universität wien

MASTER THESIS

Titel der Master Thesis / Title of the Master's Thesis

„Artificial Intelligence: A Comparison of EU and US Regulation“

verfasst von / submitted by

Tina Fokter

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Master of Laws (LL.M.)

Wien, 2022 / Vienna 2022

Studienkennzahl lt. Studienblatt / Postgraduate programme code as it appears on the student record sheet:	UA 992 548
Universitätslehrgang lt. Studienblatt / Postgraduate programme as it appears on the student record sheet:	Europäisches und Internationales Wirtschaftsrecht / European and International Business Law
Betreut von / Supervisor:	Dr. Roland Vogl

Abstract

In light of recent trends in AI development, many countries faced the potential implications of this new technology and decided to increasingly regulate it. Consequently, the EU decided to for a harmonized approach and spent years developing the concepts of human-centric and trustworthy AI, which finally culminated in the Draft Artificial Intelligence Act in April 2021. While the EU strives to be the leader in AI regulation, the US also showed its willingness to regulate AI. However, the US decided for a laissez-faire approach that particularly underlines the importance of innovation which could be stifled by over regulation. Hence, the US so far relied more on an agency-based approach as well as adapting the already existing rules and shaping them so that they also fit the new AI requirements. Unlike EU, the US did not decide for a federal AI regulation approach. The risk in both countries is internal fragmentation if new measures are not introduced. This could lead to uneven protection or treatment of the citizens.

Keywords: artificial intelligence, regulation, AI Act

Abstrakt

Angesichts der neuesten Trends in der KI-Entwicklung haben sich viele Länder mit den potenziellen Auswirkungen dieser neuen Technologie auseinandergesetzt und beschlossen, sie zunehmend zu regulieren. Infolgedessen entschied sich die EU für einen harmonisierten Ansatz und verbrachte Jahre damit, die Konzepte der menschenzentrierten und vertrauenswürdigen KI zu entwickeln, was schließlich im April 2021 im Entwurf des Gesetzes über künstliche Intelligenz gipfelte. Während die EU danach strebt, bei der Regulierung von KI führend zu sein, haben auch die USA ihre Bereitschaft zur Regulierung von KI gezeigt. Die USA haben sich jedoch für einen Laissez-faire-Ansatz entschieden, der insbesondere die Bedeutung der Innovation unterstreicht, die durch eine Überregulierung abgewürgt werden könnte. Daher setzten die USA bisher eher auf einen behördenbasierten Ansatz sowie auf die Anpassung der bereits bestehenden Vorschriften und deren Ausgestaltung, damit sie auch den neuen KI-Anforderungen gerecht werden. Im Gegensatz zur EU haben sich die USA nicht für einen föderalen KI-Regulierungsansatz entschieden. In beiden Ländern besteht die Gefahr einer internen Fragmentierung, wenn keine neuen Maßnahmen eingeführt werden. Dies könnte zu einem ungleichen Schutz oder einer ungleichen Behandlung der Bürger führen.

Schlüsselwörter: künstliche Intelligenz, Verordnung, KI-Gesetz

Table of Contents

1. Introduction *1*

2. Artificial Intelligence *3*

2.1 Definition: What is AI? **3**

2.2 Ethical and legal implications of AI **6**

2.2.1 Data bias 6

2.2.2 Data abuse 7

2.2.3 Lack of transparency 8

2.2.4 Undermining human autonomy 8

2.2.5 Insufficient trust and human oversight 8

2.2.6 AI accountability 9

2.2.7 Privacy 9

3. Regulation of artificial intelligence *10*

3.1 General **10**

3.2 The European Union **11**

3.2.1 Evolution of the AI regulation idea in the EU 12

3.2.2 Further efforts 14

3.2.3 Artificial Intelligence Act (AIA) 16

3.3 The United States **30**

3.3.1 Federal level 30

3.3.2 State level 37

4. Conclusion *40*

5. Bibliography *42*

Table of Abbreviations

ADM:	Automated Decision-Making technologies
AI:	Artificial Intelligence
AIA:	Artificial Intelligence Act
IPAI:	International Panel on Artificial Intelligence
EU:	European Union
EP:	European Parliament
GDPR:	General Data Protection Regulation
HLEG:	High Level Expert Group
MSA:	Market Surveillance Authority
US:	United States of America

1. Introduction

The term artificial intelligence is not new. In fact, it has been around since the 1950s. However, AI gained more popularity in the recent years. From robot vacuum cleaners to automated vehicles and sci-fi TV series, it is almost impossible to escape it in one form or another.

The perks of AI are immense. Requiring no breaks and never getting tired, AI has the potential to completely change our view of the workspace. The way AI can analyze enormous chunks of data is unprecedented and unmatched. The fact that AI, if trained properly, can be incomparably more accurate than humans might even seem scary at times. Consequently, AI already is and is predicted to be used in a large number of sectors, which will completely change our way of life. There are big words used when talking about the future and AI, such as a new revolution and AI's disruptive effects. Fact of the matter is that AI can be used in important sectors and can greatly contribute to our development. For example, there are numerous benefits of AI in healthcare and medicine. It can be used to reduce human error and diagnose illnesses faster than any doctor. Countless cases show how effective this technology can be, for example, in diagnosing cancer much faster and more accurately than doctors.[1] In banking, AI can prove useful in areas such as risk management, credit score calculations or managing revenue. There have been significant improvements in the automated vehicle or self-driving automobiles sector, a large part of the progress is due to AI advancements. Moreover, chatbots or virtual assistants have become ordinary in our daily life, and many people rely on their virtual assistants, such as Siri or Alexa.

However, as always, there are also downsides to any new and disruptive technology. One of the majorly discussed disadvantages is the ethics problem with AI. This is also the reason why so many countries focused on ethics and principles in their initiatives and desire to regulate AI. The ethics problem can be further broken down into more concrete cases. For example, AI poses a very real threat to privacy. If we take the example of facial recognition, the concept of mass surveillance that follows such widespread facial recognition without people being aware of it can paint quite a daunting image. There have already been examples of countries using facial recognition systems for political imprisonment, such as arresting people that were

[1] Babak Ehteshami Bejnordi, Mitko Veta, Paul Johannes van Diest, et al., 'Diagnostic Assessment of Deep Learning Algorithms for Detection of Lymph Node Metastases in Women With Breast Cancer' (2017) JAMA 318 <https://jamanetwork.com/journals/jama/fullarticle/2665774> accessed 29 June 2022.

protesting.[2] Furthermore, another large downside is AI bias, which can cause, for example, countless CVs to be overlooked in job application processes.

Accordingly, more and more countries have taken the step of recognizing AI as the inevitable future, in whatever way it may come, and preparing themselves accordingly. After years and years of discussions, guidelines and principles, the EU finally came up with a comprehensive draft of Artificial Intelligence Act in 2021. On the other hand, the US has thus far still not expressed any real progress in AI regulation, even though it is currently the leader in AI development.[3]

This thesis examines the notion of AI and the development or lack of regulation in EU and USA. To aid in this endeavor, it employs a comparative law approach in order to seek a deeper understanding of what such regulations means for the world, the EU and the USA legal systems.

The thesis progresses from discussing AI and the ethical and legal implications it may cause to a closer look at EU regulation. From the very beginnings, the EU focused mostly on ethics and human-centered approach, specifically the notion of human oversight. In this regard, the thesis starts from the beginnings of AI regulations, consisting mostly of guidelines and principles in EU, and culminates with the Draft Artificial Intelligence Act, which is analyzed in depth.

Next, the thesis turns to US regulation. Having a more fragmented legislation, the US still has no unified approach to AI regulation on a federal level up to date. However, some states, such as California, have adopted specific laws that are discussed in the third part of the thesis and compared with similar laws in the EU.

---

[2] Gleb Stolyarov, Gabrielle Tétrault-Farber, '"Face Control": Russian police go digital against protesters' (*Reuters*, 11 February 2021) <https://www.reuters.com/article/us-russia-politics-navalny-tech-idUSKBN2AB1U2> accessed 26 June 2022.

[3] Daniel Castro, Michael McLaughlin, Eline Chivot, 'Who is winning the AI race: China, the EU or the United States?' (2019) <https://datainnovation.org/2019/08/who-is-winning-the-ai-race-china-the-eu-or-the-united-states/> accessed 26 June 2022.

2. Artificial Intelligence

2.1 Definition: What is AI?

The term artificial intelligence was coined and firstly used in 1956 by John McCarthy, who defined AI as "the science and engineering of making intelligent machines".[4] Most commonly, AI refers to a program that can mimic the thought processes of the human brain. A common connection with AI is the Turing test, formed in 1950 by Alan Turing, who came up with a test that tried to determine whether a machine is able to imitate human intelligence.[5] Even though it faced many criticisms, the test proved important for the future of AI.

However, the definition of artificial intelligence is a highly contested concept that has proven very difficult to define. The technological advances are progressing so fast that no definition has remained the same for long. In digital governance glossary, AI is defined as "the ability of machines to perform tasks that would normally be attributed to people".[6] Others added the ability of the machines to act, plan, and even reason to this broad definition.[7] Rich, for example, defined AI as "the study of how to make computers to do things at which, at the moment, people are better".[8]

Interestingly, the term artificial intelligence has been connected to the word robot. However, "robot" is also an elusive word, used to describe many entities, from superhumanly intelligent software systems to simple pneumatic machines.[9] Some authors even argue that in the future, both robots and people could be understood as hybrids and their meaning will lie somewhere in between.[10] For example, self-driving cars developed by human beings today are not truly driving themselves but have some, or even major, input from the people sitting behind the wheel. Even though such vehicles are autonomous in many ways, they are still far from what

---

[4] Christopher Manning, 'Artificial Intelligence Definitions' (*Stanford University Human-Centered Artificial Intelligence*, 2020) <https://hai.stanford.edu/sites/default/files/2020-09/AI-Definitions-HAI.pdf> accessed 30 June 2022.
[5] Adam Turing, 'Computing machinery and intelligence' (1950) 59(236) Mind 433 http://www.jstor.org/stable/2251299 accessed 30 Jun 2022.
[6] Jeremy Swinfen Green, Stephen Daniels, *Digital Governance: Leading and Thriving in a World of Fast-Changing Technologies* (Routeledge 2019) 132.
[7] Davide Carneiro, Paulo Novais, Jose Neves, *Conflict Resolution and its Context: From the Analysis of Behavioural Patterns to Efficient Decision-Making* (Springer 2020) 32.
[8] Elaine Rich, 'Artificial Intelligence and the Humanities', (1985) 19(2) Computers and the Humanities 117 <http://www.jstor.org/stable/30204398> accessed 30 June 2022.
[9] Mark Lemley, Brian Casey, 'You might be a robot' (2020) Cornell Law Review 7 <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3327602> accessed 30 June 2022.
[10] Mark Lemley, Brian Casey, 'You might be a robot' (2020) Cornell Law Review 5 <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3327602> accessed 30 June 2022.

one usually imagines under the term "robot". On the other hand, it is already hard to state that human drivers are completely autonomous because they rely on the functions that the technology has provided them with, such as maps, apps and various electronic systems that help them drive safely.[11]

As demonstrated, AI is still a moving target that changes frequently, also because of recent trends in technology. Even though not one unanimous definition is in place, all of them have an underlying theme: machines that are able to do things normally attributable to people. Additionally, the definition varies depending on the sector of use because of the magnitude of fields with the potential of benefiting from AI.

Furthermore, there are differences between symbolic (rules-driven) AI and statistical (machine learning-based) AI. Symbolic AI was prominent in the first decades of conception and is also known under the term GOFAI, which stands for Good Old Fashioned AI. ML-based AI became more popular in recent years. GOFAI or symbolic AI uses "human-readable symbols that represent real-world entities or concepts as well as logic (the mathematically provable logical methods) in order to create rules for the concrete manipulation of those symbols, leading to a rule-based system".[12] In other words, GOFAI uses models that represent concepts from the real world. The machine can make decisions based on the facts it memorized. The rules that the machine follows in order to reach a conclusion are put together by logic. On the other hand, ML systems learn from a large database in order to come to a solution that is sufficient (e.g., with 99% accuracy). The machines are provided with large amounts of data, from which they themselves learn patterns and logic. Consequently, GOFAI models are cited as being more transparent because they are understandable and explainable, whereas, as demonstrated in the examples below, ML can often produce results that are not explainable by humans, even those working with the machines. For example, should AI be used in legal fields, such as a machine advising judges and lawyers, the GOFAI version would mean that a machine is analyzing the rules and using legal reasoning that humans "taught" it, or, in the case of ML based AI, the machine could be fed with large amount of data and would employ an algorithm to seek what the rules are and then apply them.

---

[11] *ibid*.
[12] Estelle Verani, 'Symbolic AI vs Machine Learning in natural language processing' (*Inbenta*, 4 March 2020) <https://www.inbenta.com/en/blog/symbolic-ai-vs-machine-learning/> accessed 30 June 2022.

To further complicate things, the terms machine learning, deep learning and neural networks are closely related to AI and, in some cases, even used synonymously. However, they all have a different meaning. Machine learning is a subfield of AI and signifies systems that learn from data. By doing so, the performance of the machine improves. Some AI definitions consider machine learning as a fundamental part of AI. For example, the European Parliament defines AI as "a combination of machine learning techniques used for searching and analyzing large volumes of data; robotics dealing with the conception, design, manufacture and operation of programmable machines; and algorithms and automated decision-making systems (ADMS) able to predict human and machine behavior and to make autonomous decisions."[13] Apparently, the EP did not consider GOFAI under their AI definition.

Furthermore, the name neural networks comes from the similarity with the human neural system, by which the concept was inspired. The process essentially signifies a series of algorithms that try to recognize underlying relationships in a set of data.[14]

Neural networks are widely used for facial or speech recognition and are the main pillar of deep learning. A neural network with more than three layers can be considered as a deep learning algorithm.[15] Deep learning is also a machine learning technique. Through deep learning, a computer can learn directly from images, sound or text. Computers learn by using big sets of data and neural network architectures.

While the efforts in defining AI are still ongoing, finding the definition is only part of the issue. There might not even be such a thing as the correct definition of AI due to rapid changes in technology, which cause every definition to already be underinclusive or irrelevant at the time of conception. A better answer might therefore be to come to terms with the fact that the exact definition will remain elusive, but through case-by-case approach, the courts and regulators can build and accordingly adapt the definitions in time, when they seem to match the reality.[16] Most importantly, a bad definition should not cause a bad outcome.

---

[13] European Parliament, 'EU Guidelines on ethics in artificial intelligence: Contex and implementation' (2019) 2 <https://www.europarl.europa.eu/RegData/etudes/BRIE/2019/640163/EPRS_BRI(2019)640163_EN.pdf> accessed 30 June 2022.

[14] James Chen, 'Neural Network' (*Investopedia*, 8 December 2021) <https://www.investopedia.com/terms/n/neuralnetwork.asp> accessed 30 June 2022.

[15] Eda Kavlakoglu, 'AI vs Machine Learning vs Deep Learning vs Neural Networks: What is the difference?' (*IBM*, 27 May 2020) <https://www.ibm.com/cloud/blog/ai-vs-machine-learning-vs-deep-learning-vs-neural-networks> accessed 30 June 2022.

[16] Mark Lemley, Brian Casey, 'You might be a robot' (2020) Cornell Law Review 5 <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3327602> accessed 30 June 2022.

2.2 Ethical and legal implications of AI

As will be discussed in the following chapter, regulation (or a lack thereof) has recently proved to be a difficult and comprehensive issue worldwide. Different countries have taken different approaches to the problem of AI regulation. Some countries are still in the phase of asking themselves whether it is even sensible to regulate AI, while others have taken major advances, especially in the last year. Nonetheless, there are common principles and values that were considered in both more and less regulated legal systems concerning use and development of AI systems. Arguably, ethics are discussed as much if not more than the law itself when it comes to AI. This has led to critics stating "ethics-washing" as one of the reasons for omitting regulation.[17] However, ethics is given the role of an AI navigator in development and regulation, making sure that AI development as well as the law remain in a certain direction that is still acceptable. Nevertheless, ethics and guidelines are not legally binding unless there are certain laws passed next to or in addition to the supporting ethical bases. Therefore, only vague principles with no foundation might be of little practical use. What is more, there is no way to enforce such rules in a court of law, which makes the effect of ethics guidelines severely lacking. However, some of the most common ethical issues found in the majority of organizations considering AI implications and regulation are the following: data abuse, data bias, lack of transparency, undermining human autonomy, insufficient human oversight, inadequate accountability, and privacy issues.

2.2.1 Data bias

Every AI system needs something to learn or train from because AI cannot learn about something it has never seen. What the AI system consequently produces is based on what it learns from. Therefore, when a system is learning from such a large data pool, the most common vulnerability is data bias because a system can be trained on flawed, unrepresentative data or even data that is biased. For example, when a self-driving vehicle is learning to recognize traffic lights, it learns that red light means it must stop; hence, it would not be able to understand if suddenly a green light would signify it has to stop because the system only learned from data that classifies red light as a prohibition.

---

[17] Emre Bayamlioglu, Irina Baraliuc, Liisa Janssens, Mireille Hildebrandt, *Being profiled: Cogitas ergo sum: 10 years of profiling the European citizen* (Amsterdam University Press 2018) 84–89.

Therefore, the machine learning training data needs to be carefully monitored. There are numerous instances where AI showed significant data bias. For example, Google's image recognition algorithms categorized black people as gorillas because the data was trained with white people.[18] Moreover, there are instances of data trained with predominantly male population, which can cause an algorithm to discriminate against women. For example, Amazon's AI recruiting tool that was reviewing job applications was discovered to be doing so. Candidates applying for certain jobs, especially in the technical sector, were rejected if they were female because the algorithm was trained to observe patterns in previously submitted resumes. Since traditionally men were applying for technical positions, Amazon's system taught itself to favor male candidates. Accordingly, some words in the resumes, such as "women's", were automatically downgrading the resumes and never giving the female candidates a chance. [19]

Lastly, the bias could also come from the person developing the system. Since the developers are the ones providing the learning data pool to the system, their personal biases can affect the provided data.

2.2.2 Data abuse

Every AI system requires a fair amount of learning, which means that a big amount of data needs to be provided in order for an AI system to function. Countless samples need to be collected and analyzed form large datasets, which then make it possible, for example, for a driverless car to recognize a pedestrian or a stop sign. However, these datasets are often linked and exchanged with one another, also from different companies, for example to identify and reidentify individuals in biometrics. These big amounts of data can lead to significant data abuse, where individuals do not know about or do not give consent for their data processing. Additionally, data abuse is closely linked to privacy issues. Furthermore, data abuse goes hand in hand with data security and potential cyberattacks on these data, forming the base of another set of related issues.

---

[18] James Vincent, 'Google fixed its racist algorithm by removing gorillas from its image-labeling tech' (*The Verge* 12 January 2018) <https://www.theverge.com/2018/1/12/16882408/google-racist-gorillas-photo-recognition-algorithm-ai> accessed 30 June 2022.

[19] Jeffrey Dastin, 'Amazon scraps secret AI recruiting tool that showed bias against women' (Reuters, 11 October 2018) <https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G> accessed 30 June 2022.

2.2.3 Lack of transparency

The AI transparency problem is twofold. First of all, transparency is connected to the general public and people not understanding how an AI system reached a certain decision that directly affected them. For example, as demonstrated above, an AI system can give biased results. Consequently, a person can be rejected from a job position. Therefore, numerous guidelines and ethical principles state AI transparency as a principal obligation in AI development, requiring AI to reach transparent decisions that are overseen by humans. Some guidelines also underline the importance of the end user being aware and in possession of such data, claiming that they should be given a chance to know if they were rejected based on a rating conducted by a person or a machine.

Furthermore, transparency can also be a problem for the developers of the AI system. In some cases, not even developers understand the system's output and how it reached a certain conclusion. Some guidelines therefore place human oversight and overruling ability to the front to mitigate this issue.

2.2.4 Undermining human autonomy

AI systems can personalize certain types of information. Hence, it is possible for a system to make sure special data is only showed to some people based on specific characteristics. Such a characteristic can be, for example, the ability to pay. Consequently, an AI system can show people with certain financial background personalized shopping information. This can affect human behavior and limit human choice, as well as decrease information diversity.[20]

2.2.5 Insufficient trust and human oversight

The objective to create trustworthy AI systems is one of the key drivers behind AI regulation and closely connects to human oversight. Trustworthy AI means that there needs to be human oversight available and present at all times of development and later on. This signifies that a human can review automated decisions, ensuring there is a way to remedy potential errors in results. Furthermore, many documents on AI ethics underline the importance of opting out, where people are given the opportunity to decide to not be subject to decisions made by AI

[20] Michiel Fierens, Stephanie Rossello, Ellen Wauters, *Artificial Intelligence and the Law* (Cambridge University Press 2021) 53.

systems. This notion has mainly been discussed in the UK.[21] The most common principle, however, is the principle of human control. This means that AI system is designed in a way that always enables a human to intervene in the process.

2.2.6 AI accountability

AI accountability principles mostly address regulations and laws that would ensure responsibility, having an appropriate governing body monitoring advances in the sector, as well as legal liability and responsibility. For example, some strategies identify an accountable system as a system that is able to effectively prevent distortion, discrimination, manipulation, and other forms of improper use.[22] Furthermore, by creating a monitoring body, such a standard would ensure some sort of overseeing standards and a body that would hold system designers responsible in cases of error. Finally, legal responsibility principles signify that an AI system causing harm could be held accountable.

2.2.7 Privacy

Especially due to large amounts of data behind it, AI is commonly used in areas such as surveillance, advertising, and similar. Consequently, privacy is considerably impacted by such use, not only directly by these fields, but also in development and training of the AI systems. Most of the guidelines and ethical principles place a strong importance to consent, meaning that a person's data should never be used without explicit consent. Furthermore, it is also expressed that every data subject should have some form of control over the use of their data. However, these notions vary by countries. The EU is strongly covered in this sense, especially compared to other countries, where there have even been talks of redefining what consent means, in order to exploit all the possibilities of AI development.[23] There is also increasingly more talk about privacy by design, which would essentially require AI developers and

---

[21] More specifically, as described by House of Lords, Select Committee on Artificial Intelligence: "It is important that members of the public are aware of how and when artificial intelligence is being used to make decisions about them, and what implications this will have for them personally. This clarity, and greater digital understanding, will help the public experience the advantages of AI, as well as to opt out of using such products should they have concerns."

[22] Jessica Fjeld, Nele Achten, Hannah Hilligos, Adam Nagy, Madhulika Srikumar, 'Principled Artificial Intelligence: Mapping Consensus in Ethical and Rights-Based Approaches to Principles for AI' (2020) Berkman Klein Center Research Publication 2020(1) 30 <http://dx.doi.org/10.2139/ssrn.3518482> accessed 30 June 2022.

[23] *ibid* 22.

operators to make sure data privacy is incorporated in the system's construction and development.

3. Regulation of artificial intelligence

3.1 General

To ensure that all ethical dilemmas described above have been assessed and considered, AI systems need to comply with certain legal frameworks. Laws enforce obligatory rules and standards that ensure individual rights are respected. If laws are successful, they can minimize risks and not hinder regulation in the process. As mentioned, it is generally agreed that AI offers many risks and benefits, but there is no general agreement on how AI should be regulated. In some countries, there is still an ongoing debate whether AI should even be regulated at all.

The underlying objective of different AI regulatory efforts is to address risks and situations that are currently not adequately addressed by existing laws and regulations.[24] Due to fast progress in the advances of AI and these systems becoming more and more autonomous, lawmakers and courts will soon have to deal with difficult questions when trying to apply the existing laws to AI. There are two possible solutions to this problem: countries and regulators will either have to manage and update the existing laws or new laws specifically intended for AI will be necessary. According to the majority of researchers, the question is no more whether laws that regulate AI should be passed, but rather when this is going to happen.[25] Another topic of concern is that at this moment in time, AI is still undoubtedly under human control, and there are laws and procedures in place that regulate human behavior, such as liability. The more dangerous problem will come to surface when AI becomes increasingly autonomous and uses techniques that are beyond human understanding. When that happens, there will be a need to have practical solutions in place.

---

[24] Miriam Buiten, *Towards intelligent regulation of artificial intelligence* (2019) European Journal of Risk Regulation 10(1) 41 <https://www.cambridge.org/core/journals/european-journal-of-risk-regulation/article/towards-intelligent-regulation-of-artificial-intelligence/AF1AD1940B70DB88D2B24202EE933F1B> accessed 30 June 2022.

[25] Mark Lemley, Brian Casey, 'You might be a robot' (2020) Cornell Law Review 6 <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3327602> accessed 30 June 2022.

Today, the law already indirectly impacts AI because it regulates how humans, who use AI, must behave. This, however, does not mean that such laws are optimal. There are already multiple occurrences of AI working independently from human beings; therefore, human conduct regulations will not always be able to cover legal disputes and issues that arise from AI. Some countries became increasingly progressive and aware of the rising issues of AI. In 2018, the governments of France and Canada published the mission of the International Panel on Artificial Intelligence (IPAI), whose goal was "to support and guide the responsible adoption of AI that is human-centric and grounded in human rights, inclusion, diversity, innovation and economic growth" as well as to "facilitate international collaboration in a multistakeholder manner with the scientific community, industry, civil society, related international organizations, and governments".[26] Many countries decided to follow these initiatives. There are currently over 30 countries that published strategies and incentives concerning AI. Furthermore, the EU expressed the desire and goal to become the leader in human-centric AI and strives to be an example for other countries. As a result of years of development, the EU Draft Artificial Intelligence Act was published in 2021, and its goal is to become law in 2024–2025.

At this point, it is important to refer to the introductory part of the thesis and mention the difference between GOFAI and ML based AI. Most of the below mentioned discussions in regulation do not differentiate or even discuss GOFAI because they implicitly refer to ML-based AI. However, it is important to note that both GOFAI and ML based AI should be addressed when discussing their legal regulation.

3.2 The European Union

The EU is currently only one of the few countries who wish to regulate AI. Many countries published initiatives and strategies, but most of those are a work in progress and only strategically plan for the future without having any real implications. In the past five years, the EU has increasingly developed multiple strategies, which culminated in the Draft AI Act. The EU expressed its intent to be a leader in AI regulation, which has received both praise and criticism. On the one hand, it seems inevitable that AI must be regulated, and the EU produced

[26] Prime Minister of Canada, 'Mandate for the international panel on artificial intelligence' (Mission statement) (2018) <https://pm.gc.ca/en/news/backgrounders/2018/12/06/mandate-international-panel-artificial-intelligence> accessed 30 June 2022.

numerous guidelines and strategies defending this position. On the other hand, EU's big regulatory ambitions have led to critics stating that the main export is still the technology itself and not its regulation.

3.2.1 Evolution of the AI regulation idea in the EU

The discussions concerning AI in the European Union started in 2018 with the publication of the Declaration of Cooperation on Artificial Intelligence.[27] Since then, the EU established numerous commissions and institutions in order to examine the role of AI in the society. The European Parliament made various attempts at trying to deal with the progress of technology and regulation in the AI area. Some of those attempts were considered unsuccessful, such as the endeavor of electronic personality. The so-called "electronic personhood" status proposed by the Parliament would ensure that the most progressive forms of AI had their own rights and responsibilities. Furthermore, according to the proposal, the legal status of robots was going to be analogous to legal personhood that is normally given to companies. The European Parliament based the initiative on the fact that robotics has a visible impact on our lives, and the need to ensure that robotics remain "in the service of humans".[28] The initiative was mainly considered because of the then emerging questions, such as what if AI machines develop consciousness, as this would cause severe moral and legal consequences. It was concluded that the talk of giving any AI any sort of personhood is extremely premature and, consequently, the EU countries firmly rejected the recommendation of electronic personality in 2018. Nevertheless, the idea of a legal personhood for AI was not something that was only considered in the EU. In 2017, Saudi Arabia was the first country to give citizenship to a robot named Sophia.[29] Sophia was also named the first Innovation Champion of the United Nations Development Programme, which made the robot the first non-human to be given a UN title. Nevertheless, there are still some authors who believe that the liability and legal status of robots

---

[27] Commission, 'EU Declaration on Cooperation on Artificial Intelligence' (Declaration) COM (2018) <https://ec.europa.eu/jrc/communities/en/node/1286/document/eu-declaration-cooperation-artificial-intelligence> accessed 30 June 2022.

[28] Alex Hern, 'Give robots personhood status, EU committee argues' *The Guardian* (12 January 2017) <https://www.theguardian.com/technology/2017/jan/12/give-robots-personhood-status-eu-committee-argues> accessed 30 June 2022.

[29] Andrew Griffin, 'Saudi Arabia grants citizenship to robot for the first time ever' *The Independent* (26 October 2017) <https://www.independent.co.uk/life-style/gadgets-and-tech/news/saudi-arabia-robot-sophia-citizenship-android-riyadh-citizen-passport-future-a8021601.html> accessed 30 June 2022.

should be reconsidered, especially considering that the more advanced some robots and general AI become, the less we can still see them merely as tools.

In 2019, due to the European Parliament's call to update the existing framework, the European Commission introduced several guidelines and ethics on artificial intelligence.[30] Even though there were various codes of ethics already available, the Guidelines dealt specifically with an EU human-centric approach to AI, which, as it turns out, became one of the staples of AI regulation in the EU and underlined the respect of European values and principles.[31] Additionally, to assist policy makers and developers, Council of Europe published Guidelines on Artificial Intelligence and Data Protection where they addressed the importance of ensuring that AI applications do not undermine the right to data protection.[32] Furthermore, the Ethics guidelines for trustworthy AI[33] were published by the European Commission, but they remained non-binding. There were also calls for further legislative proposals in this area by the European Commission.

The main premise of the EU legislation connected with AI has always been the human-centric approach, which means that AI should be developed, monitored and used with the primary goal remaining the respect of fundamental human rights that should constantly be ensured in the process. This approach is different to the one adopted in the US, where AI is still mostly developed through private sector initiatives and is mainly self-regulated. The EU emphasizes a higher standard of protection against the social risks that come with AI, such as the risks affecting privacy. One of the key requirements in the EU standards is the presence of human agency and oversight, which is prominent in all publications issued by the EU so far. The reasoning for human oversight is mainly to ensure that a machine is never in full control and to offer the possibility of a human overriding the system at any time.

---

[30] European Parliament, 'EU Guidelines on ethics in artificial intelligence: context and implementation' (Briefing) (2019) <https://www.europarl.europa.eu/RegData/etudes/BRIE/2019/640163/EPRS_BRI(2019)640163_EN.pdf> accessed 30 June 2022.

[31] Resolution of the European Parliament (2017), Recommendations to the Commission on Civil Law Rules on Robotics. https://www.europarl.europa.eu/doceo/document/TA-8-2017-0051_EN.html . Retrieved June 26, 2022.

[32] Council, 'Consultative Committee of the convention for the protection of individuals with regard to automatic processing of personal data: Guidelines on artificial intelligence and data protection' (2019) <https://rm.coe.int/guidelines-on-artificial-intelligence-and-data-protection/168091f9d8> accessed 30 June 2022.

[33] Commission, 'Ethics guidelines for trustworthy AI' (Guidelines) COM (2019) <https://op.europa.eu/en/publication-detail/-/publication/d3988569-0434-11ea-8c1f-01aa75ed71a1> accessed 30 June 2022.

3.2.2 Further efforts

In 2018, twenty-five European countries signed the Declaration of Cooperation on AI, where they decided that the best way forward would be to cooperate within the EU. The Declaration gave birth to the so-called European approach to AI, which was later on more elaborately discussed in the High Level Expert Group (HLEG). HLEG consisted of over 50 experts that advised with AI regulation. They published numerous guidelines and established the concept of Trustworthy AI, which became a staple in European AI regulation initiatives, next to the concept of human-centric AI. According to HLEG's Ethics Guidelines, AI is trustworthy if it is:

- lawful (respecting all applicable laws and regulation),
- ethical (respecting ethical principles and values)
- and robust (from a technical perspective but also taking social environment into account).

The HLEG also presented seven key requirements for trustworthy AI. Among those are human agency and oversight, as well as transparency; specifically, that humans must always be informed of the fact that they are dealing with an AI machine. The HLEG also published Policy and Investment Recommendations for Trustworthy AI that gather thirty-three recommendations supposed to guide trustworthy AI to sustainability, growth, competitiveness and inclusion while benefiting and protecting human beings, and were presented at the first European AI Assembly in June 2019.[34] Finally, the HLEG published a document called Sectoral Considerations on the Policy and Investment Recommendations[35] which represents the possibilities of implementing the recommendations, especially in the areas of public sector, healthcare and the internet of things. HLEG's work was crucial for further legislative developments of AI in the EU. Their recommendations were a base of many policies and initiatives taken on by the Commission and the Member States, the White Paper and the AIA itself.

---

[34] Commission, 'High-Level expert group on artificial intelligence' (Policy) COM (2020) <https://ec.europa.eu/digital-single-market/en/high-level-expert-group-artificial-intelligence> accessed 30 June 2022.

[35] Commission, 'Sectoral considerations on the policy and investment recommendations for trustworthy artificial intelligence' (HLEG) (2020) <https://futurium.ec.europa.eu/en/european-ai-alliance/document/ai-hleg-sectoral-considerations-policy-and-investment-recommendations-trustworthy-ai> accessed 30 June 2022.

Another big step towards AI regulation is the White Paper on AI, which was adopted by the European Commission in 2020. It was the first document where the Commission proposed that, in addition to possible amendments to current legislation, a new legislation on AI will be necessary to avoid a fragmented legislation across the EU. The underlying theme of the White Paper was again trustworthiness of AI. The Paper further connected European approach to AI with other European goals and principles, such as the European Green Deal.

The White Paper was primarily a document presenting policy options for AI development in EU. It was based on the policy framework describing measures to align efforts at European, national and regional level and creating a regulatory framework for the EU that would be compliant with all EU laws, especially concerning fundamental and consumer rights. The main goal of the White Paper was to establish EU as the leader in human-centric AI. Concerning the regulatory framework for AI, the Paper specifically addressed the concerns forming around citizens' security, their fundamental rights and their fear of malicious practices with AI. The Paper named general key requirements from the HLEG group, and further built on the value of trustworthy AI, stating that all European values, such as consumer protection, product safety, liability, data protection and privacy, must remain intact when coming up with a regulatory plan for AI.

The Commission further recognized the specific features of AI that will cause the regulation to be more difficult. At that point, Member States were already pointing out their fear of a fragmented market, especially due to numerous calls for regulation and, in some cases, already developing solutions themselves because of the regulatory absence at the EU level. Malta, for example, already introduced a voluntary classification system for AI.[36] If the EU wished to avoid a fragmented internal market, a common regulatory framework was necessary.

The White Paper proposed a solution for these risks in the form of a new regulatory framework for AI at the EU level. The Commission explicitly stated that such framework should take a risk-based approach.

[36] Malta Digital Innovation Authority, 'AI ITA Guidelines' (Guidelines) (2019) 26 <https://mdia.gov.mt/wp-content/uploads/2019/10/AI-ITA-Guidelines-03OCT19.pdf> accessed 30 June 2022.

3.2.3 Artificial Intelligence Act (AIA)[37]

The Proposal for Artificial Intelligence Act was published in April 2021. It is currently still a draft act, the goal of which is to present harmonized rules for AI systems across the EU. It is the first attempt in history that implements a horizontal regulation of AI in the world.

3.2.3.1 General

The AIA draws heavily from the White Paper and is essentially a legal proposal for establishing AI in the EU. It foresees legal intervention only in cases where there is a "justified case for concern" or "where such concern can be reasonably anticipated in the near future".[38] Doing so should avoid stifling innovation in the AI sphere. Furthermore, the Act completely prohibits certain AI practices and sets forth specific requirements for high-risk systems, as well as some rules concerning market monitoring and surveillance.

Additionally, the Act proposes a new definition of AI:

"'Artificial intelligence system' (AI system) means software that is developed with one or more of the techniques and approaches listed in Annex I and can, for a given set of human-defined objectives, generate outputs such as content, predictions, recommendations, or decisions influencing the environments they interact with."[39]

This definition notably differs from the previous definition in the Coordinated Plan stating that AI is any machine that displays "intelligent behavior by analyzing their environment and taking actions – with some degree of autonomy – to achieve specific goals".[40] However, defining AI has proven difficult in the past, so it is questionable if or for how long the new definition will be accurate. Some would even argue that defining an elusive word such as AI or robot cannot

---

[37] Commission, 'Proposal for a Regulation of the European Parliament and of the Council laying down harmonized rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts' (Proposal) COM (2021) <https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1623335154975&uri=CELEX%3A52021PC0206> accessed 30 June 2022.
[38] *ibid* para 1.1.
[39] Article 3 (1) of the Artificial Intelligence Act.
[40] Commission, 'Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions: Coordinated Plan on Artificial Intelligence' (Plan) COM (2018) <https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:52018DC0795&rid=3> accessed 30 June 2022.

even be done, especially because of so many interconnected words and blurred lines between the terms like robot, machine, computer and humans.[41]

Furthermore, the proposed regulatory framework in the AIA sets the following objectives as the basis:

- The AI systems placed on the market in the EU are safe and respect the existing law.
- Legal certainty is ensured at all times to facilitate investment and innovation in AI.
- The enforcement of EU fundamental rights and safety requirements concerning AI systems is effective.
- The single market development is facilitated and consequently market fragmentation is avoided.

Concerning risk approach, the AIA follows the approach already presented in the White Paper. Those AI systems that can be defined as high-risk have to be compliant with special requirements for trustworthy AI and will need to follow conformity assessment procedures before being placed on the market.

The proposal further elaborates a governance system on the level of Member States, which will be built on the already existing framework. Moreover, governance will be enhanced by an establishment of a European Artificial Intelligence Board, which will be comprised of representatives from Member States as well as the Commission and will smooth the implementation of the Act and collect best practices among the Member States.

3.2.3.2 Risk-based approach

The Act classifies different types of risk levels:

- unacceptable risks,
- high risks,
- limited or minimal risks,
- transparency risks.

---

[41] Mark Lemley, Brian Casey, 'You might be a robot' (2020) Cornell Law Review 5 <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3327602> accessed 30 June 2022.

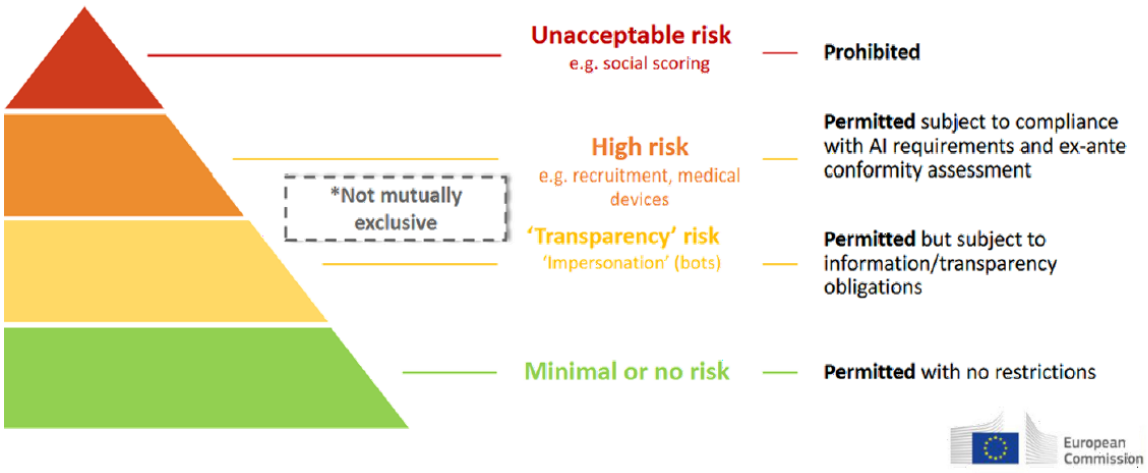

Figure 1: Classification of risks in the AIA proposal (European Commission).

The various types of risks and their differences will be further discussed in the following section.

3.2.3.3 Unacceptable risks

Unacceptable AI practices are prohibited across the EU. Among these are AI systems that go against Union values, for example because they violate fundamental rights. The unacceptable risks are listed as four prohibitions, covering practices such as manipulation beyond people's consciousness or targeting vulnerable groups, such as people with disabilities or children. The three entirely prohibited practices are:

- distortion of data,
- manipulation and
- known or predicted personality characteristics (so-called social scoring[42]).

Furthermore, real-time or remote biometric systems are also generally prohibited, but can be allowed only under strict exceptions, namely for law enforcement purposes. Nonetheless, even under exceptional circumstances, the use of biometric systems must be authorized in advance by either a judicial authority or an independent administrative authority of a Member State.[43]

a) Manipulative and exploitative systems

[42] Social scoring is a term used to describe systems that can judge people based on their behavior. Some AI tools can be used to classify people into groups, for example by gender, race or another distinguishing factor.
[43] Article 5(3) of the Artificial Intelligence Act.

Even though including a section with unacceptable risks can generally be considered as a good idea, some authors suggest that the provisions are too limited in scope and should be reconsidered and amended appropriately.[44] The text of the provision on manipulative systems states that it is prohibited

> (i) "the placing on the market, putting into service or use of an AI system that deploys subliminal techniques beyond a person's consciousness in order to materially distort a person's behaviour in a manner that causes or is likely to cause that person or another person physical or psychological harm;
> (ii) the placing on the market, putting into service or use of an AI system that exploits any of the vulnerabilities of a specific group of persons due to their age, physical or mental disability, in order to materially distort the behaviour of a person pertaining to that group in a manner that causes or is likely to cause that person or another person physical or psychological harm."

Because the provision requires "vulnerabilities" or "subliminal techniques" behind any kind of manipulation, it might result in the scope being limited since there is also the requirement for such AI system to "cause or is likely to cause […] physical or psychological harm". The current wording sets a high threshold for manipulative AI systems because it suggests that an AI system could also exploit or materially distort a person's behavior in a way that is not harmful. Of course, any AI system designed in a way that exploits people's behavior must cause harm as it is targeted to affect the behavior. The recent Facebook scandal is only one of the examples.[45] Adding the wording "that causes or is likely to cause […] harm" unnecessarily elevates the threshold for prohibition, when such a system should be prohibited solely based on going against fundamental Union rights.

Furthermore, limiting vulnerabilities only to age and physical or mental disability seems unnecessarily restrictive. It would seem reasonable to include all the vulnerabilities generally included in EU law, such as gender, race and ethnicity, sexual orientation, social or economic status, migration status, and others, in order to really achieve the limitation of vulnerabilities. This is especially important because AI systems have already been proven to frequently display bias against certain vulnerable groups. A study observing machine learning limitations

---

[44] Michael Veale, Frederik Zuiderveen Borgesius, 'Demystifying the Draft EU Artificial Intelligence Act' (2021) Computer Law Review International 99 <https://arxiv.org/pdf/2107.03721.pdf> accessed 30 June 2022.
[45] Ben Gilbert, 'Facebook's own research showed Instagram to be toxic for teens, which it denies — then it pulled Instagram Kids anyway' *Business Insider* (27 September 2021) <https://www.businessinsider.com/facebook-instagram-toxic-for-teens-pulls-intagram-kids-2021-9> accessed 30 June 2022.

concluded that while the professional risk assessment tool algorithms for predicting juvenile recidivism was fair in general, the machine learning models expressed a severe tendency to discriminate against male defenders, foreigners, or people of specific national groups.[46] The conclusion of this study was that foreigners were almost twice as likely to be wrongfully classified as high-risk by machine learning models compared to the nationals of that same country. This is problematic because the current wording of the AIA would not include such a bias since the prohibition focuses solely on age, physical or mental disability. A reconsideration of the current wording in the AIA should therefore be considered if all the disabilities currently protected under the EU law are to be included and in order to avoid restricting vulnerabilities only to age and physical or mental disability.

b) Social scoring

Next, article 5 (1) (c) prohibits social scoring. According to the article, it is prohibited:

"the placing on the market, putting into service or use of AI systems by public authorities or on their behalf for the evaluation or classification of the trustworthiness of natural persons over a certain period of time based on their social behavior or known or predicted personal or personality characteristics, with the social score leading to either or both of the following:

(i) detrimental or unfavorable treatment of certain natural persons or whole groups thereof in social contexts which are unrelated to the contexts in which the data was originally generated or collected;

(ii) detrimental or unfavorable treatment of certain natural persons or whole groups thereof that is unjustified or disproportionate to their social behavior or its gravity."

As explained in Recital 17, social scoring may lead to discrimination and exclusion of certain groups, which can violate the right to dignity and non-discrimination, as well as impact the values of justice and equality.

---

[46] Songul Tolan, Marius Miron, Emilia Gomez, Carlos Castillo, 'Why Machine Learning May Lead to Unfairness: Evidence from Risk Assessment for Juvenile Justice in Catalonia' (2019) ICAL 83 <https://doi.org/10.1145/3322640.3326705> accessed 30 June 2022.

It is unclear why social scoring is not prohibited as a whole but is instead limited to the use by public authorities or on their behalf. In recent years, there has been a lot of talk about the social credit system in China, where people and companies would be assessed and monitored about almost every aspect of their life and consequently be forced shape their behavior appropriately.[47] Social scoring would be a kind of incentive mechanism that could either reward individuals and companies or punish them. Subject to strict exceptions, the AI Act prohibits social scoring by public authorities; however, social scoring can also be a tool exploited by private entities. For example, a bank or an insurance company could use social scoring to determine the risk of certain individuals. Therefore, to ensure protection, the provision should be amended to include general purpose social scoring as well as government social scoring, especially considering that the aim of the Act is a trustworthy approach to AI. Social scoring is not contravening EU values only when exploited by governments, but also when used in general.

Some legal scholars found that the reasons why private entities are excluded from the AIA wording could be attributable to the GDPR rules.[48] According to them, social scoring by private entities would only be possible if consumers would give their consent under 6 (1) (a) GDPR[49], and even then, if the consent is to be considered free, the data subject must not suffer any detrimental effects from such a withdrawal.

While this reasoning is legitimate, it needs to be underlined that the EU strives to be a pioneer in the field of AI regulation, which is mentioned throughout the AI proposals and publications. Should other nations follow EU's example, any ambiguity should be avoided and carefully reconsidered. The wording as it is now can be misinterpreted if the reader does not possess sufficient legal knowledge of EU law and the GDPR. Moreover, the GDPR is a legal framework that is not comparable to any privacy law from other countries and could therefore be easily missed or misinterpreted. Finally, stating the same prohibition twice because it violates fundamental rights does not seem redundant, especially when the AIA aims to be the first such law in existence.

---

[47] Eunsun Cho, 'The Social Credit System: Not Just Another Chinese Idiosyncrasy' (2020) Journal of Public and International Affairs <https://jpia.princeton.edu/news/social-credit-system-not-just-another-chinese-idiosyncrasy> accessed 30 June 2022.

[48] Christiane Wenderhorst, 'The Proposal for an Artificial Intelligence Act COM(2021) 206 from a Consumer Policy Perspective' (2021) 64 <https://ssrn.com/abstract=4087402> accessed 30 June 2022.

[49] Article 6(1)(a) of the GDPR states that "processing shall be lawful only if and to the extent that at least one of the following applies: the data subject has given consent to the processing of his or her personal data for one or more specific purposes".

Additionally, it is unclear why the prohibition includes social scoring lasting "over a certain period of time". Any kind of limitation could potentially be viewed as arbitrary and used for such purposes. Since social scoring is prohibited, it should be prohibited regardless of the time frame of social scoring.

c) Biometric systems

Biometric identification is used to identify a person using their unique characteristics, such as physical, physiological or behavioral characteristics. These can include voice tone, shape of face, irises, and similar. Identification is performed by comparing one to many, meaning that their identity is established by matching stored templates with live templates and without the person's cooperation or even against their will.[50]

The use of biometric systems is only allowed under certain conditions stated in Article 5 (1) (d) of AIA. The use must be based on either of the following:

- a targeted search for specific potential victims of crime,
- prevention of an imminent threat or a terrorist attack or
- detection, localization, identification or prosecution of a suspect of a crime with a maximum sentence of at least 3 years.

While the intent behind the provision looks promising, it has been noted that the provision is not drafted carefully enough and has certain important limitations.

First of all, the provision only affects real-time systems, which are defined as capture, comparison and identification of data that occurs "without a significant delay".[51] Such limitation leaves room for loopholes in interpreting real-time systems and would mean that capturing and identifying biometric footage later in time, such as identifying individuals at protests days after they occur, for example, would not be included according to the wording of the provision as it is now. As follows, a threat to fundamental rights is not covered by adding the real-time limitation and the biometric identification should not be restricted in time if the provision really strives to protect fundamental rights.

Secondly, some authors suggest that from an ethical perspective, the biometric identification is only a small part of the concern. The primary concern is storing the biometric templates in

[50] *ibid*.
[51] Article 3 (37).

the first place and mass surveillance that follows from such data storage. [52] Biometric identification is only one way of achieving mass surveillance and is therefore of less importance. If the EU wants the fundamental rights to be respected, no storing or mass surveillance should be allowed in the first place. Therefore, the provision should be amended and expanded to include more than just biometric identification.

3.2.3.4 High-risk AI systems

Following unacceptable risks, the next risk level are high-risk AI systems.

The Act generally applies to all AI systems. However, the subcategory of high-risk AI systems only applies to two categories of systems. Article 6 of the Act states that regardless of whether an AI system is placed on the market or is put into service independently, both of the following conditions must be fulfilled cumulatively for an AI system to be considered as high-risk:

a) The intention of AI system is to be used as a product or safety component of products already covered by Union health and safety harmonization legislation. This constitutes, among others, toys, machinery, lifts or medical devices.[53]

b) AI system is by itself a product and is used in one of eight areas that are specified in Annex II. These include:
    i. biometric identification (in this case, both occurring in real-time as well as after the event),
    ii. management and operation in critical infrastructure,
    iii. educational and vocational training,
    iv. employment, worker management and access to self-employment,
    v. access to and enjoyment of essential services and benefits,
    vi. law enforcement,
    vii. migration, asylum and border management,
    viii. administration of justice and democracy.

While the Commission can add subareas to the existing fixed areas, in order to be added, the application must pose a similar risk. Additionally, updating the list is subject to Parliament and

[52] *ibid*.
[53] Michael Veale, Frederik Zuiderveen Borgesius, 'Demystifying the Draft EU Artificial Intelligence Act' (2021) Computer Law Review International 102 <https://arxiv.org/pdf/2107.03721.pdf> accessed 30 June 2022.

Council veto.[54] What is more, entirely new areas cannot be added to the list. This is a significant limitation.

Still, the Commission has the power to amend the Annex with an AI system where the next two conditions are fulfilled[55]:

(a) the AI system is intended to be used in one of the areas mentioned above[56] and
(b) the AI system is a risk to health and safety or a risk of having adverse impact on fundamental rights.

The Commission must take specific criteria into account when assessing the presence of a risk of having an adverse impact on fundamental rights. The criteria are specified in Article 7 and include, among others, the intended use of the AI system and the extent to which the AI system will likely be used, the extent of potential harm or impact of AI system and the extent to which the system has already caused harm. While there is also a measure stating that the Commission shall consider the extent to which the outcome produced with the AI system is easily reversible, and specifically notes that having the impact on health and safety of persons will not be considered as easily reversible, such a criterion seems lacking. On the one hand, it recognizes the possibility of an AI system having an impact that is not easily reversible, but on the other hand, it leaves a question of who will evaluate whether an impact is easily reversible. Furthermore, most of the obligations in the Act are imposed on providers (developers) and not on users of high-risk AI systems. Even though providers design high-risk AI systems, it should be underlined that they are not the users of such systems and consequently not those who are familiar with all the risks once the product is placed on the market. While they must undergo the conformity assessment procedure, it seems more prudent to impose such obligations (also) on the users because they are the ones who determine the use of the AI systems in the end.

Some initiatives suggest an approach where "users" (or deployers) of AI system would be required to include a fundamental rights impact assessment, meaning that they would have to specify all the individuals or groups that would be affected by the AI system and assess the system's impact on fundamental rights.[57] Furthermore, the users of AI systems could be obliged to verify if the AI system is compliant with the AI Act and respects all the rules. While

---

[54] *ibid.*
[55] Article 7.
[56] Points 1 to 8 of Annex III.
[57] An EU Artificial Intelligence Act for Fundamental Rights: A Civil Society Statement (2021) 3 <https://edri.org/wp-content/uploads/2021/12/Political-statement-on-AI-Act.pdf> accessed 30 June 2022.

that would suggest a double check, performed both by the provider and the user, this would only be happen in cases of high-risk AI systems. Since the AIA sets fundamental rights and trustworthiness of AI at the center, double verification might not be too excessive to ensure respect of fundamental rights. Furthermore, it has also been suggested as beneficial for public transparency to include the obligation of the users to upload the assessment information to an EU database.[58]

Providers and users of high-risk AI systems

A provider is defined in Article 3 (2) as "natural or legal person, public authority, agency or other body that develops an AI system or that has an AI system developed with a view to placing it on the market or putting it into service under its own name or trademark, whether for payment or free of charge". According to Article 28, any distributor, importer, user or other third-party is considered a provider if they either:

- place the AI system on the market under their name or trademark,
- modify the intended purpose of a high-risk AI system that is already on the market,
- or make a substantial modification to the high-risk AI system.

As mentioned above, AIA allocates most of the obligations to providers. For example, providers of high-risk AI systems are obliged to follow a quality management system, which includes techniques and systematic actions for development and quality assurance, examination, test and validation procedures, which have to be carried out also after the development of a high-risk system; implementation and maintenance of post-market monitoring system; procedures to report serious incidents and malfunctions, and others. Moreover, providers must draw up technical logs and keep logs automatically generated. The logging must be facilitated by providers and allow traceability that is appropriate to the system's risk. Furthermore, providers must ensure that the high-risk AI systems undergo a conformity assessment procedure before being placed on the market. If the providers suspect a non-conformity of a product they have placed on the market, such a product must immediately be either brought into conformity or recalled. Additionally, when a high-risk AI system presents a risk known to the provider, a national competent authority must immediately be notified of such a risk. Whenever a competent authority requests, the providers must prove the

[58] *ibid*.

conformity of the AI system with the requirements. If requested, providers must also provide the authority with access to AI system's automatically generated logs.

Moreover, providers must always keep human oversight in mind and AI systems must always remain under effective human supervision. This requirement will be especially hard to implement in practice, considering that automated decision-making systems make numerous decisions every day. The purpose of Article 14 is that the individuals performing the oversight can understand the limitations and capabilities of the AI system and can monitor the operation of such a system in a way that will catch anomalies or dysfunctions. The people responsible for oversight must be aware of the automation bias – of potentially overly relying on the system, and therefore need to be able to interpret the output of the system correctly.

Finally, the responsible individuals must be able to decide not to use the AI system and disregard or override the system's output and always be able to intervene with the system by stopping its function completely.

It is somewhat concerning that AIA imposes no obligations to users, especially concerning human oversight. The only requirement is that users must follow the instructions and report any incidents or malfunctioning to providers. Some authors also noticed that the leaked version of the AIA, published in January 2021, did require providers to specify "organizational measures", meaning that the overseers could decide to not use the high-risk AI systems in any situation without any reason to fear negative consequences.[59] Such a requirement has been omitted from the draft Act.

Transparency

Concerning public transparency, Article 60 regulates an EU database for standalone high-risk AI systems. The provision is a promising addition to increase transparency with individuals and society in general, but as the wording of the provision stands now, the database only covers information from high-risk systems registered by providers and gives no information on the context of use. It has been expressed that public transparency aim would be better realized if the public was able to identify where the system was used, how it was used and its purpose. According to the signatories of the Civil Society Statement for EU AI Act for Fundamental

[59] Michael Veale, Frederik Zuiderveen Borgesius, 'Demystifying the Draft EU Artificial Intelligence Act' (2021) Computer Law Review International 99 <https://arxiv.org/pdf/2107.03721.pdf> accessed 30 June 2022.

Rights, such an approach is further lacking because AIA only requires notification of individuals impacted by systems that pose transparency risks in Article 52 (bots, emotion recognition systems and deep fakes), but requires no such notification when people are impacted by using high-risk AI system in Annex III (AI systems used for biometric identification, road traffic, heat or electricity supply, recruiting candidates in job vacancies, evaluation of job performance and monitoring of persons in their work environment, evaluating creditworthiness of natural persons, etc.).[60]

Post-market monitoring, enforcement and consumer interests

Providers are required to establish post-market monitoring that will collect and analyze data provided by users or otherwise collected by high-risk AI systems. Consequently, the provider will be able to evaluate compliance of AI systems. The post-market monitoring must be based on a monitoring plan that is a part of the provided technical documentation.

The AIA gives an important role to market surveillance authorities (MSAs). They oversee penalties, obtaining information, and withdrawing products, among others. Usually, they will consist of regulatory agencies or government departments. Each Member State will be required to establish a competent authority that will ensure implementation and application of the Act. The competent authority will also function as a notifying authority and market surveillance authority. The function of such national competent authority will be mainly to provide guidance and advice concerning the application of the Act.

Space for improvement

Even though the Draft Act sets forth many rules and rights, a prominent thing that is missing is the right to complain. Many similar acts (such as the GDPR) gave individuals the right to complain and seek judicial remedy, but no such provision can be found in the AIA. Consequently, the Act currently foresees no legal right to sue a user or a provider for failure to meet the requirements of the AIA. Some argue that it is precisely the complaint mechanisms which have been the crucial point in determining the EU case law, especially where regulators

[60] An EU Artificial Intelligence Act for Fundamental Rights: A Civil Society Statement (2021) 3 <https://edri.org/wp-content/uploads/2021/12/Political-statement-on-AI-Act.pdf> accessed 30 June 2022.

are concerned.[61] The fact that such an instrument is currently missing from the AIA could potentially be a problem when holding the regulators accountable if their enforcement is lacking.

Turning to consumer interests, it seems at first that the Draft Act does not deal with them. This is based on the fact that the majority of the systems listed in Annex III that define high-risk AI systems are normally used by professional users or public authorities and therefore exclude consumers. It strikes as odd that consumer law is not specifically included in the AIA, especially since the White Paper on AI explicitly proposed a regulatory framework compliant with all EU laws, specifically naming consumer law. Being central in product safety legislation, AIA is of significant importance for consumers in Europe. Consumers will be affected by AIA anywhere where the bought product will contain embedded AI or whenever an AI system will be used by consumers in any way. As per the current wording of the Act, some consumer examples are covered but a lot of them are omitted. This creates a big chasm that needs to be amended. For example, a lawnmower robot containing AI components would qualify as high-risk AI system and would consequently have to fulfil specific obligations, such as appropriate level of robustness and accuracy covered in Article 15. However, an autonomous shopping assistant that helps consumers identifying best offers would not qualify as high-risk AI because such software is not listed in Annex III of the AIA and therefore does not qualify as high-risk AI, even though the risks of such an assistant can be evaluated as high because the assistant could push consumers into expensive deals. The problem with AIA is that most of the provisions only apply to high-risk AI systems. As mentioned above, it would seem like an appropriate solution to expand the list qualifying high-risk systems, especially making sure the list includes consumers as well. Some scholars suggest a semi-horizontal product safety regime that would apply specifically for software and would automatically include high-risk software products in the scope of Title III; but for a shorter term, a new area could be added to the list in Annex III. Furthermore, Point 5 of Annex III currently includes AI systems intended to be used to evaluate creditworthiness of natural persons or establish their credit score. Such a provision seems too narrow since it is very limited in scope. For example, AI systems used for risk assessment by insurance companies are not covered, and neither is personalized pricing.[62]

---

[61] Michael Veale, Frederik Zuiderveen Borgesius, 'Demystifying the Draft EU Artificial Intelligence Act' (2021) Computer Law Review International 103 <https://arxiv.org/pdf/2107.03721.pdf> accessed 30 June 2022.
[62] Christiane Wenderhorst, 'The Proposal for an Artificial Intelligence Act COM(2021) 206 from a Consumer Policy Perspective' (2021) 100 <https://ssrn.com/abstract=4087402> accessed 30 June 2022.

It follows that Point 5 of Annex III should be amended and include a lot more than it currently does because a lot of applications have a compared to credit scoring when it comes to fundamental rights. It does not seem justified that everything except credit scoring would be excluded from the high-risk list.

Finally, users of synthetic content disclosure (deep fakes[63]) are required to disclose the artificial nature of the content. The exceptions are crime prevention and freedom of expression or freedom of arts and sciences.

3.2.3.5 Conclusions

All in all, the Proposal for an Artificial Intelligence Act is certainly a landmark document and the first of its kind. No country in the world has attempted to regulate AI on such a level. While AIA shows some good incentives and propositions, such as differentiating AI systems by risk level, enabling societal involvement by creating a transparency approach, and introducing some valid prohibitions, it also has severe risks and weaknesses. It allows for certain loopholes that need to be addressed and amended and leaves room for improvement and comprehensiveness.

[63] Deep fake is a term defining AI systems that generate or manipulate image, audio or video content that appreciably resembles existing persons, objects, places or other entities or events and would falsely appear to a person to be authentic.

3.3 The United States

The US has been slow in regulating AI compared to other countries. Unlike the EU, which decided for a comprehensive regulatory framework in the form of AIA, which will regulate AI in all of the Member States, the US moved forward with more specific regulatory guidelines that are mostly based on agency and not on a federal level. There is no coordinated national strategy yet aimed to increase AI regulation or respond to challenges brought on by AI, only ideas. There is a laissez-faire approach to the governance in AI technologies, especially under the assumption that limiting regulation is the best way of achieving innovation. Since there has been no joint effort to regulate AI on a federal level, there is still debate across the US whether there should be specific laws regulating AI or if the best approach is to adapt the already existing rules on privacy, discrimination, cybersecurity and others, and shape them accordingly so they will fit the new AI era. There have been some developments and initiatives, but the current reality is still mostly ideas and proposals, rather than actual laws. Furthermore, the current legislation is divided by states and different sectors. The following part will analyze the current state of regulation in the US on both the federal and state level.

3.3.1 Federal level

The beginnings

In 2016, the White House published three reports that represented the basis of US AI strategy. The first report Preparing for the Future of Artificial Intelligence[64] set forward specific recommendations related to AI regulation. The National Artificial Intelligence Research and Development Strategic Plan[65] mostly indicated a plan for public funding of research and development in AI. The third report, Artificial Intelligence, Automation and the Economy[66], assessed the impact of automation.

---

[64] Executive Office of the President, 'National Science and Technology Council Committee on Technology: Preparing for the Future of Artificial Intelligence' (2016) <https://obamawhitehouse.archives.gov/sites/default/files/whitehouse_files/microsites/ostp/NSTC/preparing_for_the_future_of_ai.pdf> accessed 30 June 2022.

[65] National Science and Technology Council, 'The National Artificial Intelligence Research and Development Plan' (2016) <https://obamawhitehouse.archives.gov/sites/default/files/whitehouse_files/microsites/ostp/NSTC/national_ai_rd_strategic_plan.pdf> accessed 30 June 2022.

[66] Executive Office of the President, 'Artificial Intelligence, Automation, and the Economy' (2016) <https://obamawhitehouse.archives.gov/sites/whitehouse.gov/files/documents/Artificial-Intelligence-Automation-Economy.PDF> accessed 30 June 2022.

In 2018, the White House organized a Summit on AI, where academia, industry and government representatives were invited to participate.

In February 2019, the American AI Initiative was established under the Trump Administration via Executive Order 13859. The Initiative addressed five main topics which were codified into law in the National AI Initiative Act in 2021. The five main topics were increasing AI research investment, unleashing Federal AI computing and data resources, setting AI technical standards, building America's AI workforce and engaging with international allies.[67] However, the AI initiative is no longer updated due to change in presidency.

Guidance for Regulation of Artificial Intelligence Applications

In 2020, the Executive Office of the President issued a Guidance for Regulation of AI applications in the form of a Memorandum for the heads of executive departments and agencies.[68] The aim of the Memorandum is to provide guidance to all Federal agencies and inform of the regulatory as well as non-regulatory approaches concerning AI systems. It should be noted that government use of AI is outside the scope of the Memorandum. While the Memorandum is not a binding legislative document, it does set the tone and direction in which current and future regulatory endeavors will be headed in the US.

Firstly, it is observed that in order to ensure the positioning of the US as a global leader in AI development, Federal agencies must avoid any actions that would disrupt growth and innovation. The Guidance also enumerates the regulatory and non-regulatory approaches for the stewardship of AI applications that should be taken into account when formulating any advance in AI systems. The regulatory principles are the following[69]:

- public trust in AI,
- public participation,
- scientific integrity and information quality,

---

[67] The White House, 'Artificial Intelligence for the American People: An Executive Order on AI' (2019) <https://trumpwhitehouse.archives.gov/ai/executive-order-ai/> accessed 30 June 2022.

[68] Russell Vought, 'Memorandum for the Heads of Executive Departments and Agencies: Guidance for Regulation of Artificial Intelligence Applications' (2020) <https://www.whitehouse.gov/wp-content/uploads/2020/01/Draft-OMB-Memo-on-Regulation-of-AI-1-7-19.pdf> accessed 30 June 2022.

[69] Russell Vought, 'Memorandum for the Heads of Executive Departments and Agencies: Guidance for Regulation of Artificial Intelligence Applications' (2020) <https://www.whitehouse.gov/wp-content/uploads/2020/01/Draft-OMB-Memo-on-Regulation-of-AI-1-7-19.pdf> accessed 30 June 2022.

- risk assessment and management,
- benefits and costs,
- flexibility,
- fairness and non-discrimination,
- disclosure and transparency,
- safety and security,
- interagency cooperation.

While some of the endeavored approaches and elements can be found both in the EU approach to AI and the Memorandum, there are numerous differences. The AIA sets a strong focus on human-centered and trustworthy AI, while the Memorandum puts development and innovation front and center. There is a big focus on limiting any regulatory overreach that would stifle innovation. Additionally, there is no separation or classification of specific AI systems being riskier than others in the Memorandum. While the Memorandum acknowledges that risk assessment and risk management should be consistent in various agencies and technologies, it envisions a risk-based approach that would determine acceptable and unacceptable risks. In other words, if a harm would have more costs than the expected benefits, it should be considered as higher risk.[70] Meanwhile, the AIA set a clear risk-based approach by classifying various AI systems into different risk groups. However, the US has shown some initiatives on the risk side. For example, the National Institute of Standards and Technology (NIST) is currently in the process of developing a framework that would reshape risk management associated with AI.

Additionally, the Memorandum stresses the importance of trustworthy AI by giving examples of privacy, safety, public security and other risks, which should be properly addressed. Yet, there are no clear examples or guidelines as to what kind of regulatory approach would be necessary to support the development of the AI while promoting trustworthiness. Additionally, it is specifically stated that agencies should carefully consider the regulatory environment before taking additional measures for disclosure and transparency because public disclosure and transparency are context specific and depend on potential harms and potential benefits. It has to be noted that there is a surprisingly high number of AI use in the federal government.

---

[70] *ibid*.

According to a report about AI in Federal Administrative Agencies, half of the agencies in their survey have already experimented with AI or ML.[71] It has also been recognized in the US that AI poses serious accountability challenges, especially concerning decisions that affect the public officials. When decisions made by public officials affect the rights of citizens, such decisions would have to be explained under the usual circumstances. However, many of the advanced AI tools are not explainable and consequently non-transparent in these decisions.[72] Furthermore, even when an AI system is transparent, that transparency can be difficult to turn into accountability. That can be achieved through different methods, such as reviewing AI-based decisions (for example, judicial use of AI systems), notice and correction rights, or even a prohibition or a licensing requirement in specific cases (for example, FDA drug approvals).[73]

The Artificial Intelligence Bill of Rights

In 2021, the White House Office of Science and Technology Policy started to work on the so-called Artificial Intelligence Bill of Rights, the aim of which is to protect consumers from potential harm caused by AI technology. A statement from the White House named common problems connected to AI, such as underrepresentation of the American society in the data sets used to train the AI systems.[74] For example, virtual assistants that do not understand Southern accent, or facial recognition systems that had serious discriminatory outcomes leading to false arrest and imprisonment due to bad facial recognition match are only some of the problems causing bias and discriminatory outcomes.[75] There are many other discriminatory and prejudicial behaviors displayed by AI systems. The statement also accentuates the importance of privacy and transparency. AI failures that have so far occurred are, at least in the US, "unintentional and unacceptable".[76] Therefore, in search for a remedy, the idea was to create

---

[71] David Freeman Engstrom, Daniel E. Ho, Catherine M. Sharkey, Mariano-Florentino Cuellar, 'Government by Algorithm: Artificial Intelligence in Federal Administrative Agencies' (2020) NYU School of Law 75 <https://ssrn.com/abstract=3551505> accessed 30 June 2022.
[72] *ibid.*
[73] *ibid.*
[74] The White House Briefing Room, 'Americans Need a Bill of Rights for an AI powered world' (2021) <https://www.whitehouse.gov/ostp/news-updates/2021/10/22/icymi-wired-opinion-americans-need-a-bill-of-rights-for-an-ai-powered-world/> accessed 30 June 2022.
[75] Kashmir Hill, 'Another Arrest, and Jail Time, Die to a Bad Facial Recognition Match' *The New York Times* (29 December 2020) <https://www.nytimes.com/2020/12/29/technology/facial-recognition-misidentify-jail.html> 30 June 2022.
[76] The White House Briefing Room, 'Americans Need a Bill of Rights for an AI powered world' (2021) <https://www.whitehouse.gov/ostp/news-updates/2021/10/22/icymi-wired-opinion-americans-need-a-bill-of-rights-for-an-ai-powered-world/> accessed 30 June 2022.

an AI bill of rights, a document that would elaborate the freedoms and rights that all AI technologies should respect, such as ensuring unbiased and accurate AI systems that are trained on representative data sets and discourage discrimination, no discriminatory surveillance, and the right to a remedy in case of a harm caused by an AI system. As a next step, the White House Science and Technology Policy intends to partner with academia, private sector, government and the society to draft and develop such a bill of rights. In November 2021, the Office of Science and Technology Policy published that it plans to launch a series of listening sessions and events to engage the American public in the process of developing a Bill of Rights for an Automated Society. The goal is to ensure that the new and emerging data-driven technologies abide by the enduring values of the American democracy.[77] Furthermore, they disclosed a plan with six events organized to promote public education and engagement in areas where AI influences American lives, such as criminal justice system, healthcare, consumer rights and protections, democratic values, and other topics.

There are certainly some features that many deem important when establishing a new set of rules for the future of AI, such as relying on fundamental and democratic values to be respected, producing AI systems that are fair and non-discriminatory, ensuring transparency and explainability of AI, and guaranteeing AI accountability.

Regarding democratic and fundamental values, the US legislators are aware that leaving AI technology to develop without any oversight could be very dangerous. Consequently, there have already been calls to update the existing regulations and enhance oversight, especially considering the vast implications of AI on privacy and mass surveillance in the US.[78] Since the commercial sector has been the primary leader of AI innovation in the US, incorporating democratic and human-centered values in the AI development is the crucial next step.[79]

Furthermore, researchers are stressing the importance of understanding how AI systems work and operate beneath the surface to ensure public trust and scrutiny. For the sake of transparency, the public should be aware of why some data analysis produces errors and how AI systems

---

[77] *ibid*.
[78] Karl Manheim, Lyric Kaplan, 'Artificial Intelligence: Risks to Privacy and Democracy' (2019) Yale Journal of Law and Technology 182 <https://yjolt.org/sites/default/files/21_yale_j.l._tech._106_0.pdf> accessed 30 June 2022.
[79] Michele Elam, Rob Reich, 'Stanford HAI Artificial Intelligence Bill of Rights: A White Paper for Stanford's Institute for Human-Centered Artificial Intelligence' (2022) 4 <https://hai.stanford.edu/sites/default/files/2022-03/HAI%20Policy%20White%20Paper%20-%20Stanford%20HAI%20Artificial%20Intelligence%20Bill%20of%20Rights.pdf> accessed 30 June 2022.

operate.[80] Yet, such information is usually kept secret by companies. Especially when it comes to privacy matters, such as recording employee performance by watching their eye movements or facial expressions, or employee surveillance, employees need to have a right to understand how the AI system works and get an explanation when such a system makes a decision about their performance.[81]

While the supporters of the AI bill of rights have welcomed a new attempt to regulate AI in the US on a national level, there has been some backlash, especially among policy analysts who have pointed out that an AI bill of rights is nothing more than a mere distraction.[82] The original Bill of Rights from 1791 was created to keep the newly formed government in check, but today's big tech already has implemented the systems to keep their influence limited, precisely in the government and the constitution. Therefore, existing privacy laws should take care of AI surveillance fears. Moreover, preventing AI-biased decisions should be the primary concern for non-discrimination laws. Consequently, an AI bill of rights is not necessary and merely distracts from the real problem without proposing sufficient solutions.

Congressional efforts

The 116th Congress in the US was, so far, the Congress which historically focused on AI the most. The number of AI mentions in various forms is more than triple that of the previous Congress.[83]

The National Defense Authorization Act (NDAA), a bill regulating budget, expenditures and policies of the US Department of Defense, includes numerous provisions that are directly related to AI. One of the proposed measures establishes a new National Artificial Intelligence Initiative Office that would be under the White House and would coordinate AI research and policy strategies. Furthermore, the National Institute of Standards and Technology (NIST) was

---

[80] Karl Manheim, Lyric Kaplan, 'Artificial Intelligence: Risks to Privacy and Democracy' (2019) Yale Journal of Law and Technology 182 <https://yjolt.org/sites/default/files/21_yale_j.l._tech._106_0.pdf> accessed 30 June 2022.

[80] Michele Elam, Rob Reich, 'Stanford HAI Artificial Intelligence Bill of Rights: A White Paper for Stanford's Institute for Human-Centered Artificial Intelligence' (2022) 4 <https://hai.stanford.edu/sites/default/files/2022-03/HAI%20Policy%20White%20Paper%20-%20Stanford%20HAI%20Artificial%20Intelligence%20Bill%20of%20Rights.pdf> accessed 30 June 2022.

[81] *ibid*.

[82] Hodan Omaar, 'Letter: Creating an AI Bill of Rights is a distraction' *Financial Times* (22 October 2021) <https://www.ft.com/content/3069b0ac-56c0-4c40-817f-f2d180f1ebc7> accessed 30 June 2022.

[83] Jack Clark, Ray Perrault, 'Artificial Intelligence Index Report 2022' (2022) Stanford University Human-Centered Artificial Intelligence 12 <https://aiindex.stanford.edu/report/> accessed 30 June 2022.

ordered to develop a risk management framework for trustworthy AI systems. Also, the NDAA created a National Security Commission on Artificial Intelligence, which had to review and advise on the competitiveness of US in AI. The Commission issued various reports. In the last report, submitted to Congress and the President, however, the report warned how the US is not sufficiently well-prepared to defend against the possible threats and risks of AI.[84] The report calls for a strategy to make US ready for AI and names requirements for improving transparency, protecting privacy, and civil rights.

Part of the NDAA is also the National AI Research Resource Task Force Act, which is part of the National AI Initiative. The main function of the act is the establishment of the National AI Research Resource (NAIRR).

As a response to the NAIRR and their work, the Stanford University Human-Centered Artificial Intelligence (HAI) published a White Paper which explains the multidisciplinary research on the study and feasibility of NAIRR.[85] They spotted three main themes:

- complementarity between compute and data (data is what renders high-performance computing important and the biggest risk is privacy),
- making sure AI research is also funded in public sector (current state of AI research being highly dependent on private sector would ensure future stability), and
- addressing short- and long-term vision of the NAIRR.

Other Departments

The Department of Commerce established the National Artificial Intelligence Advisory Committee, which will be tasked with advising the President and federal agencies on issues related to AI, such as US competitiveness and leadership in AI, AI workforce issues and other problems that could be encountered with AI.

Also, the Federal Trade Commission (FTC) published a memo titled Aiming for truth, fairness, and equity in your company's use of AI, where they specifically state they will use their authority to pursue the use of biased algorithms. Moreover, they warn companies to hold

---

[84] National Security Commission, 'Final Report from National Security Commission on Artificial Intelligence' (2021) 11 <https://www.dwt.com/-/media/files/blogs/artificial-intelligence-law-advisor/2021/03/nscai-final-report--2021.pdf> accessed 30 June 2022.

[85] Daniel E Ho, Jennifer King, Russell C. Wald, Christopher Wan, 'Building a National AI Research Resource: A Blueprint for the National Research Cloud' (2021) Stanford HAI 10 <https://hai.stanford.edu/sites/default/files/2021-10/HAI_NRCR_2021_0.pdf> accessed 30 June 2022.

themselves accountable for the performance of their algorithms, or the FTC will. The FTC gives an example of algorithm whose results discriminate against a protected class, which is already considered a violation of the FTC Act. Companies are expected to be transparent and truthful with their users about the use of data, and test algorithms to prevent discrimination, both before the start of use and additionally after.

3.3.2 State level

The scope of this thesis is too small to include all the state level AI laws. Hence, only those relating to comparison with EU will be discussed.

Facial Recognition

On the federal level, there were discussions about the Facial Recognition and Biometric Technology Moratorium Act in 2021, which would introduce a prohibition on most biometric image analytics and other technologies, such as other physical characteristics. Furthermore, the proposal would ban federal funds for biometric surveillance systems and propose remedies for violations against the Act.[86] However, the opposers of the Act mainly advocated that introducing such a legislative work would aid counterterrorism investigations and would not be helpful in reuniting human trafficking victims with their families.[87] For this and other various reasons, the Act never progressed out of committee and all further talks have been haltered.

Nonetheless, on the state level, the legislation concerning facial recognition is widely fragmented. Some states have passed laws connected to facial recognition. Virginia, for example, enacted a law that prohibits local law enforcement agencies from using facial recognition technology unless it is specifically authorized. Likewise, California, New Hampshire and Oregon restricted the use of facial recognition by law enforcement.[88] On the

---

[86] Gibson Dunn, '2021 Artificial Intelligence and Automated Systems Annual Legal Review', (2022) <https://www.gibsondunn.com/2021-artificial-intelligence-and-automated-systems-annual-legal-review/#_ftn13> accessed 30 June 2022.

[87] 'Facial Recognition and Biometric Technology Moratorium Act Reintroduced in Congress' (2021) SMD Magazine <https://www.sdmmag.com/articles/99560-facial-recognition-biometric-technology-moratorium-act-reintroduced-in-congress> accessed 30 June 2022.

[88] Deborah George, 'Virginia Law Bans Local Police Use of Facial Recognition Technology' (2021) National Law Review <https://www.natlawreview.com/article/virginia-law-bans-local-police-use-facial-recognition-technology> accessed 30 June 2022.

other hand, Texas and Florida, for example, still widely employ the use of facial recognition systems.

Bots

In 2018, California passed the Bolstering Online Transparency (BOT) Act. In general, bots are defined as "software applications that complete tasks automatically".[89] According to the Act, a bot is defined as an "automated online account where all or substantially all of the actions or posts of that account are not the result of a person". The Act tackles the issue of bots presenting themselves as human beings and hiding their identities. It is therefore unlawful if an entity or a person uses a bot to communicate or interact with a Californian in order to either make a sale, a transaction or to influence a vote in an election. It must be disclosed that the bot is doing the communication. Even though the law ensures fines for transgressions, it does not provide a right of action for the people affected by the bot.

Compared to the European Draft Artificial Intelligence Act, the AIA also includes a transparency obligation and requires that a natural person must be informed of interacting with an AI system. Chatbots are classified in the limited-risk section in the AIA.

Privacy

There is a strong connection between data privacy and AI because privacy is highly impacted by AI, especially due to high volumes of data processing. Unlike the GDPR, there is no national data privacy law in the US. Instead, legislation is very sectoral and fragmented by state. The most advanced is currently the California Consumer Privacy Act (CCPA) that came into force in 2018 and tries to fill the absence of a state law regulating privacy. Additionally, the California Privacy Rights Act (CPRA) was enacted in 2020 and is the first data privacy law in the US.

There are some similarities that can be observed both in GDPR and CPRA. Both legal texts talk about automated decision-making (ADM) technologies, which are not directly

[89] Barry Stricke, 'People v. Robots: A Roadmap for Enforcing California's New Online Bot Disclosure Act' (2020) Vanderbilt Journal of Entertainment and Technology Law 847 <https://scholarship.law.vanderbilt.edu/cgi/viewcontent.cgi?article=1032&context=jetlaw> accessed 30 June 2022.

synonymous with AI, but – at least in Europe – a lot of AI systems are in fact covered by this definition and are *de facto* influenced by law. That is why it is likely that the same will occur with the CPRA, which also uses the language of ADM technologies.[90] Moreover, privacy law in some cases directly affects AI. As demonstrated above when discussing the GDPR connection with AI, especially concerning data processing, some language in the AIA is heavily dependent on GDPR and some provisions need to be read and understood together. Additionally, a provision in CCPA describing the collection and reuse of data also directly impacts AI because it prohibits companies from using data outside the scope of the original data collection.

Although discrimination is not directly connected to privacy, it is frequently rooted in personal factors, such as race, skin color or sexual identification. If such personal information is used with AI systems, it automatically falls under privacy fields. That is why it is increasingly important to guarantee transparency in AI and consequently ensure there are no discriminatory outcomes.

All in all, the current state of AI in the US varies from state to state and there is not one rule that fits all. Because of the fragmented situation, it is likely that AI regulation in the US will vary based on geography, much like the current privacy legislation. Consequently, could happen that different states will differently regulate various aspects connected with AI. While it can be observed that the US as a leader in AI spends a lot of time and effort in discussing regulation of AI, such as implementing the NAIRR or talking about the AI bill of rights, it could easily happen that the progress of AI will be too fast for the regulation to follow.

However, there is also another possibility. Depending on when the EU will pass the Artificial Intelligence Act and what the final version will look like, it is possible that the AIA will *de facto* affect US regulation. This would not be unprecedented because a similar transition occurred with the GDPR. Even though the GDPR only affects EU citizens, it requires US companies that do business anywhere in the EU to comply with the rules. Equally, if the AIA imposes rules for the EU territory, it will affect and impact the US as well.

---

[90] Eli MacKinnon, Jennifer King, 'Regulating AI Through Data Privacy' (2022) Stanford HAI <https://hai.stanford.edu/news/regulating-ai-through-data-privacy> accessed 30 June 2022.

4. Conclusion

There is no doubt that AI is a disruptive technology and it has affected so many aspects of daily life that machine learning is no longer science fiction. However, different countries approached the questions of regulation in a different manner.

The EU and its legal system are the result of unity and a wish to have internal market regulated by its own rules. While the US aims to be the leader in AI, the EU made a strong stance in being the first one to regulate it. How successfully, is yet to be determined. Also, it is questionable what is EU's real export: the technology itself or its regulation. It might have been more valuable to accept from the beginning that there is no such thing as one definition that fits AI, and thus the initial and ongoing struggles of finding the one perfect definition could have been avoided by accepting that the regulatory rules should adapt and not be based solely on the definition. Furthermore, while the discussions on ethical dilemmas are necessary to establish and decide what (or who) needs protection (and from what), it is even more necessary to make the actual protections reflected in the laws. However, there is still a large chasm between establishing these concepts in theory and their practical use. Litigation concerning AI's consequences is only at the beginning. While the AIA proposal in the EU and the multiple talks and proposals in the US provide a high-level overview, the real work and progress lie ahead. Depending on the quality and practical usefulness of the AIA, it could happen that the Act would be used also in the US, much as the GDPR.

Moreover, even though AIA is the first attempt at horizontal AI regulation and contains some positive ideas, such as the risk level differentiation or numerous prohibitions, it also displays options for improvement, most notably leaving out items it does not consider as high risk and offering no further option for those to be included either in the future or by the Member States themselves because of EU law supremacy.

On the other hand, the laws in the US regulate AI very differently across the states, which, in turn, could cause certain states to become AI-friendly. Time has probably already passed for the US to be the pioneer at regulating AI on a federal level, but only discussing the ethical implications and including them in an AI bill of rights might not be enough, especially because other countries started regulating AI more seriously. What is true for both the US and the EU is that the continued risk for both is internal fragmentation if new measures are not introduced. To some extent, it already started happening and will in the future most likely lead to uneven protection or treatment of citizens.

AI is indeed a disruptive technology, and it is already affecting our lives. It is, nonetheless, important to realize that this technology, with all its positive and negative sides, is here to stay, in one form or another. Therefore, it is crucial that countries are prepared to deal with it and that the regulations that come into place think and rethink what is truly lying on the other side of the legal definitions and how these implicate people's lives.

5. Bibliography

Academic Works

Bayamlioglu E, Baraliuc I, Janssens L, Hildebrandt M, *Being profiled: Cogitas ergo sum: 10 years of profiling the European citizen* (Amsterdam University Press 2018)

Buiten M, *Towards intelligent regulation of artificial intelligence* (2019) European Journal of Risk Regulation 10(1) 41 <https://www.cambridge.org/core/journals/european-journal-of-risk-regulation/article/towards-intelligent-regulation-of-artificial-intelligence/AF1AD1940B70DB88D2B24202EE933F1B> accessed 30 June 2022

Carneiro [D], Novais [P], Neves [J], *Conflict Resolution and its Context: From the Analysis of Behavioural Patterns to Efficient Decision-Making* (Springer 2020) 32

Castro D, McLaughlin M, Chivot E, 'Who is winning the AI race: China, the EU or the United States?' (2019) <https://datainnovation.org/2019/08/who-is-winning-the-ai-race-china-the-eu-or-the-united-states/> accessed 26 June 2022

Cho E, 'The Social Credit System: Not Just Another Chinese Idiosyncrasy' (2020) Journal of Public and International Affairs <https://jpia.princeton.edu/news/social-credit-system-not-just-another-chinese-idiosyncrasy> accessed 30 June 2022

Clark J, Perrault R, 'Artificial Intelligence Index Report 2022' (2022) Stanford University Human-Centered Artificial Intelligence 12 <https://aiindex.stanford.edu/report/> accessed 30 June 2022

Dastin J, 'Amazon scraps secret AI recruiting tool that showed bias against women' (Reuters, 11 October 2018) <https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G> accessed 30 June 2022

Elam M, Reich R, 'Stanford HAI Artificial Intelligence Bill of Rights: A White Paper for Stanford's Institute for Human-Centered Artificial Intelligence' (2022) 4 <https://hai.stanford.edu/sites/default/files/2022-03/HAI%20Policy%20White%20Paper%20-

%20Stanford%20HAI%20Artificial%20Intelligence%20Bill%20of%20Rights.pdf> accessed 30 June 2022

Ehteshami BE, Veta M, van Diest PJ, 'Diagnostic Assessment of Deep Learning Algorithms for Detection of Lymph Node Metastases in Women With Breast Cancer' (2017) JAMA 318 <https://jamanetwork.com/journals/jama/fullarticle/2665774> accessed 29 June 2022

Fierens M, Rossello S, Wauters E, *Artificial Intelligence and the Law* (Cambridge University Press 2021) 53

Fjeld J, Achten N, Hilligos H, Nagy A, Srikumar M, 'Principled Artificial Intelligence: Mapping Consensus in Ethical and Rights-Based Approaches to Principles for AI' (2020) Berkman Klein Center Research Publication 2020(1) 30 <http://dx.doi.org/10.2139/ssrn.3518482> accessed 30 June 2022

Freeman Engstrom D, Ho D, Sharkey C, Cuellar MF, 'Government by Algorithm: Artificial Intelligence in Federal Administrative Agencies' (2020) NYU School of Law 75 <https://ssrn.com/abstract=3551505> accessed 30 June 2022

George D, 'Virginia Law Bans Local Police Use of Facial Recognition Technology' (2021) National Law Review <https://www.natlawreview.com/article/virginia-law-bans-local-police-use-facial-recognition-technology> accessed 30 June 2022

Gilbert B, 'Facebook's own research showed Instagram to be toxic for teens, which it denies — then it pulled Instagram Kids anyway' *Business Insider* (27 September 2021) <https://www.businessinsider.com/facebook-instagram-toxic-for-teens-pulls-intagram-kids-2021-9> accessed 30 June 2022

Griffin A, 'Saudi Arabia grants citizenship to robot for the first time ever' *The Independent* (26 October 2017) <https://www.independent.co.uk/life-style/gadgets-and-tech/news/saudi-arabia-robot-sophia-citizenship-android-riyadh-citizen-passport-future-a8021601.html> accessed 30 June 2022

Hern A, 'Give robots personhood status, EU committee argues' *The Guardian* (12 January 2017) <https://www.theguardian.com/technology/2017/jan/12/give-robots-personhood-status-eu-committee-argues> accessed 30 June 2022

Hill K, 'Another Arrest, and Jail Time, Die to a Bad Facial Recognition Match' *The New York Times* (29 December 2020) <https://www.nytimes.com/2020/12/29/technology/facial-recognition-misidentify-jail.html> 30 June 2022

Ho D, King J, Wald R, Wan C, 'Building a National AI Research Resource: A Blueprint for the National Research Cloud' (2021) Stanford HAI 10 <https://hai.stanford.edu/sites/default/files/2021-10/HAI_NRCR_2021_0.pdf> accessed 30 June 2022

Lemley M, Casey B, 'You might be a robot' (2020) Cornell Law Review 7 <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3327602> accessed 30 June 2022

MacKinnon E, King J, 'Regulating AI Through Data Privacy' (2022) Stanford HAI <https://hai.stanford.edu/news/regulating-ai-through-data-privacy> accessed 30 June 2022

Manheim K, Kaplan L, 'Artificial Intelligence: Risks to Privacy and Democracy' (2019) Yale Journal of Law and Technology 182 <https://yjolt.org/sites/default/files/21_yale_j.l._tech._106_0.pdf> accessed 30 June 2022

Manning C, 'Artificial Intelligence Definitions' (*Stanford University Human-Centered Artificial Intelligence*, 2020) <https://hai.stanford.edu/sites/default/files/2020-09/AI-Definitions-HAI.pdf> accessed 30 June 2022

Omaar H, 'Letter: Creating an AI Bill of Rights is a distraction' *Financial Times* (22 October 2021) <https://www.ft.com/content/3069b0ac-56c0-4c40-817f-f2d180f1ebc7> accessed 30 June 2022

Rich E, 'Artificial Intelligence and the Humanities', (1985) 19(2) Computers and the Humanities 117 <http://www.jstor.org/stable/30204398> accessed 30 June 2022

Stolyarov G, Tétrault-Farber G, 'Face Control: Russian police go digital against protesters' (Reuters, 11 Feburary 2021) <https://www.reuters.com/article/us-russia-politics-navalny-tech-idUSKBN2AB1U2> accessed 26 June 2022

Stricke B, 'People v. Robots: A Roadmap for Enforcing California's New Online Bot Disclosure Act' (2020) Vanderbilt Journal of Entertainment and Technology Law 847 <https://scholarship.law.vanderbilt.edu/cgi/viewcontent.cgi?article=1032&context=jetlaw> accessed 30 June 2022

Swinfen [GJ], Daniels [S], *Digital Governance: Leading and Thriving in a World of Fast-Changing Technologies* (Routeledge 2019) 132

Tolan S, Miron M, Gomez E, Castillo C, 'Why Machine Learning May Lead to Unfairness: Evidence from Risk Assessment for Juvenile Justice in Catalonia' (2019) ICAL 83 < https://doi.org/10.1145/3322640.3326705> accessed 30 June 2022

Turing A, 'Computing machinery and intelligence' (1950) 59(236) Mind 433 <http://www.jstor.org/stable/2251299> accessed 30 June 2022

Veale M, Zuiderveen BF, 'Demystifying the Draft EU Artificial Intelligence Act' (2021) Computer Law Review International 99 <https://arxiv.org/pdf/2107.03721.pdf> accessed 30 June 2022

Verani E, 'Symbolic AI vs Machine Learning in natural language processing' (*Inbenta*, 4 March 2020) <https://www.inbenta.com/en/blog/symbolic-ai-vs-machine-learning/> accessed 30 June 2022

Vincent J, 'Google fixed its racist algorithm by removing gorillas from its image-labeling tech' (*The Verge* 12 January 2018) <https://www.theverge.com/2018/1/12/16882408/google-racist-gorillas-photo-recognition-algorithm-ai> accessed 30 June 2022

Vought R, 'Memorandum for the Heads of Executive Departments and Agencies: Guidance for Regulation of Artificial Intelligence Applications' (2020)

<https://www.whitehouse.gov/wp-content/uploads/2020/01/Draft-OMB-Memo-on-Regulation-of-AI-1-7-19.pdf> accessed 30 June 2022

Wenderhorst C, 'The Proposal for an Artificial Intelligence Act COM(2021) 206 from a Consumer Policy Perspective' (2021) 64 <https://ssrn.com/abstract=4087402> accessed 30 June 2022

Laws and Cases

European Parliament, 'EU Guidelines on ethics in artificial intelligence: Contex and implementation' (2019) 2 <https://www.europarl.europa.eu/RegData/etudes/BRIE/2019/640163/EPRS_BRI(2019)640163_EN.pdf> accessed 30 June 2022

Commission, 'EU Declaration on Cooperation on Artificial Intelligence' (Declaration) COM (2018) <https://ec.europa.eu/jrc/communities/en/node/1286/document/eu-declaration-cooperation-artificial-intelligence> accessed 30 June 2022

Commission, 'Ethics guidelines for trustworthy AI' (Guidelines) COM (2019) <https://op.europa.eu/en/publication-detail/-/publication/d3988569-0434-11ea-8c1f-01aa75ed71a1> accessed 30 June 2022

Commission, 'High-Level expert group on artificial intelligence' (Policy) COM (2020) <https://ec.europa.eu/digital-single-market/en/high-level-expert-group-artificial-intelligence> accessed 30 June 2022

Commission, 'Sectoral considerations on the policy and investment recommendations for trustworthy artificial intelligence' (HLEG) (2020) < https://futurium.ec.europa.eu/en/european-ai-alliance/document/ai-hleg-sectoral-considerations-policy-and-investment-recommendations-trustworthy-ai> accessed 30 June 2022

Commission, 'Proposal for a Regulation of the European Parliament and of the Council laying down harmonized rules on artificial intelligence (Artificial Intelligence Act) and

amending certain Union legislative acts' (Proposal) COM (2021) <https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1623335154975&uri=CELEX%3A52021PC0206> accessed 30 June 2022

Commission, 'Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions: Coordinated Plan on Artificial Intelligence' (Plan) COM (2018) <https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:52018DC0795&rid=3> accessed 30 June 2022

Council, 'Consultative Committee of the convention for the protection of individuals with regard to automatic processing of personal data: Guidelines on artificial intelligence and data protection' (2019) <https://rm.coe.int/guidelines-on-artificial-intelligence-and-data-protection/168091f9d8> accessed 30 June 2022

European Parliament, 'EU Guidelines on ethics in artificial intelligence: context and implementation' (Briefing) (2019) <https://www.europarl.europa.eu/RegData/etudes/BRIE/2019/640163/EPRS_BRI(2019)640163_EN.pdf> accessed 30 June 2022

Malta Digital Innovation Authority, 'AI ITA Guidelines' (Guidelines) (2019) 26 <https://mdia.gov.mt/wp-content/uploads/2019/10/AI-ITA-Guidelines-03OCT19.pdf> accessed 30 June 2022

Websites

An EU Artificial Intelligence Act for Fundamental Rights: A Civil Society Statement (2021) 3 <https://edri.org/wp-content/uploads/2021/12/Political-statement-on-AI-Act.pdf> accessed 30 June 2022

Chen J, 'Neural Network' (*Investopedia*, 8 December 2021) <https://www.investopedia.com/terms/n/neuralnetwork.asp> accessed 30 June 2022

Executive Office of the President, 'National Science and Technology Council Committee on Technology: Preparing for the Future of Artificial Intelligence' (2016) <https://obamawhitehouse.archives.gov/sites/default/files/whitehouse_files/microsites/ostp/NSTC/preparing_for_the_future_of_ai.pdf> accessed 30 June 2022

Executive Office of the President, 'Artificial Intelligence, Automation, and the Economy' (2016) <https://obamawhitehouse.archives.gov/sites/whitehouse.gov/files/documents/Artificial-Intelligence-Automation-Economy.PDF> accessed 30 June 2022

'Facial Recognition and Biometric Technology Moratorium Act Reintroduced in Congress' (2021) SMD Magazine <https://www.sdmmag.com/articles/99560-facial-recognition-biometric-technology-moratorium-act-reintroduced-in-congress> accessed 30 June 2022.

Gibson Dunn, '2021 Artificial Intelligence and Automated Systems Annual Legal Review', (2022) <https://www.gibsondunn.com/2021-artificial-intelligence-and-automated-systems-annual-legal-review/#_ftn13> accessed 30 June 2022

Kavlakoglu E, 'AI vs Machine Learning vs Deep Learning vs Neural Networks: What is the difference?' (*IBM*, 27 May 2020) <https://www.ibm.com/cloud/blog/ai-vs-machine-learning-vs-deep-learning-vs-neural-networks> accessed 30 June 2022

National Science and Technology Council, 'The National Artificial Intelligence Research and Development Plan' (2016) <https://obamawhitehouse.archives.gov/sites/default/files/whitehouse_files/microsites/ostp/NSTC/national_ai_rd_strategic_plan.pdf> accessed 30 June 2022

National Security Commission, 'Final Report from National Security Commission on Artificial Intelligence' (2021) 11 <https://www.dwt.com/-/media/files/blogs/artificial-intelligence-law-advisor/2021/03/nscai-final-report--2021.pdf> accessed 30 June 2022

Prime Minister of Canada, 'Mandate for the international panel on artificial intelligence' (Mission statement) (2018) <https://pm.gc.ca/en/news/backgrounders/2018/12/06/mandate-international-panel-artificial-intelligence> accessed 30 June 2022

The White House, ‘Artificial Intelligence for the American People: An Executive Order on AI’ (2019) <https://trumpwhitehouse.archives.gov/ai/executive-order-ai/> accessed 30 June 2022

The White House Briefing Room, ‘Americans Need a Bill of Rights for an AI powered world’ (2021) <https://www.whitehouse.gov/ostp/news-updates/2021/10/22/icymi-wired-opinion-americans-need-a-bill-of-rights-for-an-ai-powered-world/> accessed 30 June 2022