TECHNISCHE UNIVERSITÄT WIEN
Vienna | Austria

Navigating the AI Regulatory Landscape: A Meta-Narrative Approach to Assessing the Impact of Regulation on Innovation

Diploma Thesis

carried out for the purpose of obtaining the degree of Master of Science (MSc or Dipl.-Ing. or DI)

in

Mechanical Engineering – Management

by

Niklas Blum BSc.

Mat.Nr.: 01427044

submitted at TU Wien, Faculty of Mechanical and Industrial Engineering,

under the supervision of

Univ.-Prof. Mag. Dr. Sabine T. Köszegi

Institute of Management Science

Vienna, September 2023

.....	.....
Niklas Blum	Sabine T. Köszegi

Affidavit

I declare in lieu of oath, that I wrote this thesis and performed the associated research myself, using only literature cited in this volume. If text passages from sources are used literally, they are marked as such.

I confirm that this work is original and has not been submitted elsewhere for any examination, nor is it currently under consideration for a thesis elsewhere.

I acknowledge that the submitted work will be checked electronically-technically using suitable and state-of-the-art means (plagiarism detection software). On the one hand, this ensures that the submitted work was prepared according to the high-quality standards within the applicable rules to ensure good scientific practice "Code of Conduct" at the TU Wien. On the other hand, a comparison with other student theses avoids violations of my personal copyright.

Vienna, September 2023 *Niklas Blum*

Acknowledgements

I would like to take this opportunity to firstly thank my wife, who, despite all the circumstances, always supported me and had my back during the preparation of this work. You have also helped me to broaden my horizon with your open, loving and caring nature and for that I will be eternally grateful.

I would also like to express my gratitude to my daughter, who did not see the light of day during the writing of this thesis, but who nevertheless motivated me beyond measure. Once we finally meet, I am sure we will never drift apart.

I extend my heartfelt gratitude to my parents for their unwavering support and encouragement throughout this study, without the both of you I couldn't be where I am now, thank you for that.

Furthermore, I would also like to thank my supervisor Univ.-Prof. Mag. Dr. Sabine T. Köszegi, who helped me find my way at the TU Vienna during my studies. You are and will be forever an important role model for myself and surely many other students, that have the privilege to work with you.

Lastly, also thanks to all the other companions I met during my studies at TU Wien. It was a blast, and I could not have done it without all of you.

Abstract

Artificial intelligence is ubiquitous nowadays. For several years, there has been a massive increase in technical applications, media coverage, and scientific publications on the topic of artificial intelligence. There are more and more examples of futuristic applications and promises for a better future with, and especially through AI. At the same time, however, a countermovement is gaining more and more momentum, arguing for more regulated research and application of AI. The aim of this paper is to show how regulation of AI affects innovation. For this purpose, scientific publications from different disciplines, in the period from May 2018 to December 2022 are examined semi-systematically. This semi-systematic literature review will be used to provide a meta-narrative, interdisciplinary arch, on the impact of regulating artificial intelligence on innovation, and to map the current scientific consensus and latest findings. The time-limited publications forming the start of this semi-systematic literature review are additionally provided with metadata to map possible trends.

An interdisciplinary consensus is emerging that innovations are constrained by regulations, depending on quality and quantity thereof. However, an equally clear picture is that there is also consensus on the need to regulate AI, to promote sustainable innovation, avoid catastrophic implications, and to avoid reactive over-regulation by regulatory entities intervening too late. There is also a need to reformulate existing regulations and new regulatory measures, as the nature of AI cannot be regulated by traditional measures alone.

Kurzfassung

Seit einigen Jahren ist ein massiver Anstieg an technischen Anwendungen, medialer Berichterstattungen, und wissenschaftlicher Veröffentlichungen zum Thema künstliche Intelligenz zu beobachten. Es gibt immer mehr Beispiele futuristischer Anwendungen und Versprechen für eine bessere Zukunft mit-, und vor allem durch künstliche Intelligenz. Gleichzeitig nimmt aber auch eine Gegenbewegung zunehmend an Fahrt auf, die sich für eine stärker regulierte Forschung, sowie Anwendung künstlicher Intelligenz ausspricht. Das Ziel dieser Arbeit ist es aufzuzeigen, inwiefern sich die Regulierung künstlicher Intelligenz auf Innovation auswirkt. Dabei werden wissenschaftliche Veröffentlichungen unterschiedlicher Disziplinen, im Zeitraum vom Mai 2018 bis zum Dezember 2022 semi-systematisch untersucht. Diese semi-systematische Literaturrecherche wird genutzt, um einen meta-narrativen, interdisziplinären Bogen, über den Einfluss von Regulierung künstlicher Intelligenz auf Innovation zu spannen und den derzeitigen wissenschaftlichen Konsens und neueste Erkenntnisse abzubilden. Die den Start dieser semi-systematischen Literaturrecherche bildenden, zeitlich begrenzten, Veröffentlichungen werden zusätzlich mit Metadaten versehen, um einen möglichen Trend abzubilden.

Es zeichnet sich ein interdisziplinärer Konsens ab, dass Innovationen durch Regulierungen, je nach Qualität und Quantität dieser, eingeschränkt werden. Allerdings zeichnet sich ebenfalls ab, dass auch ein Konsens über die Notwendigkeit der Regulierung künstlicher Intelligenz, zur Förderung nachhaltiger Innovation, Vermeidung katastrophaler Auswirkungen, sowie der Vermeidung reaktiver Überregulierung durch zu spätes Eingreifen regulatorischer Entitäten, besteht. Außerdem bedarf es einer Umformulierung bestehender Verordnungen und neuer regulatorischen Maßnahmen, da künstliche Intelligenz nicht mit herkömmlichen Maßnahmen alleine reguliert werden kann.

Table of Contents

1. Introduction and Motivation

"… the long term growth of an advanced economy is dominated by the behaviour of technical progress" (Weitzman, 1998). According to McKinsey the technical progress for the upcoming years is heavily dependent on applications of Artificial Intelligence (AI), as can be seen in their Technology Trend Outlook 2022. Mc Kinsey stated in their report that applied AI has the highest innovation score of all observed technology trends and is the only trend also having a meaningful association with every other observed industry sector (*McKinsey Technology Trends Outlook 2022 | McKinsey*). However, by the start of this thesis the regulation of AI is still in its children's shoes. The European Union (EU) has drafted a concept for the regulation of AI by March 2018, as seen in Figure 14: Timeline of the EU-AI milestones. And the United Kingdom (UK) published a white paper on the regulation of AI with the title "*AI Regulation: A Pro-Innovation Approach*" on March 29th, 2023[1]. Both of these drafts take on a so-called risk-based approach, which is based to a certain degree on the precautionary principle which will be further explained in the main findings of this thesis. Other countries, such as the United States have yet to introduce their vision for a regulatory landscape in their country[2].

There surely are several reasons for the lack of implemented regulations. One factor making it very difficult to implement regulation for AI is its broad application. As stated earlier applied AI is an integral part of every industry trend observed by McKinsey, ranging from Aerospace and Defense, Education to Retail and Telecommunication. But AI also differs in ways of

[1] https://www.gov.uk/government/publications/ai-regulation-a-pro-innovation-approach/white-paper, Accessed 10.05.2023.

[2] https://www.goodwinlaw.com/en/insights/publications/2023/04/04_12-us-artificial-intelligence-regulations, Accessed 10.05.2023.

appearance, from natural language processing (NLP), decision supporting systems (DSS) and many more. And of course, other reasons, such as the lack of international consensus, ethical and societal implications and the resistance from the industry and their lobbies. And finally, AI is very dynamic, and laws and regulations tend to be more static. To implement laws on an international or even national level, many bureaucratic hurdles must be overcome. Whilst it is in the nature of AI to accumulate pace in regards of development over time, the higher complexity slows down the regulatory process. One example would be the draft of the European Union-AI Act (EU-AIA) that seems to evolve significantly slower than the technology it is set out to regulate.

Nevertheless, the importance of AI-regulation seems to become more and more visible. Many leading scientists recently demanded in an open letter[3] to put the further development of AI on hold as *"[artificial intelligence] ... pose profound risks to society and humanity, as shown by extensive research and acknowledged by top AI labs"*.

Historically, there have been a number of unwanted phenomena, both due to the lack of regulation and due to regulation itself. And it is important not to repeat those mistakes made in the past. An example for an unwanted phenomenon is the prohibition in the United States in the 1920´s. The goal was to reduce crime and improve public health. However, neither was achieved since illegal distilleries thrived and unchecked Methanol heavy drinks were brought into circulation, poisoning Americans as well as driving them into criminality (Jones, 1975). Contrary, an example for the lack of regulation is the great smog in London of 1952, killing thousands due to environmental pollution, because of the unbridled use of a former disruptive technology, namely Watt´s Steam engine, filling the streets of London with air pollutants[4].

[3] https://futureoflife.org/open-letter/pause-giant-ai-experiments/, Accessed 10.05.2023.

[4] https://www.metoffice.gov.uk/weather/learn-about/weather/case-studies/great-

According to Erik Brynjolfsson, Jacob Turner and other influential thinkers who deal with the topic of AI, we are again at the brink of such a disruptive technology.

Since the effect of regulation on AI will affect nearly all aspects of our lives and in various fields, it cannot be described in one thesis alone. The effects could be of environmental nature, compared to the Industrial Revolution, economical nature and/or of societal nature. The aim of this paper will therefore only be to address the state-of-the-art theories about the effect on innovation by regulation of AI and to take up the prevailing key narrative regarding this topic.

All over the world there are different approaches regarding the regulation of emerging technologies. In the mainstream media the prevalent opinion regarding regulations and policy measures on technology seems to be, that regulation has a negative effect on the innovation capability of companies or at the least is hindering research and development. Leading research institutes and scientists in Europe published an open letter[5] on 28th of April 2023 calling for a revision of the EU-AIA. One of their critiques being reduced competition between model drivers and a drive of investments overseas. Also, Alexander Wrabetz, head of the Expert Group Digitalisation and AI of the Think Tank Future Vienna of SWV WIEN, argues: "*If the Act is realized, global AI development will not be stopped, it just won't happen in Europe*[6]*"*. This seems to be in line with the often called for approach "*Innovate, then regulate, or do not regulate at all*[7]*"*. Surely this approach can lead to an accumulation of technology startups, as some key numbers in the 2022 Report of the Tech Scale up Silicon Valey show in Figure 1:

---

smog#:~:text=A%20fog%20so%20thick%20and,to%20death%20in%20the%20fields, Accessed 10.05.2023.

[5] https://laion.ai/notes/letter-to-the-eu-parliament/; Accessed 10.05.2023.

[6] https://www.ots.at/presseaussendung/OTS_20230227_OTS0060/arigewrabetz-eu-artificial-intelligence-act-bedroht-innovation-in-europa, Accessed 10.05.2023.

[7] https://www.tagesspiegel.de/politik/uberregulierung-versus-unwirksamkeit-wie-soll-europa-mit-kunstlicher-intelligenz-umgehen-9601610.html, Accessed 10.05.2023.

Silicon Valley VS. Global Tech Startups.

Figure 1: Silicon Valley VS. Global Tech Startups (Karim, 2022).

But is the sheer number of new Startups an indicator for innovation? Or is there more to innovation than quantity?

Another narrative claims that a more rigor approach is appropriate. Often substantiated by the fear of unanticipated effects of AI or to pre-empt possible unintended consequences. A solution to deal with this is the so-called risk-based approach, as seen by the way the European Commission (EC) is trying to regulate AI applications, by classifying them into four different risk categories, as seen in Figure 2: A Risk-based approach. Framework for the AI Regulation in the European Union. These approaches are fundamentally different. It is not reasonable, or indeed possible, to favor one over the other, without the consideration of an interdisciplinary context. Any innovation can have a terrific economical outcome for some stakeholder, whilst its impact on a societal level can be very different. Even if the question is more precise, i.e., regarding the influence of these regulatory measures or policies on innovation, factors like short, medium, and long term have to be taken into consideration.

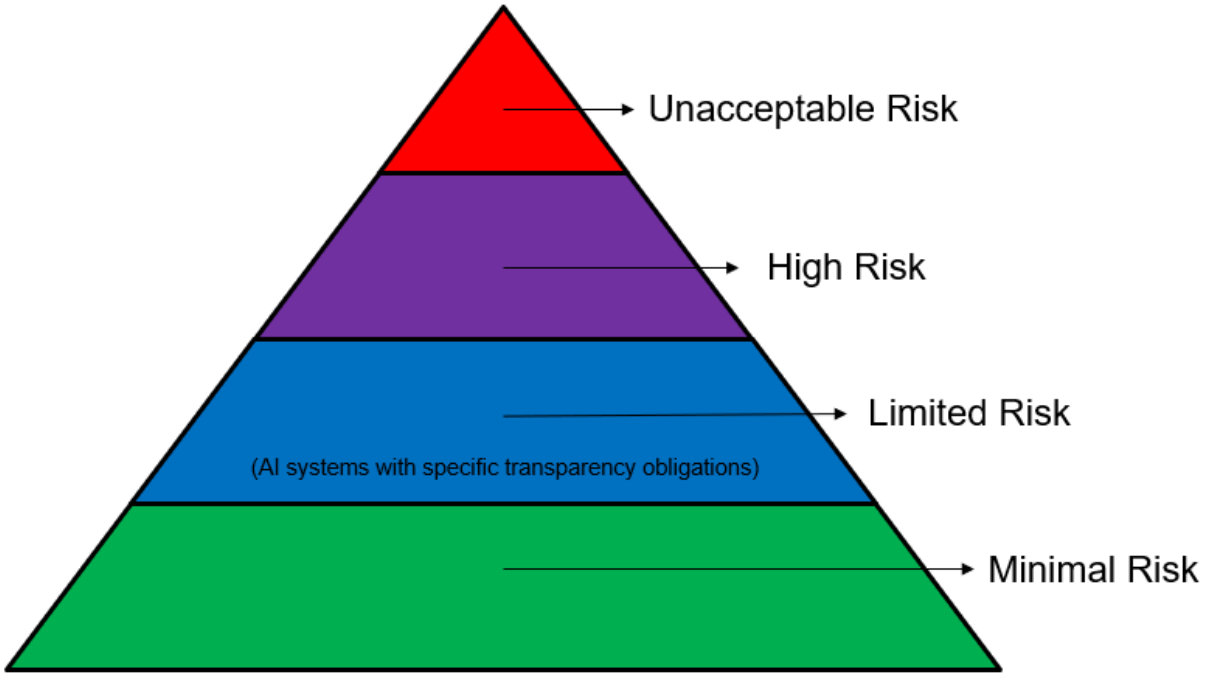

Figure 2: A Risk-based approach. Framework for the AI Regulation in the European Union, adapted from Shaping Europe´s Digital Future (2023). (Regulatory Framework Proposal on Artificial Intelligence | Shaping Europe's Digital Future, 2023)

However, the impact of technology regulation on the capability, capacity and sustainability of innovation should be, and via this thesis will be put under scientific observation. To study these effects, a meta-narrative literature review on the impact of regulation on innovation is the basis of this thesis. Thus, trying to give a comprehensive summary on what should be taken into consideration when formulating such regulations or policy measures, whilst trying to include various disciplines and combine them with an overarching narrative. A meta-narrative review highlights the diverse and complimentary approaches that academics have used to study the same or a related topic in order to provide light on a heterogeneous topic area (Wong et al., 2013). Making this the appropriate form of review to investigate the influence of regulation on innovation.

2. Theoretical Foundations

In this chapter, the necessary theoretical foundations for understanding the work are explained. These include general principles related to innovation, regulation, AI and conducting a meta-narrative literature review.

2.1. Innovation

To get a grasp of the widely used term of innovation, here are some of the most used definitions. Schumpeter has given one of the earliest and most often used definitions of innovation: "*[...] the market introduction of a technical or organisatorial novelty, not just its invention. ...the carrying out of new combinations* (Schumpeter, 1934)". In his firstly published theory regarding innovation, in the "Theory of economic development" in 1911, Schumpeter argued that entrepreneurs are the drivers of economic growth (Michael Filzmoser, 2022W). And he also coined the term of creative destruction, an economic theory emphasizing the importance of an economic dynamism. Creative destruction is best described as a process where new innovations replace existing ones, making the previous obsolete over time.

Later in Schumpeter's career, he shifted from entrepreneurs being the drivers of innovation towards larger companies -intrapreneurs- who mechanized innovation.

Another scholar, Michael E. Porter coined the term innovation as "*a new way of doing things (termed an invention by some authors) that is commercialized* (Porter, 1990)". Thus, combining innovation itself with the process of bringing new products, or services to the market, adding another layer to innovation.

The term general purpose technology (GPT) is a category of technological innovations that have the potential to significantly impact various sectors of the economy and society as a whole. GPT was coined by Bresnahan and Trajtenberg already assigning the three traits pervasiveness, technological dynamism and innovational complementarities to ideas or products that have the

potential for many important impact on many sectors of the economy (Bresnahan & Trajtenberg, 1995). A few years later economists, among them Gavin Wright labelled GPT as innovations that are recognized by being persuasive, improving over time and able to spawn additional innovations. Wrigth defined the term GPT as *"deep new ideas or techniques that have the potential for important impact on many sectors of the economy* (Wright, 2000)*"*. AI does meet all of the criteria and appears to be the next GPT, and thus heralds in a new era for recombination's and consequently innovations. Other examples of GPT are the steam engine introduced by James Watt and electricity (Brynjolfsson & McAfee, 2014). It is imminent to underline the difference between any new technology and GPT. In contrast to "regular" new technologies, a GPT is enabling the new combination of innovations. And therefore, acts as an accelerator for a large number of innovations.

2.2. Regulation

The control of an activity or process, usually through rules, is referred to as regulation. Baldwin et al. divided regulation into three different modalities (Baldwin et al., 2011):

- Regulation can be a set of specific commands – binding obligations applied to a body devoted to this purpose.
- Regulation can refer to a state of influence including financial and other incentives.
- Regulation can be used to denote all forms of social or economic suasion, including market forces.

The theory of smart regulation furthermore adds that regulatory functions can not only be carried out by institutions of governments, but also by professional associations, standard setting bodies and advocacy groups (Gunningham & PN, 1998).

Furthermore regulation can in addition to prohibiting unwanted conduct -red light regulation- enable or facilitate positive activities -green light regulation- (Harlow & Rawlings, 2009).

2.3. Artificial Intelligence

The term artificial intelligence was first mentioned in the scientific world at Dartmouth, when John McCarthy and three colleagues wrote a research proposal to the Rockefeller Foundation (J. Mc Carthy et al., 1955):

> *"We propose that a 2 month, 10 man study of artificial intelligence be carried out during summer of 1956… The study is to proceed on the basis of the conjecture that every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it. An attempt will be made to find how to make machines use language, form abstractions and concepts, solve kinds of problems now reserved for humans, and improve themselves. We think that a significant advance can be made in one or more of these problems if a carefully selected group of scientists work on it together for a summer."*

After this event the first definitions for AI had a human centric approach. The goal was to define intelligence referring to human like intelligence. Until today AI is far away from reaching this human level of intelligence. However, notably there are some AI programmes that, in specific tasks, exceed "human intelligence". Examples for such AI´s range from applications in medicine to the Fintech area (Grace et al., 2018), but also in more narrow fields like the Chinese board game Go[8].

One of the most famous examples for an AI definition was given by Ray Kurzweil: *"[…] the art of creating machines that perform functions that require intelligence when performed by people* (Kurzweil, 1990)". Later on, scientists moved from a human centric approach to more

[8] https://www.deepmind.com/research/highlighted-research/alphago#:~:text=AlphaGo%20is%20the%20first%20computer,strongest%20Go%20player%20in%20history. Accessed, 23.07.2023.

rational definitions by focussing on thinking or being rational. An example being: *"[Artificial Intelligence is]…that quality that enables an entity to function appropriately and with foresight in its environment* (Nilsson, 2010)."

Today, we differentiate between a so called "weak" AI and "strong" AI. Li and Zhang (2017) categorized strong AI as following: *"In the category of strong AI, AI system is considered to have human-like high level cognition ability, such as common sense, self-awareness and creativity, while weak AI simulates human intelligent processes passively without real understanding* (p.416)." A weak AI therefore, is built to serve a specific purpose or to accomplish a specific task. Or in other words, strong AI refers to systems that self-identify data from, possibly unstructured patterns and weak AI refers to a set of correct answers that accompany the data input (Truby et al., 2020). This can range from playing chess, recognize images and many more options. A weak AI is dependent on human interference, defining the parameters of its learning algorithm and providing training data to maximize accuracy.[9]

With many different terminologies available and the frequent use of incorrect terms, especially in literature and films, many terms are used very interchangeable. Currently, when people are referring to AI, they refer mostly to machine learning. In reality, machine learning is only a subfield of AI and can be further divided into unsupervised and supervised machine learning. A more structured overview of AI is given in the following section.

Weak AI, or so called Narrow Intelligence, is an umbrella term for all technologies (so far) deriving from Artificial Intelligence. Meaning, Artificial Intelligence is the outermost shell followed by Machine Learning, Neutral Networks and lastly Deep Learning. This structure can be seen in Figure 3: Hierarchical Structure of AI.

[9] https://www.ibm.com/cloud/learn/strong-ai#toc-strong-ai--YaLcx8oG, Accessed 28.08.2023.

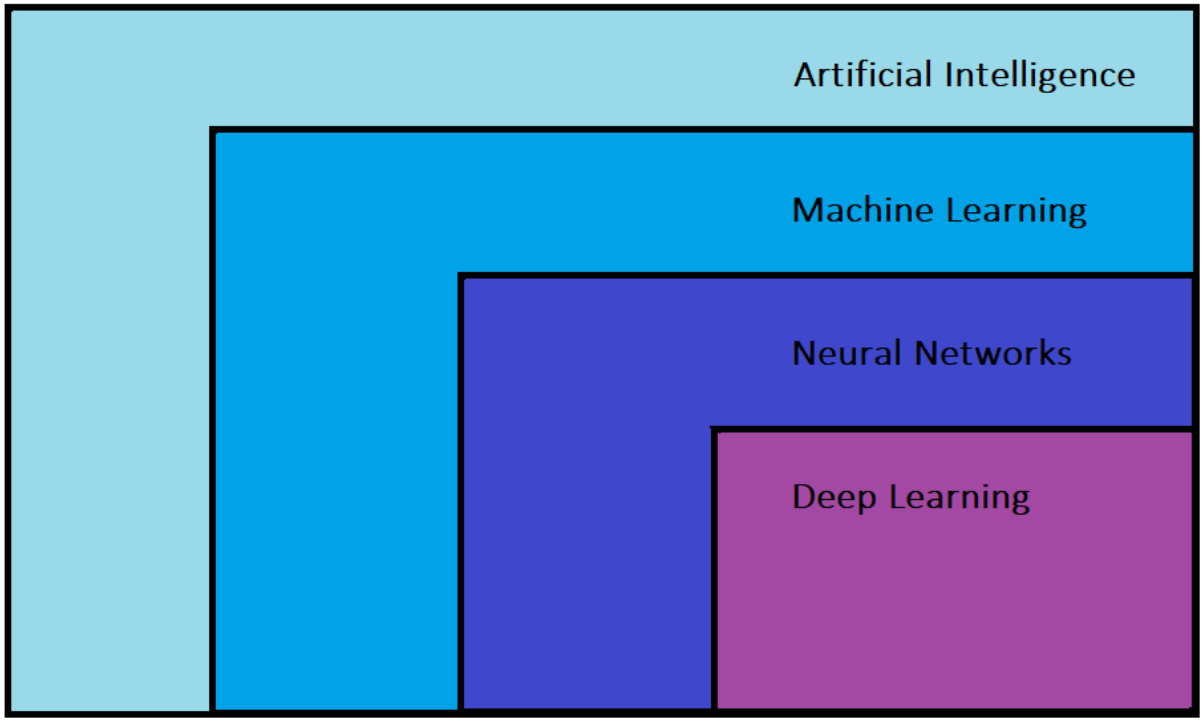

Figure 3: Hierarchical Structure of AI.

The method of Machine Learning can also be further divided into three different categories. These three categories are Supervised Machine Learning, Unsupervised Machine Learning and finally Semi-supervised Machine Learning. IBM, as one of the leading companies developing applications with AI, is describing those as follows:

***Supervised learning**, also known as supervised machine learning, is defined by its use of labelled datasets to train algorithms that to classify data or predict outcomes accurately. As input data is fed into the model, it adjusts its weights until the model has been fitted appropriately. This occurs as part of the cross validation process to ensure that the model avoids overfitting or underfitting* (What Is Supervised Learning? | IBM, 2021).

***Unsupervised learning**, also known as unsupervised machine learning, uses machine learning algorithms to analyse and cluster unlabelled datasets. These algorithms discover hidden patterns or data groupings without the need for human intervention. Its ability to discover similarities and differences in information make it the ideal solution for exploratory data analysis, cross-selling strategies, customer segmentation, image and pattern recognition. It's*

also used to reduce the number of features in a model through the process of dimensionality reduction; principal component analysis and singular value decomposition are two common approaches for this. Other algorithms used in unsupervised learning include neural networks, k-means clustering, probabilistic clustering methods, and more (What is Unsupervised Learning? | IBM, 2021).

Semi-supervised *learning offers a happy medium between supervised and unsupervised learning. During training, it uses a smaller labelled data set to guide classification and feature extraction from a larger, unlabelled data set. Semi-supervised learning can solve the problem of having not enough labelled data (or not being able to afford to label enough data) to train a supervised learning algorithm* (Delua, 2021).

Neural Networks have been the purpose of study for a long time. Already in 1943, Warren S. McCulloch and Pitts published a paper called "*A logical calculus of the ideas immanent in nervous activity* (McCulloch & Pitts, 1943)". Published in the Bulletin of Mathematical Biophysics, Warren laid the groundwork for Neural Networks by trying to represent the human brain activity through the use of Boolean logic – zero and one; true and false -. Fifteen years later Frank Rosenblatt was able to teach a computer to distinguish cards marked either on the right or on the left side (Rosenblatt, 1958). In the year 1989 Yann LeCun published his paper on how neural network architecture can be used to train algorithms. His research led to a breakthrough in the postal service, when he designed a computer that could recognize handwritten zip-codes digits in the United States Postal Service (LeCun et al., 1989).

To better understand what exactly Neural Networks are nowadays, IBM has published an online article stating the following:

Artificial neural networks are comprised of a node layers, containing an input layer, one or more hidden layers, and an output layer. Each node, or artificial neuron, connects to another and has an associated weight and threshold. If the output of any individual node is above the

specified threshold value, that node is activated, sending data to the next layer of the network. Otherwise, no data is passed along to the next layer of the network (What Are Neural Networks? | IBM, 2021).

Lastly, in this sequence is Deep Learning. Deep Learning can be further divided into more applications such as speech recognition, pattern recognition, contextual recommendations and even tasks like fact checking can be achieved by Deep Learning algorithms.[10] This gives a first indication of how many areas AI is used in nowadays without us expecting it. In very simplified terms, Deep Learning is a multi-layered Neural Network. These networks aim to copy the behaviour in a human brain, meaning to learn from input data. Of course, Deep Learning is as of now far away from being comparable to a human brain in terms of matching it´s ability. The added layers of Neural Networks allow the machine to optimize and refine accuracy.[11]

[10] https://www.ibm.com/cloud/learn/strong-ai#toc-strong-ai--TCZC11gq, Accessed 29.08.2023.

[11] https://www.ibm.com/cloud/learn/deep-learning, Accessed 29.08.2023.

2.4. Literature Review

The amount of data and the accumulated knowledge on any given topic is constantly increasing. In order to reflect the actual state of the art, literature research can be used as a scientific method. However, the way a literature review is conducted, depends on the topic to be researched and the desired outcome. A classification of different types of literature research can be found in Table 1: Approaches to literature reviews (Snyder, 2019).

Approaches to literature reviews.

Approach	Systematic	Semi-systematic	Integrative
Typical purpose	Synthesize and compare evidence	Overview research area and track development over time	Critique and synthesize
Research questions	Specific	Broad	Narrow or broad
Search strategy	Systematic	May or may not be systematic	Usually not systematic
Sample characteristics	Quantitative articles	Research articles	Research articles, books, and other published texts
Analysis and evaluation	Quantitative	Qualitative/quantitative	Qualitative
Examples of contribution	Evidence of effect	State of knowledge	Taxonomy or classification
	Inform policy and practice	Themes in literature	Theoretical model or framework
		Historical overview	
		Research agenda	
		Theoretical model	

Table 1: Approaches to literature reviews (Snyder, 2019).

Contrary to other forms of literature reviews, the semi systematic literature review, or narrative literature review can have five different outcomes (Baumeister & Leary, 1997):

1. Theory development.
2. Theory evaluation.
3. Survey the state of knowledge.
4. Problem identification.
5. Providing a historical account of the development of the theory.

Semi systematic reviews are suitable when topics have been approached interdisciplinary, because the complexity of building a suitable systematic approach for a scientific paper, taking different approaches of different disciplines into account, is too high (Wong et al., 2013). In other words, a quantitative approach is either not possible due to the number of publications, or due to the combination of different methods and results. Therefore, a mix of qualitative and quantitative is advisable for the investigation of the impact of regulation on innovation.

Another suitable scope of application for semi systematic reviews is to give a historical overview of the topic (Snyder, 2019). Especially when writing a qualitative review this can help to showcase how a topic has evolved over time and how it was perceived by different disciplines, creating a interdisciplinary overarch.

To fulfill a certain scientific standard, it is advisable to conduct a literature review using an existing guideline. After deciding on a specific approach, based on the desired outcomes of the literature review, according to Table 1: Approaches to literature reviews (Snyder, 2019), a guideline can be chosen. Snyder also provides some examples of these guidelines, which can be seen in Table 2: Examples of existing guidelines for conducting a literature review (Snyder, 2019).

Examples of existing guidelines for conducting a literature review.

Authors	Discipline	Type of literature review	Contribution
Baumeister and Leary (1997)	Psychology	Narrative review	• Overviews reasons for conducting a review • Discusses common mistakes for conducting a review
Tranfield et al. (2003)	Management	Systematic review	• Compares management and healthcare research • Highlights the challenges of conducting a systematic review in management research • Provides guidelines for conducting a systematic literature review in management research
Torraco (2005)	Human Resources	Integrative review	• Defines the integrative literature review • Provides guidelines and examples for integrative literature reviews • Discusses contributions of a integrative literature review
Liberati et al. (2009)	Medicine	Systematic review and meta-analysis	• Provides guidelines for conducting and reporting systematic reviews and meta-analysis
Wong et al. (2013)	Medicine	Semi-systematic review	• Provides guidelines for conducting a meta-narrative review
Davis et al. (2014)	Social Sciences	Systematic review and meta-analysis	• Synthesizes guidelines for systematic literature reviews • Provides guidelines for conducting a systematic review and meta-analysis in social sciences
Palmatier et al. (2018)	Marketing	Review papers and systematic reviews	• Provides guidelines for publishing review papers in the Journal of the Academy of Marketing Science

Table 2: Examples of existing guidelines for conducting a literature review (Snyder, 2019).

2.4.1. Meta Narrative Literature Review - RAMESES

RAMESES is an abbreviation for “Realist And Meta-narrative Evidence Synthesis: Evolving Standard” and is the first publication standard derived for the execution of an meta-narrative literature review (Wong et al., 2013). The goal of this meta-narrative method is to combine different scientific approaches from an heterogenous field of studies on a similar topic. As an inspiration for this meta-narrative review Thomas Kuhn´s work is cited, which itself is grounded in a constructivist philosophy of science (Kuhn, 2012).

Even though this is a qualitative method, the RAMESES standard provides a list of items to ensure the scientific quality of the review. This list can be seen in Tables 3 and 4.

TITLE		
1		In the title, identify the document as a meta-narrative review or synthesis
ABSTRACT		
2		While acknowledging publication requirements and house style, abstracts should ideally contain brief details of: the study's background, review question or objectives; search strategy; methods of selection, appraisal, analysis and synthesis of sources; main results; and implications for practice.
INTRODUCTION		
3	Rationale for review	Explain why the review is needed and what it is likely to contribute to existing understanding of the topic area.
4	Objectives and focus of review	State the objective(s) of the review and/or the review question(s). Define and provide a rationale for the focus of the review.
METHODS		
5	Changes in the review process	Any changes made to the review process that was initially planned should be briefly described and justified.
6	Rationale for using meta-narrative review	Explain why meta-narrative review was considered the most appropriate method to use.
7	Evidence of adherence to guiding principles of meta-narrative review	Where appropriate show how each of the six guiding principles (pragmatism, pluralism, historicity, contestation, reflexivity and peer review) have been followed.
8	Scoping the literature	Describe and justify the initial process of exploratory scoping of literature.
9	Searching processes	While considering specific requirements of the journal or other publication outlet, state and provide a rationale for how the iterative searching was done. Provide details on all the sources accessed for information in the review. Where searching in electronic databases has taken place, the details should include (for example) name of database, search terms, dates of coverage and date last searched. If individuals familiar with the relevant literature and/or topic area were contacted, indicate how they were identified and selected.
10	Selection and appraisal of documents	Explain how judgements were made about including and excluding data from documents, and justify these.
11	Data extraction	Describe and explain which data or information were extracted from the included documents and justify this selection.

Table 3: List of items to be included when reporting a meta-narrative review, part 1 (Wong et al., 2013).

12	Analysis and synthesis processes	Describe the analysis and synthesis processes in detail. This section should include information on the constructs analysed and describe the analytic process.
RESULTS		
13	Document flow diagram	Provide details on the number of documents assessed for eligibility and included in the review with reasons for exclusion at each stage as well as an indication of their source of origin (for example, from searching databases, reference lists and so on). You may consider using the example templates (which are likely to need modification to suit the data) that are provided.
14	Document characteristics	Provide information on the characteristics of the documents included in the review.
15	Main findings	Present the key findings with a specific focus on theory building and testing.
DISCUSSION		
16	Summary of findings	Summarise the main findings, taking into account the review's objective(s), research question(s), focus and intended audience(s).
17	Strengths, limitations and future research directions	Discuss both the strengths of the review and its limitations. These should include (but need not be restricted to) (a) consideration of all the steps in the review process and (b) comment on the overall strength of evidence supporting the explanatory insights which emerged. The limitations identified may point to areas where further work is needed.
18	Comparison with existing literature	Where applicable, compare and contrast the review's findings with the existing literature (for example, other reviews) on the same topic.
19	Conclusion and Recommendations	List the main implications of the findings and place these in the context of other relevant literature. If appropriate, offer recommendations for policy and practice.
20	Funding	Provide details of funding source (if any) for the review, the role played by the funder (if any) and any conflicts of interests of the reviewers.

Table 4: List of items to be included when reporting a meta-narrative review, part 2 (Wong et al., 2013).

3. Methodology

The procedure for obtaining the body of literature for the literature review is described in this chapter, following – as far as permissible - the RAMESES publication standard (Wong et al., 2013).

3.1. Changes in Review process

Firstly, to get an overview of the topic Google, as well as Google Scholar are used to get a first impression of the availability of the topic.

After having identified the relevant databases and a general idea of the scientific point of view, regarding the topic of regulating technology and its impact on innovation, a decision was made to only search the databases listed in Table 5: Databases for Literature Research.

Due to a non-editable number of results, the search had to be narrowed down by not only using Boolean operators “AND” and “OR”, but also define additional inclusion and exclusion criteria.

The database IEEE Explore has been excluded, as no relevant results were found here with the search terms.

After having done the formal literature review and reviewing the relevant content, a forward and backward searching method was applied to gather further, useful information. This gathered sources will not be included in the RAMESES procedure but are mentioned in the text and the bibliography at the end of the thesis.

3.2. Rationale for using this method

Firstly, the approach of executing a literature review needs to be determined. According to Snyder a semi-systematic approach is the fitting approach for a multidisciplinary view on such a broad topic. As can be seen from Table 1: Approaches to literature reviews (Snyder, 2019), the semi-systematic approach is very broad, can include qualitative and quantitative analysis and evaluation, and it´s examples of contribution include themes in literature, theoretical

models as well as a historical overview.

Moreover, Snyder refers to Geoff Wong et al. for an existing guideline on executing a semi systematic review, as can be seen in Table 2: Examples of existing guidelines for conducting a literature review (Snyder, 2019), namely "RAMESES publication standards: meta-narrative reviews" (Wong et al., 2013).

3.3. Evidence of adherence to guiding principles of meta-narrative review

The meta-narrative review currently expresses six guiding principles (Wong et al., 2013):

The principle of pragmatism.

... The reviewer must be guided by what will be most useful to the intended audience(s).

Principle of pluralism

... the topic should be illuminated from multiple angles and perspectives, using the established quality criteria appropriate to each.

Principle of historicity

... research traditions are often best described as they unfolded over time, highlighting significant individual scientists, events and discoveries which shaped the tradition.

Principle of contestation

... conflicting data' from different research traditions should be examined to generate higher order insights.

Principle of reflexivity

... throughout the review, reviewers must continually reflect, individually and as a team, on the emerging findings.

Principle of peer review

... emerging findings should ideally be presented to an external audience and their feedback used to guide further reflection and analysis.

Where it is applicable this thesis adheres to those principles, especially by the principle of contestation, with the claim to state differences of opinion, when deemed scientifical.

3.4. Scoping the Literature

To include research results from several disciplines, it is important to include multidisciplinary databases. In Table 5: Databases for Literature Research, a listing of the chosen databases is provided, with a description of their covered disciplines.

Database	**Disciplines (an excerpt of the most relevant)**
ABI INFORM COLLECTION[12]	• Business • Economic conditions • Corporate strategies • Management theory • Management techniques • Business trends • Competitive landscape and product information • Accounting • Finance
ACM Digital Library[13]	• Artificial Intelligence, Machine Learning, Computer Vision, Natural language processing • Applied Computing: Industry/Business, Physical Sciences, Life Sciences, Education, Law, Forensics, Arts/Humanities, Entertainment • Architecture, Embedded Systems and Electronics, Robotics • Graphics and Computer-Aided Design • Hardware, Power and Energy

[12] https://proquest.libguides.com/abiinformglobal#:~:text=The%20ABI%2FINFORM%20Collection%20contains,of%20them%20in%20full%2Dtext, Accessed 24.04.2023.

[13] https://dl.acm.org, Accessed 24.04.2023.

	• Human Computer Interaction • Information Systems, Search, Information Retrieval, Database Systems, Data Mining, Data Science • Web, Mobile and Multimedia Technologies • Networks and Communications • Software Engineering and Programming Languages • Security and Privacy • Society and the Computing Profession • Computational Theory, Algorithms and Mathematics
Science Direct[14]	• Agricultural and Biological Sciences • Biochemistry, Genetics and Molecular Biology • Business, Management and Accounting • Chemical Engineering • Chemistry • Computer Science • Decision Science • Economics • Energy and Power • Engineering and Technology • Environmental Science • Health Sciences • Pharmacology, Toxicology and Pharmaceutics • Psychology • Social Sciences

[14] https://www.sciencedirect.com/, Accessed 24.04.2023.

Scopus[15]	• Architecture, Civil Engineering and Surveying • Electrical engineering, measurement, and control technology • Energy, Environmental Protection, Nuclear Engineering • History • Computer science • Mechanical Engineering • Mathematics • Medicine • Philosophy • Physics • Psychology • Sociology • Process Engineering, Biotechnology, Food Technology • Economics

Table 5: Databases for Literature Research.

In Addition to literature from these sources a decision has been made on using books on the topic to adapt a narrative style, as well as overarch the various interdisciplinary fields, that give insight to the topic.

The following books have been used:

- The Surge of A.I – The rise of intelligent machines and its implications for the future of humanity, by *Emmanuel Akinnodi.*
- The AI Economy – Work, wealth and welfare in the robotic age, by *Rodger Bootle.*

[15] https://www.scopus.com/search/form.uri?display=basic#basic, Accessed 24.04.2023.

- We, the Robots?: Regulating Artificial Intelligence and the Limits of the Law – *Simon Chesterman.*
- Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies – *Erik Brynjolfsson & Andrew McAfee.*
- Machine, Platform, Crowd: Harnessing Our Digital Future – *Erik Brynjolfsson & Andrew McAfee.*
- Robot Rules: Regulating Artificial Intelligence – *Jacob Turner.*

3.5. Literature Searching Process

For the first round of the systematic gathering of literature the following combination of Boolean Operators have been used:

1. "innovation" AND ("regulation" OR "regulatory framework" OR “regulatory compliance" OR "regulatory barriers")
2. "technology innovation" AND "policy"
3. "regulation" AND "innovation diffusion"
4. "regulatory impact" AND "technological innovation"
5. "regulatory environment" AND "innovation adoption"
6. "innovation policy" AND "regulatory challenges"
7. "regulatory sandbox" AND "innovation“
8. ("regulation" OR "policy") AND ("innovation" OR "technological change")

Because of the high number of results al of the single combinations 1 to 8 have been combined into one search input:

("innovation" OR "technology innovation" OR "innovation diffusion") AND ("AI" OR "Artificial Intelligence") AND ("regulation" OR "regulatory framework" OR “regulatory compliance" OR "regulatory barriers" OR "regulatory impact" OR "regulatory sandbox" OR "regulatory challenges")

By using inverted commas, results can be searched for, which contain the exact term listed in inverted commas anywhere in the article. By using parentheses and Boolean operators, only results that meet all three criteria are returned. All results deal with innovation, AI and

regulation, or a variation of the terms.

To further narrow down the results, additional parameters are used for the advanced search. All results must be written in English and must have been published between the 25.05.2018 and 31.12.2022. The time restriction is due to the implementation of the general data protection regulation (GDPR) in 2018.[16]

Since every database has slightly different setting possibilities, they will be discussed in more detail below.

ABI Inform Collection

In total 24.412 results have been found, with the setting, that the keywords may occur anywhere in the text. After reviewing the first 200 results, it became apparent that most results have no relevance regarding the topic, even decreasing. Therefore, the filter was adapted to only include results having the three criteria in the abstract and/or the summary text. With this adaptation the results have been narrowed down to 34.

ACM Digital Library

In total 1.042 results have been found, with the settings described above. Since no additional filters can be applied, the results are manually screened until the relevance of the output is no longer given. From the initial 100 results only 15 remain after the initial screening.

Science Direct

Since the first two databases have shown to only provide relevant results when the filter is strict enough, already in the first round of the searching process, the three broad criteria, separated by the Boolean operators, must be found in title, abstract or in the keywords. This results in a total number of 43 results.

[16] https://edps.europa.eu/data-protection/data-protection/legislation/history-general-data-protection-regulation_de, Accessed 02.05.2023.

Scopus

Similar to the before mentioned database searches the three broad criteria have been used, resulting in a total of 143 results. Additionally, only open access and English papers have been taken into account. After reviewing title and abstract only 22 remain.

3.6. Selection and appraisal of documents

For the selection the abstracts, summary or discussion had to mention, to some capacity, the impact of regulations and/or policies of AI on innovation or innovation processes. Abstracts that discussed innovation management were also deemed relevant for the purpose of this review, however it was not a strict inclusion criterion. With these inclusion and exclusion criteria, a total of 59 papers are left after reviewing all titles and abstracts. The further sorting out of duplicates leads to a number of 50 papers remaining.

If the papers were not freely accessible, or accessible with the use of the "Technische Universität (TU) Wien Datenbank Infosystem", they have been sorted out as well, which has been the case for another 7 papers.

One other paper turned out to be only a magazine article and one to be a book chapter and therefore had also been sorted out. Resulting in a final number of 41 papers for the initial full text review.

3.7. Data Extraction

To give a quantitative overview of the accumulated resources the following characteristics were of interest:

Contribution

How intensively has the impact of regulation on innovation been considered?

Minor consideration	Publications that only mention the correlation of regulation and innovation briefly.
Major consideration	Publications that address the correlation of regulation and innovation.

Minor contribution	Publications that add a new aspect, or further emphasize known aspects to the regulation of technologies were classified as a minor contribution. Such could be an empirical study, which is not solely about the impact of regulation on innovation.
Major contribution	Publications that add important new aspects to the impact of regulation on innovation. These can be conceptual papers, but also reviews that fill a research gap or empirical studies that focus on the correlation of regulation and innovation.

Analytical approach

Wheater the work was of a conceptual or empirical nature.

Conceptual	Conceptual research includes thoughts and ideas. Publications that discuss existing literature or raise new questions have been deemed conceptual.
Empirical	Empirical research includes the phenomena, that are observable and can be measured. If research have both, conceptual and empirical parts, it has been deemed empirical.

Total citation count

The total number of citations according to google scholar until 30.05.2023.

Name of first author

Name of the first author. Due to simplification the names of the other contribution authors have not been taken into account.

Place of origin

The country in which the paper was published.

Year of publication

The year in which the paper was published.

Discipline

The academic discipline of the first author of the publication. When the author had more than one discipline the first mentioned one was considered.

3.8. Analysis and synthesis process

Within the analyzed papers the references that gave insight regarding the research question have been further examined and their contributions have been added to the main findings, adhering to the principle of pragmatism mentioned in chapter 3.3.

Because of the high count of different disciplines and therefore different approaches, similar topics have been combined in several paragraphs. Thus, trying to give an overarching narrative regarding the topic of innovation regulation and its implications, adhering to the principle of pluralism. Especially the interplay of tensions regarding different theories, i.e., the precautionary principle was mapped in direct succession, adhering to the principle of contestation. In addition, whenever possible, a chronological sequence of the theories that have emerged has been followed, adhering to the principle of historicity.

4. Results

4.1. Document Flow Diagram

Figure 4: Document Flow Diagram, adapted from Xiao & Watson shows the procedure of the identification, screening and inclusion of papers. The process is described in more detail in chapter 3.5 Literature Searching Process.

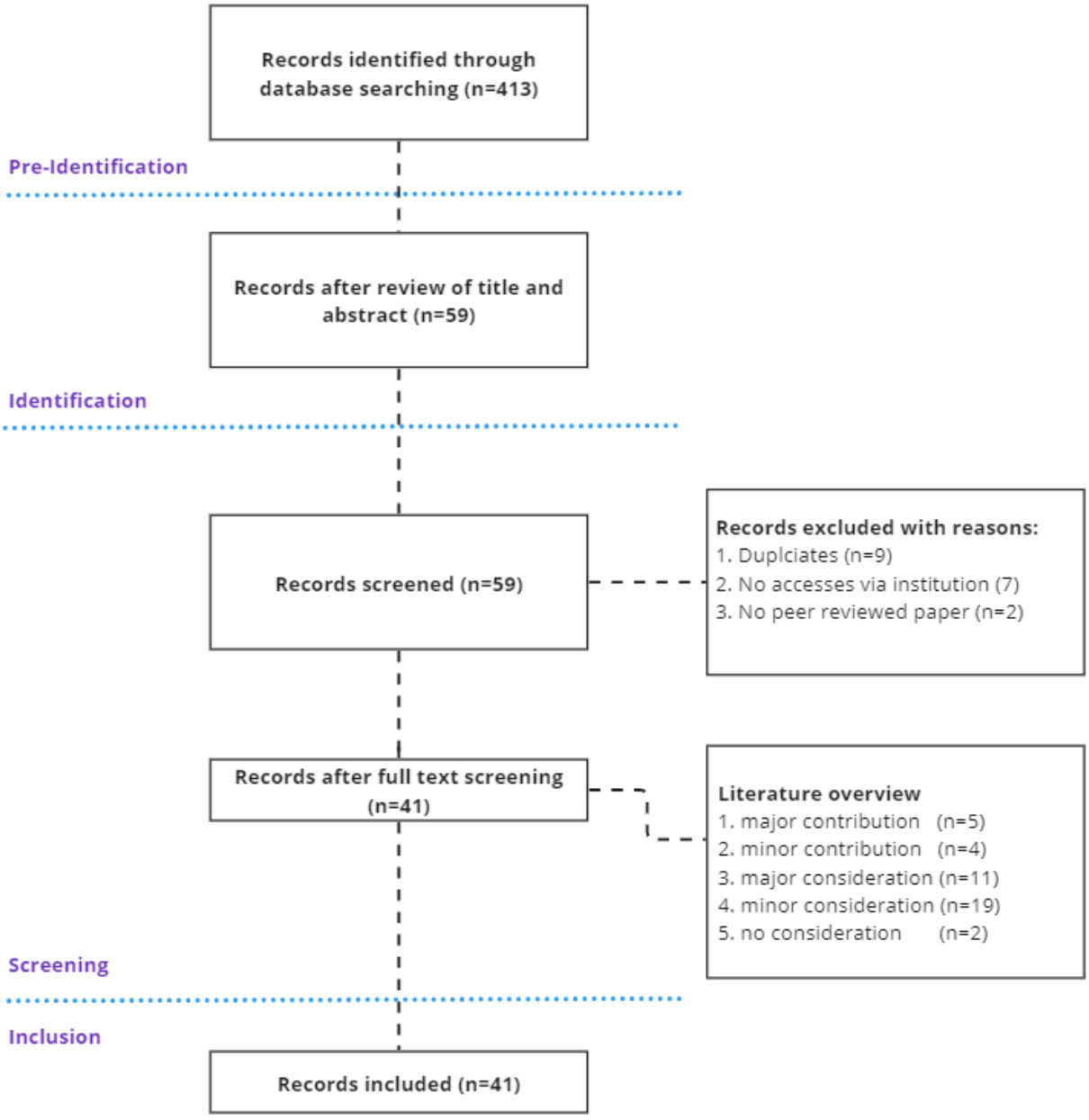

Figure 4: Document Flow Diagram, adapted from Xiao & Watson (2019).

4.2. Document Characteristics

The included papers from Figure 4: Document Flow Diagram, adapted from Xiao & Watson, have been worked through and the characteristics described in more detail in Chapter 3.7, as well as the initial source of the paper described in chapter 3.4, were assigned and can be seen in Table 6 to 13.

Database Distribution

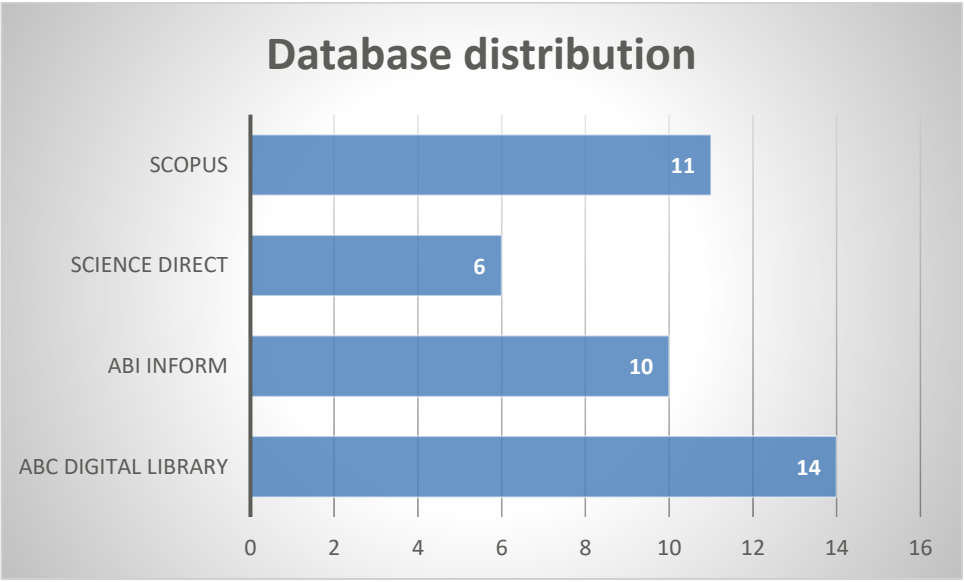

Table 6: Database distribution.

The database distribution does not show any significant differences in contributions of databases. The highest count is from the database ABC Digital Library with 14 (34,15%) and the lowest count is from Science Direct with 6 (15,63%). Major considerations and contributions, as seen in Table 7: Contribution correlated with database, can also be found throughout all the databases.

Contribution	Database	Count for each database
major contribution	ABI Inform	2
major contribution	Science Direct	2
major contribution	Scopus	1
major consideration	ABC Digital Library	2
major consideration	ABI Inform	3
major consideration	Science Direct	2
major consideration	Scopus	4

Table 7: Contribution correlated with database.

Contribution

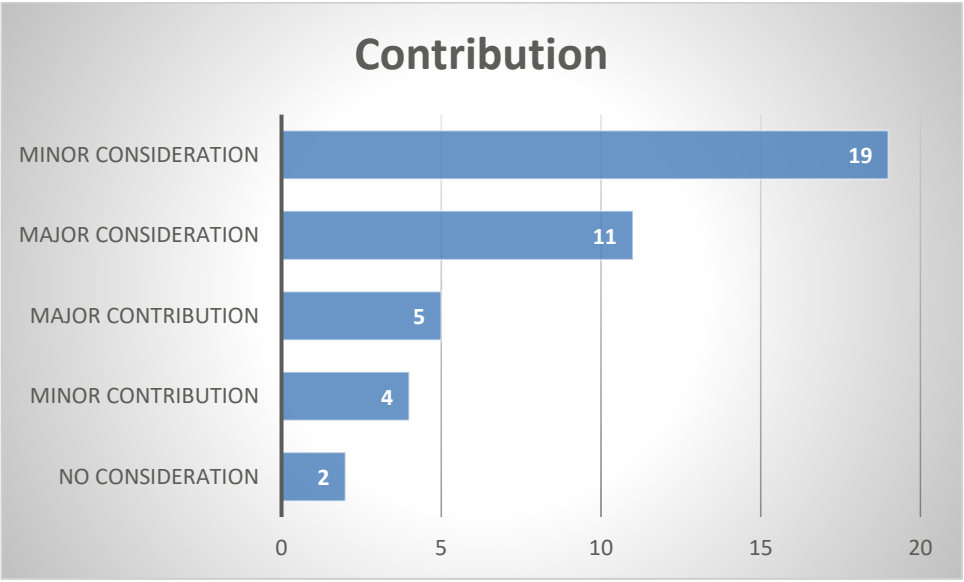

Table 8: Contribution.

Of the examined papers, 19 (46.34%) have only a minor consideration of the effect on regulation on innovation, mostly mentioning them in the abstract without going into detail within the paper. The second highest count are 11 major contributions (26,83%), where at least one correlation between regulation or governance in any with innovation is addressed, without gaining new insights or filling a research gap. 4 (9,76%) papers do add a minor contribution, but are not solely about the impact of regulation on innovation and in a niche area, such as data access challenges in the interface of AI Companies and Hospitals (Kemppainen et al., 2019).

Of special interest are the 5 (12,2%) papers which have a major contribution for the regulation impact on innovation. They are both of conceptual (3) and empirical (2) nature.

Place of Origin

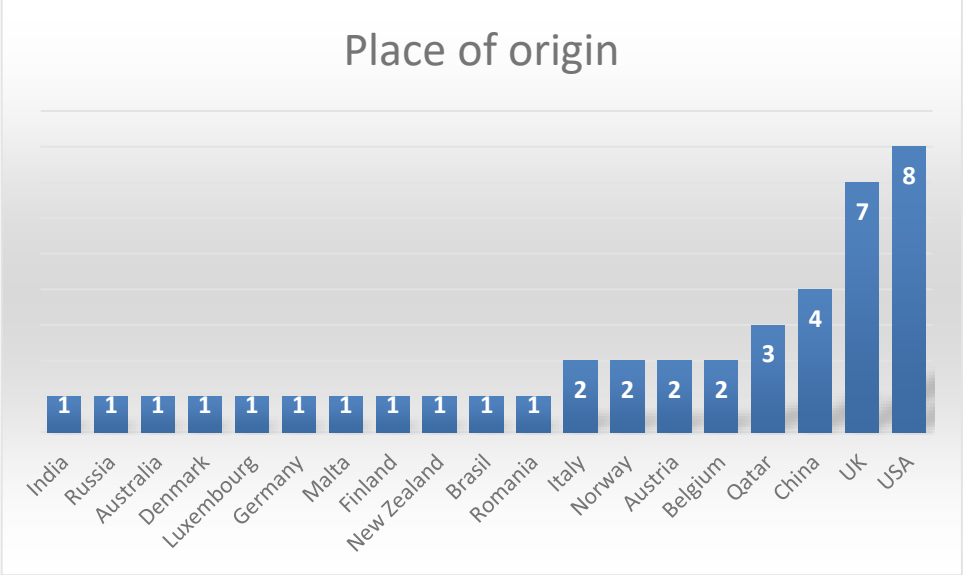

Table 9: Place of Origin

The place of origin shows most interestingly that the nations which are regarded as leading nations in the development of AI appear to have a higher count of publications. It should also be noted that all publications from Qatar are published by only one author, John Truby, who completed his PhD in Newcastle, England.[17]

[17] http://qufaculty.qu.edu.qa/jon-truby/dr-jon-truby/, Accessed 17.08.2023.

Year of publication

As seen in Table 10: Year of the publication and Table 11: Year of publication count, there has been steady number of publications from 2019 to 2021, with just one publication lower in 2020. In the year 2022 the number then more than doubles from 7 to 17 publications.

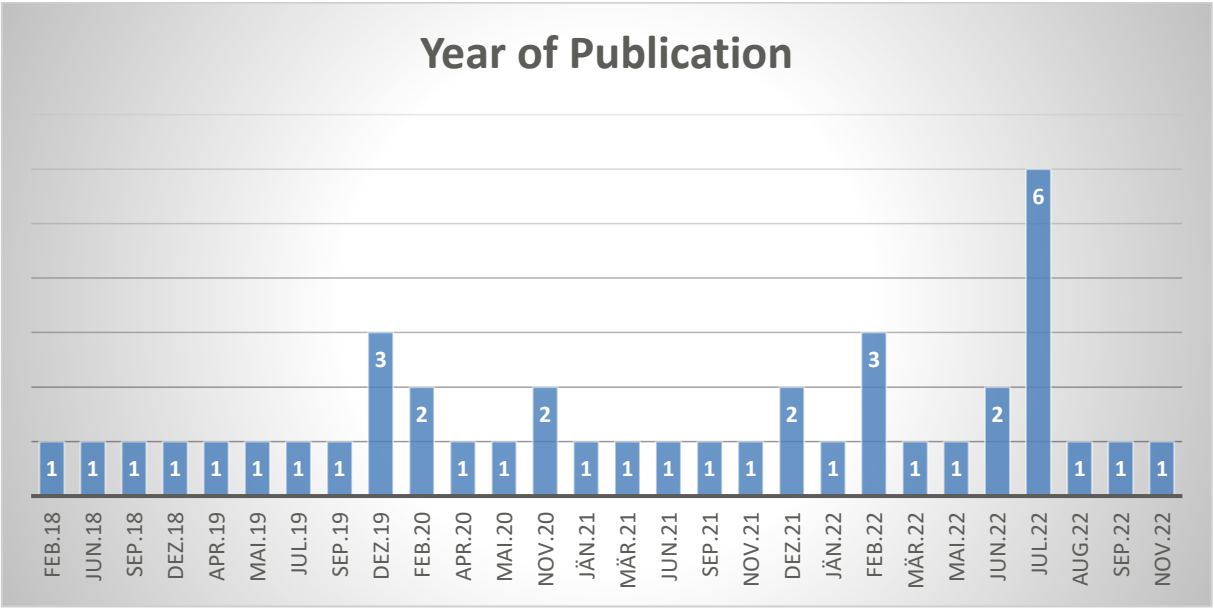

Table 10: Year of publication.

Year of Publication	Count of Contributions
2018	4
2019	7
2020	6
2021	7
2022	17

Table 11: Year of publication count.

Disciplines

In total there have been 21 different disciplines in the field of law (10), economy (4) and computer science (4) are the most publications with a total of 18 (43,9%) out of 41. Since regulation, with the exemption of soft regulation (which is described further in the main findings) has to be done by a legislative entity, this was to be expected. However, the high number of different disciplines is an indicator on how interdisciplinary the topic of regulating AI and its impact on innovation is.

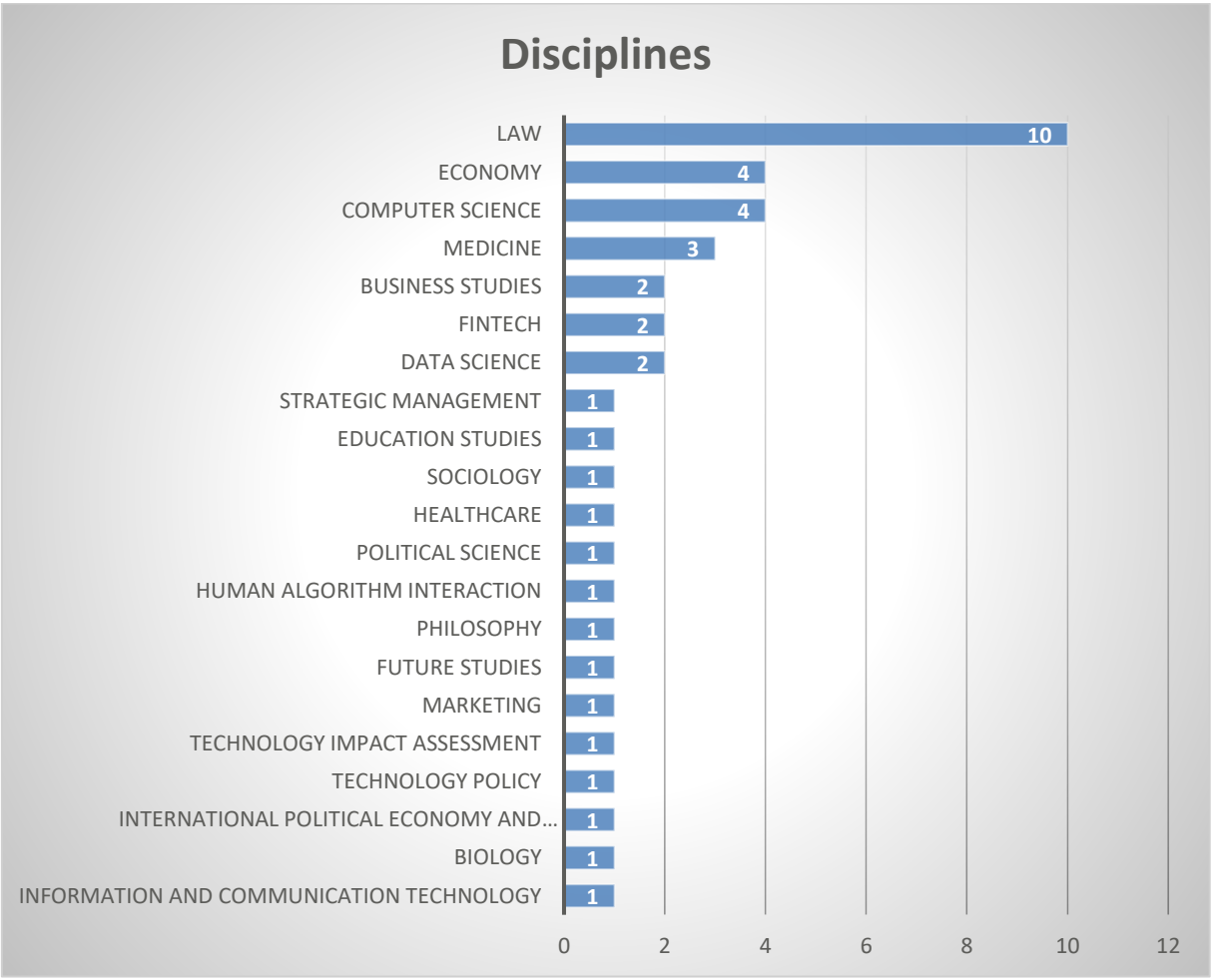

Table 12: Disciplines.

Analytical Approach

The analytical approach is more than twice as often (29 compared to 12 times) of conceptual nature. This circumstance is also not surprising, as the effects are difficult to study, especially since effects can usually only be researched after regulation has been implemented. Table 13: Analytical approach combined with Contribution, shows the distribution of the analytical approaches, combined with the contribution of the respective papers.

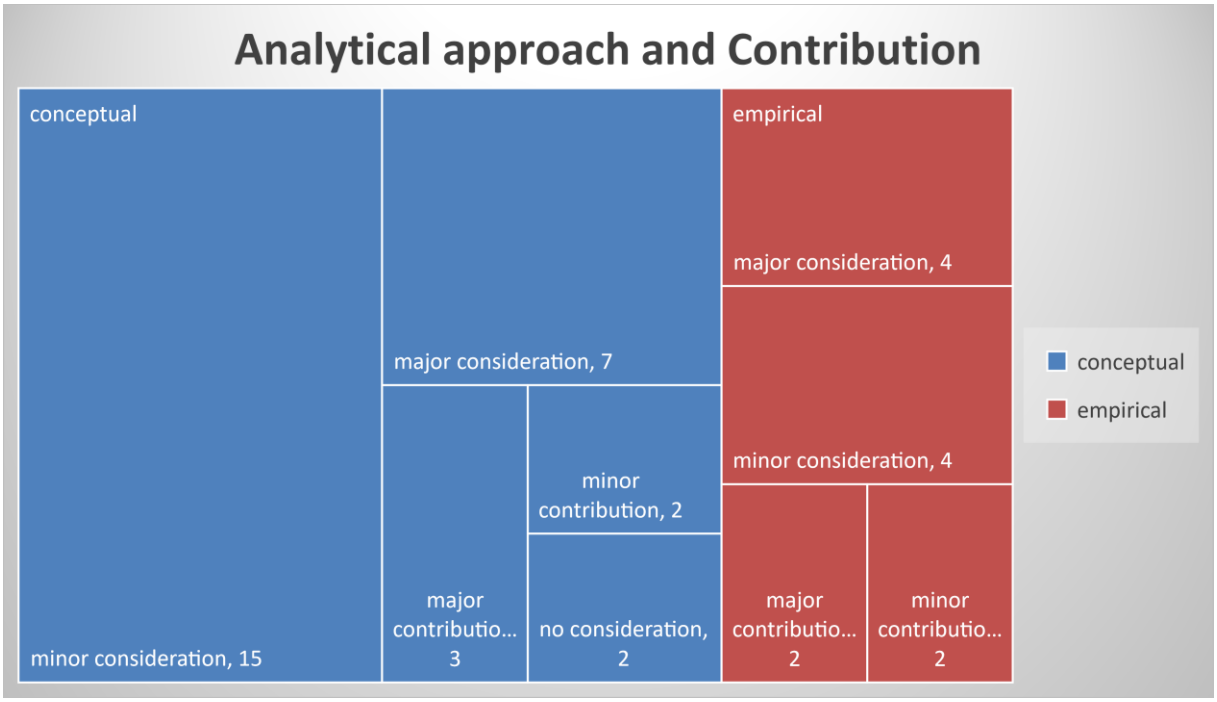

Table 13: Analytical approach combined with Contribution.

4.3. Main findings

In order to structure the main findings in a comprehensive way, they have been divided into different dimensions of regulatory impact, after an excursion into the definition dilemma of AI. The first dimension to be considered deals with the fundamental nature of regulation and whether it promotes or inhibits innovation, with special focus on GPT, liability, startups and the transnational handling. The second dimension brings the factor time into play and provides information about the effects of regulation in the medium and long term. The third dimension examines the different approaches of regulation. The fourth and final dimension deals with the interplay of trust in AI and regulation.

Why bother regulating? The general conception of the public seems to be, that they are convinced of the importance of AI regulation, which can be observed in Figure 12: Perceptions of AI governance challenges in the U.S. and around the world. Already the AI governance challenge with the lowest score (Critical AI systems failure), scores on a Likert scale of 0 to 3 a value of 2,475 (3= very important, 0= Not at all important), which is already in the highest quarter of the scale. This is furthermore underlined by the high account of news articles regarding regulation of AI.

When we are talking about regulating AI, it should first be clarified what is meant by AI. Definitions of AI do come in various different forms and each single definition slightly differs from the latter. AI is not new, as already discussed in the theoretical foundations of AI. Appropriate definitions of AI today, may range from and are dependent on an applications point of view (chatbot, image recognition) but also to subfields in academia (computer science, or engineering) such as reasoning (logic), learning (neural networks) autonomous behaviour (driving and robotics) and much more (Prem, 2019). The European Commission currently defines AI based on *the objective to create human like behaviour in machines for perception, reasoning, and action* (AI for Europe, 2018). However, with this definition arises a new

problem. Namely, what is human-like behaviour? A cynic may argue that making mistakes is as human like as it may get. These unclarities within definitions do complicate the goal of giving a precise working definition. Especially in the context of law, the definition of AI is often far too short sighted and can be too openly worded, making it possible to find loopholes, or it can be too narrow, therefore not covering enough legal ground. An attempt to define AI, suitable for use in the legislative field is provided by Jacob Turner in his book Robot Rules - Regulating Artificial Intelligence: "*Artificial Intelligence Is the Ability of a Non-natural Entity to Make Choices by an Evaluative Process"* (Turner, 2019). This definition avoids the use of words, which in turn provide a scope which could be interpreted differently in the legal context themselves.

The Organization for Economic Cooperation and Development (OECD) provides an even more precise definition than the EC, or Turner, whilst also offers little room for interpretation: "*An AI system is a machine-based system that can, for a given set of human-defined objectives, make predictions, recommendations, or decisions influencing real or virtual environments*."[18]

[18] Recommendation of the Council on Artificial Intelligence, OECD/ LEGAL/0449, Adopted on: May 22, 2019, https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449. Accessed 04.07.2023.

4.3.1. Regulation's Dual Role: Impeding and Encouraging Innovation

The difference between new technology and general-purpose technology

Innovation will be majorly driven by smart technologies in the upcoming years (Lee & Trimi, 2016) and many scholars, Agraval amongst them see and praise the innovation potential of AI (Agrawal et al., 2019). This insight in combination with the acknowledgement of AI as GPT paints a clear picture of the innovation potential for AI, compared to regular new technology. Scholars argue that in contrast to any new technology, GPT´s are pervasive, improving over time and have to be able to spawn new innovations (Jovanovic & Rousseau, 2005). These criteria are all met by AI. Generally, GPT´s need several years to develop their full potential, leading to a high number of innovation by spawning complementary innovations (Clarke & Whittlestone, 2022) and increasing productivity (Brynjolfsson & McAfee, 2014) over time. Therefore, we can expect to soon enter this phase of complementary innovation. The nature of those complementary innovations highly depends on the regulatory framework that is set at the time of their discovery. By setting a strict regulatory corset, innovation will be hampered on the one hand, but also be driven towards the desired direction.

As AI can not only be considered as a GPT but also a disruptive innovation it offers the chance for evolutionary change if regulated properly. This is only possible if the implemented regulations do not only rely on past mechanisms but seek to evolve with AI. Finding the right balance between essential control and pointless innovation restraints.

Liability's Dual Impact on Innovation: Enabler and Deterrent

Liability and its interpretation in the eyes of the law is something that often comes up in connection with innovation and regulation. Clearly, the liability issue does not affect every sector or every innovation to the same extent. AI has the potential to be used in all possible sectors, as described in earlier chapters. Especially in areas like medicine, or those that have a large impact on the environment, a lot of hope is placed on AI. However, especially in exactly those areas' innovations are hampered by unclear liability regulations. Thus, making liability law one of the important pillars, regarding the impact regulation of AI has on innovation.

If not in an already heavily regulated sector like pharmaceuticals and the financial system, generally any innovation is possible and the person who brings this innovation into circulation also bears the responsibility in the eyes of the law. In a study carried out for their own purposes, the EU Parliament came to the conclusion that a lack of civil liability at European level would: "*[…] potentially discourage innovation, increase prices for consumers, substantially increase administrative costs for public administrations and judicial bodies and ultimately even challenge the social desirability of the overall liability system* (Evas, 2020)*"*. There are several papers underlining the importance of a precise liability regulation in the EU (Commission Report on Safety and Liability Implications of AI, the Internet of Things and Robotics, 2020; White Paper on Artificial Intelligence, 2020). Making it therefore very important to develop a liability framework which promotes innovation. Currently the EP favors the strict liability approach, especially for high-risk AI systems, which will be described shortly. First, some issues regarding the complexity of assigning liability are given.

The first problem for the liability of AI is, that AI can take on various forms. From a component within a product or larger system, as well as a stand-alone entity in unlimited environments and purposes. Secondly, again, the criticism of the unclear definition resurfaces. Furthermore, there is still no consensus how to deal with the liability question regarding AI when facing

phenomenon like the black-box problem. The problem hereby is, as Truby (et. al. 2022) argues, that opaque neural networks, whose inputs and operations are visible neither to the user nor to any other interested party are impenetrable. Making it impossible for intent and causation tests. These intent and causation tests, which currently are used to assess what is foreseeable, are ineffective when it comes to black-box AI (Bathaee, 2018). However, it is crucial that the law also makes a distinction that something might be a black box problem for some, but no for all people, for example the developer of an AI system and its user (Reed, 2018). This is important because the reasonable knowledge at the time of application, or liability event plays a role in many current legislations.

There are currently two dominant approaches when it comes to liability, namely the fault-based liability -manly applied in the USA- and the strict liability -manly applied in the European Union and across the Atlantic (Truby et al., 2022).

In essence the fault-based approach, also called negligence liability, is based on the idea, that if someone is careless and causes harm to others, they should be held responsible depending on their level of carelessness. The level of carelessness is thereby set by a judiciary body, which should also take the potential social benefit of a technology into account. Hence officials often cannot have the knowledge depth of those developing new technologies, private companies' knowledge should therefore be considered when determining what risks are acceptable (Zech, 2021). In the case of technologies for which possible negative effects are already known, the fault-based approach incentivizes developers to take more precautions to prevent possible harm. In case of new technology this incentive may not apply. Zech criticizes, that when there are no options for avoidance for the plaintiff, as in the case with ubiquitous systems, and the defendant acted within their duties, the defendant is not liable because of the general risk of living. In the case of development risks which are yet unknowable based on the state of scientific and technical knowledge, fault-based liability fails. *There is no duty to avoid the unknowable* (Zech,

2021). As a result, fault-based liability in many new technologies leads to fostering technologies at the expense of those affected. This ultimately results in the lack of trust in new technologies (2021).

Summarized, the fault-based approach has several critiques, firstly, there appears to be an information asymmetry between courts and producers. Business insider, developer, and producers of AI systems within the industry will have a much deeper understanding of these systems. Secondly, new technologies with unforeseeable risks for the future would just not meet the requirements of legal causation. And lastly, leads to a weakening of trust in new technologies.

The alternative is strict liability, the approach mentioned earlier as favored by the EP. However, this approach also has some down sights. Strict liability means that, the person who caused the damage must compensate those who suffer losses or damage, regardless of whether the person causing the harm was careless during the event (Reed, 2018). Strict liability provides a high degree of legal certainty because it doesn't hinge on proving the defendant's fault. Additionally, there is the strict product liability -which is reasonable similar throughout EU Member States- which declares that the manufacturer or the seller of any product is liable for any harm caused by any defect of that product. If now for example an AI system causes harm, the manufacturer of this technology would be liable for the incurred damage. Truby argues, that the strict liability approach potentially has a chilling effect when it comes to innovation, especially for Startups and small and medium enterprises (SME), due to the high compliance costs and barriers to entry such a strict liability regime and the potential risk of being punished regardless of the conduct (Truby et al., 2022). A historic example for hampering of innovation through strict liability, are the early ages of the Internet. At first internet service providers were held responsible for anything users have posted on the internet. Consequently there have been wide-ranging liability immunities to prevent a stifle in innovation (Lemley, 2007). Moreover, in most of the liability

case's fault does play an important role. Reed points out that even if the cause can be identified, allocating the responsibility for this cause can get exceptionally hard in case of AI technology (2018).

In summary, the key difference between fault-based liability and strict liability is the requirement to prove fault or negligence. Fault-based liability depends on demonstrating that the defendant acted wrongfully or negligently, while strict liability holds the defendant accountable regardless of fault. Strict liability provides more legal certainty and can promote safety in product manufacturing and distribution.

In a study conducted by Prem, experts from the AI sector have been asked about the main barriers in the AI sector. An excerpt of his findings can be seen in Figure 15: Expert interviews, barriers and challenges. The interviewed experts mentioned the unclear responsibilities for overall systems behavior as one of the challenges, which is in line with the findings from this chapter. Many of the provided examples are also addressed by suggestions for the regulation AI by the ITU and the IGF mentioned later on in the thesis.

According to Reed (2018), transparency could go a long way in resolving the liability issue, making it easier to allocate misconduct. But transparency is yet another term that has to be defined properly. Zarsky already demonstrated that transparency can take a range of meanings, as well as being justified by different reasons (Zarsky, 2013). Furthermore, a distinction between *ex ante* and *post ante* has to be made. It is simply much easier to justify a decision made by an AI afterwards than to evaluate the decision beforehand. And another problem, especially with increasingly complex neural networks, is the amount of data that has led to the decision. Above a certain level of complexity, it is simply no longer possible for humans to fully comprehend a decision. In his paper, "How the machine ´thinks: understanding opacity in machine learning algorithms", Burrell puts it that way:

> *"[...] opacity that stems from the mismatch between mathematical optimization in highdimensionality characteristic of machine learning and the demands of human-scale reasoning and styles of semantic interpretation* (Burrell, 2016)."

Reed suggests that the following transparency framework for AI (2018):

- Complete lack of any transparency
    - Acceptable if the society benefits overall and the loss to individuals could be compensated (i.e., malfunctioning central heating system).
- Ex post transparency
    - Acceptable if the society benefits overall and the loss to individuals is legally compensable (i.e., personal injury caused by car accidents).
- Ex ante transparency
    - Firstly, where the AI imposes the risk of a breach of a fundamental right.
    - Secondly, where an implementation of an AI technology would not be accepted by society at large, without explaining the decision-making process. (i.e., trolley problem).

Startups and Monopolies: Competing Forces in Innovation

Already in 1934 Schumpeter argued that the likelihood of increasing innovation is higher in upstarts than in incumbent companies (Schumpeter, 1934). He argues that: *"New combinations are*," - which are the basis for innovation – "*...as a rule, embodied ... in firms which generally do not arise out of old ones. ... It is not the owner of a stagecoach who builds railways*.". Which in turn implies that the promotion of start-ups, or not over-regulating them, will lead to an increase in the number of innovations. Arrow investigated existing businesses, compared to Schumpeter who investigated upstarts and has expressed the position that prevailing businesses lack the potential of new ones. Arrows school of thought proposes the thesis, that monopolists have a lower incentive to innovate (Arrow, 1972), because entrepreneurs that are still competing to gain momentum in the market have to innovate, to establish themselves as the new monopolist, whilst the established monopolist could only replace itself. Another scientist, Erich Prem conducted a more recent empirical study to investigate possible strategies to support measures for AI based innovation. Prem concludes, that "*An opinion prevails among those interviewed that startups have a vital role to play in both the application and deployment of AI innovations. They are considered the main leaders and competence carriers in AI technology and are praised for their flexibility compared to large industry actors* (Prem, 2019)." Which further underpins the importance of start-ups and the regulation thereof.

These theories do not stand without contestation. They are contradicted by the mechanization of innovation by larger companies, by the later work of Schumpeter, arguing that larger companies can mechanize innovation in their research departments (Michael Filzmoser, 2022W). Empirical studies have explored the relationship between market concentration and innovation, coming to the conclusion that a U-Shape is the most likely truth (Aghion et al., 2005; Hashmi, 2013). This suggests that there are truths to be found in both schools of thought. Either way, recent empirical studies investigating the number of start-up foundations conclude,

that more start-ups settle in regions that are less heavily regulated (Karim, 2022). In a European study researchers Klapper, Laeven and Rajan have additionally strengthened this theory, that higher regulation leads to a reduced startup activity (Klapper et al., 2006). This could be influenced by the fact that start-ups do not have the same scope or power as established large companies with their own lobby groups. Which makes dealing with a high number of regulations much more difficult and carries a financial risk. This correlates with the often so called "pebble in the stream effect", coined by economist Michael Mandel. Mandel suggests that it is not all about the nature of the regulation, but rather that the number of regulations accumulate to a large number of pebbles in a stream, ultimately slowing it down (Gold, 2017; Mandel & Carew, 2013). Furthermore, there is the interplay between multiple regulations that interact in obvious and non-obvious ways, raising costs for businesses. And the last regulatory accumulation effect is the behavioural overload. Meaning that, if the number of regulations is getting too high, businesses are forced to focus on the prioritization of those, rather than focusing on growth and innovation (Regulatory Reduction Efforts in Ohio, 2022). These three types of regulatory accumulation effects, pebbles in a stream, interaction between regulations and behavioural overload, contribute to the innovation-inhibiting nature of regulation.

However, there are also contradicting studies, which suggest that if regulation and legislation are created specifically to support startups and SME, in particular to simplify the process for requesting support and to effective determinants in facilitating the growth of startups and SME (Wisuttisak, 2020), regulation shows to have a positive effect on the number of new foundations. A further perspective contrary to the regulatory accumulation effect, advocated by Alexander Hilton (2019) is, that only bad practice and bad policy stifles innovation, not just the number of regulations. In a great many cases this *bad regulation* derives due to misinformation of policymakers and mistrust of the public regarding AI, which is addressed later on in this thesis.

The former mentioned effects show that regulation has an, at least indirect, effect on the innovation potential of companies. The regulatory accumulation is more difficult for smaller companies to deal with because of the resources they must invest. It inevitably follows that competition law also has a significant influence on innovation. At the moment there are some technology companies that have an overwhelming share of AI related business, becoming, or already being digital super companies, examples are Alibaba Group, Alphabet, Amazon, Microsoft and OpenAI[19]. Subhashish Gupta argues that a problem of the current competition law is that it is still based on the "old" economy. Companies from the "new economy" are much more dynamic. With the rise of Internet and Communication Technology (ICT), AI and the Internet of things (IoT) amongst others, there should be a rethinking of competition law (Gupta, 2022). Over enforcing could lead to stopping innovation and underenforcing could lead to an unstoppable accumulation of power of monopolies. Scholars agree that new and innovative ways of regulations need to be found and as important is, that regulatory bodies are not too wary of old ways of regulation. However, apart from investigating carefully whether the acquisition of smaller companies is purely for the purpose of eliminating competition, concrete suggestions are still missing (Shapiro, 2019; Wu, 2018).

Related to innovation and creative destruction, there is also a example of a new business model that is slowly emerging. Namely AI as a Service (AIaaS). Interestingly this new business model takes over, respectively arises from classical business consulting. Many companies realized, in order to stay on top, they need to incorporate AI services, including newer firms like "Accilium

[19] https://www.suedkurier.de/ueberregional/rundblick/kuenstliche-intelligenz-ki-unternehmen-technologie-globaler-wettstreit;art1373253,11428840, Accessed 09.08.2023.

GmbH" and old-established firms like "Boston Consulting Group" and "IBM".[20] Prem argues that within AIaaS a new AI profession seems to be arising, namely a "AI trainer". An AI trainer needs to be an expert in the computer application domain with competencies in data analytics, whereas the former is the more important domain (2019). Most interestingly, there is evidence that startups that work with big technology firms show more ethical behavior (Bessen et al., 2022). Bessen explains this phenomenon by arguing that big established firms often have created a rulebook, a code of conduct or a similar guiding principle, because there is a lack of governance and regulation from legislative bodies. These guidelines are then adopted from startups. An empirical study by Winecoff (2022) showed, that in case of medical device startups in the United States, onerous Food and Drug Administration (FDA) approval processes are stifling innovation. However, the same startups argue that such regulations were necessary to protect users from harmful products. Privacy regulations are perceived with aligning with important values such as personal freedom and autonomy, whilst FDA regulations were seen as a barrier to innovation and entrepreneurial autonomy (Winecoff & Watkins, 2022).

[20] https://www.forbes.com/sites/bernardmarr/2022/08/17/what-are-the-10-best-ai-consulting-firms/, Accessed 26.06.2023.

Transnational nature of AI regulation

Many, if not most nations have by now concluded that the impact AI will have on the economy and thus to some degree the pecking order amongst them is tremendous. The same also applies for the companies operating within and amongst them. This acceptance of the fourth upcoming industrial revolution, as Brynjolfsson (2014) and Schwab[21] call the emerging of AI technologies, has led to an immense flow of money into AI connected sectors and an effort to be on top of the wave when it comes to AI. China hopes to facilitate a 150-billion-dollar industry by 2023[22], but also most European Nations push to be on top of the AI playing field in the future[23]. Therefore, the conventions interacting with each other must also be rewritten.

There are already some examples of international applied laws and guidelines, one of the most prominent ones is the GDPR. With the implementation of the GDPR in May 2018 the European Union has set a high standard when it comes to the data protection of their citizens. One of the intentions the policymakers had, was to ensure, that scientific research is not hindered by the implementation of the GDPR. Thus, fostering innovation by stating exemptions especially for research. The GDPR therefore lists three categories that result in an exemption for data processing for scientific research:

1) Exemptions from data processing principles and lawful grounds for processing.

2) Exemptions from the data subjects' rights.

3) The Member States can implement further research exemptions.

With these very broad exemptions for scientific research, the question needs to be answered, what exactly is scientific research? According to the GDPR scientific research is defined very

[21] https://www.weforum.org/about/the-fourth-industrial-revolution-by-klaus-schwab, Accessed 04.07.2023.

[22] https://www.gov.cn/zhengce/content/2017-07/20/content_5211996.htm, Accessed 04.07.2023.

[23] https://medium.com/politics-ai/an-overview-of-national-ai-strategies-2a70ec6edfd, Accessed 04.07.2023.

broadly as:

"[…] processing of personal data for scientific research purposes should be interpreted in a broad manner, including, for example, technological development and demonstration, fundamental research, applied research and privately funded research."[24]

An important remark is, this is only found in the recital of the GDPR, thus not legally binding. Meszaros argues that when it comes to restricting access to data for research purposes, there should be distinguished between basic or fundamental research and applied research. Basic research can be of experimental or theoretical nature and purses the acquisition of new insights (Meszaros & Ho, 2021). Applied research on the other hand is according to the EC, "*the planned research or critical investigation aimed at the acquisition of new knowledge and skills for developing new products, processes or services, or for bringing about a significant improvement in existing products, processes or services* (EC - Community Framework, 2006)". However, both basic and applied research can be allocated to academic and commercial research, as seen in Figure 5: The comparison of academic and commercial research. Unfortunately, the GDPR does not make this distinction for its exemptions.

	Academic research	Commercial research
Focus	Basic and applied research	Applied research
Basic rationale	Advance knowledge	Increase efficiency
Aim	New ideas	Profit
Characteristics	Idea-centred	Practical, product-centred
Schedule	Open-ended, longer periods	Tight, predetermined
Recognition	Scientific honours	Payment
Framework	Open	Close, confidential
Evaluation	Peer-review	By the leaders
Dissemination	Through academic publishing (e.g., journals, books)	Internal reports, professional conferences
Oversight	Rigorous institutional safeguards (e.g., ethics committee)	Less strict overview

Figure 5: The comparison of academic and commercial research (Meszaros & Ho, 2021).

[24] https://www.privacy-regulation.eu/en/recital-159-GDPR.htm, Accessed 21.08.2023.

Summarized, the aim of the GDPR regarding data protection and research was, that scientific findings and innovation are not hindered by the introduction of the GDPR. Problematically, however, no distinction has been made as to the origin of these findings, namely whether they are of an academic or commercial nature. The problem is that commercial research does not usually meet the same ethical standards and requirements.

An example of a blurring of boundaries regarding research is the Facebook – Cambridge Analytica case, where academic research results, have been converted into unethical tools, with the aim of manipulating people (*ICO*, 2018). Yet another example is the Google DeepMind Case from 2016, when the Royal Free London National Health Service Foundation Trust ("Royal Free") uploaded 1.6 million records of medical patients data, for the development of an AI diagnosis application, without informing the subjects, nor de-identifying the data (Powles & Hodson, 2017).

Meszaros (2021) suggests three measures to protect data whilst not slowing down innovation:

> *(1) The harmonised application of GDPR research exemption on AI research in the EU.*
>
> *(2) Commercial AI research should not benefit from the GDPR research exemption without public interest and similar safeguards as academic research.*
>
> *(3) Oversight and enforcement by the EU and Member State authorities (e.g., DPAs and related authorities responsible for scientific research) from the start of AI research until the application of final products and services.*

With the case of the GDPR in mind, regulation on AI should be international applicable and strive to level the AI playing field worldwide. This international approach is supported by many scholars who also emphasize the importance of an international, worldwide regulatory framework for coordinating and streamlining any regulation efforts and can be applied across sectors (Cihon et al., 2020; Erdélyi & Goldsmith, 2018).

Banking and finance will be, and to some extend are already an industry sector that is heavily impacted by AI as well as being transnationally connected. Current areas that face the need of regulation in the financial sector combined with AI, range from third party vendor management, data ownership, privacy ownership rights and cybersecurity.[25] With this wide range of application areas the variety of AI powered applications is obvious. Examples are algorithmic investigation, insurance, and credit assessment, as well as robo-advising. Currently, Truby classifies financial regulation as complex, imperfect and often reactive. A big part of this complexity is the interconnection of technologies. Regulators must examine the multiple points of contact between financial services and technological innovation, whereas the biggest three are: AI, Big Data and cybersecurity (Truby et al., 2020). The possibly greatest potential for innovation in the banking sector through AI lies in the area of advisory services, including the ability to deliver valuable, tailored and informed financial services to customers, thus also accommodate the biggest risk (Zetzsche et al., 2020). Truby argues that an unregulated approach promotes a Wild West environment. Meaning that there is a prevailing climate of high uncertainty and high risk. Therefore, AI regulation in the finance sector must adhere to the core principles transparency, accountability, data protection and privacy.

[25] https://emerj.com/ai-sector-overviews/ai-cybersecurity-finance-current-applications/, Accessed 28.06.2023.

4.3.2. Temporal Dynamics of Regulatory Influence on Innovation

Regulations implemented now, would influence how complementary innovations look like in the future and therefore co-influence the effect they may have on innovation. When Clark et al. surveyed the potential impact of AI, they specifically focused on the long term impact AI could have. Long term impact is described as *"[...] impacts that would be felt not only by our generation or the next, but by many future generations who could come after us* (Clarke & Whittlestone, 2022)". Within this survey a rather dark picture is painted, potentially increasing international conflicts, and further increasing the unequal distribution of power. But they also argue that AI has the potential for enabling a flourishing future for humanity if the right governance choices and strategies are employed.

Timely regulation is even more integral when it comes to the regulation of GPT´s and disruptive innovations as already mentioned earlier, due to the unforeseeable number of complementary innovations. Simultaneously, AI increases the pace of innovation, making the regulation of thereby incurred technologies much more difficult. This outpacement refers to another phenomenon, the pacing problem coined by Larry Downes in his book "The laws of Disruption". Not only is the technology adaption and its consequences difficult to predict, an additional problem is that technology changes exponentially, but social, economic, and legal systems change incrementally (Downes, 2009).

Another well-known problem when it comes to regulating new technologies is the Collingridge dilemma. The thought experiment was presented in the book "*The Social Control of Technology*" in 1980. Collingridge argues that a major problem when it comes to regulating technologies is the dual challenge of information and time. Any technology's ability to be controlled decreases as its implications are realized during deve lopment. Or as Collingridge stated: "*When change is easy, the need for it cannot be foreseen; when the need for change is apparent, change has become expensive, difficult, and time-*

consuming (Collingridge, 1980)". It is very difficult to hit the sweet spot of the ideal area of intervention. This argument is visualized in Figure 6: The Collingridge Dilemma, Control vs. Predictability. Notably, Collingridge himself argued for an approach to design technology in a way that its flexible or reversible, rather than to anticipate risks (p. 23-43).

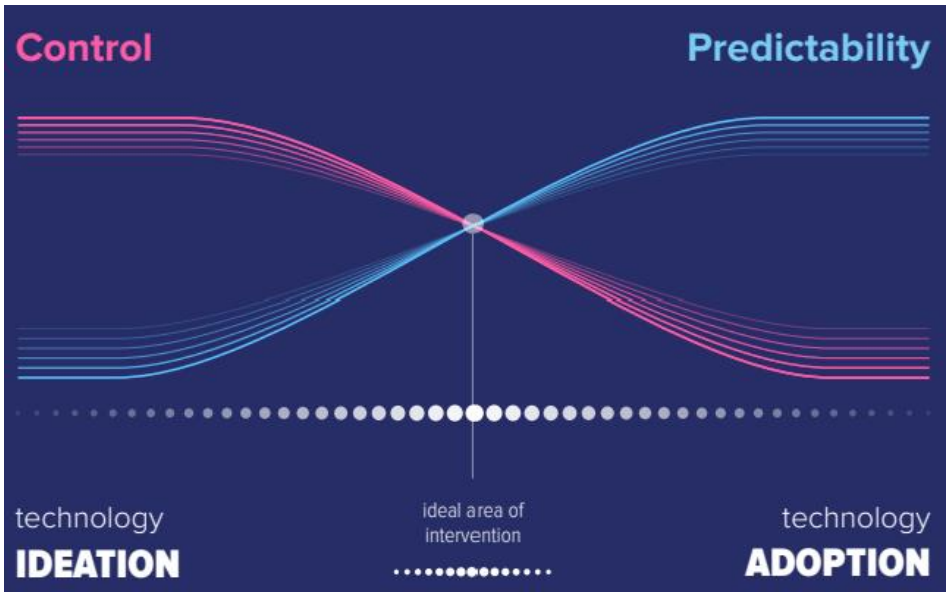

Figure 6: The Collingridge Dilemma. Control vs. Predictability (Besti & Samorè, 2018).

An example given by Chesterman, to visualize the Collingridge dilemma, is the commercialization of cars and the climate change it has favored. In 1906, the Royal Commission was commissioned to investigate the potential risks of this new technology (Royal Commission, 1906). One of the biggest concerns at the time was the dust blown up behind the cars. Today, cars and the transportation sector account for about a quarter of the world's energy-related CO2 emissions (Chesterman, 2021).

To prevent a "knee-jerk" regulatory backlash against AI, which would result in a stifle in mid to long term innovation, Truby (2020) suggests addressing AI technology in the financial sector early on with focused, proactive, but balanced regulatory approaches. This knee-jerk reaction is likely to take place in any industry sector and is not bound to the banking sector. The requested regulations by Truby ought to be in line with newly established, widely

acknowledged principles, for instance those outlined in the "Ethics Guideline for Trustworthy AI"[26] by the Artificial Intelligence High-Level Expert Group (AI HLEG).

With the pacing problem in mind there is a consensus among experts that the speed of technology advancement outpaces the creation of legal frameworks (El-Gazzar & Stendal, 2020; Herian, 2018). An interesting question therefore is if AI is currently in compliance with the GDPR? According to El-Gazzar the answer is clearly no. As seen in Figure 13: Compliance of AI in GDPR, there are nine Articles in the GDPR that are not met. One of those conflicts (Article 5 (2)) is that the conflict of AI and GDPR that manifests in the autonomy of AI systems leading to compliance issues with the accountability principle of the GDPR (El-Gazzar, 2020). A suggested workaround could be to treat AI systems as natural persona, awarding them equal legal obligations and rights, so that they can be held accountable for their own autonomous decisions (Butterworth, 2018).

Already with the temporal impacts in mind, the aim of the UK´s National AI Strategy is to encourage innovation and investments, whilst protecting the fundamental values of the UK. What highlights the strategy of UK is a chronological division of their roadmap, which is for example missing in the later described Malta approach. The UK´s national AI strategy distinguishes between three time periods and what is important within those periods (Kazim et al., 2021):

- *Short term*
    - *Public assurance roadmap, consultation on issues of data protection and governance, international AI activity.*

- *Medium term*

[26] "Artificial Intelligence: Commission Takes Forward its Work on Ethics Guideline", https://europa.eu/rapid/press-release_IP-19-1893_en.htm?locale=en (Published 8 April 2019) Accessed 28.06.2023.

- *Analysis of algorithmic transparency, piloting an AI standards Hub, increase governments awareness of AI safety.*

- *Long term*
  - *Technical standards, work globally on R&D challenges, update guidance on ethics and safety, understand what public sector actions can safely advance AI and mitigate catastrophic risks.*

To control the impact of implemented regulations over time, the regulations should be measured and regularly assessed for their effect and the dynamic nature of the technology. Susar and Aquaro (2019) note in their paper on achieving the Social Development Goals (SDG) through AI a list of existing tools for measuring AI in the public sector. These could be used by policy makers as well, to monitor developments and intervene and adjust regulations if necessary. In Table 14: Existing measures and index tools, according to Susar and Aquaro, the existing measurement tools are highlighted with a short explanation.

2017 MGI Research	*The 2017 MGI research uses a simulation to examine the pote ntial effects of AI on global economic activity at the country, sector, and worker levels.*
Research on development of AI in China 2018	*Measures the deployment of AI in four dimensions:* • *technical output and talent input,* • *industry development and market application,* • *development strategy and policy environment* • *social cognition and comprehensive influence.*
New generation AI white paper: industry growth point judgement 2018	*Objective to assess the development and growth point of the AI industry. Two evaluation systems:*

	• *One was to assess the stage of new generation AI industry development on 20 items i.e., innovation ability scale merit, capital operation etc.* • *The second divides new generation representative AI enterprises into three groups based on a value estimate from five dimensions.*
GOVERNMENT AI READINESS INDEX" Published on Oxford Insight (2018)	*Index regarding three factors:* • *public service reform (innovation, digital public services, government effectiveness),* • *economy and skills (digitization, technology skills, AI start-ups),* • *digital infrastructure (quality of data, available data and data capability).*
Stanford's Human Centered Artificial Intelligence Institute (HAI)	*Index regarding three main factors:* • *volume of activity* • *technical performance* • *other measures.*

Table 14: Existing AI measurement and index tools, according to Susar and Aquaro (2018).

4.3.3. A Spectrum of AI Regulatory Approaches

Ethical approach

Regulatory approaches often derive out of ethical considerations. Niels van Dijk combines the governance of AI and ethical considerations in a term he calls the ethification phenomenon. Through the integration of those ethical considerations, research and development of AI should be forced onto a sustainable path. Van Dijk defines ethification as: " *[...] the proliferation of invocations of 'ethics' and 'ethical principles and values' in legal, policy, academic and corporate discourses about the governance of technology and the growing institutional importance of ethical expertise through bodies such as ethical committees, advisory groups and boards, as well as through procedures enshrined in methodological documents such as ethical guidelines and checklists* (van Dijk et al., 2021)*"*. The use of ethical principles in research and development can also be viewed as soft regulation, a much more fluid and less rigor approach of regulation. Especially in international contexts, where the formulation of binding laws is difficult and takes a long time, soft laws offer a possibility to act swiftly. However, soft laws are not an all-purpose solution and some aspects are heavily criticized, especially for the lack of legal enforcement (DiMatteo, 2013; Turner, 2019), as well as the lack of the fact that rules formulated by corporate ethics will never reach the level of legitimacy that a government could provide[27].

Ethics can appear in different conceptions in the regulation context, Research Ethics, Innovation Governance Ethics and Industry Ethics. Matching these conceptions, the EU funding program Horizon 2020 can be divided into different pillars, Excellent science, Industrial Leadership and

[27] https://techcrunch.com/2017/10/04/deepmind-now-has-an-ai-ethics-research-unit-we-have-a-few-questions-for-it, Accessed 25.08.2023.

Societal Challenges.[28] Each of which is addressed by van Dijk with a different conception of ethics as seen in Table 15: Horizon 2020 pillars with corresponding Ethics conception.

Excellent Science	Research Ethics
Industrial Leadership	Innovation Governance Ethics
Societal Challenges	Industry Ethics

Table 15: Horizon 2020 pillars with corresponding Ethics conception.

Firstly, van Dijk addresses Research ethics, which is shortly summarized as ethics concerning the screening of project proposals, research activities and protection against unsound and non-integer science practices (2021).

Of particular interest is the second pillar, corresponding with Innovation Governance Ethics. Here, ethics occur as a comprehensible, repeatable instruction. It can take the form of a rulebook on how research design should look like or a code of conduct for the researchers to adhere. Via this method any socio-technical change can be governed, without being as restraining by a legal binding force. According to van Dijke Innovation Governance Ethics still has the ability to advise policy makers or serve as a reference for soft laws, like standardization or innovation programmes (2021). Both conceptions have been mapped out in Table 16: Mapping of Research and Innovation Governance Ethics (van Dijk et al., 2021). In the column Relation to AI/ICT, Innovation Governance Ethics is listed as a potential flexible way to regulate AI. A beneficial factor of regulating via ethics is the convergence of soft and hard sciences making it an interdisciplinary approach, trying to reach a common goal.

[28]https://era.gv.at/horizon-europe/horizon-2020-new/#:~:text=Horizon%202020%20was%20the%20EU,tackling%20seven%20%22Societal%20Challenges%22. , Accessed 09.08.2023.

Ethics type	Institutional/Societal Setting	Disciplines/Practices involved	Methods used and documents delivered	Relation to data protection	Relation to AI/ICT	Relation to (hard) law
Research Ethics	Ethics committees EU Institutions: REA, DG RTD Ethics Experts working for the EC	Research integrity Data protection law Social sciences	Ethics screening Assessment (check)lists Ethics guidelines	GDPR concepts framed as ethical issues in research	Research on AI should be bound by 'good' research practices	Legally enforceable in EC's research programmes regulations Ethics as part of the projects' Grant Agreement
Innovation Governance Ethics	EC expert groups: AI HLEG and EGE EC policy units: DG Connect (Unit A.1) EDPS: EAG on Digital Ethics	Law Politics Policy Business	Guidelines, frameworks, statements Principle-based ethics Assessment (check)lists Public consultations	GDPR concepts framed as ethical principles Ethics goes beyond data protection law	Ethics as a flexible way to regulate AI Advisory role for policy makers and the public	Opinions/ guidelines as part of a regulation

Table 16: Mapping of Research and Innovation Governance Ethics (van Dijk et al., 2021).

An example of such an interdisciplinary ethics board, giving advice on how to cope with AI is the EU AI High Level Expert Group (HLEG). Some of the beneficial factors of ethics, but not an exhaustive list is also provided by van Dijke (2021):

- Ethics promotes research accountability.
- Ethics supports innovation.
- Ethics is fast, flexible and anticipates unintended harms.
- Ethics enhances the legislation process.
- Ethic enhances public discussion.
- Ethic as experimental discourse on common European values and identity.

In the past, more precisely during the later part of the 20th century "ethification" processes already took place in the regulation of Biotechnology and the estimation of its potential implications. To have an expert group guiding the EC in governing issues around the field of Biotechnology they implemented the Group of Advisors on the Ethical Implications of Biotechnology (GAEIB), consisting of nine members.[29] Besides consulting the EC on related

[29] https://cordis.europa.eu/article/id/7643-ethical-implications-of-biotechnology, Accessed 09.08.2023.

issues, the GAEIB also generates and adopts opinions on its own initiative, they consult outside experts and meet on average once a month. Two examples of GAEIB´s influence on the EC legislation are, the labelling of foods derived from modern biotechnology (on request), and the legal protection of biological inventions (on own initiative) (GAEIB, 1997).

Innovation governance approach

Another approach to regulation, which is not (solely) based on ethical considerations, but on the effects of regulation directly on innovation, are the innovation governance approaches, as seen in Figure 7: Governance approaches to innovation, adapted from Hemphill, (2021).

- The Precautionary Principle
- Responsible Innovation
- Permissionless Innovation
- The Innovation Principle

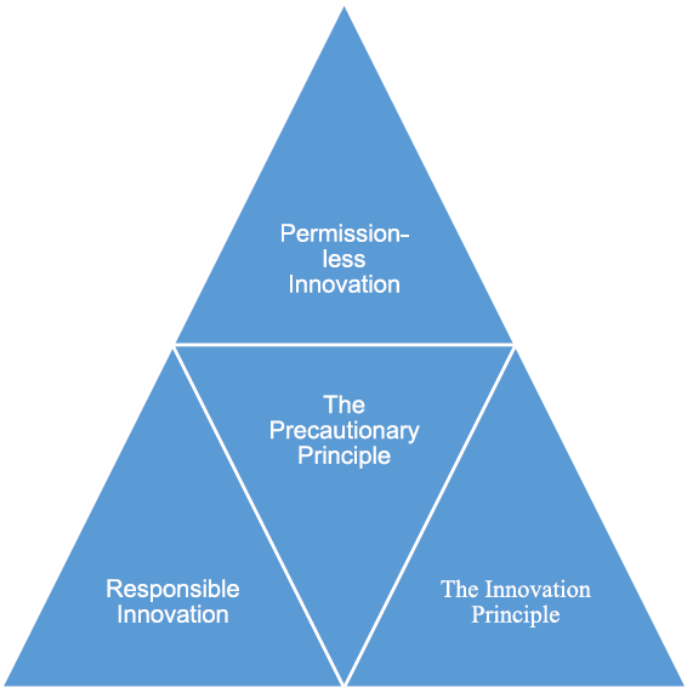

Figure 7: Governance approaches to innovation, adapted from Hemphill, (2021).

The most seminal approach of these principles nowadays is the precautionary principle (PP). Even though it is not completely clear when the PP was first introduced, the European Parliament Think Tank[30], as well as Beyerlin & Marauhn (2011) argue it was most likely during

[30] https://www.europarl.europa.eu/thinktank/en/document/EPRS_IDA(2015)573876, Accessed 10.08.2023.

the 1970´s in Sweden. For the Environmental Protection Act of 1969[31], the Swedish government decided to reverse the burden of proof for any environmentally hazardous activity. Since then, other nations have implemented this principle as well as in the Treaty on the Functioning of the European Union in Article 191.[32] Furthermore, the UNESCO World Commission on the Ethics of Scientific Knowledge and Technology advocates for the PP and gives the following working definition for the PP (COMEST, 2005):

"*When human activities may lead to morally unacceptable harm that is scientifically plausible but uncertain, actions shall be taken to avoid or diminish that harm. Morally unacceptable harm refers to harm to humans or the environment that is:*

- *threatening to human life or health, or*
- *serious and effectively irreversible, or*
- *inequitable to present or future generations, or*
- *imposed without adequate consideration of the human rights of those affected.*

The judgement of plausibility should be grounded in scientific analysis. Analysis should be ongoing so that chosen actions are subject to review. Uncertainty may apply to, but need not be limited to, causality or the bounds of the possible harm. Actions are interventions that are undertaken before harm occurs that seek to avoid or diminish the harm. Actions should be chosen that are proportional to the seriousness of the potential harm, with consideration of their positive and negative consequences, and with an assessment of the moral implications of both action and inaction. The choice of action should be the result of a participatory process."

[31] https://www.ilo.org/dyn/natlex/natlex4.detail?p_lang=&p_isn=29700, Accessed 10.08.2023.

[32] https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:12012E/TXT, Accessed 10.08.2023.

A second definition that is widely used was formulated in 1988 by a group of interdisciplinary scholars at Wingspread Wisconsin[33]:

"When an activity raises threats of harm to human health or the environment, precautionary measures should be taken even if some cause-and-effect relationships are not fully established scientifically. In this context the proponent of an activity, rather than the public, should bear the burden of proof. The process of applying the precautionary principle must be open, informed and democratic and must include potentially affected parties. It must also involve an examination of the full range of alternatives, including no action."

The PP can be further divided into two manifestations, the strong PP and the weak PP. The major differences from the strong and weak variant of the PP is the assertiveness of regulatory action, the triggers for regulation, the degree of prudence, and the burden of proof. The strong variant emphasizes immediate and assertive regulation based on potential threats, while the weak variant emphasizes managing risks in the context of scientific uncertainty. Table 17: Comparison of the strong and weak precautionary principle, adapted from Hemphill, summarizes the individual differences in more detail:

	Strong PP	**Weak PP**
Scope of Regulation	Regulation should be applied when an activity or product poses serious threats to human health or the environment, even if there is scientific uncertainty about the nature or extent of those threats.	Empowering regulators to address risks in situations of scientific uncertainty, even before a full understanding of the risk's nature or extent is achieved.

[33] https://www.asmalldoseoftoxicology.org/blog/2018/3/8/the-precautionary-principle-the-wingspread-statement, Accessed 10.08.2023.

Trigger for Regulation	Regulation is triggered when there are serious threats to human health or the environment, regardless of the level of scientific understanding about those threats. Focusing on the potential severity of harm.	Regulation is triggered when there's scientific uncertainty about risks, regardless of the immediate severity of the threats. Focusing on the potential for unknown risks.
Application of Regulation	Regulation should be presumptively applied in cases where there are serious threats, regardless of the level of scientific uncertainty. The burden of proof is placed on those advocating for the activity or product to demonstrate its safety.	Regulators should address risks in situations of scientific uncertainty, implying that regulations can be implemented even before the full extent of risk is understood, but it doesn't necessarily require the same level of presumptive regulation as the strong version.
Degree of Caution	A more cautious approach, advocating for swift action to prevent potential harm even in the absence of complete scientific understanding	While cautious, this variant is more flexible, allowing for regulatory action in cases of scientific uncertainty, but not necessarily demanding immediate measures as the other version.

Table 17: Comparison of the strong and weak precautionary principle, adapted from Hemphill (2020).

Naturally, the strong as well as the weak PP have been criticized. The former for being too rigid and nipping innovations in the bud, i.e. by Cass R. Sunstein in his article "*The paralyzing principle* (2002)". And the latter for being too shallow, as well as unprecise in proclaiming how

and when to take measures[34].

As seen in Figure 7: Governance approaches to innovation, adapted from, besides from PP are three other governance principles. Even though the PP is the most advocated form, they shall be described shortly in the next paragraphs. A more detailed comparison of the individual approaches is shown on Table 18: Innovation governance approaches and can be found in the list of figures at the end of the paper.

Responsible Innovation (RI) rather describes a process to follow or gives a framework to use when developing new innovations. Stilgoe provides four dimensions for RI - *reflexivity, responsiveness, inclusion and anticipation* (Stilgoe et al., 2013). All of those need to be addressed during the process of creating innovations. Schomberg, who coined the phrase responsible innovation also talked about a process in his initial definition: *"[…] a transparent, interactive* ***process*** *by which societal actors and innovators become mutually responsive to each other with a view to the (ethical) acceptability, sustainability and societal desirability of the innovation process […]* (Schomberg, 2013)". This shows that sustainable innovation does not come about through one-off mechanisms, but through constant steering and questioning during a process. At the same time RI does not imply a hurdle for all innovation, on the contrary. Valdivia argues that, "*RI […] do not seek to slow down innovation because they do not constrain the set of options for researchers and businesses, they expand it* (Valdivia & Guston, 2015)".

Even though inclusion is one of the pillars mentioned in connection with RI, critics of RI argue that the social responsibility falls short with RI. Especially such problems, that derive out of unrestricted growth such as climate change and income equality (Blok & Lemmens, 2015).

[34] https://techliberation.com/2017/05/18/does-permissionless-innovation-even-mean-anything/, Accessed 10.08.2023.

The next of the three contestants to PP is Permissionless Innovation (PI), which was coined by the scholar Adam Thierer. In his book *"Permissionless Innovation: The continuing case for comprehensive technological freedom",* Thierer advocates that innovation should be permitted by default and any problems that might, or might not arise, can be addressed in the future (Thierer, 2014). Furthermore, Thierer questions if entrepreneurs and researchers must be permitted by officials to develop and deploy innovations. To support his argument, Thierer provides five key points to consider (2014):

- *First, technological innovation is the single most important determinant of long-term human well-being.*
- *Second, there is real value to learning through continued trial-and-error experimentation, resiliency, and ongoing adaptation to technological change.*
- *Third, constraints on new innovation should be the last resort, not the first. Innovation should be innocent until proven guilty.*
- *Fourth, as regulatory interventions are considered, policy should be based on evidence of concrete potential harm and not fear of worst-case hypotheticals.*
- *Fifth, and finally, where policy interventions are deemed needed, flexible, bottom-up solutions of an ex post (responsive) nature are almost always preferable to rigid, top-down controls of an ex ante (anticipatory nature).*

PI also have been criticized, one of the critiques is, similar to AI and PP, that scholars cannot agree on the exact definition of what exactly "technological harm" is (Hemphill, 2020). PI is also mentioned as an "*unadulterated liberal disposition towards innovation* (Valdivia & Guston, 2015)". Furthermore, the pacing problem is mentioned in this context and lastly that, some impacts of innovations take time to become tangible[35].

[35] https://www.learnliberty.org/blog/permissionless-innovation-the-fuzzy-idea-that-rules-our-lives/, Accessed

The last of the three contestant is the Innovation Principle (IP). IP is an idea born by executives of 22 big firms (including Bayer AG, IBM Europe and BASF) with the intent to *"to ensure that whenever policy or regulatory decisions are under consideration the impact on innovation as a driver for jobs and growth should be assessed and addressed* (Hemphill, 2020)". The most intensive examination of IP comes from the European Risk Forum (self-described as a non-profit, expert led think tank), which defines IP as: "*[…] whenever policy or regulatory decisions are under consideration, the impact on innovation should also be fully assessed and addressed*[36]*". A*dditionally the European Risk Forum argues that, the IP was not invented to cancel out or undermine the PP, but rather to be additional and supportive. Nevertheless, there are some strong critics against the IP. Nina Holland from the Corporate Europe Observatory argues that, in contrast to the stated goal of the IP, it aims to "*weaken EU chemicals regulation (REACH) as well as EU rules for novel foods, nano materials, pharmaceuticals, medical devices and biotechnologies*"[37], and to create a counterweight to the PP, which is already embedded in the EU. Holland receives approval from the Greens, regarding her critique in the European Parliament[38].

All four of the just described governance approaches have been revised and clustered into an affinity diagram by Hemphill (2020) which can be found in the chapter figures and tables.

15.08.2023.

[36] https://www.eriforum.eu/uploads/2/5/7/1/25710097/monograph_innovation_principle.pdf, Accessed 15.08.2023.

[37] https://corporateeurope.org/sites/default/files/attachments/briefing_innovation_principle_final.pdf, Accessed 15.08.2023.

[38] https://www.greens-efa.eu/en/article/news/the-innovation-principle-is-a-regulatory-trojan-horse-from-the-industry, Accessed 15.08.2023.

Analyzing Table 18: Innovation governance approaches, as well as the different approaches, offers the insight, that the PP approach is the one with the highest risk management possibility. However, it needs to be differentiated between the strong and the weak approach, whereas only the strong PP is characterized by its strong risk management. RI and PI on the other hand are the two approaches with the highest potential for technological innovation, with lesser focus on the potential harm.

Innovation strategy approach

When it comes to innovation strategies, there are a lot of different options. From open source, standardization, to the use of intellectual property. Especially in the software industry, where AI development can be assigned to in most cases, open source and standardization played a big role in the past, i.e., in the development of the programing language Python, or the operating system Linux. The beneficial aspect of open-source software, as well as a standardization approach, is that the software will be spread further, giving a wider range of developers the ability to test, improve or maintain the software (Benkler, 2002). As well as lessen the need for reinvention of solutions for the same problem. Especially in software development the possibility to achieve the same goal with different approaches can lead to a high number of reinventions for the same solution (Kim et al., 2017). The implementation of standardization communities can reduce this number, while at the same time promote recombinant innovations. Another beneficial aspect of standardization organizations is the possibility of educating stakeholders such as policy makers, startups and SME. They can act as a hub of knowledge, gathering tacit knowledge, building cross industry networks and steer the institutional knowledge into the desired trajectory (Blind & Böhm, 2019). Unlike intellectual property, standardization also promotes innovation, while intellectual property is labelled by some researchers as inhibiting innovation. This is due to the fact that intellectual property registration is a rather costly undertaking and is therefore mainly carried out by larger companies, which in

turn encourages monopolies (Boldrin & Levine, 2008).

In an empirical study, investigating 50 different startups entering the European data and AI economy through DMS Accelerator (A hub comprised of a team of coaches, lawyers, consultants, researchers and communicators, established in 2019 in light of the European Union´s Horizon 2020 research and innovation program[39]) Maria Priestley and Elena Simperl examined how open innovation programs can have the most positive impact (2022):

- Case studies and stories to promote best practices.
    - As already undertaken by the EU, the use of descriptive stories and best practice examples can promote the operational, organizational benefit as well as the benefit of data-driven technologies. Priestly suggests that a similar format could be effective in conveying tacit knowledge related to the value of standardization, intellectual property, GDPR and process innovation.
- Demonstrate the extent to which existing motivations are covered by these best practice examples.
    - Standardization can provide access to insider information, networking opportunities, and a way to guarantee platform interoperability.
    - Utilizing intellectual property to mark their value to investors and maintain their position in the market.
    - In order to foster leadership and trust, compliance with the General Data Protection Regulation is necessary.
- Innovation program managers have to utilize the monitoring data and their experience to report regarding the changing needs of startups.
    - These insights might be of particular interest for policy makers.

[39] https://www.datamarketservices.eu/about/, Accessed 12.07.2023.

- A distinction between process and product innovation.
    - This will offer a greater precision to assess how AI technologies are assimilated by companies. Such knowledge can aid public funders in distributing funds to a wider range of inventive activities.

Regulatory Flexibility in AI: Soft Law vs. Binding Legislation

Since 2016 there has been a sudden increase of guidelines, frameworks and principles that focused on AI (Chesterman, 2021), some of them being published by conferences and industry associations like the Partnership on AI´s Tenets, Bejing Academy of Artificial Intelligence´s Bejing AI principles , Future of Life Institute´s Asilomar AI Principles and most recently the IEEE´s Ethically Aligned Design. And others published by companies from Microsoft´s Responsible AI Principles, IBM´s Principle for Trust and Transparency and Google´s AI Principles. In accordance with the pacing problem, governments reacted slower, but laws and guidelines have also been passed by many governments and intergovernmental institutions. These include, but are not limited to the Malta AI Strategy[40], Singapore Model AI Governance Framework[41], Australia´s AI Ethics Principles[42], China´s AI Governance Principles[43], the G7 Charlevoix Common Vision for the Future of Artificial Intelligence[44], the EU Guideline for Trustworthy AI[45] and the OECD Recommendation of the Council on Artificial Intelligence[46].

[40] https://www.mdia.gov.mt/malta-ai-strategy/, Accessed 31.08.2023.

[41] https://www.pdpc.gov.sg/help-and-resources/2020/01/model-ai-governance-framework, Accessed 31.08.2023.

[42] https://www.industry.gov.au/publications/australias-artificial-intelligence-ethics-framework/australias-ai-ethics-principles, Accessed 31.08.2023.

[43] https://www.most.gov.cn/kjbgz/202109/t20210926_177063.html, Accessed 31.08.2023.

[44] G7 Charlevoix Common Vision for the Future of Artificial Intelligence, Accessed 31.08.2023.

[45] https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai, Accessed 31.08.2023.

[46] https://legalinstruments.oecd.org/en/instruments/oecd-legal-0449, Accessed 31.08.2023.

Despite of the range of the proposals regarding the governance and regulation of AI, the single suggestions are not contradicting each other, but are to be seen as building blocks that can be put together to achieve a holistic framework of regulation. It is not easy to combine such a plethora of information. Nevertheless, Gasser tried to design a holistic governance framework, as seen in Figure 8: Layered model for AI governance. Gasser and Almeida combine different efforts from Singapore's accountability-based framework to the IEEE's Ethically Aligned Design Principles. They also emphasize, that it is not a comprehensive structure by the mere fact, that many aspects of AI governance are still to be examined and the list of proposals is likely to further expand.

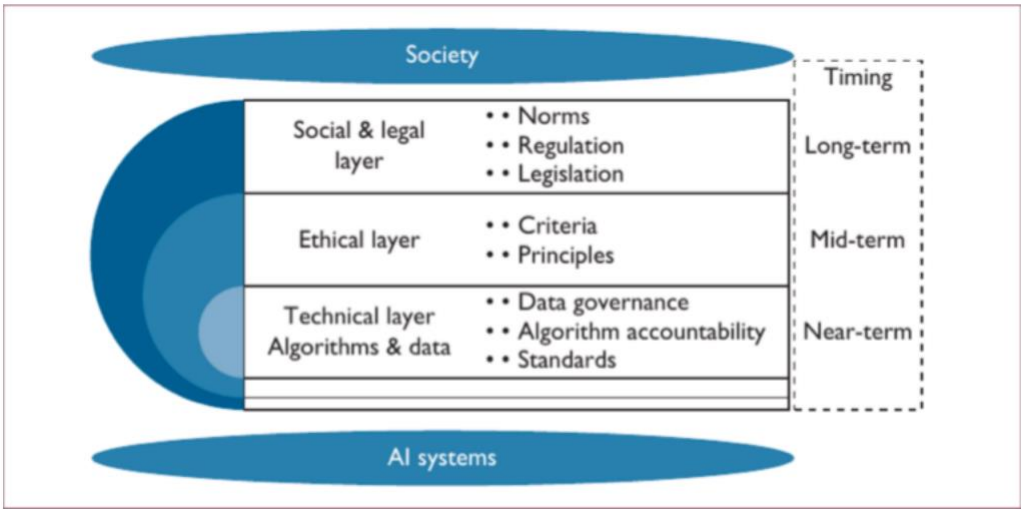

Figure 8: Layered model for AI governance (Gasser & Almeida, 2017).

Another approach which shows the many different aspects that should be taken into consideration, is provided by the Center for long-term Cybersecurity (CLTC) and can be seen in Figure 9: AI security domains, adapted from the CLTC Whitepaper, by Jessica Cussions Newman. Newman et al. argue, that it is essential for policy makers to include technology experts to recognize, manage and limit the risks that AI poses in the physical/digital, political, economic, and social sphere (Newman, 2019). The CLTC notes furthermore, that not necessarily every actor must consider every topic, but rather that every topic is considered in an AI ecosystem. For example, Canada focusing on AI talent attraction, Germany on AI

manufacturing and the UK on ethical AI development and application.

DIGITAL / PHYSICAL	POLITICAL	ECONOMIC	SOCIAL
RELIABLE, VALUE-ALIGNED AI SYSTEMS	PROTECTION FROM DISINFORMATION AND MANIPULATION	MITIGATION OF LABOR DISPLACEMENT	TRANSPARENCY AND ACCOUNTABILITY
AI SYSTEMS THAT ARE ROBUST AGAINST ATTACK	GOVERNMENT EXPERTISE IN AI AND DIGITAL INFRASTRUCTURE	PROMOTION OF AI RESEARCH AND DEVELOPMENT	PRIVACY AND DATA RIGHTS
PROTECTION FROM THE MALICIOUS USE OF AI AND AUTOMATED CYBERATTACKS	GEOPOLITICAL STRATEGY AND INTERNATIONAL COLLABORATION	UPDATED TRAINING AND EDUCATION RESOURCES	ETHICS, FAIRNESS, JUSTICE, DIGNITY
SECURE CONVERGENCE / INTEGRATION OF AI WITH OTHER TECHNOLOGIES (BIO, NUCLEAR, ETC.)	CHECKS AGAINST SURVEILLANCE, CONTROL, AND ABUSE OF POWER	REDUCED INEQUALITIES	HUMAN RIGHTS
RESPONSIBLE AND ETHICAL USE OF AI IN WARFARE AND THE MILITARY	PRIVATE-PUBLIC PARTNERSHIPS AND COLLABORATION	SUPPORT FOR SMALL BUSINESSES AND MARKET COMPETITION	SUSTAINABILITY AND ECOLOGY

Figure 9: AI security domains, adapted from the CLTC Whitepaper, by Jessica Cussions Newman.

Another promising institution, a best practice forum for policy makers, is the Internet Governance Forum (IGF)[47]. Goal of this platform is to establish and provide a best practice forum on topics like IoT, Big Data and AI. The forum is focused on the lack of knowledge that consumers and especially policy makers have about the benefits of these technologies. They don't reinvent these things, but rather follow recommendations or bring them together from different committees and experts, like Gasser and Almeida. Among them are the Harvard Berkman Center and ITS Rio initiative, the Oxford Internet Institute, AI Now Institute , IEEE AI and Ethics Initiative, Internet Society IoT Security for Policymakers and many more.[48] The proposed best practices are:

1. *Define your terms narrowly.*

[47] https://www.intgovforum.org/en, Accessed 27.06.2023.

[48] https://www.intgovforum.org/en/filedepot_download/6733/1438, Accessed 27.06.2023.

2. *Be ecumenical about technology (or "Strive to be technology-neutral").*
3. *Collaborate.*
4. *Consider ethics and human rights when applying IoT, Big Data, and AI.*
5. *Watch out for bias and incomplete data sets.*
6. *Make privacy and transparency a policy goal and a business practice.*
7. *Ensure systems are adequately secured before they get to the market.*
8. *Foster technologies and business practices that empower SMEs.*

Points 3, Collaboration and 7 adequate security, in particular address innovation. The former point emphasizes the importance of open communication when it comes to the impact new technologies can have on the public and existing networks. Through this collaboration then create ways of working together when developing future-looking policies. And the latter calls for an industry self-regulation to avoid stalling innovation, alongside clear liability regulations. These two points combine different dimensions of the impact regulation has on innovation addressed in chapter 4.3.2 and chapter 4.3.3.

Other international standard bodies also came forth with recommendations on regulating AI. The origin of such standards stems from the field of robotics, more precisely from the Engineering and Physical Sciences Research Council, which aims at designing robotics with human-centric purposes (Boden et al., 2017). Also, the Institute of Electrical and Electronics Engineers (IEEE) advertise these human-centric approach in particular but not exclusively for AI (Chatila et al., 2017).

Similar to the IGF, most likely to some extend due to the IGF, the OECD also agreed on "Principles for responsible stewardship of trustworthy AI" including the following five aspects (Truby & Fenster, 2020):

1. *Inclusive growth, sustainable development, and well-being.*
2. *Human-centered values and fairness.*
3. *Transparency and explainability.*
4. *Robustness, security, and safety.*
5. *Accountability.*

Dismissing any such principles when regulating AI would create eminent threats to sustainable development. These risks have been analyzed by Vinuesa and can be seen in Figure 10: Survey of Positive and Negative Impacts of AI.

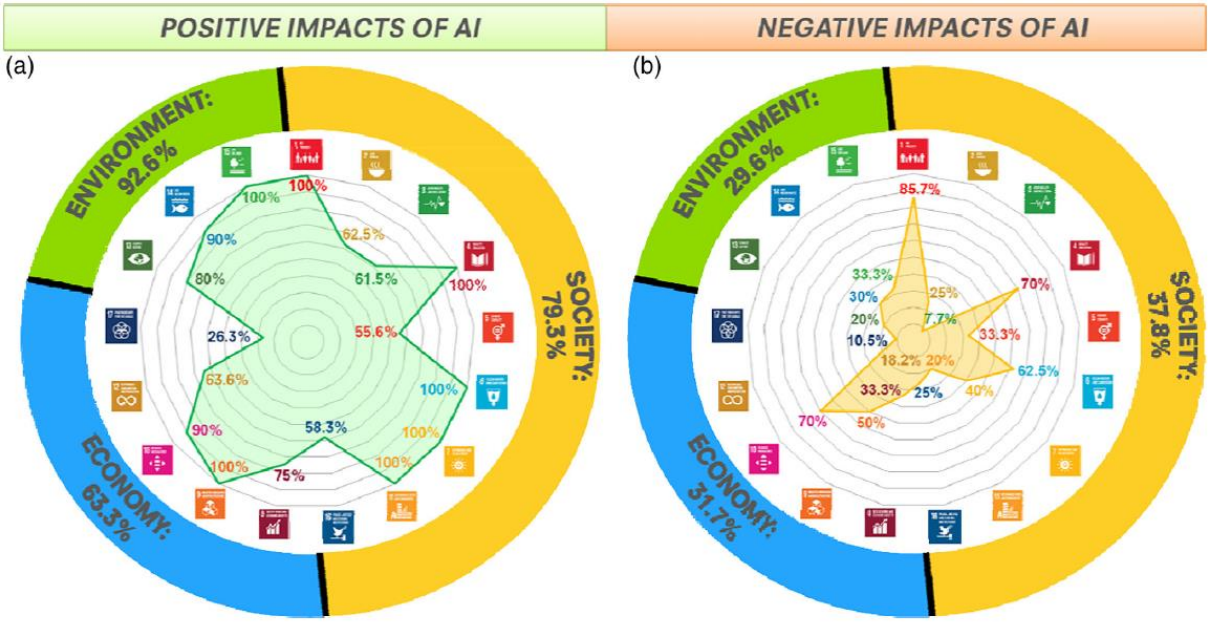

Figure 10: Survey of Positive and Negative Impacts of AI (Vinuesa et al., 2020)

As presented, many suggestions on governing AI have already been published. The International Telecommunication Union (ITU) published a guideline for policy makers to

design such positive policies, by addressing points to consider.[49] The ITU also comes to the conclusion, that an over-regulation might torment innovation. They also state that one guiding principle should be "*the focus on entrepreneurship and innovation not on AI (p.17)".* In order to achieve that, the availability and the buildout of physical infrastructure needs to be increased and the local ecosystems of entrepreneurs and startups needs to be supported. This support can be achieved by various means such as technology incubators, government programs, as well as commercial transfer of research.

The AI HLEG, as another instance providing aspects to consider when regulating AI, advocates to establish procedures to ensure that AI operates within the limits of accountability, transparency, explainability and fairness to ensure trustworthiness. To act within these limits, the AI HLEG, but also the OECD promote certain core elements, which are (Truby et al., 2020):

1. *Human agency and oversight.*
2. *Robustness and safety.*
3. *Privacy and data governance.*
4. *Diversity, non-discrimination and fairness.*
5. *Transparency.*
6. *Societal and environmental well-being.*
7. *Accountability.*

[49] https://www.itu.int/en/ITU-D/Conferences/GSR/Documents/GSR2018/documents/AISeries_GovernanceModule_GSR18.pdf , Accessed 27.06.2023.

Country Comparison

Malta was the first country to implement a regulatory AI framework among national authorities (Ellul et al., 2021). Because of the stifling nature of regulation, the Malta Digital Innovation Authority (MDIA) has designed a framework for regulating AI, that is primarily voluntary, but can be demanded based upon the sector of the AI application, as well as with the associated risk the AI application poses. Furthermore, the MDIA focuses on regulating AI technology based on the state of the art possibilities, meaning the regulations apply to narrow AI, and not general AI (Bostrom & Yudkowsky, 2011). The MDIA will also continue to monitor any developments and if necessary, update their guidelines. Therefore, trying to address the dynamic nature of AI innovation. Joshua Ellul assesses the Malta approach in his paper "*Regulating Artificial Intelligence: A Technology Regulator's Perspective*" and makes a case for voluntary assurances of AI and mandatory assurances of regulated and critical activities. To promote his argumentation Ellul gives the example of an AI identifying user preferences. Whilst this AI system could act in a relatively harmless sector as providing movie recommendations, the same AI framework could also be used for a social network targeting campaign to influence the outcome of an election (2021). The underlying infrastructure is application agnostic, raising the question if the infrastructure should be regulated or much rather the software, or activity.

Other scholars agree, that just regulating all forms of AI would stifle innovation (Gurkaynak et al., 2016). In many cases the call for regulation of AI stems from the fear of malicious forms, or unintended developments of general AI. Therefore, the Malta approach and the proposition of Ellul is to differentiate between Artificial General Intelligence (AGI) and Artificial Narrow Intelligence (ANI) in terms of regulation. Furthermore, Ellul proposes that mandatory regulation should distinct between sectors/activities and not technology per se. This argument is supported by other scholars like Matthijs M. Maas, who also investigated possible AI accidents, to propose how to regulate AI. In one of his papers, he suggests mapping out sectors,

which are prone to accommodate features that increase the possibility of AI-accidents and how to mitigate those. Possible solutions could be to reduce opacity or increased "slack", as well as newer generations of AI, that are capable of mapping signatures of failure and could be used as monitor and fail safe systems (Maas, 2018).

The EU also does take into account that AI cannot be regulated on its form alone (e.g., being part of a larger system, or a stand-alone product), similar to the approach Malta approach. To solve this problem the EU-AIA proposes to regulate AI on the level of risk the AI application may pose. The so-called risk-based-approach. The EU specifies four layers of risks (EU-AI Act, 2021):

1) **Unacceptable risk**. AI posing risks to the fundamental rights is supposed to be banned and must not enter the market. e.g. "*Practices that have a significant potential to manipulate persons through subliminal techniques beyond their consciousness or exploit vulnerabilities of specific vulnerable groups such as children or persons with disabilities*…".
2) **High risk**. AI systems that pose significant risks to the health and safety of fundamental rights of persons. High-risk AI types are subject to mandatory standards; these are evaluated to see if they comply, and if they do, they are judged acceptable.
3) **Limited risk.** AI systems that are included in this category impose requirements for transparency to inform the user that the interaction is with a machine.
4) **Minimal risk.** AI systems that are included in this category are allowed to be legally developed.

4.3.4. The Interplay Between Regulation and Trust in AI

When AI comes into play, the ball is passed back and forth, creating a whole new interaction. An interaction in which it is incredibly important how AI is seen by society, and how society interacts with AI. This interaction depends largely on trust and anxiety regarding AI. This anxiety and fear people have regarding AI can be due to numerous facts. One of them being that the definition of AI is quite unclear, as already addressed. Secondly, fear can be fuelled by dystopian sci-fi representation in books and television, and some people might just think about losing their jobs due to AI automation and AI robotics (Gallup and Northeastern University, 2018). Another aspect why people struggle to grasp the concept of AI, therefore increasing anxiety towards AI can be explained by the AI effect. Once a research technology arisen from AI becomes a standard repertoire of computer science, it is often no longer considered as AI by the public (Stone et al., 2022). Carabantes argues: "*Even without being aware of it, millions of decisions per second that affect our lives are made by computer systems equipped with [machine learning], and the tendency is to increase that delegation of tasks in computers* (2019)". The AI effect thus leads to the perception that only new, inexplicable areas of application for AI are seen as such, while AI applications that are already familiar are not recognized as such and in truth AI is already ubiquitous and part of our daily lives. Although this recognition is precisely the approach -using and getting familiar with AI- which could actually lead to a higher level of trust.

However, at least in America, according to a 2018 survey, most of the population thinks AI will affect their lives positively, as seen in Figure 10: Survey of Positive and Negative Impacts of AI. In the content of anxiety, regulation of AI can even further increase the trust people have in AI. Regulating and classifying AI systems will most likely build up more trust, just as the GDPR

in Europe has done[50]. As already mentioned in chapter 4.3.1, strict liability is one of those ways to create public acceptance of AI. This is due to the fact that companies have to deal with potential problems more intensively in advance in order to avoid taking a financial risk. In addition to creating public acceptance clear liability law also promotes societal trust in new AI applications (Zech, 2021).

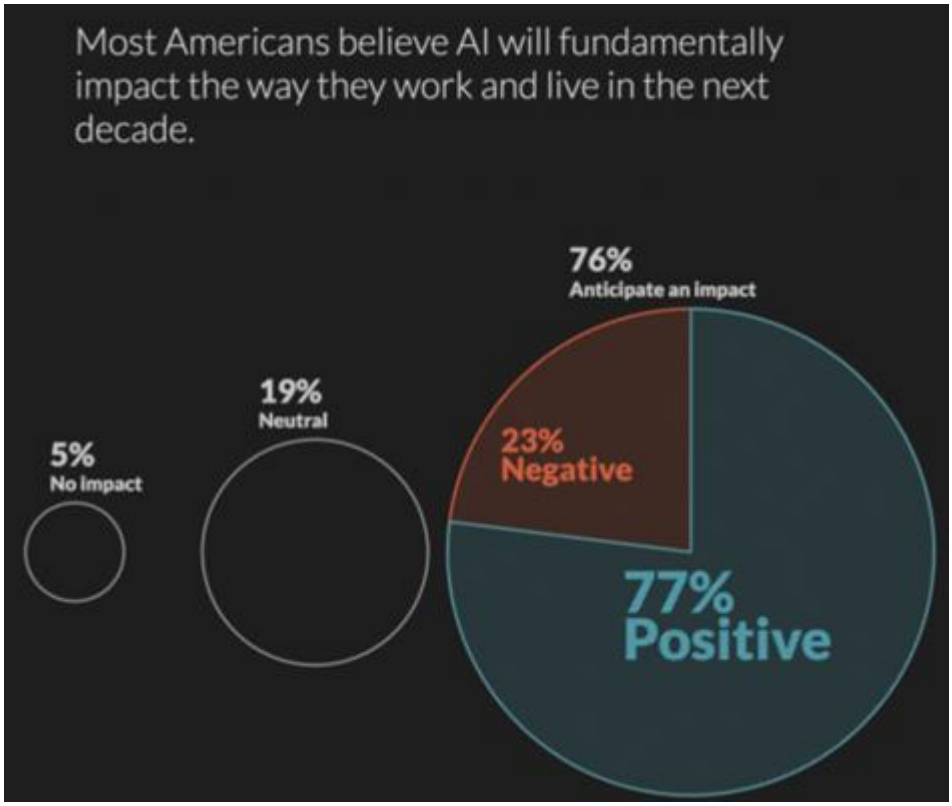

Figure 11: Views on the Impact of Artificial Intelligence and Higher Educations response, Gallup 2018.

Information campaigns and labelling AI applications is a good way to generally increase trust in AI. An even more effective way can be to train and educate people in the field of AI and show them how to use AI applications. This works best for the broader masses through low-threshold access and free content. Some companies, universities, governments and even

[50] https://dma.org.uk/uploads/misc/5af5497c03984-gdpr-consumer-perspective-2018-v1_5af5497c038ea.pdf, Accessed 27.06.2023.

individuals are already trying to implement this. For example, Andrew Ng's deep learning.ai[51], Siraj Raval's School of AI[52], but also big players like IBM's Cognitive Class [53] and the world renown University, the Massachusetts Institute of Technology (MIT)[54]. Equally important is the training of policy makers regarding AI, as demanded by institutions such as the IGF. As already mentioned in chapter 4.3.1, Alexander Hilton believes, that it is not the number of regulations, but rather the quality of regulations that promotes or hinders entrepreneurs, which in turn affects the potential for innovation in general.

Finally, besides policy makers and the broad mass of people, those who work on the forefront of developing AI applications should also receive training regarding AI. In contrast to the former groups, AI developers should, on the recommendation of the OECD receive training in regards of bias, diversity, and inclusion. Furthermore, there should be a certificate for those that have undergone such a training. This is not a new practice. In many other professions such as doctors, lawyers, and engineers this is common practice.

[51] https://www.coursera.org/specializations/deep-learning, Accessed 15.09.2023.

[52] https://www.youtube.com/channel/UCWN3xxRkmTPmbKwht9FuE5A, Accessed, 15.09.2023.

[53] https://cognitiveclass.ai/, last Accessed 15.09.2023.

[54] https://openlearning.mit.edu/courses-programs/mit-opencourseware, Accessed 15.09.2023.

5. Discussion

5.1. Summary of findings

This thesis asks the question, how regulation of AI will affect innovation. In the beginning the fundamentals of terminology and the problems associated with imprecise definitions is addressed to set the stage for answering the question. Then the main findings are subdivided into four different dimensions to create an overarch of this complex topic. Before those dimensions the problem of defining terms is addressed. The first dimension, the regulation's dual role, addresses whether regulation is impending or encouraging innovation. This dimension is again subdivided into different branches from the regulation of GPT, the positive and negative effect of liability on innovation, the relationship between innovation startups monopolies and regulation and finally the transnational nature of regulation. The second dimension is the temporal dynamic of regulatory influence with special focus on the pacing problem, the Collingridge dilemma and the danger of overregulation through inaction. Thirdly, the spectrum of different regulatory approaches and their effect on innovation is illuminated. And the last dimension is the interplay between regulation and trust. The basis of this compiled information is the literature research initially conducted, where out of the initial 413 database hits, 59 were further investigated and 41 included. Findings of the therein present references have, if deemed appropriate, also been included.

As apparent in the main findings, the first difficulty to regulating AI is, that the terminology must be clarified more precisely. The definition for the term AI has changed throughout the past, depending on the scientific conception. Additionally, it can vary on type of application, it can depend on the level of knowledge of the user and way of learning/training. Apart from the fact that at no time was there a single clear definition for AI. Especially in the context of legislation, a precise definition is necessary and is also called for by institutions such as the

IGF, a multi-stakeholder group on governance. The terminology of any law and regulation will affect the quality of the regulation itself and thereby inevitably the impact regulations have on innovation.

Regulation's Dual Role: Impeding and Encouraging Innovation

Firstly, AI being a GPT indicates the innovation promoting nature of the technology itself. Regulation on AI therefore must have an effect on said innovation. Scholars are not in total agreement, how these effects will play out, but underline the importance of regulation AI because of the chance for evolutionary change towards a more sustainable future through AI. Undoubtedly, AI will lead to countless complementary innovations, whether it will be regulated or not, the question to answer is how sustainable this innovation will look like.

One mechanism to steer innovation onto the right path is liability law. As the EP concluded, a lack of liability law would discourage innovation. The most difficult parts in creating a liability framework for AI that encourages innovation are the various forms of AI, definition of AI and the black box phenomenon. The last one making intent and causation tests near impossible. Furthermore, the different liability approaches, like the strict approach and fault-based approach, that are prevailing do not level the playing field of AI. Both approaches have their supporters and opponents. Looking at past transgressions the strict approach seems to be the safer route, even though it might impede innovation to a higher degree than the fault-based approach. To soften the innovation hampering effect a transparency framework could be implemented.

Depending on Schumpeter's earlier or later work startups, respectively large corporations are the source of innovation. A supportive environment for one of these forms of enterprises would therefore have a direct impact on innovation. Under the umbrella term "regulatory accumulation", hide the pebbles in the stream effect, the interaction between regulations and the behavioral overload. Whilst large established companies have the means to handle these

regulatory accumulations, startups and SME more likely struggle with them. Nevertheless, this regulatory accumulation underlines the innovation hampering effect of regulation. Contrary to the accumulation is the perspective brought for by Hilton, which argues only badly designed policy stifles innovation. However, the truth probably lies somewhere in between. While the accumulation of regulation slows down innovation, bad policy also has an additional bad effect on innovation, whilst good policy cancelled out some of the negative effects.

There are many scholars who emphasize the importance of an international, worldwide regulatory framework for coordinating and streamlining any regulation efforts that can be applied across sectors. Such a framework or legal entity would discourage the current wild west environment, characterized by a climate of high uncertainty and high risk. This in turn could potentially lead to an encouragement of innovation.

Temporal Dynamics of Regulatory Influence on Innovation

The way in which regulation now intervenes in the development and use of AI will have a significant impact on how AI develops, making a timely regulation paramount. Clark et al. emphasize the long-term ramifications of AI, projecting potential international conflicts and heightened power imbalances. However, as already addressed, AI has also the potential for creating a flourishing future for humanity, if governed properly.

Regulating GPT and disruptive innovations poses a unique challenge due to the unpredictable nature and the ever-increasing pace of complementary innovations facilitated by AI. This acceleration underscores the pacing problem, highlighting the discrepancy between the rapid evolution of technology and the incremental changes in social, economic, and legal systems. The Collingridge Dilemma further exemplifies the importance of regulating technologies, emphasizing the dual challenge of time and information, while the cost and complexity of implementing change increase as the technology adaption progresses. To prevent a knee-jerk regulatory backlash that could stifle innovation, a proactive, balanced approach to regulation is recommended. This knee-jerk reaction is to be expected when regulatory bodies do not

intervene early enough, leading to exploitation of the public by corporations. In turn the public is likely to demand a timely regulation which will be most likely not well thought out.

To effectively navigate the impact of regulations over time, regular measurement and assessment of regulations are essential as well. Hereby, established AI measurement and index tools could be used.

A Spectrum of AI Regulatory Approaches

Regulation in AI often aligns with ethical considerations. Niels van Dijk introduced the ethification phenomenon, integrating ethical principles into AI governance to steer research and development responsibly. Ethical integration could serve as soft regulation, offering flexibility and interdisciplinary collaboration, especially in the international context. However, critics highlight the lack of legal enforcement and legitimacy which is associated with governmental regulation.

Innovation governance approaches, beyond ethics, focus on regulatory effects on innovation. The most seminal approaches are the PP, RI, PI, and the IP. The PP advocates preemptive action even with scientific uncertainty regarding potential harm. RI emphasizes a transparent, interactive process ensuring innovation aligns with ethical considerations. PI supports a default permission for innovation, addressing problems as they arise. The IP assesses innovation's impact on jobs and growth in policy and regulatory decisions.

Besides innovation governance there are also innovation strategies like, open source, standardization, and intellectual property, that play crucial roles. Open source and standardization promote spreading software and knowledge more widely. Intellectual property is seen as inhibiting innovation, and potentially leading to monopolies.

Guidelines and principles have rapidly emerged since 2016, offering the possibility for a holistic approach, when combined properly. Agencies and governments try to incorporate emphasizing human agency, transparency, safety, innovation promotion and societal well-being.

In summary, AI regulatory approaches encompass ethics-driven governance, innovation-focused strategies, and risk-based frameworks. Combining these approaches into an effective regulatory system is crucial to foster responsible AI development while promoting innovation and societal well-being. A perfect regulation however is yet to be written.

The Interplay Between Regulation and Trust in AI

The expansion of AI raises concerns and uncertainties, partly due to the vague definition of AI and negative sci-fi depictions. Fear of job loss to AI automation amplifies this unease. Furthermore, the AI-effect blurs recognition of AI applications as they become mainstream. Regulation plays a vital role in demystifying AI and enhancing trust. The GDPR in Europe demonstrates how regulation can positively influence public trust. Clear liability laws, like strict liability, foster further acceptance and trust in AI applications, especially in the long run. Efforts such as information campaigns and proper labeling of AI applications can boost trust as well. Education on how to use AI applications, made accessible to the public, are a further possibility to promote trust. Besides teaching the broad public about AI, it is just as important to teach policymakers about AI for informed decision-making. Lastly, AI developers should receive training on bias, diversity, and inclusion, as recommended by the OECD. Certifying programmers who complete this training is in alignment with existing professional practices and ensures responsible AI development.

In summary, building trust and lessen AI-related anxiety involves clear definitions, effective regulation, education on AI's potential, and training for AI developers.

5.2. Strengths, limitations and future research directions

The strength of this review lies within the inclusion of interdisciplinary approaches of regulating AI and the resulting impact. That being said, it is also its weakness due to the immense amount of information that cannot be completely displayed within on master thesis and the difficulty of relating them to each other. Additionally, there is little research investigating the direct impact of regulation on innovation, but much research that shows an indirect correlation. Thus, aggravating an accurate allocation and evaluation of regulation an innovation.

From the initial search in different databases to the use of the therein referenced publications the number of sources to be investigated is not within the feasibility range of just one person.

Another aspect is the timing of the thesis, since there are numerous publications emerging in an ever-increasing speed regarding AI, regulation and its impacts, it is not possible to display all latest findings.

In line with the strength of interdisciplinarity it may pose benefits to include more than one person and even better persons with different scientific backgrounds, in the examination of this topic. The most interesting combination of backgrounds would probably be, philosophy, computer science, law, engineering, social studies and economy.

Future research directions could be the monitoring and measuring the impact of the regulations described, as well as the deviation of the predicted development of AI from the actual one. Of course, the latter is only possible retrospectively and can therefore only happen in the future.

6. Conclusion & Recommendation

After concluding all the main findings, a precise forecast how regulating AI will impact the innovation of this extremely diverse technology is impossible to give. There seems to be a unity within scholars, that regulation of AI will lead to an uncertain degree of stifling innovation (Gurkaynak et al., 2016; Vallverdú, 2014). However, scholars also agree that the regulation of AI is needed in order to foster and ensure sustainable innovation (Susar & Aquaro, 2019; Truby & Fenster, 2020), and good regulation can actually increase innovation (Harlow & Rawlings, 2009; Wisuttisak, 2020). If AI will not be regulated soon, there ought to be misconduct from some organizations. Those could be of governmental nature, but also from the private sector. Sadly, when looking at the past this seems inevitable. Many big-tech firms have repeatedly breached the trust of their customers and the general population[55] [56] [57]. As seen in Figure 10: Survey of Positive and Negative Impacts of AI, the further development of AI does not only have positive impacts, but might also have severe negative impacts for reaching the SDG´s. This further brings up the question, do we want unregulated unsustainable innovation in the first place, or do we want to take the chance this technology poses, to foster positive and sustainable innovation and to reach the goals set by the UNESCO and other more regulatory bodies. Furthermore, in agreement with Truby and Fenster (2020) a proactive regulation seems to mitigate the risk of a knee jerk regulatory response. Something that will most likely appear, after AI shows severe negative impacts after which the public will inevitably call for regulation. In other words, if AI goes unregulated for too long, there will be a poorly designed

[55] https://www.ft.com/content/ea028300-a88d-11e9-984c-fac8325aaa04, Accessed 04.07.2023.

[56] https://www.ft.com/content/1388c544-a812-11e9-984c-fac8325aaa04, Accessed 04.07.2023.

[57] https://www.ft.com/content/79b56392-dde5-11e8-8f50-cbae5495d92b, Accessed 04.07.2023.

overregulation at some point. Implementing regulations aimed at being sustainable and ethical now, will ensure that future innovations are designed in a beneficial way.

In my opinion it is clear that the recommendations for any regulatory body, have to be very close to proposals and guidelines already provided by competent and independent institutions like the IEEE, the AI HLEG, the OECD and others, as mentioned in the thesis. Many of those proposal do not specifically aim to foster innovations. However, they foster an innovation- and more importantly sustainable innovation environment. My recommendations for regulatory bodies therefore are very close to what other scholars suggested.

Firstly, we have to aim for transparency. The before mentioned black-box phenomenon bears the risk of repeating and reinforcing institutional biases, as well as causing a lack of control and therefore blinding any controlling bodies, which should supervise developers, distributors, and users of AI. The transparency framework provided by Reed (2018) seems to be of reasonable nature as well as actionable.

Secondly, and at the same time interconnected to the first point, is explainability. In contrast to transparency which aims to unravel the inner workings of AI systems, explainability provides explanations for the decisions made by AI. Both are crucial concepts in ensuring fairness as well as building trust in AI (Truby et al., 2020). We already, at least in the EU, have to oblige to the GDPR, which states that anybody within the EU has the right to know how an automated decision about themselves has been made. Nevertheless, there are already severe breaches of this act, as seen in Figure 13: Compliance of AI in GDPR. Hence, no black-box scenario can be compliant with the GDPR. Making human oversight, or human judgement mandatory.

Thirdly, acknowledging these rules on an international level. In other words, the more countries combine their efforts to regulate AI the more impact the regulation could have. In other industries history has shown, that when influential enough countries demand that imported goods are up to a certain standard, even if this standard is not needed in the producing country,

as in the automobile industry (Truby & Kratsas, 2017), they will comply in fear of losing market share. Since already the EU, America, Japan and other countries have published similar guidelines for AI regulation, the possibility to get other countries to oblige is not unsubstantial. In September 2015 193 nations pledged to comply to the SDG´s.[58] If this was not just lip service, these nations have the opportunity and the obligation to focus on sustainable innovation with AI. An international regulatory body, with the means to act in case of infringement of set rules, would be the most preferable option. This international body could be built following the example of the International Atomic Energy Agency (IAEA). A core idea of the IAEA is to distribute beneficial purposes of atomic technology, whilst ensure that those beneficial purposes are the only ones pursued (Chesterman, p. 196, 2021).

As a mean of monitoring the AI development, there are several possibilities listed in the main findings. In Agreement with David C. Benton, it seems apparent that the regulation of a disruptive technology such as AI is an opportunity for evolutionary change and therefore necessary to implement. Regulatory agencies need to obtain the ideal balance of necessary controls and unnecessary restraints of innovation (Benton et al., 2020). However, the monitoring of regulation is heavily dependent on the previous elaborated international body.

Furthermore, in agreement with other scholars like Miller et al., I am convinced that a culture change is needed as well to foster sustainable innovation (Miller et al., 2018). A measure that could be taken to accelerate this transformation would be the training of AI professionals, especially the coders. Not only the code has to be regulated, but the coders as well. Additionally, AI training should also be provided for policymakers and with regards to the future basic knowledge should be already thought in schools.

To accelerate innovation best practice stories and examples should be provided (Priestley &

[58] https://www.sdgwatch.at/de/ueber-sdgs/, Accessed 05.07.2023-

Simperl, 2022), as well as the provision of physical infrastructure for independent research agencies and businesses that do not have the needed resources to build such infrastructures themselves.

Lastly, in a similar manner to the GAEIB, the EC could implement an Ethical Implication Committee regarding issues of AI. This Committee should be approachable by different stakeholders with their concerns, but also provide recommendations on its own initiative.

The number of recommendations, each affecting different areas, argues for a holistic approach, similar to the approach of Gasser and Almeida described in the paper.

Figures and Tables

Innovation governance approaches.

Conceptual Typologies	Precautionary Principle	Responsible Innovation	Permissionless Innovation	Innovation Principle
Risk	The "cause-and-effect" relationship between an action and effect is not fully established scientifically, thus defers to a "worst case" scenario.	This approach takes an "anticipatory" approach to societal benefit and costs of technological innovation, recognizing the social costs that markets do not internalize.	Experimentation with technology is encouraged, and should generally be permitted by default.	Protect society and the environment, while simultaneously protecting Europe's ability to attract and benefit from technological innovation.
Risk Level	Credible threat of catastrophic or irreversable harm to human health or the environment is "morally unacceptable."	Characterized as "much less" constraining, i.e., accepting (and managing) risk, than the precautionary principle.	Innovation should be allowed to continue unabated, unless a "compelling case" is made that it will bring serious harm to society.	The innovation principle should be utilized in conjunction with the precautionary principle; it is not intended to be used to undermine the importance of the precautionary principle.
Burden of Proof	The proponent of innovation bears the burden of proof – and not the public.	Characterized as not constraining "the set of options for researchers and businesses", but is concerned with "potential risks, dangers, and public worries."	Constraints on the innovator should be "the last resort", based on the legal concept of "innocent until proven quilty."	It is identical to the precautionary principle i.e., the proponent of innovation bears the burden of proof – and not the public.
Regulatory Response	Public regulation is the default position.	It is not "a doctrine of regulation"; however, regulation is an acceptable response when deemed necessary for public safety and to mitigate environmental harm.	As a last resort, regulation imposed should be based on evidence of "concrete" potential harm, not worst-case hypotheticals." However, such regulation should emphasize "flexible, bottom-up" solutions (of an *ex post* nature) versus "rigid, top-down controls (of an *ex ante* nature).	Should avoid "uneccesary regulation" … move from prescriptive to dynamic and adaptive regulation … consider both risks and benefits … and weigh risks and benefits of alternative solutions (rather than negative asects of a single approach or solution).
Governance Perspective	There are two variants: the "strong" precautionary pricciple, involving an outright ban on the technological innovation, and the "weak" precationary principle, involving regulatory constraints on the technological innovation.	This approach is process-oriented and holistic, viewed as an integral part of democratic life, social order and economic prosperity. It includes an innovator's awareness of values and ethics; stakeholder diversity; and societal desirability, acceptability, and sustainability.	There is real value to learning through "trial-and error, resiliency, and ongoing adaptation to technological change."	This approach "aims to stimulate investment in innovation by increasing the confidence of innovation in the regulatory system."

Table 18: Innovation governance approaches, (Hemphill, 2021).

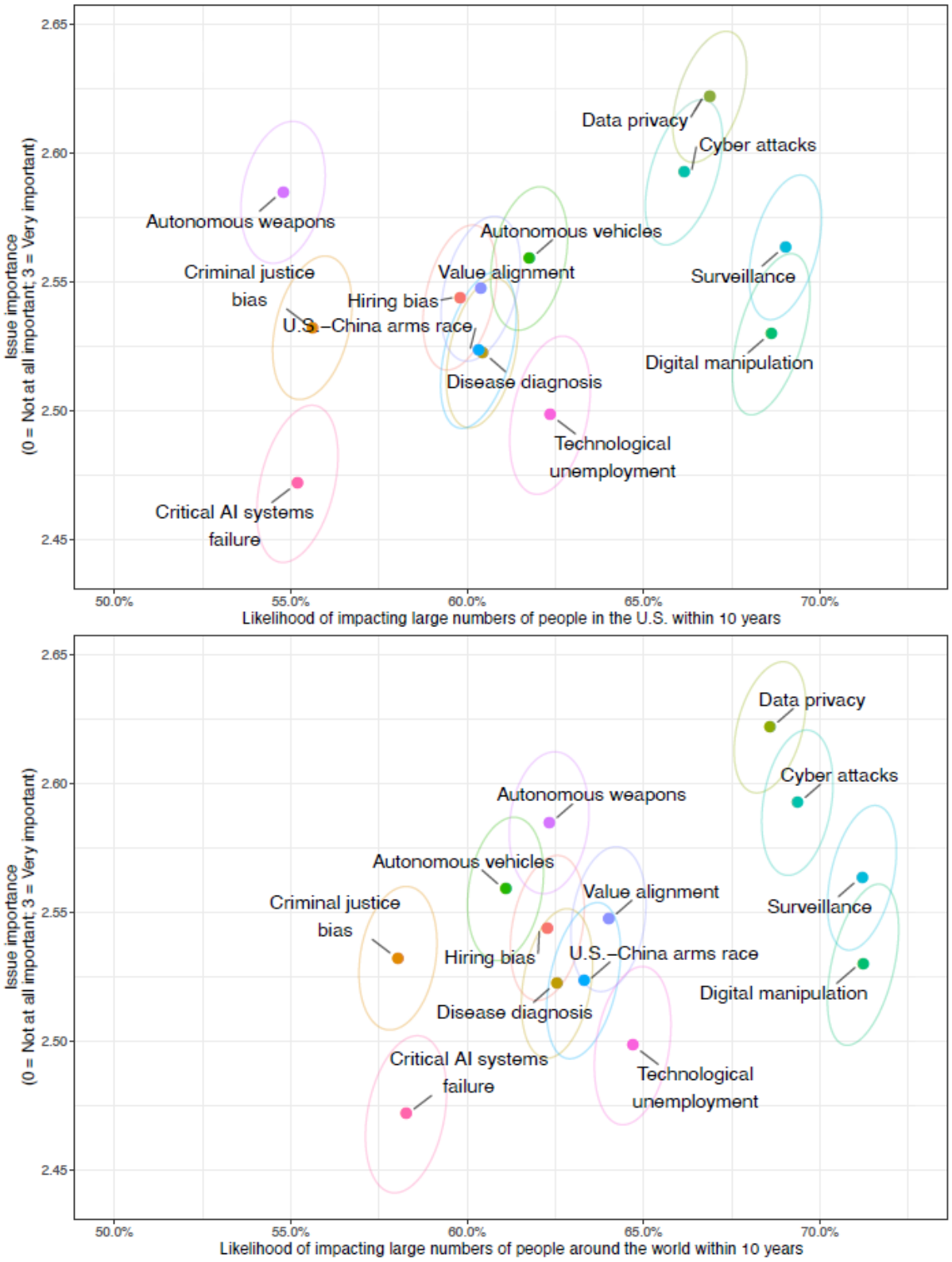

Figure 12: Perceptions of AI governance challenges in the U.S. and around the world (Zhang & Dafoe, 2020).

AI characteristics	Complies?		GDPR Articles	Justification
	Yes	No		
Autonomy		X	Article 5 (2)	It makes it difficult to hold AI systems accountable for the harm they may cause to the data subject.
Automation and ML algorithms		X	Article 13 (2)f, Article 14 (2)g, Article 15 (1) h, Article 22 (1)	It makes it difficult to explain the complex ML algorithm logics used to process the personal data undermines fairness and transparency of processing. Thus, the data subjects are unable to choose not to be subject to automated decision-making and profiling, as they do not know the consequences of it on them.
	X		Article 12, Article 13, Article 14	1) It enables controllers to conduct an automated legal analysis, using ML algorithms, of the privacy policies of their online platforms and services and refine them to, indirectly, comply with their information obligations under GDPR. 2) It enables automatic summarization of the controllers' privacy policies, so that data subjects can exercise their rights to be informed about data processing activities and make informed decisions about disclosing their personal data.
		X	Article 5 (1)a	3) ML algorithms may produce results that are discriminatory because the training data provides a biased representation of the reality, which indicates unfair processing of the personal data.
ML algorithms and big data		X	Article 5 (1)b	ML algorithms may process personal data for ambiguous and illegitimate purposes and generate new data, which makes it unclear for data subjects what will be the purpose for using all these data.
		X	Article 5 (1)c	ML algorithms tend to collect big personal data and repurpose them, which makes it unclear for data subjects if all these data are adequate and relevant for the purpose they are processed for.
		X	Article 4 (11)	Processing and repurposing big personal data make it difficult to obtain an informed consent based on a statement or a clear affirmative action.

Figure 13: Compliance of AI in GDPR (El-Gazzar & Stendal, 2020).

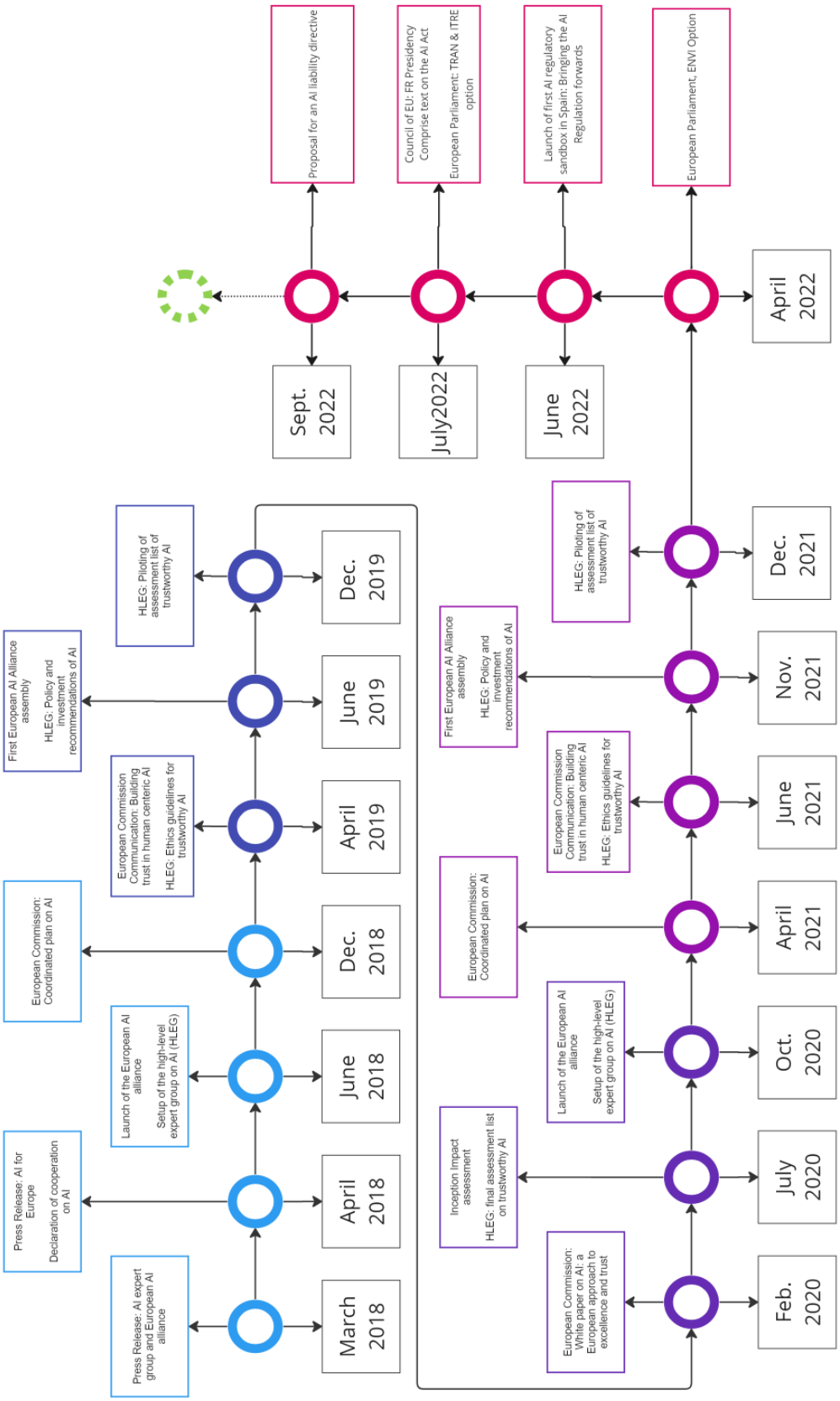

Figure 14: Timeline of the EU-AI milestones. https://digital-strategy.ec.europa.eu/en/policies/european-approach-artificial-intelligence; Accessed on 10.04.2023.

Barriers and challenges	*Examples*
Lack of qualified staff	IT-experts (general), IT-staff with AI expertise, data scientists, specialists and generalists, software developers, AI experts
Costs	Know-how creation, development costs, long development times (trial and error for innovative solutions), hardware costs for robotics
Lack of knowledge	Insufficient information about AI (general), unrealistic expectations, insufficient competence in AI (with not even the definition being clear)
Credibility of AI solutions	Unrealistic claims regarding AI and disappointment
Technical aspects	Lack of explainability for learning systems, lack of data - a strong limitation for AI, especially for SMEs
Regulation	Current legal regulation, e.g. in the health sector; unclear responsibilities for overall system behaviour
Hesitation	Executives hesitate, especially in SMEs; but so do customers
Hype	Risk that the current hype about AI hampers its development, because it blurs the vision of real opportunities and creates wrong expectations

Figure 15: Expert interviews, barriers and challenges (Prem, 2019).

List of Figures

List of Tables

List of Abbreviations

AGI	Artificial General Intelligence
AI	Artificial Intelligence
ANI	Artificial Narrow Intelligence
AIaaS	Artificial Intelligence as a Service
CLTC	Center for long-term Cybersecurity
DSS	Decision Supporting System
EC	European Commission
EP	European Parliament
EU	European Union
EU-AIA	European Union Artificial Intelligence Act
FDA	Food and Drug Administration
GAEIB	Group of Advisors on the Ethical Implications of Biotechnology
GDPR	General Data Protection Regulation
GPT	General Purpose Technology
HLEG	High Level Expert Group
IAEA	International Atomic Energy Agency
ICT	Internet and Communication Technology
IEEE	Institute of Electrical and Electronics Engineers
IGF	Internet Governance Forum
IoT	Internet of Things
IP	Innovation Principle
ITU	International Telecommunication Union
MDIA	Malta Digital Innovation Authority
MIT	Massachusetts Institute of Technology
NLP	Natural Language Processing
OECD	The Organization for Economic Cooperation and Development
PI	Permissionless Innovation
PP	Precautionary Principle
RAMESES	Realist And MEta-Narrative Evidence Syntheses: Evolving Standards
RI	Responsible Innovation

SDG	Social Development Goals
SME	Small and Medium Enterprises
TU	Technische Universtität
UK	United Kingdom
UNESCO	United Nations Educational, Scientific and Cultural Organization

Bibliography

Aghion, P., Bloom, N., Blundell, R., Griffith, R., & Howitt, P. (2005). Competition and Innovation: An Inverted-U Relationship. *The Quarterly Journal of Economics*, *120*(2), 701–728.

Agrawal, A., Gans, J., & Goldfarb, A. (2019). Economic Policy for Artificial Intelligence. *Innovation Policy and the Economy*, *19*(1), 139–159.

Arrow, K. J. (1972). Economic Welfare and the Allocation of Resources for Invention. In C. K. Rowley (Hrsg.), *Readings in Industrial Economics: Volume Two: Private Enterprise and State Intervention* (S. 219–236). Macmillan Education UK. https://doi.org/10.1007/978-1-349-15486-9_13

Baldwin, R., Cave, M., Lodge, M., Baldwin, R., Cave, M., & Lodge, M. (2011). *Understanding Regulation: Theory, Strategy, and Practice* (Second Edition, Second Edition). Oxford University Press.

Bathaee, Y. (2018). *The Artificial Intelligence Black Box and the Failure of Intent and Causation*. *31*.

Baumeister, R., & Leary, M. (1997). Writing Narrative Literature Reviews. *Review of General Psychology*, *1*, 311–320. https://doi.org/10.1037//1089-2680.1.3.311

Benkler, Y. (2002). Intellectual property and the organization of information production. *International Review of Law and Economics*, *22*(1), 81–107. https://doi.org/10.1016/S0144-8188(02)00070-4

Benton, D. C., Scheidt, L., & Guerrero, A. (2020). Regulating Disruptive Technologies: Oxymoron or Essential Evolution? *Journal of Nursing Regulation*, *11*(1), 24–28. https://doi.org/10.1016/S2155-8256(20)30057-0

Bessen, J., Impink, S. M., & Seamans, R. (2022). The Cost of Ethical AI Development for AI

Startups. *Proceedings of the 2022 AAAI/ACM Conference on AI, Ethics, and Society*, 92–106. https://doi.org/10.1145/3514094.3534195

Besti, F., & Samorè, F. (2018). *Responsibility driven design for the future self-driving society*.

Beyerlin, U., & Marauhn, T. (2011). *International Environmental Law*. Bloomsbury Publishing.

Blind, K., & Böhm, M. (2019, November 6). *The Relationship Between Open Source Software and Standard Setting*. JRC Publications Repository. https://doi.org/10.2760/163594

Blok, V., & Lemmens, P. (2015). The Emerging Concept of Responsible Innovation. Three Reasons Why It Is Questionable and Calls for a Radical Transformation of the Concept of Innovation. In *Responsible Innovation 2: Concepts, Approaches, and Applications* (Bd. 2, S. 19–35). https://doi.org/10.1007/978-3-319-17308-5_2

Boden, M., Bryson, J., Caldwell, D., Dautenhahn, K., Edwards, L., & Kemper, S. (2017). *Principles of robotics: Regulating robots in the real world*. https://www.tandfonline.com/doi/epdf/10.1080/09540091.2016.1271400?needAccess=true&role=button

Boldrin, M., & Levine, D. K. (2008). *Against Intellectual Monopoly*. Cambridge University Press. https://doi.org/10.1017/CBO9780511510854

Bostrom, N., & Yudkowsky, E. (2011). *The Ethics of Artificial Intelligence*.

Bresnahan, T. F., & Trajtenberg, M. (1995). General purpose technologies 'Engines of growth'? *Journal of Econometrics*, *65*(1), 83–108. https://doi.org/10.1016/0304-4076(94)01598-T

Brynjolfsson, E., & McAfee, A. (2014). *The second machine age: Work, progress, and prosperity in a time of brilliant technologies* (S. 306). W W Norton & Co.

Burrell, J. (2016). *How the machine 'thinks': Understanding opacity in machine learning algorithms.* https://journals.sagepub.com/doi/full/10.1177/2053951715622512

Butterworth, M. (2018). The ICO and artificial intelligence: The role of fairness in the GDPR framework. *Computer Law & Security Review*, *34*(2), 257–268. https://doi.org/10.1016/j.clsr.2018.01.004

Carabantes, M. (2019, April 12). *Black-box artificial intelligence: An epistemological and critical analysis | SpringerLink*. https://link.springer.com/article/10.1007/s00146-019-00888-w

Chatila, R., Firth-Butterflied, K., Havens, J. C., & Karachalios, K. (2017). The IEEE Global Initiative for Ethical Considerations in Artificial Intelligence and Autonomous Systems [Standards]. *IEEE Robotics & Automation Magazine*, *24*(1), 110–110. https://doi.org/10.1109/MRA.2017.2670225

Chesterman, S. (2021). *We, the Robots?: Regulating Artificial Intelligence and the Limits of the Law*. Cambridge University Press. https://doi.org/10.1017/9781009047081

Cihon, P., Maas, M. M., & Kemp, L. (2020). Should Artificial Intelligence Governance be Centralised? Design Lessons from History. *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*, 228–234. https://doi.org/10.1145/3375627.3375857

Clarke, S., & Whittlestone, J. (2022). A Survey of the Potential Long-term Impacts of AI: How AI Could Lead to Long-term Changes in Science, Cooperation, Power, Epistemics and Values. *Proceedings of the 2022 AAAI/ACM Conference on AI, Ethics, and Society*, 192–202. https://doi.org/10.1145/3514094.3534131

Collingridge, D. (1980). *The Social Control of Technology*. St. Martin's Press.

COMEST, U. W. C. on the E. of S. K. and T. (2005). *The Precautionary Principle*. https://unesdoc.unesco.org/ark:/48223/pf0000139578

Commission Report on safety and liability implications of AI, the Internet of Things and Robotics. (2020, Februar 20). https://commission.europa.eu/publications/commission-report-safety-and-liability-implications-ai-internet-things-and-robotics-0_en

COMMUNICATION FROM THE COMMISSION Artificial Intelligence for Europe, {SWD (2018) 137 final}, European Commission (2018). https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM%3A2018%3A237%3AFIN

Community framework for State aid for research and development., European commission, C 323/1. (2006). https://eur-lex.europa.eu/EN/legal-content/summary/community-framework-for-state-aid-for-research-and-development.html

Delua, J. (2021, März 12). Supervised vs. Unsupervised Learning: What's the Difference? *IBM Blog*. https://www.ibm.com/blog/supervised-vs-unsupervised-learning/

DiMatteo, L. A. (2013). Soft Law and the Principle of Fair and Equitable Decision Making in International Contract Arbitration. *The Chinese Journal of Comparative Law*, *1*(2), 221–255. https://doi.org/10.1093/cjcl/cxt013

Downes, L. (2009). *The Laws of Disruption: Harnessing the New Forces that Govern Life and Business in the Digital Age*. Basic Books.

El-Gazzar, R., & Stendal, K. (2020). *Examining How GDPR Challenges Emerging Technologies*. *10*, 237–275. https://doi.org/10.5325/jinfopoli.10.2020.0237

Ellul, J., Pace, G., McCarthy, S., Sammut, T., Brockdorff, J., & Scerri, M. (2021). Regulating artificial intelligence: A technology regulator's perspective. *Proceedings of the Eighteenth International Conference on Artificial Intelligence and Law*, 190–194. https://doi.org/10.1145/3462757.3466093

Erdélyi, O. J., & Goldsmith, J. (2018). Regulating Artificial Intelligence: Proposal for a Global Solution. *Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society*, 95–101. https://doi.org/10.1145/3278721.3278731

Gallup and Northeastern University. (2018). *Optimism and Anxiety: Views on the Impact of Artificial Intelligence*. Northeastern University, GALLUP. https://camd.northeastern.edu/visualization/optimism-and-anxiety-views-on-the-

impact-of-artificial-intelligence/

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

Gold, S. (2017, November). *Removing the Pebbles in the Regulatory Stream*. IndustryWeek. https://www.industryweek.com/the-economy/regulations/article/22024652/removing-the-pebbles-in-the-regulatory-stream

Grace, K., Salvatier, J., Dafoe, A., Zhang, B., & Evans, O. (2018). *When Will AI Exceed Human Performance? Evidence from AI Experts* (arXiv:1705.08807). arXiv. http://arxiv.org/abs/1705.08807

Gunningham, N., & PN, G. (1998). *Smart Regulation: Designing Environment Policy*.

Gupta, S. (2022). The interaction between technology, business environment, society, and regulation in ICT industries. *IIMB Management Review*, *34*(2), 103–115. https://doi.org/10.1016/j.iimb.2022.07.001

Gurkaynak, G., Yilmaz, I., & Haksever, G. (2016). Stifling artificial intelligence: Human perils. *Computer Law & Security Review*, *32*(5), 749–758. https://doi.org/10.1016/j.clsr.2016.05.003

Harlow, C., & Rawlings, R. (2009). *Law and Administration* (3. Aufl.). Cambridge University Press. https://doi.org/10.1017/CBO9780511809941

Hashmi, A. R. (2013). Competition and Innovation: The Inverted-U Relationship Revisited. *The Review of Economics and Statistics*, *95*(5), 1653–1668.

Hemphill, T. A. (2020). "The innovation governance dilemma: Alternatives to the precautionary principle". *Technology in Society*, *63*, 101381. https://doi.org/10.1016/j.techsoc.2020.101381

Herian, R. (2018). *Regulating Disruption: Blockchain, Gdpr, and Questions of Data Sovereignty* (SSRN Scholarly Paper 3241228).

https://papers.ssrn.com/abstract=3241228

Hilton, A. D. (2019). Artificial intelligence: The societal responsibility to inform, educate, and regulate. *AI Matters*, *5*(3), 70–76. https://doi.org/10.1145/3362077.3362088

Investigation into the use of data analytics in political campaigns. (2018). Information Commissioners Office.

J. Mc Carthy, N. Rochester, M. Minsky, & C. Shannon. (1955). *A PROPOSAL FOR THE DARTMOUTH SUMMER RESEARCH PROJECT ON ARTIFICIAL INTELLIGENCE | BibSonomy*. http://www-formal.stanford.edu/jmc/history/dartmouth/dartmouth.html

Jessica C. Newman. (2019). *Toward AI Security: Global Aspirations for a More Resilient Future*. CLTC UC Berkeley Center for Long-Term Cybersecurity. https://cltc.berkeley.edu/towardaisecurity/

Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence*, *1*. https://doi.org/10.1038/s42256-019-0088-2

Jones, B. C. (1975). Prohibition and Prosperity, 1920-1930. *Social Science*, *50*(2), 78–86.

Jovanovic, B., & Rousseau, P. L. (2005). Chapter 18—General Purpose Technologies. In P. Aghion & S. N. Durlauf (Hrsg.), *Handbook of Economic Growth* (Bd. 1, S. 1181–1224). Elsevier. https://doi.org/10.1016/S1574-0684(05)01018-X

Karim, A. (2022, Oktober 17). Tech Scaleup Silicon Valley. *Mind the Bridge*. https://mindthebridge.com/tech-scaleup-silicon-valley-report-2022/

Kazim, E., Almeida, D., Kingsman, N., Kerrigan, C., Koshiyama, A., Lomas, E., & Hilliard, A. (2021). Innovation and opportunity: Review of the UK's national AI strategy. *Discover Artificial Intelligence*, *1*(1). https://doi.org/10.1007/s44163-021-00014-0

Kemppainen, L., Pikkarainen, M., Hurmelinna-Laukkanen, P., & Reponen, J. (2019). Connected Health Innovation: Data Access Challenges in the Interface of AI Companies and Hospitals. *Technology Innovation Management Review*, *9*(12), 43–55.

Kim, D., Lee, H., & Kwak, J. (2017). Standards as a driving force that influences emerging technological trajectories in the converging world of the Internet and things: An investigation of the M2M/IoT patent network. *Research Policy*, *46*(7), 1234–1254. https://doi.org/10.1016/j.respol.2017.05.008

Klapper, L., Laeven, L., & Rajan, R. (2006). Entry regulation as a barrier to entrepreneurship. *Journal of Financial Economics*, *82*(3), 591–629. https://doi.org/10.1016/j.jfineco.2005.09.006

Kuhn, T. S. (2012). *The Structure of Scientific Revolutions*. University of Chicago Press.

Kurzweil, R. (1990). *The Age of Intelligent Machines*. MIT Press.

LAYING DOWN HARMONISED RULES ON ARTIFICIAL INTELLIGENCE (ARTIFICIAL INTELLIGENCE ACT) AND AMENDING CERTAIN UNION LEGISLATIVE ACTS, {SEC(2021) 167 final}-{SWD(2021) 84 final}-{SWD(2021) 85 final}, European Commission (2021). https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A52021PC0206

LeCun, Y., Boser, B., Denker, J. S., Henderson, D., Howard, R. E., Hubbard, W., & Jackel, L. D. (1989). Backpropagation Applied to Handwritten Zip Code Recognition. *Neural Computation*, *1*(4), 541–551. https://doi.org/10.1162/neco.1989.1.4.541

Lee, S., & Trimi, S. (2016). Innovation for creating a smart future. *Journal of Innovation & Knowledge*, *3*. https://doi.org/10.1016/j.jik.2016.11.001

Lemley, M. A. (2007). *RATIONALIZING INTERNET SAFE HARBORS*. *6*.

Li, X., & Zhang, T. (2017). An exploration on artificial intelligence application: From security, privacy and ethic perspective. *2017 IEEE 2nd International Conference on Cloud Computing and Big Data Analysis (ICCCBDA)*, 416–420. https://doi.org/10.1109/ICCCBDA.2017.7951949

Maas, M. 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*, 223–228. https://doi.org/10.1145/3278721.3278766

Mandel, M., & Carew, D. G. (2013). *Regulatory Improvement Commission:*

McCulloch, W. S., & Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *The Bulletin of Mathematical Biophysics*, *5*(4), 115–133. https://doi.org/10.1007/BF02478259

McKinsey Technology Trends Outlook 2022 | McKinsey. (2023). https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/the-top-trends-in-tech

Meszaros, J., & Ho, C. (2021). AI research and data protection: Can the same rules apply for commercial and academic research under the GDPR? *Computer Law & Security Review*, *41*, 105532. https://doi.org/10.1016/j.clsr.2021.105532

Miller, F., Katz, J., & Gans, R. (2018). *AI x I = AI2: The OD imperative to add inclusion to the algorithms of artificial intelligence*. *50*.

Nilsson, N. (2010). *The quest for artificial intelligence: A history of ideas and achievements*. https://doi.org/10.1017/CBO9780511819346

Porter, M. E. (1990). *The competitive advantage of nations*. Free Press.

Powles, J., & Hodson, H. (2017). Google DeepMind and healthcare in an age of algorithms. *Health and Technology*, *7*(4), 351–367. https://doi.org/10.1007/s12553-017-0179-1

Prem, E. (2019). Artificial Intelligence for Innovation in Austria. *Technology Innovation Management Review*, *9*(12), 5–15.

Priestley, M., & Simperl, E. (2022). Open innovation programmes related to data and AI: How do the entrepreneurial orientations of startups align with the objectives of public funders? *Data & Policy*, *4*, e16. https://doi.org/10.1017/dap.2022.8

Reed, C. (2018). How should we regulate artificial intelligence? *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *376*(2128), 20170360. https://doi.org/10.1098/rsta.2017.0360

Regulatory framework proposal on artificial intelligence | Shaping Europe's digital future. (2023, März 24). https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai

Regulatory Reduction Efforts in Ohio: Can It Work? | INSIDE COMPLIANCE. (2022, April 21). https://blogs.luc.edu/compliance/?p=4647

Resolution on the mandate of the Group of Advisers on the Ethical Implications of Biotechnology to the European Commission, European Commission, B4-0484/97 (1997). https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:51997IP0484

Rosenblatt, F. (1958). The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review*, *65*(6), 386–408. https://doi.org/10.1037/h0042519

Schomberg, R. (2013). *A vision of responsible innovation*.

Schumpeter, J. A. (1934). *The Theory of Economic Development: An Inquiry into Profits, Capital, Credit, Interest, and the Business Cycle*. Harvard University Press.

Shapiro, C. (2019). Protecting Competition in the American Economy: Merger Control, Tech Titans, Labor Markets. *Journal of Economic Perspectives*, *33*(3), 69–93. https://doi.org/10.1257/jep.33.3.69

Snyder, H. (2019). Literature review as a research methodology: An overview and guidelines. *Journal of Business Research*, *104*, 333–339. https://doi.org/10.1016/j.jbusres.2019.07.039

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

https://doi.org/10.1016/j.respol.2013.05.008

Stone, P., Brooks, R., Brynjolfsson, E., Calo, R., Etzioni, O., Hager, G., Hirschberg, J., Kalyanakrishnan, S., Kamar, E., Kraus, S., Leyton-Brown, K., Parkes, D., Press, W., Saxenian, A., Shah, J., Tambe, M., & Teller, A. (2022). *Artificial Intelligence and Life in 2030: The One Hundred Year Study on Artificial Intelligence* (arXiv:2211.06318). arXiv. https://doi.org/10.48550/arXiv.2211.06318

Sunstein, C. R. (2002). The Paralyzing Principle. *Regulation*, *25*, 32.

Susar, D., & Aquaro, V. (2019). Artificial Intelligence: Opportunities and Challenges for the Public Sector. *Proceedings of the 12th International Conference on Theory and Practice of Electronic Governance*, 418–426. https://doi.org/10.1145/3326365.3326420

Tatjana Evas. (2020). *Civil liability regime for artificial intelligence* [Study]. think Tank European Parlaiment. https://www.europarl.europa.eu/thinktank/en/document/EPRS_STU(2020)654178

Thierer, A. (2014). *Permissionless Innovation: The Continuing Case for Comprehensive Technological Freedom* (1st Aufl.). Mercatus Center at George Mason University.

Truby, J., Brown, R. D., Ibrahim, I. A., & Parellada, O. C. (2022). A Sandbox Approach to Regulating High-Risk Artificial Intelligence Applications. *European Journal of Risk Regulation*, *13*(2), 270–294. https://doi.org/10.1017/err.2021.52

Truby, J., & Fenster), L. zu externer W. (Anzeige in neuem. (2020). Governing Artificial Intelligence to benefit the UN Sustainable Development Goals. *Sustainable Development*, *28*(4), 946–959. https://doi.org/10.1002/sd.2048

Truby, J., Fenster), L. zu externer W. (Anzeige in neuem, Brown, R., Fenster), L. zu externer W. (Anzeige in neuem, & Dahdal, A. (2020). Banking on AI: Mandating a proactive approach to AI regulation in the financial sector. *Law and Financial Markets Review*,

14(2), 110–120. https://doi.org/10.1080/17521440.2020.1760454

Truby, J., & Kratsas, G. (2017). VW's 'Defeat Devices' and Liability for Claims for Lost Emissions Tax Revenue. *Global Journal of Comparative Law*, *6*, 1–24. https://doi.org/10.1163/2211906X-00601001

Turner, J. (2019). *Robot Rules: Regulating Artificial Intelligence*. Springer International Publishing. https://doi.org/10.1007/978-3-319-96235-1

Valdivia, W. D., & Guston, D. H. (2015). Responsible innovation: A primer for policymakers. *Center for Technology Innovation at Brookings*.

Vallverdú, J. (2014). Governance, Regulation and Innovation: Theory and Evidence from Firms and Nations edited by Mehmet Ugur. *Science and Public Policy*, *41*(5), 697–698. https://doi.org/10.1093/scipol/scu045

van Dijk, N., Casiraghi, S., & Gutwirth, S. (2021). The 'Ethification' of ICT Governance. Artificial Intelligence and Data Protection in the European Union. *Computer Law & Security Review*, *43*, 105597. https://doi.org/10.1016/j.clsr.2021.105597

Vinuesa, R., Azizpour, H., Leite, I., Balaam, M., Dignum, V., Domisch, S., Felländer, A., Langhans, S. D., Tegmark, M., & Fuso Nerini, F. (2020). The role of artificial intelligence in achieving the Sustainable Development Goals. *Nature Communications*, *11*(1), Article 1. https://doi.org/10.1038/s41467-019-14108-y

Weitzman, M. (1998). Recombinant Growth. *Quarterly Journal of Economics*, *113*(2), 331–360.

What are Neural Networks? | IBM. (2021). https://www.ibm.com/topics/neural-networks

What is Supervised Learning? | IBM. (2021). https://www.ibm.com/topics/supervised-learning

What is Unsupervised Learning? | IBM. (2021). https://www.ibm.com/topics/unsupervised-learning

White Paper on Artificial Intelligence: A European approach to excellence and trust. (2020,

Februar 19). https://commission.europa.eu/publications/white-paper-artificial-intelligence-european-approach-excellence-and-trust_en

Winecoff, A. A., & Watkins, E. A. (2022). Artificial Concepts of Artificial Intelligence: Institutional Compliance and Resistance in AI Startups. *Proceedings of the 2022 AAAI/ACM Conference on AI, Ethics, and Society*, 788–799. https://doi.org/10.1145/3514094.3534138

Wisuttisak, P. (2020). *Comparative Study on Regulatory and Policy Frameworks for Promotion of Startups and SMEs in Japan, the Republic of Korea, Malaysia, and Thailand*.

Wong, G., Greenhalgh, T., Westhorp, G., Buckingham, J., & Pawson, R. (2013). RAMESES publication standards: Meta-narrative reviews. *BMC Medicine*, *11*(1), 20. https://doi.org/10.1186/1741-7015-11-20

Wright, G. (2000). General Purpose Technologies and Economic Growth. In *Journal of Economic Literature* (Bd. 38, Nummer 1, S. 161–162). American Economic Association. https://www.proquest.com/docview/213185950/abstract/4826D6320F564563PQ/1

Wu, T. (2018). After Consumer Welfare, Now What? The „Protection of Competition" Standard in Practice. *Competition Policy International, 2018; Columbia Public Law Research Paper No. 14-608*. https://scholarship.law.columbia.edu/faculty_scholarship/2291

Zarsky, T. (2013). *Transparent Predictions* (SSRN Scholarly Paper 2324240). https://papers.ssrn.com/abstract=2324240

Zech, H. (2021). Liability for AI: Public policy considerations. *ERA Forum*, *22*(1), 147–158. https://doi.org/10.1007/s12027-020-00648-0

Zetzsche, D. A., Arner, D. W., Buckley, R. P., & Tang, B. (2020). *Artificial Intelligence in Finance: Putting the Human in the Loop* (SSRN Scholarly Paper 3531711).

https://papers.ssrn.com/abstract=3531711

Zhang, B., & Dafoe, A. (2020). U.S. Public Opinion on the Governance of Artificial Intelligence. *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*, 187–193. https://doi.org/10.1145/3375627.3375827