 Conclusion

In the introduction to this paper, it was argued that an industrial policy might be the means to unleash value promises of the technology narrow AI for the manufacturing sector and to address barriers to AI scale-up. This variable was modelled in the paper as a "policy mix" -system of innovation enabling and counterbalancing dimensions that ensure a human-centric, safe, and ethical deployment of the technology in the manufacturing value chain.

The value propositions are manifold and retrievable in every part of the manufacturing value chain. Some of them might include a reduction of human headcount, should this not be outbalanced by an increasing employment gap. The creation of a precautionary social and skills policy was therefore found to be of utmost importance. Regulation problematics arising out of AI usage are context-dependent and should be treated as such with initial light-touch approaches by policymakers in early maturity state of the technology.

Many of the barriers to AI scale-up, however, must be resolved by the manufacturers themselves. The industrial policy can be supportive based on the creation of a favourable ecosystem. This could be accompanied by occasional direct incentives for firms should the value of AI be partially externalized and hardly be calculable in a business case, for example in sustainability contexts. If technical standardization and data sharing guidelines can set further incentives for manufacturers to pool their data and allow algorithms to be trained on real-case datasets, this would further accelerate AI deployment in the manufacturing sector. Finally, a coordinated governance approach for the complexity of the technology is important.

Some characteristics of AI cause risks for policymaking. First, the emerging, uncertain, and dynamic character of the technology makes regulation difficult and bears the risk of regulatory misfit. Due to AI's transversal character, policymakers might easily lose focus in designing their instruments and forget to include context, such as the industry for an AI system's deployment. The systemic character comprising subcomponents like the model, data, and supervisory systems make policymaking more complex. Finally, AI is a second-tier optimization, as it requires in many cases a pre-investment into infrastructure for unleashing value; this could cause an innovation policy inaction on AI.

The empirical validation of the four-wing model has highlighted its usefulness as method to analyse executive policies for artificial intelligence. It can also be applied as a framework for comparisons of national policy systems for their AI readiness, that might allow for the creation of a typology of such national systems. The underlying taxonomizing logic of the model could be further developed by adding a scoring system for the provided dimensions, which may then be used as a government scorecard. Flanagan et al. (2011, p. 704) writes: “It is hard enough to see how any policy actor operating within a system of policy systems can at the same time step outside the system and take a rational and objective overview.” The four-wing model can serve as a tool, which allows the policymaker to perform this task with transparency and to administer the future AI deployment in the manufacturing value chain. Many of its principles might be applicable to other emerging technologies and other sectors with a focus on their deployment rather than development.

References

References with an asterisk (*) were included in the systematic literature review.

Accenture (2019). Manufacturing the future: Artificial Intelligence will fuel the next wave of growth for industrial equipment company. [pdf], Available at: https://www.accenture.com/_acnmedia/pdf-74/accenture-pov-manufacturing-digital-final.pdf#zoom=50 [Accessed 28-07-2019].

*Acemoglu, D. and Restrepo, P. (2018). Artificial intelligence, automation, and work. In: A. Agrawal, J. Gans and A. Goldfarb (eds.), *The Economics of Artificial intelligence: An Agenda*. National Bureau of Economic Research, pp. 197-236.

*Aghion, P., David, P. and Foray, D. (2009). Science, technology and innovation for economic growth: Linking policy research and practice in 'STIG Systems'. *Research Policy*, 38(4), pp. 681-693.

*Agrawal, A., Gans, J., and Goldfarb, A. (2019). Economic policy for artificial intelligence. *Innovation Policy and the Economy*, 19(1), pp. 139-159.

*Agre, P. and Rotenberg, M. (1998). *Technology and Privacy: The New Landscape*. Cambridge, MA: MIT Press.

*Alcácer, V. and Cruz-Machado, V. (2019). Scanning the industry 4.0: A literature review on technologies for manufacturing systems. *Engineering Science and Technology*, 22(3), pp. 899-919.

*Ananny, M. and Crawford, K. (2018). Seeing without knowing: Limitations of the transparency ideal and its application to algorithmic accountability. *New Media & Society*, 20(3), pp. 973-989.

*Asaro, P. (2016). The liability problem for autonomous artificial agents. *2016 AAAI Spring Symposium Series*, pp. 190-194.

Atkinson, R. (2019). Don't fear AI. [pdf], Available at: https://www.eib.org/attachments/big _ideasdontfearaien.pdf [Accessed 28-07-2019].

*Austin, L. (2003). Privacy and the question of technology. *Law and Philosophy*, 22(2), pp. 119-166.

*Autor, D. (2015). Why are there still so many jobs? The history and future of workplace automation. *Journal of Economic Perspectives*, 29(3), pp. 3-30.

*Aziz, S. and Dowling, M. (2019). Machine learning and AI for risk management. In: T. Lynn, J. Mooney, P. Rosati and M. Cummins (eds.), *Disrupting Finance*. Cham: Springer, pp. 33-50.

*Barocas, S. and Selbst, A. (2016). Big data's disparate impact. *California Law Review*, 104(3), pp. 671-732.

Barton, A. and Lazarsfeld, P. (1955). Some functions of qualitative analysis in social research. *Frankfurter Beiträge zur Soziologie*, pp. 321-361.

*Baum, S. (2017). On the promotion of safe and socially beneficial artificial intelligence. *AI & Society*, 32(4), pp. 543-551.

Baum, S. (2018). Reconciliation between factions focused on near-term and long-term artificial intelligence. *AI & Society*, 33(4), pp. 565-572.

*Benrimoh D., Israel S., Perlman K., Fratila R., Krause M. (2018). Meticulous transparency. An evaluation process for an agile AI regulatory scheme. In: M. Mouhoub, S. Sadaoui, O. Ait Mohamed and M. Ali (eds.), *Recent Trends and Future Technology in Applied Intelligence*. Cham: Springer, pp. 869-880.

*Bertolini, A. (2013). Robots as products: the case for a realistic analysis of robotic applications and liability rules. *Law, Innovation and Technology*, 5(2), pp. 214-247.

*Bessen, J., (2018). AI and jobs: The role of demand. In: A. Agrawal, J. Gans and A. Goldfarb (eds.), *The Economics of Artificial Intelligence: An Agenda*. National Bureau of Economic Research, pp. 291-307.

*Bodansky, D. (1991). Law: Scientific uncertainty and the precautionary principle. *Environment: Science and Policy for Sustainable Development*, 33(7), pp. 4-44.

*Boesl, D.B.O. and Bode, B.A.M. (2016). Technology governance. *IEEE International Conference on Emerging Technologies and Innovative Business Practices for the Transformation of Societies (EmergiTech)*, pp. 421-425.

*Boesl, D.B.O. and Bode, B.A.M. (2017). Roboethics and robotic governance. A literature review and research agenda. *Iberian Robotics Conference*, pp. 140-146.

Boland, A., Cherry, M. and Dickson, R. (2013). *Doing a Systematic Review: A Student's Guide*. London: Sage Publications.

*Borgogno, O. and Colangelo, G. (2019). Data sharing and interoperability: Fostering innovation and competition through APIs. *Computer Law and Security Review*, Forthcoming.

*Bostrom, N., Dafoe, A. and Flynn, C. (2018). Public policy and superintelligent AI: A vector field approach. Governance of AI program, Future of Humanity Institute, University of Oxford. Forthcoming in: S. Liao (ed.): *Ethics of Artificial Intelligence*. Oxford: Oxford University Press.

*Boyd, R. and Holton, R. (2018). Technology, innovation, employment and power: Does robotics and artificial intelligence really mean social transformation? *Journal of Sociology*, 54(3), pp. 331-345.

Branscomb, L. and Keller, J. (1999). *Investing in Innovation: Creating a Research and Innovation Policy that Works*. Cambridge, MA: MIT Press.

*Bresnahan, T. and Trajtenberg, M. (1995). General purpose technologies 'Engines of growth'? *Journal of Econometrics*, 65(1), pp. 83-108.

Brinkmann, S. and Kvale, S. (2015). *InterViews: Learning the Craft of Qualitative Research Interviewing*, 3rd ed. Los Angeles: Sage.

*Brundage, M., Avin, S., Clark, J., Toner, H., Eckersley, P., Garfinkel, B., Dafoe, A., Scharre, P., Zeitzoff, T. and Filar, B. (2018). The malicious use of artificial intelligence: Forecasting, prevention, and mitigation. [pdf], Available at: https://arxiv.org/ftp/arxiv /papers/1802/1802.07228.pdf [Accessed 20-07-2019].

*Brynjolfsson, E., Mitchell, T. and Rock, D. (2018). What can machines learn, and what does it mean for occupations and the economy? *AEA Papers and Proceedings*, 108, pp. 43-47.

*Burrell, J. (2016). How the machine 'thinks': Understanding opacity in machine learning algorithms. *Big Data & Society*, 3(1), pp. 1-12.

*Calo, R. (2017). Artificial intelligence policy: A primer and roadmap. *U.C. Davis Law Review*, 51(2), pp. 399-436.

*Cath, C., Wachter, S., Mittelstadt, B., Taddeo, M. and Floridi, L. (2018). Artificial intelligence and the 'good society': The US, EU, and UK approach. *Science and Engineering Ethics*, 24(2), pp. 505-528.

*Čerka, P., Grigiene, J. and Sirbikyte, G. (2015). Liability for damages caused by artificial intelligence. *Computer Law and Security Review*, 31(3), pp. 376-389.

Cimoli, M., Dosi, G. and Stiglitz, J. (2009). *The Political Economy of Capabilities Accumulation: The Past and Future of Policies for Industrial Development*. LEM Working Paper Series, 2008/15.

*Clarke, S. (2005). Future technologies, dystopic futures and the precautionary principle. *Ethics and Information Technology*, 7(3), pp. 121-126.

*Crane, D., Logue, K. and Pilz, B. (2016). A survey of legal issues arising from the deployment of autonomous and connected vehicles. *Michigan Telecommunications and Technology Law Review*, 23(2), pp. 191-320.

Deloitte (2018). Distinctive traits of digital frontrunners in manufacturing. Embracing the Fourth Industrial Revolution. [pdf], Available at: https://www2.deloitte.com/insights/us/en/focus/industry-4-0/digital-leaders-in-manufacturing-fourth-industrial-revolution.html [Accessed 28-07-2019].

Deloitte (2019). Future in the balance? How countries are pursuing an AI advantage. [pdf], Available at: https://www2.deloitte.com/cn/en/pages/technology-media-and-telecommunications/articles/how-countries-are-pursuing-an-ai-advantage.html [Accessed 28-07-2019].

Easterby-Smith, M., Thorpe, R. and Jackson, P. (2012). *Management Research*. 4th ed., London: Sage.

Edmondson, A. and McManus, S. (2007). Methodological fit in management field research. *Academy of Management Review*, 32(4), pp. 1246-1264.

Edquist, C. (2013). *Systems of Innovation: Technologies, Institutions and Organizations*. London: Routledge.

European Commission (2018a). Communication on Artificial Intelligence. COM(2018) 237 final. [pdf], Available at: https://ec.europa.eu/digital-single-market/en/news/communication-artificial-intelligence-europe [Accessed 30-07-2019].

European Commission (2018b). Coordinated Plan on Artificial Intelligence. COM(2018) 795 final. [pdf], Available at: https://ec.europa.eu/digital-single-market/en/news/coordinated-plan-artificial-intelligence [Accessed 30-07-2019].

European Commission (2019). Building Trust in Human-Centric Artificial Intelligence. COM(2019) 168 final. Available at: https://ec.europa.eu/digital-single-market/en/news/communication-building-trust-human-centric-artificial-intelligence [Accessed 30-07-2019].

European Commission (2021). 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. COM(2021) 206 final. Available at: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52021PC0206&from=EN (Accessed 07-10-2022).

*Felten, E., Raj, M. and Seamans, R. (2018). A method to link advances in artificial intelligence to occupational abilities. *AEA Papers and Proceedings 2018*, 108, pp. 1-4.

Flanagan, K., Uyarra, E. and Laranja, M. (2011). Reconceptualising the 'policy mix'for innovation. *Research Policy*, 40(5), pp. 702-713.

Flick, U., Rapley, T., Gibbs, G., Banks, M., Barbour, R., Angrosino, M. and Kvale, S. (2007). *The Sage Qualitative Research Kit*. London: Sage.

*Floridi, L. (2016). Mature information societies. A matter of expectations. *Philosophy & Technology*, 29(1), pp. 1-4.

*Frey, C. and Osborne, M. (2017). The future of employment: How susceptible are jobs to computerisation? *Technological Forecasting and Social Change*, 114, pp. 254-280.

*Friedewald, M., Vildjiounaite, E., Punie, Y. and Wright, D. (2007). Privacy, identity and security in ambient intelligence: A scenario analysis. *Telematics and Informatics*, 24(1), pp. 15-29.

FTI (2018). Artificial intelligence: the race is on. The global policy response to AI. [pdf], Available at: https://euagenda.eu/upload/publications/untitled-128126-ea.pdf [Accessed 28-07-2019].

*Furman, J. and Seamans, R. (2019). AI and the economy. *Innovation Policy and the Economy*, 19, pp. 161-191.

Freeman, R. E. (1984). *Strategic Management: A Stakeholder Approach*. Cambridge: Cambridge University Press.

*Gandy, O. (2010). Engaging rational discrimination: exploring reasons for placing regulatory constraints on decision support systems. *Ethics and Information Technology*, 12(1), pp. 29-42.

Gawer, A. (2014). Bridging differing perspectives on technological platforms: Toward an integrative framework. *Research Policy*, 43, pp. 1239-1249.

Glaser, B. and Strauss, A. (1967). *The Discovery of Grounded Theory: Strategies for Qualitative Research*. Hawthorne: Aldine de Gruyter.

Ghanbarnejad, F., Gerlach, M., Miotto, J. M., & Altmann, E. G. (2014). Extracting information from S-curves of language change. *Journal of The Royal Society Interface*, *11*(101), 20141044.

*Goldfarb, A. and Tucker, C. (2012). Privacy and innovation. *Innovation Policy and the Economy*, 12(1), pp. 65-90.

*Goodall, N. (2014). Ethical decision making during automated vehicle crashes. *Transportation Research Record*, 2424(1), pp. 58-65.

*Goodman, B. and Flaxman, S. (2017). European union regulations on algorithmic decision making and a "right to explanation". *AI Magazine*, 38(3), pp. 50-57.

*Goolsbee, A. (2018). Public policy in an AI economy. In: A. Agrawal, J. Gans and A.Goldfarb (eds.), *The Economics of Artificial Intelligence: An Agenda*. National Bureau of Economic Research, pp. 309-316.

Graham, O. (1994). *Losing Time: The Industrial Policy Debate*. Cambridge, MA: Harvard University Press.

*Guihot, M., Matthew, A. and Suzor, N. (2017). Nudging robots: Innovative solutions to regulate artificial intelligence. *Vanderbilt Journal of Entertainment & Technology Law*, 20, pp. 385-456.

*Gunningham, N. and Sinclair, D. (1999). Regulatory pluralism: Designing policy mixes for environmental protection. *Law & Policy*, 21(1), pp. 49-76.

*Gurkaynak, G., Yilmaz, I. and Haksever, G. (2016). Stifling artificial intelligence: Human perils. *Computer Law and Security Review*, 32(5), pp. 749-758.

Hay, C. (1999). Crisis and the structural transformation of the state: interrogating the process of change. *The British Journal of Politics & International Relations*, 1(3), pp. 317-344.

Heiberger, R. and Robbins, N. (2014). Design of diverging stacked bar charts for Likert scales and other applications. *Journal of Statistical Software*, 57(5), pp. 1-32.

*Heinen, N., Heuer, A. and Schautschick, P. (2017). Artificial intelligence and human labour: Implications for companies and economic policy. *Wirtschaftsdienst*, 97(10), pp. 714-720.

Heinonen, K., Helkkula, A., Holmlund-Rytkönen, M., Mustak, M., Jaakkola, E. and Halinen, A. (2013). Customer participation and value creation: A systematic review and research implications. *Managing Service Quality*, 23(4), pp. 341-359.

*Hevelke, A. and Nida-Rümelin, J. (2015). Responsibility for crashes of autonomous vehicles: An ethical analysis. *Science and Engineering Ethics*, 21(3), pp. 619-630.

*Himma, K. (2009). Artificial agency, consciousness, and the criteria for moral agency: What properties must an artificial agent have to be a moral agent? *Ethics and Information Technology*, 11(1), pp. 19-29.

*Hristov, K. (2017). Internet plus policy: A study on how China can achieve economic growth through the internet of things. *Journal of Science and Technology Policy Management*, 8(3), pp. 375-386.

*Iphofen, R. and Kritikos, M. (2019). Regulating artificial intelligence and robotics: ethics by design in a digital society. *Contemporary Social Science*, pp. 1-15.

Isaksen, S., Dorval, K. and Treffinger, D. (2010). *Creative Approaches to Problem Solving: A Framework for Innovation and Change*. London: Sage.

Kay, A. (2006). *The Dynamics of Public Policy: Theory and Evidence*. Northampton: Edward Elgar Publishing.

Ketokivi, M. and Mantere, S. (2010). Two strategies for inductive reasoning in organizational research. *Academy of Management Review*, 35(2), pp. 315-333.

*Kirchberger, T. (2017). European union policy-making on robotics and artificial intelligence: Selected issues. *Croatian Yearbook of European Law and Policy*, 13, pp. 191-214.

*Kroll, J., Barocas, S., Felten, E., Reidenberg, J., Robinson, D. and Yu, H. (2016). Accountable algorithms. *University of Pennsylvania Law Review*, 165, pp. 633-706.

Li, J. F. & Garnsey, E. (2014). Policy-driven ecosystems for new vaccine development. *Technovation*, 34, pp. 762-772.

*Lin, P., Jenkins, R. and Abney, K. (2017). *Robot Ethics 2.0: From Autonomous Cars to Artificial Intelligence*. Oxford: Oxford University Press.

*Maas, M. (2018). Regulating for 'normal AI accidents': Operational lessons for the responsible governance of artificial intelligence deployment. *Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society*, pp. 223-228

Magro, E. and Wilson, J. (2013). Complex innovation policy systems: Towards an evaluation mix. *Research Policy*, 42(9), pp. 1647-1656.

*Majone, G. (2002). The precautionary principle and its policy implications. *Journal of Common Market Studies*, 40(1), pp. 89-109.

McKinsey & Company (2018). Artificial intelligence: automotive's new value-creating engine. McKinsey Center for Future Mobility. [pdf], Available at: https://www.mckinsey.com/industries/automotive-and-assembly/our-insights/artificial-intelligence-as-auto-companies-new-engine-of-value [Accessed 28-07-2019].

*McNally, P. and Inayatullah, S. (1988). The rights of robots: Technology, culture and law in the 21st century. *Futures*, 20(2), pp. 119-136.

Miles, M. and Huberman, A. (1994). *Qualitative Data Analysis: An Expanded Sourcebook*. London: Sage.

Moore, J. F. (1993). Predators and prey: A new ecology of competition. *Harvard Business Review*, 71, pp. 75-86.

Mundell, R. (1962). The appropriate use of monetary and fiscal policy for internal and external stability. *IMF Staff Papers*, 9(1), pp. 70-79.

Naisbitt, J. (1982). *Megatrends*. New York: Warner Books.

Niederman, F. & March, S. (2019). The "Theoretical lens" concept: We all know what it means, but do we all know the same thing?. *Communications of the Association for Information Systems*, *44*(1), pp. 1-33.

Oqubay, A. (2015). *Made in Africa: Industrial Policy in Ethiopia*. Oxford: Oxford University Press.

Osborn, A. (1953). *Applied Imagination: Principles and Procedures of Creative Thinking*. New York: Charles Scribner's Sons.

Padmore, T., Schuetze, H., & Gibson, H. (1998). Modeling systems of innovation: An enterprise-centered view. Research Policy, *26*(6), pp. 605-624.

Padmore, T., & Gibson, H. (1998). Modelling systems of innovation: II. A framework for industrial cluster analysis in regions. Research Policy, *26*(6), pp. 625-641.

*Palmerini, E., Bertolini, A., Battaglia, F., Koops, B., Carnevale, A. and Salvini, P. (2016). RoboLaw: Towards a European framework for robotics regulation. *Robotics and Autonomous Systems*, 86, pp. 78-85.

*Park, S. (2018). The fourth industrial revolution and implications for innovative cluster policies. *AI & Society*, 33(3), pp. 433-445.

*Pasquale, F. (2015). *The Black Box Society*. Boston: Harvard University Press.

Porter, M.E. (1985). *Competitive Advantage: Creating and Sustaining Superior Performance*. New York: Free Press.

*Rouvroy, A. (2008). Privacy, data protection, and the unprecedented challenges of ambient intelligence. *Studies in Ethics, Law, and Technology*, 2(1), pp. 1-51.

*Russell, S. and Norvig, P. (2016). *Artificial Intelligence: A Modern Approach*. Edinburgh: Pearson Education Limited.

Sapsford, R. and Jupp, V. (2006). *Data Collection and Analysis*. 2nd ed., London: Sage.

Saunders, M., Thornhill, A. and Lewis, P. (2016). *Research Methods for Business Students*. 7th ed., Harlow: Pearson.

Scherb, A., Palmer, A., Frattaroli, S. and Pollack, K. (2012). Exploring food system policy: A survey of food policy councils in the United States. *Journal of Agriculture, Food Systems, and Community Development*, 2(4), pp. 3-14.

*Scherer, M. (2015). Regulating artificial intelligence systems: Risks, challenges, competencies, and strategies. *Harvard Journal of Law & Technology*, 29(2), pp. 353-400.

*Schuelke-Leech, B., Jordan, S. and Barry, B. (2019). Regulating autonomy: An assessment of policy language for highly automated vehicles. *Review of Policy Research*, 36(4), pp. 547-579.

*Schwab, K. (2017). *The Fourth Industrial Revolution*. Geneva: World Economic Forum.

*Solaiman, S. (2017). Legal personality of robots, corporations, idols and chimpanzees: a quest for legitimacy. *Artificial Intelligence and Law*, 25(2), pp. 155-179.

Sotarauta, M. & Srinivas, S. (2006). Co-Evolutionary Policy Processes: Understanding Innovative Economies and Future Resilience. *Futures*, 38(3), 312-336.

*Sparrow, R. (2016). Kicking a robot dog. *ACM/IEEE International Conference on Human-Robot Interaction*, pp. 229-229.

Staat, W. (1993). On Abduction, Deduction, Induction and the Categories. *Transactions of the Charles S. Peirce Society, 29*(2), 225-237.

*Stahl, B. and Wright, D. (2018). Ethics and privacy in AI and big data: Implementing responsible research and innovation. *IEEE Security and Privacy*, 16(3), pp. 26-33.

*Stilgoe, J., Owen, R. and Macnaghten, P. (2013). Developing a framework for responsible innovation. *Research Policy*, 42(9), pp. 1568-1580.

Strauss, A. and Corbin, J. (1990). *Basics of Qualitative Research: Grounded Theory Procedures and Techniques*. Newbury Park: Sage.

*Sunstein, C. (2002). Beyond the precautionary principle. *University of Pennsylvania Law Review*, 151(3), pp. 1003-1058.

*Székely, I., Dániel Szabó, M. and Vissy, B. (2011). Regulating the future? Law, ethics, and emerging technologies. *Journal of Information, Communication and Ethics in Society*, 9(3), pp. 180-194.

*Tabandeh, A. (1994). Characterising artificial intelligence technology for international transfer. *AI & Society*, 8(4), pp. 315-325.

Tiwana, A. (2014). *Platform Ecosystems: Aligning Architecture, Governance, and Strategy*. Waltham, MA: Morgan Kaufmann.

*Tutt, A. (2016). An FDA for algorithms. *Administrative Law Review*, 83-123.

*Van Liebergen, B. (2017). Machine learning: A revolution in risk management and compliance? *Journal of Financial Transformation*, 45, pp. 60-67.

*Villaronga, E. and Golia, A. (2019). Robots, standards and the law: Rivalries between private standards and public policymaking for robot governance. *Computer Law & Security Review*, 35(2), pp. 129-144.

*Wall, L. (2018). Some financial regulatory implications of artificial intelligence. *Journal of Economics and Business*, 100, pp. 55-63.

*Wallach, W. (2011). From robots to techno sapiens: Ethics, law and public policy in the development of robotics and neurotechnologies. *Law, Innovation and Technology*, 3(2), pp. 185-207.

Witt, U. (2003). Economic policy making in evolutionary perspective. *Journal of Evolutionary Economics*, 13(2), pp. 77-94.

*Wright, D. (2011). A framework for the ethical impact assessment of information technology. *Ethics and information technology*, 13(3), pp. 199-226.

*Wurah, A. (2017). We hold these truths to be self-evident, that all robots are created equal. *Journal of Futures Studies*, 22(2), pp. 61-74.

Yearworth, M. (2010). Inductive modelling of an entrepreneurial system. Proceedings of the *28th International Conference of the System Dynamics Society*, pp. 25-29.

*Zhong, R., Xu, X., Klotz, E. and Newman, S. (2017). Intelligent Manufacturing in the Context of Industry 4.0: A Review. *Engineering*, 3(5), pp. 616-630.

*6, P. (2001a). Ethics, regulation and the new artificial intelligence. Part I: Accountability and power. *Information Communication and Society*, 4(2), pp. 199-229.

*6, P. (2001b). Ethics, regulation and the new artificial intelligence. Part II: autonomy and liability. *Information, Communication & Society*, 4(3), pp. 406-434.

Appendix A – AI Use Cases in the Manufacturing Value Chain

Support Activities

Value chain area	Application case	Description	AI type	Proposed value*
Firm infrastructure	Facility / process planning	High number of design propositions, shorter design iteration steps and a higher number of optimal solutions.	Software	Time-to-operation cost decrease in the design phase due to lower maintenance cost.
	Automated accounting	Automated information capturing system to read structured and non-structured information from invoices; automated invoices.	Software	Employment cost reduction. Shorter lead time to sending out invoices.
	Auditing with AI	Quick and more comprehensive audits of datasets compared to other sampling approaches.	Software	Increase in audit reliability. Employment cost reduction.
	Building maintenance	Higher energy generation and sustainability improvements.	Software	Externalized: lower CO2 levels, advanced peak-load predictions. grid balancing. Lower energy costs. Reputation win.
	Asset valuation	Pattern recognition of aerial images via AI-powered drones / software.	Software/Embodied	Balance sheet adjustments depending on asset valuations. Control of a facility's functionality and security.
Human resource management	Automated interview process	Replacement of human interviewers. Auto-selection of candidates based on their CV. Handling administrative tasks.	Software	Employment cost reduction decreases administrative costs.
	Chatbot / auto-voice in HR inquiries	Communication with employees.	Software	Improvement of the HR function's service availability. Employment cost reduction.
	Worker assistance system	Replacement of a middle manager/ shift leader in the production process with an AI communication tool.	Software	Avoidance of process interruptions. Enhanced planning capabilities. Employment cost reduction. Worker well-being.
	Time management	Shift management, work-place planning, and time management, potentially via mobile phone applications.	Software	Worker well-being. Reduction of transaction costs in planning. No interruptions in the production process.
Technology development	R&D process support	Quick testing of variations. Testing of different designs or molecules.	Software/ Embodied	Decrease in costs for laboratories. Faster time-to-market.
		Testing of many product combinations with high diligence.	Software/ Embodied	Potentially better products and customer satisfaction.
	Decision support systems	Generation of hypotheses to automatically suggest decisions based on information available in the company.	Software	Better decisions in the interest of the firm.
Procurement	Automated order processing	Automated order processing, for instance with an automation of emails with suppliers.	Software	Employment cost reduction.
	Quality prediction of buy-in parts	Prediction of assembled group based on input-part information. Supplier data to predict end-product quality.	Software	Increases in end-product quality. Reduction of reject rates. Less transaction costs and regress claims for manufacturer/supplier.
	Price comparison mechanism	Adjustment of procurement processes based on supplier/price comparison.	Software	Shorter time to change suppliers. Decreasing costs for input parts.
	Stock level predictions	Improvement of stock level predictions.	Software	Reduced working capital. Better inventory management.
	Data analytics for the virtual organization	Steering and analysis of data from the ecosystem and manufacturers in a virtual organization.	Software	Decreasing transaction costs. Horizontal and vertical integration. Transparency.

Primary Activities

Value chain area	Application case	Description	AI type	Proposed value*
Inbound & outbound logistics	Improved logistics network	General improvement of the logistics network. Optimisation of routes, adjustment of delivery intervals, and the shortest path.	Software	Reduction of overall transportation costs and delivery time.
	Real-time routing	Real-time tracking of delivery vehicles.	Software	Better timing of deliveries, especially in a just-in-time or just-in-sequence production environment.
	Autonomous trucks	Use of autonomous trucks in the delivery process.	Software/ Embodied	Employment cost reduction.
	Intelligent robots for loading/ unloading	Improvement of the loading process.	Embodied	Employment cost reduction.
Operations / production	Predictive maintenance	Prediction of service interval length and risk of machine breakdown.	Software	Increase in equipment utilization. Reduction of repair costs. Fewer service intervals.
	Machine improvement	More powerful and improved machines, for instance in tool-path optimization.	Software/ Embodied	Shorter production cycles. Higher equipment utilization.
	Quality control of end product or process step	Image recognition system for quality control of the assembly with simulation or visual inspection.	Software	Replacement of manual quality control tasks. Additional reassurance system. Decreasing number of scrap parts. Better overall quality and customer satisfaction.
	Prediction of production processes based on in-process parameters (real-time)	Overall control of the process performance, predictive quality assurance of the whole process system. Analysis of system performance.	Software	Reduction in overall process volatility. Performance stabilisation. Autonomously adjusting process parameters. Right-first-time.
	Production scheduling	Improvement of production routing and product flow.	Software	Better machine utilization.
	Performance of parts under testing	Prediction of performance of parts tested under real-case scenarios.	Software	Increased quality. Right-first-time. Assurance of components' safety.
	Voice interaction with machines	Communication of workers with machines, utilization of voice recognition software and algorithms.	Software	Reduction of transaction costs for workers. Higher flexibility in process adjustments.
	Robotics	Control system and image recognition in robotic systems for higher precisions and flexibilities, for instance for gripper reorientation.	Embodied	Quality improvement. Reduced maintenance costs. Right-first-time. Avoidance of process interruption.
	Autonomous inhouse-logistics	Incremental improvement of human-less systems based on LIDAR technology.	Embodied	Employment cost reduction.
	Machine connectivity	Connection of machines. Algorithmic optimization of the machine system.	Software/ Embodied	Increase in the production flexibility. Mastering of complex scheduling tasks.
	Soft robotics / collaborative robotics	Worker-human interaction with soft robotics.	Embodied	Employment cost reduction.
	Digital twin	Digital representation of a machine or factory providing real-time tracking of developments.	Software	Transparency. Maintenance cost reduction. Performance control.
	Worker assistance	Tablets or VR/AR devices for real-time worker augmentation.	Software	Better worktime utilization. Faster learning of tasks. Real-time problem recognition and parameter adjustments.

Marketing & sales	Demand forecast based on predictions	Enhancement of manufacturing capabilities for demand predictions and production adjustments.	Software	Incremental innovation of existing regression models with higher accuracy and higher revenue.
	Advanced customer analytics / trends	Algorithmic learning from purchasing behaviour of customers and better market trend predictions.	Software	Sale of customer-centric and customized products.
	Chatbot / auto-voice in customer interface	Customer exchange with chatbots and auto-voice comparable to HR chatbots.	Software	Improvement of service availability. Employment cost reduction.
	Use of AI in produced goods	Increase of product functionality for more innovative products. Satisfaction of increasing customer demand.	Software/ Embodied	Firm differentiation from competitors. Higher sales.
Service	Product pre-configuration for dealership	Customer demand analysis for a pre-configuration of cars for products with less generic options.	Software	Brand competitive advantage. Higher order rate for products. Better equipped and more expensive products.
	Production planning for spare parts	Planning for the demand and steering of spare parts.	Software	Lower inventory costs.
	Price adjustments for spare parts	Adjustment of spare part prices according to demand.	Software	Exploitation of purchasing need/willingness.
	Maintenance of AI systems in products	Maintaining regular service or updates in products.	Software	Control of access and servicing with proprietary programming for a lucrative business model. Risk of antitrust cases.
	Prescriptive servicing	Self-selection of service intervals and autonomous scheduling.	Software	Higher customer satisfaction. Employee cost reduction.
	Fleet management (B2B)	Improved fleet management and scheduling.	Software	Customer satisfaction. Employee cost reduction. Brand reputation.

Appendix B – Barriers to AI Scale-Up in the Manufacturing Value Chain

Type	Problematic	Details	Description
Technical barriers	Data gathering	Infrastructure deployment	Sensors, cloud, data factory and edge computing are major components to be deployed.
		Data pooling	Central pool of data; becomes more difficult of a task if data is siloed in different parts of the organisation and doesn't have the same format
		Legacy systems	Potential incompatibility of legacy systems with the new AI system.
	Data cleaning	Data transformation	Data needs to have the right semantic and ontology as information structure or needs to be transformed accordingly.
		Data labelling	Data must have the rights labels according to which it is classified; a time consuming and expensive task if done manually.
		Removal of errors	Cleaning of the datasets from potential errors is necessary.
		Data ethics	Data could introduce an ethical bias , based on gender, race, or other potentially discriminatory data. Very much context- and application use case dependent.
		Accuracy decisions	A use case could focus on either preventing false positives or false negatives depending on the application; for instance in safety critical applications, the AI system might need to be overly accurate.
	AI algorithm	Method and model identification	Company or AI service provider will have to identify the right method and model to the desired application, potentially in a trial and error phase.
		Model maintenance	Maintaining the model and data based on continuous learning and retraining is required.
		System adjustments & model recalibration	Adjusting the AI system once there is an adjustment of the underlying operational process becomes necessary due to changed data sources, process parameters, process steps or new product derivatives.
		Manufacturing process awareness	High awareness and understanding of the operational process itself on which the AI system should be applied is crucial.
		Data waste	Unnecessary data production could be viewed as waste according to the lean manufacturing principles and should be avoided.
		Open-source software	Many AI algorithms rely on open-source software; internal guidelines must be found and adjusted to use it.
		Repeatability of models	Repeatability of the AI models on different types of datasets is problematic. Differently structured datasets in different companies or different functions prevent a simple transfer from one application to another.
	System control	Supervisory system	Human-in-the-loop or conventional control system need to control for errors in the AI predictions and provide feedback; could also involve a kill switch for potentially very advanced, likely cognitive systems.
		Cyber security	Provision of sufficient security of the system is necessary and the cloud needs to be secure.
Operational barriers	Value proposition	Business case	The system's value should be quantifiable in a business case, which is more difficult to measure in soft KPIs.
		Proof of concept	Usefulness of the desired AI application might not be determinable until a proof of concept and trial phase: "fail fast, be quick".
	Management accountability	Black box problem	Outcome can be better and more powerful the deeper neural networks, but the less traceable are steps as to how the system reaches a conclusion; risk-aversion of top management and liability are problematic.
	Acceptance	High internal commitment	The system needs to be accepted by the workforce – or it is otherwise introduced against the will and tolerance of workers, which would endanger social cohesion.
	Selection of data source	Privacy concerns	Present AI systems are run based-on machine data, process data, or non-personal data. Employee data could be included as input into the AI model, as well, which might invoke privacy concerns.

	Transformation of work	Coordination	AI model and data are owned and administered by a company function that coordinates the system’s deployment.
		“Marriage” of IT and domain experts	Data scientists and domain experts need to work closely together, integrated into a team, and supported by a central IT function
		Job role change	Domain experts are potentially engineers that teach the algorithm, a task which they are first to be skilled for. This could involve repetitive tasks in supervised learning.
	Health and safety	Machine certification	Machines have to be certified before usage to ensure safe operations. Once a machine learning algorithm is applied in their control software, the machine changes its properties over the lifetime, which might require a renewed certification.
		Product certification	A manufacturer needs to think ahead if an AI system should be deployed in a product; especially if this product needs to be certified in tightly regulated industries like pharmaceuticals, aerospace, and automotive.
		Edge case failures	Might impose entirely new accountability and liability issues; for example, if a machine or product is used outside the predefined parameters for the algorithm
Ecosystem barriers	Awareness	Orientation / access point	Enhancing the awareness of manufacturing firms for AI requires the provision of orientation or access points to identify suitable AI service providers
		Hype	There is a danger that companies perceive AI as non-substantial hype and try to diminish the seriousness of the technology.
		Technology diversity	Technology diversity has increased, which diffuses the potential amount of private investment into AI.
	Technology transfer	Basic research applicability	Technology transfer from basic AI research into applied AI solutions offered by start-ups is challenging, as well as from start-ups and universities to a deployment by industrial firms.
		Overcoming late adoption	Manufacturing is generally a laggard or late adopting sector for digital technologies.
	Funding	VC funding	The level of VC funding into early-stage AI start-ups rarely targets the manufacturing sector due to quick ROI expectations by investors.
		Scale-up funding	Scale-up equity is necessary in addition to VC funding.
	Data sharing	Sharing of data with AI service providers	An AI service provider can establish and runs a cloud, as well as the AI solution in the cloud. The cloud of a large service-provider, like Azure or AWS, can serve as a platform. Sharing of data with involved firms needs to be facilitated.
		Sharing of data along the supply chain	Cross-company AI optimizations require data sharing along the supply chain in a pre-defined format and with a value proposition for all involved parties.
		Lack of technical and procedural standards	Data formats and the process requirements for data sharing are hurdles, and many data types like employee data and customer data cannot be shared.
		Availability of real training data	The AI service provider should have access to real industrial datasets for training their algorithms; otherwise, the predictions are potentially inaccurate
	Internal employment effects	Workforce reduction	Employees in the company potentially lose their tasks or even jobs in certain parts of the industrial value chain if AI deployment is scaled-up. This might trigger protests by trade unions.
		Wage and contract adjustments	Training of AI systems and work with AI alters the job roles in the manufacturing sector and could lead to necessary wage and contract adjustments.
	Skills development	Upskilling	If the AI system is used to assist or to augment the human, the need to learn new skills is limited to operating the AI system only. Upskilling might be necessary for other tasks based on job rotation.
		Hire of new employees	Data scientists are required to develop and program the AI systems company-internally. They need to be attracted and hired by manufacturing firms.

.

Appendix C – Taxonomy of Policy Remedies

Dimension	Sub-dimension	Policy remedy	Discussion	Time horizon
Innovation	N/A	Laissez-faire approach	Accelerating AI deployment was identified by the majority of respondents as desirable, but it was cautioned that the application should not be supported for the sake of the technology alone.	Short-term
		Business case transparency	Manufacturing firms need to see that AI applications can create a value in the form of a business case, which could be supported by policymakers. Many interviewees were confident that few public incentives are necessary then.	Short-term
		Network creation	Policymakers should think AI innovation from a network perspective, where the education and networking of firms is important. This could be initiated by policymakers.	Short-term
	Infrastructure	Use case database	The establishment of use case databases and access points for the showcasing of best practice examples, whether from an applied research perspective or as industrial application, was considered important by stakeholders.	Short-term
		Expert consulting programs	It was stressed that in many cases, especially for SME firms, it is not sufficient to showcase solutions, but also to provide application support by consulting programs, for instance with AI trainers.	Short-term
		AI algorithm platforms	Another initiative potentially useful is the development of a platform for AI algorithms, where these can be accessed in a standardized format as a playground for firms and developers. The challenging task is the adaptation of the AI model to its semantic infrastructure though.	Short-term
	Technology transfer	Clusters	Insufficient technology transfer mechanisms were generally seen as an important access point for policymaking, based on the formation of industrial clusters that include AI service providers and start-ups. Spatial proximity between AI firms and manufacturers was agreed to be a potential mean of increasing trust between firms, scepticism remained if this is a decisive problem in the digital age.	Medium-term
		Research collaboration	Close collaboration between industry and universities should enable a better uptake of AI technologies, and industrial firms could position themselves prominently in a university network or even with an on-campus representation supported by policymakers.	Short-term
	Public funding	Tax breaks / grants	Policymakers could develop funding vehicles, which allow companies to gain a tax refund or tax break if they buy and apply an AI system. For example, this could be a “research premium” as technology-agnostic funding vehicle that allows a manufacturing firm to claim a tax refund for expenses when applying a new AI system in its value chain.	Short-term / Medium-term
		Public-private partnerships	An approach for accelerated deployment could be the creation of public-private-partnerships through which a joint undertaking for innovation between the private and public sector is envisaged.	Short-term / Medium-term
		Other funding vehicles	Further public spending would be thinkable to finance data infrastructure or to match private VC funding.	Short-term / Medium-term
Employment	Human-centricity	Ethical deployment	Ethical deployment was clearly identified by many interviewees as a prerequisite of administering this technological advancement in the manufacturing sector.	Short-term
		Guidelines	New guidelines could be developed by policymakers or existing ones might be revised. This could include the use of employee data by AI systems, and the clarification of employees’ responsibilities when teaching algorithms or controlling their outcomes and performance.	Medium-term

	Automation	Job reduction counterbalance	With artificial intelligence technologies being an automation tool, its deployment in the manufacturing value chain could also lead to further automation and workplace reductions. A job guarantee for employees provided by manufacturing firms could be voluntary or imposed by policymakers.	Medium-term / Long-term
		Social cohesion support	A reduction of headcount could be perceived as undesirable by society and policymakers. The degree of intervention that is necessary might be determined by the demographic change within a country, the rigidity of the labour market, and the permeability and attractiveness of the immigration system for high-skilled labour. The degree of adjustment of these variables depends on the pace of the development.	Medium-term / Long-term
		Fiscal reactions	A reduction of wealth in the national pension and insurance schemes might force governments to potentially adopt new tax regulation for the taxation of automated systems, such as robots, and concepts like the universal basic income could become an option.	Long-term
	Skilling	Increase of digital talent availability	The envisaged number of data scientists is difficult to find for manufacturing firms on the job market, also due to the attractivity of high-tech B2C companies. Brain drain was reported as hampering the formation of applied research institutes. A recommendation for policy remedies was the creation of a strong educational system for digital talent.	Medium-term / Long-term
		Early practical and digital education	The educational system also needs to provide digital education in areas like machine learning for students in early ages and with practical elements. Non-profit organizations could support such a system and profit from public support.	Short-term
		Vocational training	The type of skills level needed to perform future data scientist tasks is difficult to determine. It might not necessarily PhD education, but rather bachelor and master's level students alongside special vocational training programs.	Short-term
Regulation	Soft	Internal auditing of AI	A strong evidence was found that auditing of AI systems is a regulatory option. Whilst an internal auditing was generally endorsed, the exact design of such an audit should be carefully prepared by policymakers to not hamper innovation.	Short-term
		Standardization of AI models	A technical standardization of AI models might make auditing easier and provides transparency for regulators.	Medium-term
		Principles development & harmonization	The development of general principles for the deployment of AI systems was endorsed by interviewees. Guidelines' scope could range from ethical considerations to corporate governance principles Policymakers could collect principles, compare them, and develop standardized corporate governance guidelines.	Short-term
		Regulatory sandboxes	Regulatory sandboxes could be created around AI applications constrained to a certain duration. In such a period, manufacturers could experiment with the use of AI applications in their value chains under relaxed regulatory requirements.	Short-term
	Strong	External auditing of AI	External audit mechanisms could supplement mandatory guidelines. A mechanism that allows governments to collect audits in an analogy to financial audits is thinkable.	Medium-term / Long-term
		Risk classes for algorithms	A risk classification system for AI algorithms could indicate different actions and requirements for firms, depending on the risk class of the AI system.	Medium-term / Long-term
		Health and safety guidelines	Clear development of health and safety guidelines for embedded AI might be necessary.	Short-term
		Liability clarification	Legal clarification of liability issues, or sole responsibility for courts to decide on liability issues, are alternative options to resolve the liability problematic.	Short-term

		Transparency of deployment	A transparency requirement for AI algorithms could be introduced, which aims to counterbalance the black box problematic.	Short-term / Medium-term
		AI system certification	Enforced certification and validation of AI systems in the value chain is amongst those activities that lean towards a stricter regulation of AI	Medium-term / Long-term
		Awareness of autonomy	The degree of autonomy of AI systems might decide about the future need of regulatory policymaking, potentially through standard setting bodies.	Medium-term / Long-term
Data	Collaboration	Third-party intermediaries for training of algorithms on real datasets	The use of third-party intermediaries would support the data sharing process. Governmental institutions or commercial firms could provide a trustee function for industrial datasets, which could also provide access to training data for AI service providers. The data trust in the United Kingdom was repeatedly mentioned as a possible approach.	Short-term / Medium-term
		Third-party intermediaries for supply chain sharing	Third-party intermediaries, also provisioned by the government, could ensure that datasets are anonymized. Supply chain partners and AI service providers can have anonymized access to these datasets. This might only be successful if there is a value proposition for the participating parties.	Short-term / Medium-term
	Data sharing	Technical standardization of data sets	A major problematic for third-party intermediaries is the differing information semantic between companies and siloed data within a company. The repeatability of the AI model on datasets of different manufacturers is therefore limited and the value of a third-party of third-party intermediary in this regard debatable.	Short-term / Medium-term
		Guidance	NDAs as contractual mean to safeguard the sharing process is often necessary. The provision of and guidelines for such templates could be provided by governmental institutions. GDPR could be extended to industrial data sharing.	Short-term / Medium-term
Governance	Governance approach	Coordination / agency building	A coordinated policymaking approach was stressed as favourable by interviewees. In the United Kingdom, the Office for Artificial Intelligence was formed as a team between the Department for Business, Energy & Industrial Strategy (BEIS) and Department for Digital, Culture, Media and Sport (DCMS), one of the first governmental initiatives of such kind, and the United Arabian Emirates have installed the first Minister for Artificial Intelligence. However, the exact competences of such institutions are disputed.	Short-term / Medium-term
		Ethical governance	Ethical policymaking was often named as the most suitable approach for the early maturity state of the technology, and several advisory bodies in various jurisdictions have been formed. This was widely supported by interviewees, but also the need for practicality to an application in real-world industrial environments was emphasized.,	Short-term
	Industry participation	Working groups	Industrial participation based on consultations or working groups, possibly within industry federations, were recommended for any kind of policymaking on AI, but especially for a domain-knowledge dependent field like manufacturing	Short-term
		Technology fund	To further enable an industrial participation based on an increase in R&D for advanced technology application, an industrial technology fund led by a consortium of decision makers of the industry could invest in AI type of technologies, applications, start-ups and entrepreneurs; this could potentially be initiated by policymakers.	Short-term