White Paper

Regulating AI:
The Role of Co-Regulation

Find out more on our website:
www.hufeld.com

Author: Lukas Hufeld, LL.M., *Attorney at HUFELD LAW, hufeld.lukas@hufeld.com*
Published on: 1 Aug 2024

HUFELD LAW
hufeld-law@hufeld.com
Hackenstraße 2, 80331 Munich
Germany

TABLE OF CONTENTS

A. The AI Revolution

As AI applications have moved to the center of society, superlatives are no longer spared. Sundar Pichai, CEO of Google's parent company Alphabet, has described AI as the “most profound technology humanity is working on. More profound than fire, electricity, or anything that we have done in the past."[1] Geoffrey Hinton, who has been called the ‘Godfather of AI,’ left his role at Google just so he could „talk about the dangers of AI”[2] and Warren Buffet compared the impact of AI to the creation of the atomic bomb[3].

It may have been phrases like these that helped the European AI-Act see the light of day much faster than other EU legislative projects. Or perhaps it was the fact that the gap with the leading AI nations was widening. In any case, something had to be done to strengthen Europe as an AI hub. And now it's here: A bold attempt to not only close the gap with the US and China, but also to lay the foundations for the responsible and sustainable use of AI.

The US is taking a different approach and is holding back on AI regulation for the time being. Arguing that the rapidly advancing technology cannot be regulated anyway, the responsibility for self-regulation is placed in the hands of business leaders.[4] That is why, in July 2023, seven leading AI companies agreed on voluntary safeguards for the development of the technology and committed to new standards of safety, security and trust at a meeting at the White House.[5]

Indeed, all modern economies are now asking whether and how AI should be regulated. Many see the ideal solution as a delicate balance between government intervention and self-regulation. Against this backdrop, it is worth taking a closer look at the AI-Act and the co-regulatory mechanisms it provides for.

This paper aims to develop an understanding of why regulation is needed (B) and will highlight its main regulatory concepts (C). The paper then examines what co-regulatory mechanisms are provided for in the AI Act (D) and how a co-regulatory framework needs to be designed to be effective (E). The paper concludes with a summary (F) and an answer to the question of how much self-regulation the AI industry can tolerate.

AI Technology

AI systems possess the capability to analyze and organize enormous amounts of data, commonly referred to as Big Data. Their developers can "train" them to not only identify pre-existing patterns in data sets but also to draw independent conclusions and adapt to evolving circumstances.[6] As a result, AI systems can recognize novel meanings and connections in real-time, providing them with unique predictive abilities. This process of generating new knowledge from existing information is referred to as machine learning, and a specific learning technique called deep learning emulates the human brain through artificial neural networks (KNN). The multilayered structure of KNNs enables them to attain superior learning outcomes, which could potentially increase as they grow and strength in the future.[7]

B. Why is Regulation needed?

Politicians and the industry expect AI research to be the key to a (fully) automated future in which all areas of life can be improved on the basis of the constantly increasing volume of data.[8] The use of AI systems promises, for

[1] Nolan, INSIDER, Apr 17, 2023.
[2] Korn, CNN Business, May 3, 2023.
[3] D’Cruze, Business Today, May 8, 2023.
[4] Ammananth, FORTUNE, January 16, 2023.
[5] Shear et al., New York Times, July 21, 2023.
[6] Martini, Blackbox, p. 21 et seqq.
[7] Ibid.; Botta, ZfDR 2022, p. 391 (391).
[8] Botta, ZfDR 2022, p. 391 (394).

example, a lower-threshold Alzheimer's diagnosis thanks to intelligent speech recognition, or an accident-free road traffic thanks to autonomous vehicles and intelligent traffic management. Mass procedures in public administration could be fully automated through AI, enabling accelerated access to government services. Further areas of application for AI systems are emerging in the face of climate change. It almost seems as if AI offers a solution for every challenge.[9]

However, the enormous potential of learning systems not only creates opportunities but also significant risks. For example, surveillance methods such as facial recognition or profiling can be optimized using AI, which can erode civil rights. The use of AI systems also risks deepening societal divides. If the personal biases of developers (conscious or unconscious) are incorporated into the software code or the selection of datasets, discrimination can be perpetuated. For example, the Austrian labor market promotion algorithm disadvantaged mothers with dependent children in the allocation of support measures.[10] An increased risk of discrimination also arises when providers of AI systems promise mathematical accuracy, even though this technology often only shows correlations but not causality. In addition, AI systems are sometimes associated with a significant transparency deficit. Opaque technology and systems may contain errors and undesirable discrimination, leading to a lack of decision transparency and unaccountability. With learning systems, even the programmers cannot trace every step on the way to a result. Discrimination factors embedded in the algorithm are therefore difficult to identify and remedy.[11] Naive assumptions about the quality, comprehensiveness, and compatibility of data from different sources can also lead to erroneous decisions and biased outcomes. The broader implications of these issues include the undermining of human rights, the unrecallable delegation of autonomy to weak and insufficiently adaptable models of reality, the disruption of culture and work-based income distribution and the potential for the meaninglessness of human life.[12] Therefore, it is essential to develop ethical frameworks and governance mechanisms to ensure that AI development and deployment do not pose all these significant threats to society.

The multi-layered and hardly manageable field of artificial intelligence poses great challenges to common regulatory concepts. AI systems are characterized by various features that must shape the requirements for possible regulatory approaches.

Pace is one of the most significant challenges is the rapid pace of technological advancement and the diversity of AI applications. The high speed of innovation enables the daily emergence of new AI applications. At the same time, the ability to innovate is not limited to specific areas, but affects a wide variety of industries. The spread of AI applications is also happening more rapidly than with previous technologies. ChatGPT, the popular chatbot from OpenAI, reached 100 million monthly active users just two months after launch, making it the fastest-growing consumer application in history.[13] This makes it difficult for regulators to keep up with the latest developments and assess the risks associated with new AI technologies. As the innovation cycles become shorter and shorter, there is a need for flexible regulation that can

[9] Ibid.
[10] Ibid.
[11] Botta, ZfDR 2022, p. 391 (391).
[12] Clarke, Computer Law & Security Review 35/2019, 423 (426); Scherer, Harvard J Law Technol, 353 (362 et seqq.); Yampolskiy et al, arXiv 2016, p. 3 et seqq.; Botta, ZfDR 2022, p. 391 (391).
[13] Hu, Reuters, February 2, 2023.

adapt to changing markets at short notice. Furthermore, rules must be issued quickly to still be relevant when they come into force.

 Autonomy refers to the ability of AI to make decisions that are not or only minimally predetermined by humans, and these decisions can have legal consequences such as damages.[14] However, it is difficult to predict what decision an AI will make in a given situation, as there are always varying elements such as health conditions when it comes to surgery assistance or weather and road conditions when it comes to self-driving cars.[15]

 Opacity refers to the difficulty in understanding how decisions are made that result in legally relevant outcomes, particularly in complex forms of AI based on machine learning. AI uses data as input to generate an output as a result. While it is possible to generally understand how AI goes from input to output, it is difficult to explain why an AI arrives at a specific outcome and thus a legally relevant result based on certain data.[16] These Characteristics make it challenging to understand their decision-making processes and require a deep understanding of machine learning, statistics, and computer science. Regulators may not have the necessary expertise to understand the technology fully.[17]

 Regulators will also need to consider how to strike the right balance between innovation and security, as well as how to address the global nature of AI and harmonize regulations across borders. Potential conflicts arising from political competition and different values need to be recognized. Otherwise, a global landscape of AI regulation, fragmented by national policies and industry initiatives, could lead to a "race to the bottom", allowing companies to 'cherry-pick' their favorable regulation and leaving risks unaddressed.[18] Therefore, especially regarding fundamental value-based decisions, national legislation is called for.[19] While this could exacerbate fragmentation, more stringent local regulations might also establish a global standard, much like the EU's data protection law. In both scenarios, competition often takes precedence over cooperation. Nevertheless, the feasibility of cooperation will ultimately be determined by the specific regulatory frameworks at both national and international levels. While this could reinforce fragmentation, stricter local rules could also set a global benchmark, similar to EU-data protection law.[20] There is a need to develop robust ethical frameworks and governance mechanisms that ensure the responsible and transparent use of AI, while respecting human rights and privacy. To address these challenges, regulators need to work closely with industry, academia and civil society to develop comprehensive and flexible regulations that balance innovation, security and ethical considerations. For this purpose, the regulator can make use of three essential regulatory mechanisms which are to be examined below for their effectiveness.

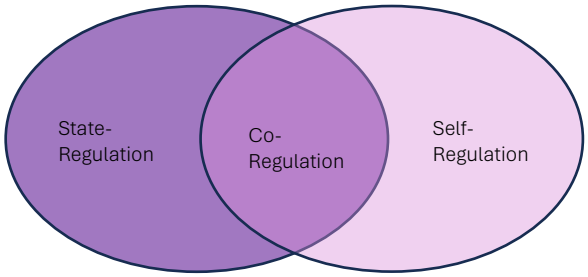
C. Regulatory concepts

The term 'regulation' has various definitions. For the purpose of this paper, we adopt a practical definition of regulation as the exertion of control over the behavior of entities.[21] This definition is intentionally broad, encompassing both deliberate policy instruments

[14] Staudenmayer, NJW 2023, 894 (895).
[15] Ibid.
[16] Ibid.
[17] Ibid.
[18] Brake, Lex Electronica, 2020, p. 20; Wagner, Amsterdam University Press 2018, p. 86.
[19] Brake, Lex Electronica, 2020, p. 20; Scherer, Harvard J Law Technol, 353 (379 et seqq); Wagner, Amsterdam University Press 2018, p. 86; Hoffmann-Riem, Regulating Artificial Intelligence, p. 8.
[20] Brake, Lex Electronica, 2020, p. 20.
[21] Clarke, Computer Law & Security Review 35/2019, p. 398 (399).

and unintentional control mechanisms. It avoids specifying the methods employed and does not address the specific goals of regulation. The objectives of regulatory schemes are often subject to debate and will therefore be elaborated in the following.

Furthermore, it should be recognized that a wide regulation over the different industries will be challenging, and no one-size-fits-all regulatory solution will be appropriate when it comes to maximizing the potential of this new technology.[22] While a one-size-fits all regulatory approach is not practicable, several universal legal and ethical themes have to be tackled by state regulation, self-regulation and a mixture of these two, co-regulation.



I. State-Regulation

1. Definition

State regulation, also known as government regulation, refers to the process by which a government or state authority establishes and enforces rules, laws, and policies to govern various aspects of society, including economic activities, public services, public health and safety, consumer protection, environmental protection, and more. In state regulation, the state intervenes directly in market processes and influences the behavior of companies through regulations aimed at achieving certain objectives in the general interest. In this context, only those sovereign behavioral restrictions are to be covered which limit the freedom of trade and contract of economic entities by means of generally applicable standards.[23] Thus, state regulation is distinguished from other market interventions, such as taxation or subsidies. State intervention is justified by market failure and the need to safeguard property rights and rights of disposal as a basis for market development.[24]

Regulation is carried out by the state organs of the legislative, executive and judicial branches as well as by institutions that are bound by instructions to the state organs. The regulations issued are binding for all players. In this way, cross-sector regulations can be made to which all must adhere. The enforcement power of the specialized supervisory authorities also ensures a sanctions regime that verifies compliance with rules. However, in order to legitimize these rules under the rule of law, a democratic procedure is first required. Here, the prescribed legislative or administrative procedure must be followed before laws, regulations or administrative acts can be enacted.[25]

2. Current Status

Several countries have already implemented various regulations that impact the use of AI. Data protection and privacy regulation like Europeans General Data Protection Regulation (GDPR), California's Consumer Privacy Act (CCPA) or Brazil's General Data Protection Law (LGPD) have large implications for AI systems that rely on personal data such as machine learning algorithms or facial recognition technology. European Union's (EU) Digital Markets Act (DMA) has also emerged as one of the significant regulatory frameworks targeting the regulation of Artificial Intelligence. While not exclusively focused on AI, it contains provisions that have implications for AI-related practices and services such as transparency requirements for large online

[22] Piepgrass et al., RTInsights, May 10, 2022.
[23] Spindler, Eckpunkte, p. 19 et seqq.
[24] Spindler, Eckpunkte, p. 19 et seqq.
[25] Hufeld, rescriptum 2017/2, p. 107 (109).

platforms with significant market power ("gatekeepers"), ensuring that they provide clear information on how they use AI algorithms. However, when it comes to AI in a broader sense, the prominent jurisdictions worldwide had not established dedicated regulatory agencies or adequately empowered the existing ones.[26] The AI Act of the European Union is now one of the first state regulative regimes that tackles AI in a broad range. Its importance cannot be underestimated which is why this thesis will be devoted to it below.

3. Criticism

While state regulation enjoys a high degree of democratic legitimacy, the democratic decision-making process is not infrequently lengthy, which can lead to legislators merely lagging the legal problems at hand. State regulations are often slow to adapt to fast-paced technological advancements. AI technologies evolve rapidly, and static regulations might not keep up with the emerging challenges and opportunities. [27] This inflexibility can result in outdated rules that no longer address the most critical issues in the AI landscape.

In addition, critics of state regulation point to knowledge and competence deficits among state decision makers, which can lead to failure to solve problems and unintended side effects. Among other things, government regulation can contribute to inhibiting the initiative, innovation and sense of responsibility of the objects of control and to triggering resistance. This resistance can consist of seeking evasion strategies and loopholes or mobilizing political pressure against regulatory initiatives, for example by arguing that jobs are at risk.[28]

Another concern that is always mentioned in connection with regulation from Brussels – including the AI Act – is that overly stringent regulations may stifle innovation and hinder the development of AI technologies. Imposing rigid rules could slow down progress, making it difficult for businesses and researchers to experiment and create new solutions. Regulations that focus heavily on risk avoidance may lead to overly conservative AI systems that miss out on potential benefits due to an excess of caution. At the same time, critics argue that costly implementation and compliance with AI regulations might disproportionately affect smaller players, potentially limiting competition and innovation.[29]

4. State-Regulating AI

The strongest and hardly refutable argument for state intervention is the dangerousness of artificial intelligence. There is a valid argument, that the more influential a technology is, the more compelling the need for proactive action by governing bodies becomes. Technologies like nuclear energy and large-scale extractive and manufacturing industries have suffered from inadequate regulation, leading to considerable pollution. These instances are often cited as examples where anticipatory action by parliaments is crucial.[30]

The "precautionary principle," articulated in the Wingspread conference in 1998[31], has been incorporated into environmental laws in some jurisdictions.[32] This principle suggests

[26] Clarke, Computer Law & Security Review 35/2019, p. 398 (403).
[27] Porket, Czech Sociological Review, p. 311 (315).
[28] Schulz, Selbstregulierung, p. A-10; Hufeld, rescriptum 2017/2, p. 107 (109).
[29]Porket, Czech Sociological Review, p. 311 (315 et seqq.)
[30] Clarke, Computer Law & Security Review 35/2019, p. 398 (401).

[31] On January 26th, 1998 at Wingspread, headquarters of the Johnson Foundation scientists, philosophers, lawyers and environmental activists, reached an agreement on the necessity of the Precautionary Principle in public health and environmental decision-making, The Science and Environmental Health Network on the Wingspread Conference, August 5, 2013.
[32] For a general analysis of the implications of this principle on legal systems and on regulation policies, see,

that if human activities have the potential to cause morally unacceptable harm that is scientifically plausible but uncertain, measures should be taken to avoid or minimize that potential harm. While the precautionary principle is primarily associated with environmental matters in specific jurisdictions, it generally serves as an ethical guideline. It states that if an action or policy is suspected of causing harm and there is a lack of scientific consensus on its safety, the burden of proof falls on those undertaking the action.[33]

As we have seen above (paragraph B. II.), AI poses identifiable and significant threats. Even if this argument is not universally accepted, the projected impact of AI, as anticipated by its proponents, is so substantial that the precautionary principle, at least in its weaker form, becomes applicable.[34] What is clear is that the quality of regulation must be closely connected to the potential harm of AI since it is the fundamental purpose of governments, to safeguard their citizens while also ensuring their own protection.

If governments perceive a credible and imminent existential threat posed by AI, they may take de facto control of AI companies and subject them to regulation as national security assets. Some governments may even consider nationalizing AI companies or banning AI research and development altogether. While these measures may seem radical, AI technology has the potential to have a significant impact on society, including issues of privacy, bias, discrimination and transparency. It is unlikely that any government would allow private companies to control a technology that, in their view, has the potential to put their citizens, themselves, and the global community at risk.[35]

A government takeover of the AI industry represents the most extreme scenario of the state regulatory spectrum. Alternatively, if society acknowledges that AI can cause harm but does not constitute an existential threat, more moderate forms of regulation become feasible. One moderate regulatory option involves criminalizing undesirable outcomes through the criminal code. Individuals deploying AI for fraudulent purposes or to harm and stalk others could be held criminally liable. Similarly, laws could be enacted to impose criminal liability on those responsible for creating harmful AI in the first place.[36]

Another possibility would be civil regulation, which entails imposing financial penalties for violations. This could potentially deter undesirable outcomes linked to AI by establishing the threat of civil liability. For instance, existing statutes that pertain to matters like race and sex discrimination could be applicable to AI makers and users, making them liable under civil law. To address the specific social wrongs associated with AI more effectively, dedicated statutes could also be introduced.[37]

A third, intermediate alternative is the creation of administrative rules and agencies with the power to formulate the necessary regulations and ensure their enforcement. This strategy relies on administrative expertise. While potential concerns exist about industry influencing the agencies, the administrative approach may be better suited to the complex and specialized field of AI, as opposed to direct congressional control.[38]

C. R. Sunstein, Beyond the Precautionary Principle, in SSRN, January 2003.

[33] Clarke, Computer Law & Security Review 35/2019, p. 398 (401).

[34] Clarke, Computer Law & Security Review 35/2019, p. 398 (401).

[35] Feldmann, The Washinton Post, April 2, 2023.

[36] Ibid.

[37] Ibid.

[38] Clarke, Computer Law & Security Review 35/2019, p. 398 (403).; Ibid.

At the other end of the spectrum, a more hands-off approach to regulation means resorting to litigation. Under the tort liability framework in the United States, developers or vendors of technology are expected to exercise "reasonable care". If they fall short of this standard, parties affected by their actions can take legal action to seek financial compensation for the harm they have suffered. This system holds technology companies legally accountable for their actions or omissions.[39]

In devising the optimal regulatory mechanisms for AI, a balance must be struck between facilitating innovation, ensuring ethical conduct, and safeguarding against potential harm. A comprehensive approach might encompass a combination of the mentioned methods, tailored to the unique complexities of the AI landscape.

II. Self-Regulation

1. Definition

An alternative to government regulation is self-regulation. It involves the establishment, application, and enforcement of rules by private entities in their own affairs without explicit legal foundations and frameworks. An individual, a private organization (e.g., a company), or a group (e.g., an industry association) voluntarily adopts a code of conduct and independently monitors its compliance and enforcement.[40] Self-regulation can include areas such as corporate governance, responsible marketing, environmental sustainability, workplace practices, and product safety. It requires individuals and organizations to proactively monitor and evaluate their actions, behaviors, and impact on stakeholders, aiming to align their operations with ethical principles and industry standards.[41]

In the high-tech field of AI, alongside such behavioral rules for corporate ethics and morals, the establishment of technical standards plays a prominent role.[42] Such acts of self-regulation derive their legitimacy from the autonomy of the parties involved. By signing a code of conduct, a self-binding commitment is made, in the sense of establishing a duty program.[43] This duty program differs from a purchase contract for a smartphone in that it creates new obligations not explicitly specified in the law, which can be adopted by a multitude of companies.[44] Especially when a leading company sets a technical standard, the rest of the industry will follow suit to maintain interoperability. By practicing self-regulation, businesses also strive to build trust, maintain a positive reputation and help to prevent the need for excessive external regulation.

2. Economic implications

The economic argument for self-regulation or rests on the belief that individuals, when acting in their own self-interest, will indirectly contribute to the overall welfare of society. This idea was famously described by Adam Smith in his seminal work "The Wealth of Nations" (1776) as the "invisible hand" guiding economic activities towards socially beneficial outcomes. Additionally, transaction cost economics, as developed by Oliver Williamson (1979), posits that self-regulating mechanisms arise naturally within economic systems as a way to reduce transaction costs and increase efficiency.[45]

However, the concept of self-regulation also has its limitations. One of the most well-

[39] Ibid.
[40] Spindler, Eckpunkte, p. 25 et seqq.
[41] Clarke, Computer Law & Security Review 35/2019, p. 398 (404).
[42] Latzer et al., Mediatiksektor, p. 31.
[43] Podszun, Corporate Social Responsibility, p. 73.
[44] Ibid.
[45] Clarke, Computer Law & Security Review 35/2019, p. 398 (401).

known examples is the "Tragedy of the Commons," as originally described by Garrett Hardin in 1968. It illustrates the problem of overexploitation of shared resources when individuals pursue their self-interest without considering the collective consequences. This scenario highlights a case where self-regulation may fail to protect the common good.[46]

Despite the faith in self-regulation by neoclassical economists, Joseph Stiglitz[47] brings forward alternative perspectives. He introduces the notion of "market irrationality," where individual behaviors in markets can deviate from rationality, leading to inefficiencies and potential market failures. For example, during periods of extreme volatility, herd behavior among investors can cause prices to deviate significantly from their fundamental values. To address such problems, Stiglitz suggests interventions such as circuit breakers to temporarily halt trading and prevent abrupt market collapses.[48]

Furthermore, Stiglitz advocates for considering "distributional justice" alongside self-regulation. While AI technology could increase the productivity of workers or change the composition of their tasks, it could also deprofessionalize jobs, as low-paid, less-skilled workers with AI systems could outperform current professionals. In the long run, this process could still increase the productivity of the economy, but the impact of labor market restructuring on wages and unemployment will depend on the distribution of productivity gains, which in turn will depend on government policy.[49] Distributional justice involves recognizing that market outcomes may not always be equitable, and certain segments of society may be left behind or unfairly treated. To address these concerns, interventions like safety nets and anti-discrimination measures are proposed to ensure a fair distribution of resources and opportunities.[50]

3. Advantages

Self-regulation offers numerous advantages over government regulation. One significant aspect is the high degree of flexibility and speed it provides. Private rule-setters are not as bound by state principles as the government legislator, allowing for a much broader scope in establishing, applying, and interpreting rules. This enables a quick adaptation to the ever-changing conditions of an information, production, and communication society.[51]

Another benefit of self-regulation is the professionalism and expertise it brings. Private actors' specialized knowledge contributes to formulating specific regulations that best suit the respective circumstances.[52] Unlike legislative "one size fits all" solutions, this approach ensures a higher level of precision for businesses and consumers. Furthermore, self-regulation can effectively achieve legal standards that might be challenging to enforce due to limited governmental regulatory capacity.[53]

Private self-regulation also stimulates creativity, initiative, and innovation among market participants. Commitment agreements have the advantage of avoiding unnecessary

[46] Ibid.
[47] Stiglitz, Government failure vs. market failure, p. 5 et seqq.
[48] Clarke, Computer Law & Security Review 35/2019, p. 398 (401).
[49] Stiglitz, The Future of Work Speech.
[50] Clarke, Computer Law & Security Review 35/2019, p. 398 (401); Stiglitz, Government failure vs. market failure, p. 5 et seqq.
[51] OECD Report, Alternatives to traditional Regulation, p. 6; Buck-Heeb/Dieckman, Selbstregulierung im Privatrecht, p. 220 et seqq.
[52] Kleinsteuber, Regulation.
[53] Buck-Heeb/Dieckman, Selbstregulierung im Privatrecht, p. 220 et seqq.

coercion, granting involved actors greater autonomy in implementation. When regulations are voluntarily followed out of conviction, acceptance of the regulation increases, strengthening the principles of freedom and liberalism.[54]

Additionally, private regulation allows for the involvement of market actors operating across borders. This becomes especially crucial in the context of AI, where modern information and communication technologies have led to globalization and the presence of international companies. State regulations often fail in such cases due to the lack of a central authority to harmonize diverse legal and social norm systems regionally. However, private self-regulatory measures, in which individual actors have a vested interest in compliance, can overcome these boundaries.

4. Criticism

On the other hand, self-regulation also has its weaknesses. There is the argument that self-regulation circumvents democratic processes and only sets its own "rules of the game" in line with economic interests. In fact, Critics note that "self-regulation has a formidable history of industry abuse of privilege"[55]and that 'voluntarism' is generally an effective regulatory element only when it exists in combination with 'command-and-control' components.[56] However, industry self-regulatory mechanisms are, by their very nature, non-binding and unenforceable under the influence of trade practices, anti-monopoly and anti-cartel laws.[57]

The lack of a binding force of private regulation also entails that complete industry coverage is rarely achieved. On one hand, this leads to the ineffectiveness of regulatory measures for consumers, and on the other hand, it can lead to significant distortions of competition.[58] While companies that comply with private regulations face considerable costs or limitations in product design and marketing, others who have not committed to such rules can still benefit from the positive effects of the regulations (the so-called "free-rider dilemma"[59]). Additionally, for small and medium-sized enterprises with insufficient personnel and financial resources, participating in self-regulatory institutions can often be impossible.

5. Self-Regulating AI

It is crucial to acknowledge that solely relying on self-regulation is not sufficient to tackle all the challenges associated with AI. The complexity and potential risks posed by AI technologies often require a comprehensive and balanced regulatory framework that combines self-regulation with targeted interventions and government oversight. Some Critics already argue that the call for self-regulation was largely ignored by the digital industry and a “great but missed historical opportunity”.[60]

While it is far too early to declare the instrument of self-regulation dead, it must be admitted, that the field of AI lacks meaningful self-regulation at various levels, including organizational, industry, and professional. The existing codes and guidelines cover only a

[54] Podszun, Corporate Social Responsibility, p. 78; OECD Report, Alternatives to traditional Regulation, p. 6.
[55] Braithwaite, Types of responsiveness, p. 124.
[56] Gunningham/Darren, Smart regulation p. 137 ff.
[57] Clarke, Computer Law & Security Review 35/2019, p. 398 (404).
[58] Spindler, Eckpunkte, p. 34 et seqq.
[59] The free rider problem is that the efficient production of important collective goods by free agents is jeopardized by the incentive each agent has not to pay for it: if the supply of the good is inadequate, one’s own action of paying will not make it adequate; if the supply is adequate, one can receive it without paying, Hardin/Cullity, The Stanford Encyclopedia of Philosophy.
[60] Floridi, Philosophy & Technology, p. 619 (622).

fraction of the necessary regulations and are often difficult to enforce.[61]

For example, in 2016, six major technology companies –Amazon, Apple, Google, Facebook, IBM and Microsoft– formed the Partnership on Artificial Intelligence to Benefit People and Society to „study and formulate best practices on AI technologies, to advance the public's understanding of AI, and to serve as an open platform for discussion and engagement about AI and its influences on people and society." [62] An initiative like this is very positive and valuable but not highly sufficient since it lacks the legitimacy the State can provide. In fact, they leave out the myriad small to medium-sized enterprises that are also developing AI.[63] In subsequent years, notable incidents like the Facebook-Cambridge Analytica scandal in 2018 and the ill-conceived Advanced Technology External Advisory Council by Google in 2019 demonstrated further difficulties and ineffectiveness of self-regulation. It became evident that companies were either unwilling or unable to address their ethical problems comprehensively. Despite the industry's attempts to respond to AI's ethical concerns by generating numerous codes, guidelines, and statements, self-regulation lacked substance and authenticity. This led to a strong and widespread impression of "blue washing,"[64] where companies merely pretended to focus on ethical considerations, rather than taking meaningful action.[65]

It became clear that certain technology companies may choose to disregard voluntary rules that do not favor them, leading to advantages for some organizations over others. Furthermore, lacking a unified framework, the proliferation of private ethics committees could result in an excess of varying sets of rules. This situation could be chaotic and hazardous, like the consequences if every large company were to develop its own AI code, just as if each private citizen could establish individual legal statutes.[66] While technology companies might be well-suited to create rules due to their expertise in the field, they are rarely in the best position to thoroughly assess democratic, moral, ethical, and legal risks. History has demonstrated the potential outcomes when the State withdraws and allows private companies to set exclusive regulatory standards. Permitting such a scenario in the context of AI would not only be risky but also reckless.[67]

In conclusion, self-regulation can be a valuable tool to reinforce natural controls and enhance responsible practices in the AI domain. However, due to the evidence of market failure and distributive justice issues, self-regulation alone may not be sufficient when it comes to regulating AI. A well-designed regulatory framework that combines self-regulation with targeted interventions is essential to ensure the responsible and ethical development of AI and to address the challenges posed by this transformative technology. Therefore policymakers, industry stakeholders, and researchers must collaborate to establish clear, enforceable, and comprehensive regulations. This form of collaboration is commonly referred to as co-regulation.

III. Co-Regulation

1. Definition

The co-regulatory approach can be described as regulating social phenomena not only through top-down government intervention

[61] Clarke, Computer Law & Security Review 35/2019, p. 398 (401).
[62] Andrés, IUS, p. 48 et seqq.
[63] Ibid.
[64] "Bluewashing" is a term used to describe deceptive marketing that overstates a company's commitment to responsible social practices.
[65] Floridi, Philosophy & Technology, p. 619 (621).
[66] Andrés, IUS, p. 51 et seqq.
[67] Ibid.

but also involving private stakeholders in the rule-making process. In contrast to the self-regulatory approach, co-regulation maintains the involvement of public authorities throughout the rule-making process. This ensures public scrutiny of the co-regulatory outcomes, leading to greater transparency, representation of public interests, and respect for fundamental rights.[68]

Co-regulation often involves the establishment of industry bodies or associations that work alongside the regulatory authority to develop codes of conduct, standards, and guidelines specific to the industry. These industry bodies take on a self-regulatory role, but their actions and decisions are overseen and monitored by the regulatory authority. By involving the regulated entities in the regulatory process, co-regulation aims to promote more effective and efficient regulation, better industry compliance, and improved outcomes for stakeholders. It can foster innovation, flexibility, and a sense of ownership within the industry while maintaining necessary oversight from the government or regulatory body.[69]

2. Co-Regulating AI

The role of the state in the governance of AI can be understood in various ways. Two analytical dimensions can be distinguished conceptually[70]: Firstly, governments can either take a proactive approach, actively participating in the development of AI technologies (strong state intervention), or a more passive stance, stepping back and granting private actors and/or markets more autonomy in AI governance (weak state intervention). Secondly, governments can focus on regulating potential AI risks (enclosure-and-control approach), or prioritize promoting AI deployment and development (stimulation approach).[71]

The self-regulation-promoting state entails the idea that industry and relevant stakeholders within a field collaborate to regulate themselves and implement self-restricting mechanisms. In this approach, the role of the state is limited to acting as a facilitator, observer, or certifier of private initiatives for AI regulation and governance, rather than engaging directly.[72] The goal is risk prevention, but this is achieved through soft regulatory instruments such as codes of conduct, industry standards, quality seals, certification bodies, ombudsmen, arbitration/mediation boards, and ethics committees.[73] Which of these different approaches is preferred depends largely on the dangers of AI technology. This risk-based approach is particularly clear in the AI Act (see below).

As we have concluded above, state regulation is considered more superior and concerned about society's security compared to private actors. Public institutions have a proven track record of securing stable compliance, which is essential for maintaining just institutions and respecting privacy regulations over time. Even individuals with good intentions can be vulnerable to dynamic preference inconsistency, where the desire to act in self-serving ways conflicts with the general preference for doing what is right.[74] A lack of regulation can encourage behavior that infringes reasonable public expectations. Cavalier organizational behavior may be driven by executives, groups and even lone individuals who perceive opportunities.[75] This can give rise to substantial direct and indirect threats to the

[68] Vigna, Co-regulation Approach.
[69] EESC, Brochure, Foreword, p. 7.
[70] Borrás/Edler, Research Policy, p. 2.
[71] Djeffal et al., Journal of European Public Policy, p. 3.
[72] Borrás/Edler, Research Policy, p. 2.
[73] Djeffal et al., Journal of European Public Policy, p. 3.
[74] Ferretti, Moral Philosophy and Politics, p. 239 (246).
[75] Clarke, Computer Law & Security Review 35/2019, p. 398 (400).

reputation of every organisation in the sector. It is therefore in each organisation's own self-interest for a reasonably amount of regulation to exist, in order to provide a protective shield against media exposés, and to avoid stimulating a public backlash and regulatory activism.[76]

In order to introduce large-scale self-binding rules in the field of AI, it is crucial to install a system of checks and balances. The government is better suited for this role than private agents for various reasons. Firstly, the government has the ability to effectively force everyone, ensuring broader reach and compliance.[77] Secondly, the government can implement legal safeguards and constraints to prevent potential abuses of power and arbitrary interference by those in charge.[78] To enhance self-regulation, it is important for the state to draw insights from natural controls, especially from economic observations. One approach could be to modify the cost/benefit/risk balance perceived by the involved parties through various mechanisms such as cost subsidies, revenue levies, and risk allocation. For example, establishing strict liability for operating an AI-driven car could incentivize more careful risk assessment and risk management practices.[79]

The failures of a purely self-regulated AI industry have become evident, and the risks posed by AI are becoming more tangible. Legislation has been catching up and is likely to do so even more in the future. Particularly in the EU, regulatory initiatives such as the General Data Protection Regulation (GDPR) in force since 2016 have paved the way for further legislative actions, including the Digital Markets Act, the Digital Services Act, and the AI Act, among others. This regulatory movement is expected to create a substantial "Brussels effect," shifting the focus from soft-regulation, which has not been very effective, to legal compliance and the enforcement of penalties. The shift towards legal frameworks and stricter enforcement is aimed at better safeguarding individual rights and addressing the complex challenges presented by AI technologies.[80]

D. AI Act

On 21 May 2024, the Council of the 27 EU member states adopted the AI Act and thus a uniform framework for the use of artificial intelligence in the European Union. The legal text[81] was published on 12 July 2024 and came into force on 1 August 2024. It is the world's first comprehensive set of regulations for Artificial Intelligence. The aim of the AI Act is to strengthen trust in AI, promote the development of innovations and ensure that the safety and fundamental rights of European citizens are safeguarded when using AI.

The AI Act is not the first piece of legislation governing digital technologies in Europe. It follows the adoption of data protection regulations, including the General Data Protection Regulation (GDPR) in 2016 (Regulation 2016/679). The Data Governance Act (Regulation 2022/868) was adopted in 2022, followed by the Data Act (Regulation 2023/2854) in 2023. In addition, the Digital Markets Act (DMA) and the Digital Services Act (DSA) (Regulation 2022/1925; Regulation 2022/2065) were adopted in 2022 with the objective of regulating online platforms. However, the AI Act takes a different approach, drawing inspiration from European product safety rules.

[76] Ibid.
[77] Simon, PS: Political Science and Politics, p. 749, (750 et seqq).
[78] Ferretti, Moral Philosophy and Politics, p. 246.
[79] Clarke, Computer Law & Security Review 35/2019, p. 398 (401).
[80] Floridi, Philosophy & Technology, p. 619 (621).
[81] 2024/1689 REGULATION (EU) 2024/1689 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 13 June 2024.

It is a European harmonisation law on product conformity and is schematically designed according to the so-called New Legislative Framework (NLF), which to a certain extent embodies the current state of regulatory technology in European product safety law.[82] It thus fits into the complex system of European product regulation, which is enforced by the authorities on the basis of the European Market Surveillance Regulation (EU) 2019/1020. The AI Act also makes software as such subject to harmonised European product regulation for the first time.

I. A risk-based approach

The AI Act follows the principle of technology neutrality with specific provisions for various application areas and includes dynamic elements to accommodate ongoing technological developments. It aims to strike a balance between the legitimate need for protection regarding a novel technology with yet unknown implications and the fundamental requirement to enable innovation.[83] The law is based on a "risk-based regulatory approach": Depending on the risk associated with each AI application concerning European fundamental rights and values, higher or lower requirements are set, such as regarding the security or transparency of the AI application's processes.[84] The AI Act differentiates between four risk levels: unacceptable risk (Article 5), high risk (Article 6), low risk (Article 50) and the non-explicitly regulated minimal risk. In addition, Article 51 lays down provisions for general-purpose models (GPAI), which, depending on their intended use, must fulfil similarly strict requirements as high-risk systems.

AI systems with an unacceptable risk are prohibited. This includes certain use of subliminal techniques or exploitation of vulnerabilities of specific groups (children, disabled persons) likely to cause psychological or physical harm, social scoring for general purposes done by public authorities as well as certain use of 'real time' remote biometric identification systems in publicly accessible spaces for the purpose of law enforcement.[85]

Systems with a high risk are subject to strict requirements. The Act identifies two main categories of high-risk AI systems: (1) AI systems using a safety component or product covered by EU Annex I legislation AND subject to third party conformity assessment under that Annex I legislation and (2) other stand-alone AI systems with mainly fundamental rights implications that are explicitly listed in Annex III, unless exceptions apply.[86] The Act allows 'high-risk' systems to be present on the European market as long as they adhere to mandatory requirements concerning data and data governance, documentation and record-keeping, transparency and information provision to users, human oversight, robustness, accuracy, security, and pre-market conformity assessment.

In Article 50, the AI Regulation provides for special transparency and user sensitisation obligations for certain AI systems. These systems can be labelled as those with limited risk. These include AI systems that are intended for direct interaction with natural persons or those that generate synthetic audio, image, video or text content. Among other things, this is intended to minimise the risks posed by deepfakes (see Article 50 (4)).

The lowest risk is unregulated (including most AI applications currently available in the EU

[82] 2024/1689 REGULATION (EU) 2024/1689 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 13 June 2024, reason for consideration 9; 46.
[83] Ebert/Spiecker, NVwZ 2021, p. 1188, (1188).
[84] Rostalski/Weiss, ZfDR 2021, p. 329, (336).
[85] 2024/1689 REGULATION (EU) 2024/1689 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 13 June 2024, Article 6.
[86] Monica, Saggi, p.145 (148 et seqq).

single market, such as AI-powered video games and spam filters). Although more and more applications are likely to slip into the riskier tiers as generative AI systems become more widespread, this approach leaves a wide range of AI-systems, with potentially serious impact unregulated. The Commission considered a number of regulatory options and ultimately favored a regulatory framework for high-risk AI systems only, with the possibility for all providers of non-high-risk AI systems to follow different types of receive regulatory support.[87] This is not the only time that a collaborative approach between government and industry is visible in the AI Act. In fact, in many ways it is a visual example of co-regulation.

II. Co-Regulation

1. Standards

The model of the AI Act focuses on the development of standards rather than requiring authorization by national and European administrations for placing products on the market. This means that private operators actively participate in the standard-setting process, similar to the REAC+ Regulation in the chemical sector. Standards become binding not only for economic operators but also for public authorities, guiding risk assessment and risk management in a consistent manner.[88] Consequently, this interplay should establish a positive feedback loop by encouraging high-risk AI providers to align with technical standards.

Article 40 of the AI Act stands out as a pivotal provision in the regulation, serving as the mechanism that offers a presumption of conformity with regulatory mechanisms upon high-risk AI system providers. The impact of Article 40 is that providers of high-risk AI systems can showcase compliance with the extensive set of requirements outlined in chapter III, Section 2 or, as applicable, with the obligations set out in of Chapter V, Sections 2 and 3, of the AI Act by adhering to officially adopted "harmonized standards" tailored to their domain.[89] These harmonized standards come into play following a standardization request (mandate) from the Commission to recognized European Standardization Organizations (ESOs) such as CEN, CENELEC, and ETSI. This request procedure entails the Commission preparing a draft standardization request through consultations with various stakeholders, including social partners, consumers, SMEs, industry associations, and EU Member States. Before its formal submission to an ESO, the draft request undergoes voting by the Committee on Standards, a committee comprised of EU Member States. The Commission maintains a repository of the standardization requests it has initiated.[90]

Private parties participate as co-decision makers in European standardization, which brings together public and private elements through non-legal instruments. Regulating through standards enables the identification of effective methods for predicting, combating, and correcting risks, encouraging compliance. As private operators are involved in the development of these standards, they are obligated to adhere to rules they had a hand in creating.[91] This co-regulative approach bridges the gap between public and private entities and results in the creation of effective, adaptable, and widely accepted standards that guide responsible AI development and deployment while meeting both societal needs and regulatory expectations.

[87] Stuurman/ Lachaud, Computer Law & Security Review, p. 3 et seqq.
[88] Monica, Saggi, p.145 (148 et seqq).
[89] McFadden et al., The Role of Standards, p. 7 et seqq.
[90] Ibid.
[91] Monica, Saggi, p.145 (148 et seqq).

2. AI regulatory sandboxes

The AI Act aims to strike a balance between openness to innovation and responsibility for innovation within the digital single market. The objective is to promote the potential of AI systems while mitigating their risks for individuals affected. In particular, the specific legal framework for so-called “AI regulatory sandboxes” (Articles 57 to 59 of the AI Act) is designed to ensure that the EU remains at the forefront of AI system development despite the new legal requirements.[92]

"Regulatory Sandboxes" are designated as protected spaces, both in the digital and physical realms, where the experimentation of novel technologies takes place under the supervision of regulatory authorities to ensure appropriate safety measures. They serve as platforms for communication between participating companies and regulatory bodies, fostering tailored guidance for companies while offering valuable real-world insights to authorities.[93] Additionally, participating companies benefit from temporary exemptions from certain regulatory requirements, enabling them to test new technologies with lower obstacles. As a result, regulatory sandboxes are widely acknowledged as fertile grounds for innovation, benefiting the entire (digital) economy.[94] According to the AI Act, AI experimentation labs are designed to establish controlled environments, facilitating the development, testing, and validation of pioneering AI systems for a limited period before their introduction or operation, following a specific plan (Article 57, paragraph 5, sentence 1 of the AI Act).

Regulatory sandboxes are a prime example of co-regulation between the state and the industry since they involve active collaboration and partnership between government authorities and private companies in the development and testing of new technologies. The concept of regulatory sandboxes reflects a balanced approach where both parties work together to achieve common goals while maintaining adequate oversight and risk management.

In sandboxes, companies do not face the full burden of strict regulations, fostering innovation and encouraging the development of cutting-edge solutions that may have otherwise been hindered by regulatory barriers. During the sandbox process, regulatory authorities are actively involved in providing guidance and supervision, engaging directly with industry players to understand technological developments and gain valuable insights into potential risks and benefits.[95] As companies progress through the sandbox, they transition from experimental to fully compliant stages, allowing adjustments and improvements to align with regulatory requirements and ensuring a smoother and safer transition to the market.

3. Cooperation with the AI Office

The AI Office was established by Commission Decision[96] with the task of developing the Union's expertise and capacities in the field of AI and contributing to the implementation of Union law in the field of AI (Article 64 AI Act). It develops the Union's expertise and capabilities in the field of AI. To this end, it is to work together with the scientific community, industry, civil society and other experts.[97]

The tasks of the office are diverse. In addition to extensive governance functions (e.g. Article 56(6); Article 75(1); Article 89(1)), it should in

[92] Botta, ZfDR 2022, p. 391 (398).
[93] Ibid.
[94] Goo/Heo, Technology, p. 15 et seqq.
[95] Engelmann/Brunotte/Lütkens, RDi 2021, p. 317 (322).
[96] Commission Decision C(2024) 390 of 24 January 2024 establishing the European AI Office.
[97] 2024/1689 REGULATION (EU) 2024/1689 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 13 June 2024, reason for consideration 111.

particular serve as a point of contact and support for industry. For example, providers of general-purpose AI models with systemic risk must inform the Artificial Intelligence Office immediately of serious incidents (Art. 55 para. 1 lit. C). At the same time, the Office promotes and facilitates compliance with the requirements of the AI Act. To this end, it provides templates in accordance with Art. 53 (1) lit. d, for example, in which the training content of AI models with a general purpose must be summarised in detail.

4. Codes of Practice

In accordance with Article 50 (7) and Article 56 (1), the AI Office also supports the development of codes of practice at Union level in order to contribute to the proper application of the AI Regulation, taking into account international approaches.

Article 56 of the AI Act describes the Codes of Practice as placeholders for compliance to bridge the gap between the entry into force of the General Purpose AI Model Provider Obligations (twelve months) and the adoption of standards (three years or more). While compliance by the GPAI model providers with the measures set out in the codes of practice is not legally binding, it is a presumption of conformity with the GPAI model provider obligations under Articles 53 and 55 until the standards come into force.[98]

In order to ensure that the codes of practice reflect the state of the art and take due account of different perspectives, the Artificial Intelligence Office shall co-operate with the relevant competent national authorities in the development of such codes of practice. The codes of practice should, in the context of systemic risks, contribute to establishing a risk taxonomy for the type and nature of systemic risks at Union level, including their causes and should also focus on specific risk assessment and mitigation measures.[99]

5. Codes of Conduct

Moreover the AI Act introduces the voluntary use of codes of conduct for non-high risk AI systems. Since most AI applications are non-high risk systems the importance of voluntary codes of conduct cannot be overstated. These codes shall aim to encourage the application of mandatory requirements designated for high-risk AI systems, as detailed in Chapter III, Section 2, to AI systems that do not fall under the high-risk category. The Act outlines the mechanisms through which these codes of conduct will operate and their significance in co-regulating AI systems with varying levels of risk.[100]

Article 95, paragraph 1 of the AI Act designates the AI Office and the Member States with the responsibility of promoting and facilitating the formulation of voluntary codes of conduct for non-high risk AI systems. These codes are intended to be built on “the available technical solutions and industry best practices.”[101] This approach recognizes that different AI systems may have distinct characteristics, necessitating flexible and context-specific compliance measures.

Additionally, Article 95 (2) introduces another category of codes of conduct that cover both high-risk and non-high risk AI systems. These codes are tailored to address specific objectives such as environmental sustainability, accessibility, stakeholder participation, and diversity within development teams. These objectives will be clearly defined and accompanied by key performance indicators to

[98] Farrell, Verhaltenskodizes.

[99] 2024/1689 REGULATION (EU) 2024/1689 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 13 June 2024, reason for consideration 116.

[100] Stuurman/ Lachaud, Computer Law & Security Review, p. 5 et seqq.

[101] Article 95 paragraph 1 of the Act.

gauge their achievement. The application of these codes remains voluntary.[102]

In terms of the process, the AI Act allows for AI system providers to develop their own codes of conduct individually, or these codes can be created by organizations representing multiple providers or a combination of both approaches.[103] Furthermore, users and other stakeholders, along with their representative organizations, are encouraged to actively participate in the formulation of these codes. Codes of conduct can be system specific or cover multiple AI systems, taking into account the similarity in their intended purposes.[104] Notably, there is a focus on accommodating the interests and requirements of small-scale providers and startups in the development of these codes, as mandated by Article 95, paragraph 4 of the Act. However, the specifics of how this accommodation will be achieved remain to be clearly defined.[105]

In essence, the AI Act recognizes the value of voluntary codes of conduct as a co-regulatory tool for non-high risk AI systems. These codes enable tailored compliance measures, foster collaboration between stakeholders, and provide a flexible framework for promoting responsible AI development and deployment across diverse contexts.

Problematic is, that the voluntary codes of conduct outlined in Article 95 of the Act do not incorporate a validation mechanism to ensure the adherence of participating entities to the prescribed code. As pointed out above, numerous examples of self-regulation relying on codes of conduct have revealed that the absence of effective enforcement mechanisms can quickly erode the confidence in the credibility of such initiatives. To counter this challenge, the incorporation of a verification process, possibly accompanied by a label, should be considered. This approach could grant entities that receive the label the opportunity to showcase their products and services based on the recognition obtained.[106]

E. A co-regulatory Framework for AI

I. Self-Regulating Principles

The implementation of principles into self-regulating mechanism (like codes of conduct or standards) within the realm of Artificial Intelligence should be guided by a comprehensive framework of principles that addresses various crucial aspects.

Foremost, entities operating within the AI landscape are tasked with conducting thorough assessments of both the potential positive outcomes and negative consequences associated with their actions. A comprehensive understanding of the broader implications of AI deployment is essential, compelling these entities to demonstrate the attainability of projected benefits, address potential drawbacks proactively, and involve relevant stakeholders in the decision-making process.[107]

In a context where concerns regarding the replacement of human decision-making with automated processes and the displacement of human workers persist, the principle of complementing humans assumes paramount importance. This principle seeks to mitigate public apprehensions by ensuring that AI-based systems work in harmony with human capabilities rather than supplanting

[102] Stuurman/ Lachaud, Computer Law & Security Review, p. 5 et seqq.
[103] Article 95 paragraph 3 of the Act.
[104] Article 95 paragraph 3 of the Act.
[105] Stuurman/ Lachaud, Computer Law & Security Review, p. 5 et seqq.
[106] Ibid, p. 6 et seqq.
[107] Clarke, Computer Law & Security Review, 411 (416).

them. Equally significant is the principle of maintaining human control, which addresses the unease surrounding the potential transfer of power to AI-based entities. It necessitates that human stakeholders remain in control of the technology and that AI processes remain transparent and understandable.[108]

Safeguarding human safety, wellbeing, and interests is a central commitment across the board. Those engaged in AI development and application are obligated to provide protective measures for all parties affected by AI systems, fostering a sense of trust and ensuring the overall welfare of society. This principle ties closely with the imperative of aligning AI systems with human values and rights, preventing negative impacts and advancing the interests of individuals at every step.[109]

Transparency and auditability form pillars of accountability, requiring entities to offer explanations that can be comprehended by humans for the decisions and actions undertaken by AI systems. By being transparent about their data practices, companies build trust and credibility, setting themselves apart from less transparent competitors. Due to an easier communication between companies, current customers and potential customers, businesses can use their creativity to raise the concept of digital self-determination to a new level for their customers and to distinguish themselves from competitors. If businesses can guide their customers in being responsible online, it helps customers trust the decisions made by computer algorithms. This journey isn't just about having rules and saying that things are clear or using attractive AI guidelines. It continues by letting customers have a say and be a big part of how companies make sure AI is used responsibly.[110]

Meanwhile, embedding quality assurance ensures that the processes, products, and outcomes associated with AI meet high standards, contributing to reliability and public trust. Additionally, the concepts of robustness and resilience underscore the importance of building AI systems that can withstand disruptions and swiftly recover from malfunctions.[111]

Accountability, discoverability, and timely resolution are fundamental principles that guarantee due process and procedural fairness. These principles enable responsible entities to address challenges and rectify issues promptly, maintaining trust within the ecosystem. Moreover, embracing enforcement mechanisms, liabilities, and sanctions underscores the gravity of adhering to these principles, promoting a culture of ethical conduct and responsible AI development.[112]

The "Ethics Guidelines for Trustworthy AI" published by the High-Level Expert Group on Artificial Intelligence (AI HLEG)[113] on April 8, 2019, also formulated many of these principles for achieving trustworthy AI. They furthermore concentrated on societal and environmental well-being as well as diversity, non-discrimination, and fairness. These principles are designed to guide the development and deployment of AI systems in a manner that aligns with human rights, ethics, and societal values.[114]

In summary, these principles should guide companies when it comes to setting up standards or codes of conduct, ensuring that

[108] Clarke, Computer Law & Security Review, 411 (416).
[109] OECD, Recommendation, p. 7; Clarke, Computer Law & Security Review, 411 (416).
[110] Koska/Filipović, Digitale Welt, p. 28 (31).
[111] OECD, Recommendation, p. 8.
[112] Clarke, Computer Law & Security Review, 411 (416).
[113] The AI HLEG is an independent expert group that was set up by the European Commission in June 2018.
[114] AI HLEG, Ethics Guidelines p. 14.

technological advancement aligns harmoniously with human values, rights, and societal well-being. In order to bring them to bear effectively, a state legal framework is needed that both promotes and monitors self-regulation.

II. Governmental framework

While AI may not raise fundamentally new ethical challenges and the technology itself is neutral, the way we choose to use it will determine its good or bad effects.[115] The challenge for government is therefore to empower industry to comply with these principles and to create a framework that canalizes the strength of government and self-regulation.

1. Incentives

One of the main problems of self-regulation lies in the fact that not all companies are willing to adhere to a privately established code. Firstly, because they are not obliged to do so due to the lack of binding effect, and secondly, because the positive impacts of self-regulatory measures (such as preventing "harsh" government regulations or gaining cross-industry reputation) do not solely benefit the companies that follow these measures (free-rider dilemma). Since self-regulation often entails significant costs, there is a need for exclusive incentives for those who submit to the rules.[116] To promote a comprehensive and transparent process of rule-making, the government could, for instance, provide financial support or establish public platforms for stakeholder exchange. However, a genuine foundation for a sector-wide implementation of private rule-setting would require a comprehensive system for the recognition of specific codes of conduct by state institutions.[117] Those who, for instance, adhere to legal requirements when setting standards could then claim approval for a code of conduct. The introduction of quality seals and labels should allow corresponding codes of conduct to be used for advertising purposes, signaling to customers that above-average requirements have been voluntarily met. Another possibility would be financial incentives, such as tax breaks or grants, to organizations that engage in self-regulation. This can encourage compliance and reward responsible behavior.

2. Standard-Setting

To compensate for the loss of legitimacy in private regulation, state law must steer the procedures of industry standard-setting in a way that procedural safeguards lead to substantive results that do not fall below certain minimum standards. In order to incorporate the interests of all those affected by the rules and achieve a consensus-based norm, all interested parties must initially be involved in the standardization process. Individual deficiencies need to be balanced, and minority rights ensured. Moreover, principles of the rule of law, such as equality and transparency, must be upheld. Based on these minimum standards, the justification or publication of decisions should be required. Furthermore, the possibility to raise objections against decisions or initiate arbitration proceedings must exist. Minimum requirements for rule-setting are also present at the international level. The EU Commission assumes a "Good Faith" from private regulators and calls for a complete information flow that should particularly enable small and medium-sized enterprises to participate in an objective regulatory process.[118] The task of various self-regulating standard-setting bodies is to develop industrial minimum standards that enable

[115] Hassabis, The History, Capabilities and Frontiers of AI Speech; Ferretti, Moral Philosophy and Politics, p. 239 (240).
[116] Spindler, Eckpunkte, p. 53; Hufeld, rescriptum 2017/2, p. 107 (110 et seqq).
[117] Spindler, Eckpunkte, p. 51; Hufeld, rescriptum 2017/2, p. 107 (110 et seqq).
[118] EU-Commission, Principles, p. 1.

interoperability. The state is responsible for promoting this unification of technical standards and the opening of fair competition already within the standard-setting process.[119] This includes global harmonization of recognized norms, defining a common vocabulary and unified software language, as well as agreements regarding quality requirements for technical standards.[120] As standards inevitably intersect with fundamental rights such as privacy, property rights, or personal freedom, the state is also obligated to make political demands on standard-setting. To simultaneously avoid hindering the innovation potential of Artificial Intelligence, industry, research initiatives, standardization bodies, and state regulatory authorities must collaborate closely and continuously review and evaluate standards and codes.[121] AI firms such as DeepMind already have an Ethics & Society unit running initiatives of this sort.[122] The AI industry can also help governments find loopholes or update obsolete legislation.[123]

3. Standard-Recognition

These (self-regulating) standards specified in codes of conduct that are industry- and function-specific must also find consideration in court by filling general clauses or contributing to the specification of ambiguous legal terms. This regulatory technique of "norm-concretizing reference" imparts a significant coercive effect that practically sets an effective incentive for norm compliance. To ensure legal certainty and counteract the "free rider dilemma," this presumption must be legally established in the AI industry. Additionally, it should be the task of standardization organizations as well as industrial policy to evaluate the existing standards, identify detailed needs for revision concerning digitalization aspects, and then initiate corresponding standardization or standard amendment procedures within these organizations.[124]

4. Simplify Enforcement

Co-regulation can only work if the private regulations are followed in practice. The foundation for this lies in the binding commitment declarations of participating organizations, expressing their intent and commitment to adhere to and implement technical or social standards. This can be conveyed through press releases or inclusion in internal organizational codes of conduct. While codes of conduct are generally voluntary, this doesn't mean that there are no avenues for enforcement. For instance, there are private enforcement mechanisms that range from mere exposure ("naming and shaming") to contractual penalties and even network access restrictions imposed by private internet providers. Additionally, claims for profit disgorgement and compensation can also be asserted under competition law as a form of private rule enforcement.[125] Furthermore, market-based sanctions are conceivable ("comply-or-explain"). However, in a data-driven market, establishing competitive infrastructures isn't straightforward, especially in the field of AI, where it could lead to a lack of competition and consumer "lock-in" effects. Consumers may no longer be able to penalize undesired company behavior through switching providers due to the absence of viable alternatives.[126] For this reason, it's important to promote soft enforcement mechanisms such as incentives for rule-compliant behavior, effective dispute resolution mechanisms, and

[119] EU-Commission, Data economy, p.16; EU-Commission, ICT, p. 3.
[120] EU-Commission, ICT, p. 11.
[121] EU-Commission, ICT, p. 3.
[122] Hassabis, The History, Capabilities and Frontiers of AI Speech.
[123] Ferretti, Moral Philosophy and Politics, p. 239 (260).
[124] DIN/DKE, Roadmap, p. 32 et seqq.
[125] Buck-Heeb/Dieckman, Selbstregulierung im Privatrecht, p. 291
[126] Buck-Heeb/Dieckman, Selbstregulierung im Privatrecht, p. 295 et seqq.

continuous monitoring to ensure functional and stable co-regulation.

5. Collaborative Platforms

To foster a culture of responsible Co-Regulation, states must also establish collaborative platforms that bring together industry stakeholders, experts, and regulators. These platforms serve as arenas for open discussions, sharing insights, and jointly developing self-regulation frameworks that align with evolving challenges. By involving diverse stakeholders, such as consumers, civil society organizations, and academia, the state enhances the legitimacy and inclusivity of self-regulatory initiatives.[127] The AI-Act already establishes a number of corresponding institutions. In addition to the AI Office (Article 64), these include a European Artificial Intelligence Board (Article 65), consisting of representatives of the Member States. An Advisory Forum (Article 67), represented by a balanced selection of stakeholders, including industry, start-ups, SMEs, civil society and academia. And additionally, a Scientific Panel of independent experts (Article 68). These "expert panels" sound promising and pave the way for a technically competent regulation for AI. It will be crucial not to neglect the voices of start-ups and think tanks in the various countries. The gap to the USA and other countries in the race for the AI revolution is already far too large. It must be a stated goal to enter dialogue with young visionaries and provide maximum freedom and resources to promising projects. If all the big AI companies move to the USA, it will not only be incredibly expensive for the European Union. They will also lose all data sovereignty and become involuntary spectators of the 21st century.

F. Conclusion

The future of artificial intelligence can be scary. It is up to today's lawmakers and industries to take the fear out of this man-made technology. State regulation will be necessary to ensure the safety of citizens. In fact, studies have found, that all types of non-state actors express concerns about AI technology and are in favor of regulating its development and use at the European level.[128] Self-regulation can therefore only play a complementary role in helping to make rules more effective and enforce them. In this context, a gap between the constantly evolving technology and the legal rules is not only unavoidable, but also necessary to avoid stifling the innovation potential of AI.

The AI sector holds a pivotal role in helping governments to implement more efficient policies. To begin with, it is essential that they meet their tax responsibilities, thereby contributing to the financial support of impactful retraining initiatives and social benefits to absorb the transition costs generated by disruptive AI technologies.[129] Self-Regulation and soft law shall contribute to the legislation itself, anticipating and experimenting with solutions that are more easily adaptable and improvable. As recognized by the AI Act, soft ethics and soft law can work as sandpits.[130] The industry has to proactively participate in legislative processes by providing insights into the AI systems and their needs.[131] If this is the case, those committed to self-regulation can indeed experience higher acceptance. Some companies are already rising to the responsible A.I. challenge: Microsoft has an Office of Responsible A.I. Use, Walmart a Digital Citizenship team, and Salesforce an Office of

[127] Ferretti, Moral Philosophy and Politics, p. 239 (259 et seqq); DIN/DKE, Roadmap, p. 34 et seqq.
[128] Tallberg et al., AI Regulation in the European Union, p. 23.
[129] Ferretti, Moral Philosophy and Politics, p. 239 (259 et seqq.).
[130] Floridi, Philosophy & Technology, p. 619 (622).
[131] Ferretti, Moral Philosophy and Politics, p. 239 (259 et seqq.).

Ethical and Humane Use of Technology. However, more companies need to quickly embrace a new era of A.I. self-regulation.[132]

For effective self-regulation to work it must be the goal to adopt a whole-of-lifecycle approach. To achieve this, processes need to be refined to establish distinct milestones that guide decision-making processes. Additionally, it is necessary to educate and train employees to actively participate in A.I. governance and to prevent bias. This education should equip them with a robust understanding of the tools, their impact, and their individual responsibilities throughout the lifecycle of the project.[133] Furthermore, it is vital to ensure alignment within the technology ecosystem encompassing edge devices, cloud platforms, sensors, and other tools. This alignment serves to prioritize the qualities of trust essential for each specific deployment.

Ultimately, the responsibility also rests on us, the users and consumers, to be vigilant in how we navigate AI tools and harness their complete capabilities in a manner that is both trustworthy and sustainable. Relying on governments and industries to act is no longer a viable approach.

[132] Ammananth, FORTUNE, January 16, 2023.

[133] Ibid.

Bibliography

Ammananth, Beena: Investors are pouring billions into artificial intelligence. It's time for a commensurate investment in A.I. governance, FORTUNE, January 16, 2023, https://fortune.com/2023/01/16/investors-billions-artificial-intelligence-self-regulation-governance-tech/.

Andrés, Moisés Barrio: Towards legal regulation of artificial intelligence, IUS Vol 15, No 48, 2021, p. 35-53.

Borrás, Susana/ Edler, Jakob.: The roles of the state in the governance of socio-technical systems' transformation. Research Policy, 2020). 49(5), 1-9.

Botta, Jonas: Die Förderung innovativer KI-Systeme in der EU, Zeitschrift für Digitalisierung und Recht: ZfDR, p. 391-412, 2022.

Brake, Johannes: Co-regulation or Capitulation? Addressing conflicts arising by AI and standardization, Lex Electronica, Volume 25, 2020, 11- 23.

Braithwaite, John: Types of responsiveness, https://press-files.anu.edu.au/downloads/press/n2304/pdf/ch07.pdf [Cited as: Braithwaite, Types of responsiveness].

Browne, Ryan: EU lawmakers pass landmark artificial intelligence regulation, CNBC, 14 June 2023, https://www.cnbc.com/2023/06/14/eu-lawmakers-pass-landmark-artificial-intelligence-regulation.html.

Buck-Heeb, Petra/Dieckmann, Andreas, Selbstregulierung im Privatrecht, Tübingen 2010, [Cited as: Buck-Heeb/Dieckman, Selbstregulierung im Privatrecht].

Clarke, Roger: Regulatory Alternatives for AI, Computer Law & Security Review 35, 398-409, 2019.

Clarke, Roger: Why the world wants controls over Artificial Intelligence, Computer Law & Security Review 35, Issue 4, 423-433, 2019.

Clarke, Roger: Roger Clarke, Principles and business processes for responsible AI, Computer Law & Security Review, Volume 35, Issue 4, 2019, 410-422.

Djeffal Christian/Siewert Markus/Wurster Stefan: Role of the state and responsibility in governing artificial intelligence: a comparative analysis of AI strategies, Journal of European Public Policy, 2022, 1-23.

DIN, DKE: German Standardization Roadmap on Artificial Intelligence (2nd edition), (2022), www.din.de/go/roadmap-ai [Cited as: DIN/DKE, Roadmap].

D'Cruze, Danny: 'AI can change everything in the world': Warren Buffett compares AI to the invention of atom bomb, Business Today, May 8, 2023,

https://www.businesstoday.in/technology/news/story/ai-can-change-everything-in-the-world-warren-buffett-compares-ai-to-the-invention-of-atom-bomb-380384-2023-05-08.

Ebert, Andreas/Spiecker, Indra: Der Kommissionsentwurf für eine KI-Verordnung der EU, NVwZ 2021, 1188-1193.

Engelmann, Christoph/Brunotte, Nico/Lütkens, Hanna: Regulierung von Legal Tech durch die KI-Verordnung, RDi 2021, 317-323.

EU Commission: Principles for Better Self- and Co-Regulation, 2014, https://ec.europa.eu/digital-single- market/sites/digital-agenda/files/CoP%20-%20Principles%20for%20better%20self-%20and%20co-regulation.pdf [Cited as: EU-Commission, Principles]

EU Commission: Building a european data economy, 2017, http://ec.europa.eu/news-room/dae/document.cfm?doc_id=41205 [Cited as: EU-Commission, Data economy]

EU Commission: ICT Standardisation Priorities for the Digital Single Market, 2016, http://ec.europa.eu/newsroom/dae/document.cfm?doc_id=15265 [Cited as: EU-Commission, ICT] p. 3.

European Economic and Social Committee (EESC), Brochure on EUROPEAN SELF-REGULATION AND CO-REGULATION, https://www.eesc.europa.eu/sites/default/files/resources/docs/auto_coregulation_en--2.pdf
[Cited as: EESC, Brochure].

Feldmann, Noah: AI Will Be Essential. And Complicated, The Washinton Post, April 2, 2023, https://www.washingtonpost.com/business/2023/04/02/regulating-ai-might-require-a-new-federal-agency/cde2deaa-d150-11ed-ac8b-cd7da05168e9_story.html.

Farrell, Jimmy: Eine Einführung in Verhaltenskodizes für das AI-Gesetz, 3. Juli 2024, https://artificialintelligenceact.eu/de/introduction-to-codes-of-practice/, [Cited as: Farrell, Verhaltenskodizes].

Ferretti, Thomas: An Institutionalist Approach to AI Ethics: Justifying the Priority of Government Regulation over Self-Regulation, Moral Philosophy and Politics, vol. 9, no. 2, 2022, pp. 239-265.

Floridi, Luciano: The End of an Era: from Self-Regulation to Hard Law for the Digital Industry, Philosophy & Technology (2021), p. 619-622.

Gunningham, Neil/ Sinclair, Darren: Smart regulation https://press-files.anu.edu.au/downloads/press/n2304/pdf/ch08.pdf [Cited as: Gunningham/Darren, Smart regulation].

Goo, Jayeoung/Heo, Joo-Yeun: Technology, Market, and Complexity 6 (2020), Issue 2, 1-18, [Cited as Goo/Heo, Technology].

Hassabis, Demis: “The History, Capabilities and Frontiers of AI.” Speech at: In You and AI, April 30, 2018, London. Retrieved from, https://royalsociety.org/science-events-and-lectures/2018/04/you-and-ai-history/ [Cited as Hassabis, The History, Capabilities and Frontiers of AI Speech].

High-Level Expert Group on Artificial Intelligence (AI HLEG): Ethics Guidelines for Trustworthy AI", April 8, 2019, [Cited as AI HLEG, Ethics Guidelines].

Hoffmann-Riem, Wolfgang: Artificial Intelligence as a Challenge for law and Regulation in Wischmeyer/Rademacher, Regulating Artificial Intelligence, Cham 2020, 1-19. Cited as [Hoffmann-Riem, Regulating Artificial Intelligence].

Hu, Krystal: ChatGPT sets record for fastest-growing user base - analyst note, Reuters, February 2, 2023, https://www.reuters.com/technology/chatgpt-sets-record-fastest-growing-user-base-analyst-note-2023-02-01/.

Hufeld, Lukas: Industrie 4.0, Zwischen Selbstregulierung und staatlicher Intervention, rescriptum 2017/2, 107-113.

Kleinsteuber, Hans: State – Regulation – Media OSCE Conference Amsterdam "Guaranteeing Media Freedom on the Internet", Aug. 27/28, 2004, [Cited as: Kleinsteiber, Regulation].

Koska, Christopher/ Filipović, Alexander: Blackbox AI — State Regulation or Corporate Responsibility?, Digitale Welt, 2019, 28-31.

Korn, Jennifer: AI pioneer quits Google to warn about the technology’s ‘dangers’, CNN Business, May 3, 2023, https://edition.cnn.com/2023/05/01/tech/geoffrey-hinton-leaves-google-ai-fears/index.html.

Latzer, Michael/Just, Natascha/Saurwein, Florian/Slominski, Peter: Selbst- und Ko- Regulierung im Mediatiksektor, Wiesbaden 2002, [cited as: Latzer et al., Mediatiksektor].

Martini, Martin: Blackbox Algorithmus – Grundfragen einer Regulierung Künstlicher Intelligenz, Springer Berlin Heidelberg, 2019, [cited as Martini, Blackbox].

McFadden Mark/ Kate Jones/ Taylor Emily/Osborn Georgia: The Role of Standards in the EU AI Regulation, Oxford Information Labs, December 2021, [Cited as: McFadden et al., The Role of Standards].

Monica, Alessia: Regulating AI and the key-role of standard in the co-regulation of ICT: EU, Members States and private entities, Saggi, p. 145-156.

Nolan, Beatrice: Sundar Pichai says AI technology could be more profound than fire or electricity, INSIDER, Apr 17, 2023, https://www.businessinsider.com/sundar-pichai-google-ai-bard-profound-tech-human-history-2023-4 .

OECD, Recommendation of the Council on Artificial Intelligence, OECD/LEGAL/0449, [Cited as: OECD, Recommendation].

OECD Report, Alternatives to traditional Regulation, https://www.oecd.org/gov/regulatory-policy/42245468.pdf .

Piepgrass, Stephen/Waltz, Daniel/Miklaszewsk, Rachel: Self-Regulation and Policymaking Guidance Regarding the Use of AI and ML, RTInsights, May 10, 2022, https://www.rtinsights.com/self-regulation-and-policymaking-guidance-regarding-the-use-of-ai-and-ml/.

Podszun, Rupprecht, in: Hilty/Henning-Bodewig (Hrsg.), Corporate Social Responsibility. Verbindliche Standards des Wettbewerbsrechts? 2014, [Cited as: Podszun, Corporate Social Responsibility].

Porket, J. L.: Reflections on the Pros and Cons of State Regulation. Sociologický časopis / Czech Sociological Review, 38(3), 2002, 311-326.

Rostalski, Frauke/Weiss, Erik: Der KI-Verordnungsentwurf der Europäischen Kommission, ZfDR 2021, 329 -256.

Russell, Hardin/ Cullity, Garrett: "The Free Rider Problem", The Stanford Encyclopedia of Philosophy (Winter 2020 Edition), Edward N. Zalta (ed.), https://plato.stanford.edu/archives/win2020/entries/free-rider/.

Shear, Michael D./ Kang, Cecilia/ Sanger, David E.: Pressured by Biden, A.I. Companies Agree to Guardrails on New Tools, New York Times, July 21, 2023, https://www.nytimes.com/2023/07/21/us/politics/ai-regulation-biden.html.

Scherer, M.U.: Regulating artificial intelligence systems: risks, challenges, competencies, and strategies, Harvard J Law Technol, 29 (2) (2016), pp. 353-400

Schulz, Wolfgang/ Held, Thorsten: Regulierte Selbstregulierung als Form modernen Regierens. Im Auftrag des Bundesbeauftragten für Angelegenheiten der Kultur und der Medien. Endbericht, Hamburg 2002 [Cited as: Schulz, Selbstregulierung].

Staudenmayer, Dirk: Die deliktsrechtliche Anpassung des europäischen Privatrechts an die Digitalisierung, NJW 2023, 894-901.

Spindler, Gerald/Thorun, Christian: Eckpunkte einer digitalen Ordnungspolitik. Politikempfehlungen zur Verbesserung der Rahmenbedingungen für eine effektive Ko-Regulierung in der Informationsgesellschaft, Berlin 2015, [Cited as: Spindler, Eckpunkte].

Stiglitz, Joseph: `Government failure vs. market failure' principles of regulation
Initiative for Policy Dialogue (2008), Working Paper #144 [Cited as: Stiglitz, Government failure vs. market failure].

Stiglitz, Joseph: “The Future of Work.” Speech at “You and AI”, September 11, 2018, London, [Cited as: Stiglitz, The future of Work Speech], https://www.youtube.com/watch?time_continue=684&v=aemkMMrZWgM&embeds_referring_euri=https%3A%2F%2Froyalsociety.org%2F&source_ve_path=Mjg2NjY&feature=emb_logo&ab_channel=TheRoyalSociety .

Simon, H. A. : Government in Today’s World of Organizations and Markets.” PS: Political Science and Politics, 2000, 33 (4): 749–756.

Stuurman, Kees/ Lachaud, Eric: Regulating AI. A label to complete the proposed Act on Artificial Intelligence, Computer Law & Security Review, Volume 44, 2022, 1-23.

Sunstein, C. R.: Beyond the Precautionary Principle, in SSRN, January 2003. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=307098.

Tallberg, Jonas/Lundgren, Magnus/ Geith, Johannes: AI Regulation in the European Union: Examining Non-State Actor Preferences, April 20, 2023, [Cited as: Tallberg et al., AI Regulation in the European Union].

The Science and Environmental Health Network on: The Wingspread Conference, August 5, 2013, https://www.sehn.org/sehn/wingspread-conference-on-the-precautionary-principle.

Vigna, Francesco: Co-regulation Approach for Governing Big Data:
Thoughts on Data Protection Law. In 15th International Conference on Theory and Practice of Electronic Governance , October 04–07, 2022, Guimarães, Portugal- [Cited as: Vigna, Co-regulation Approach].

Wagner, Ben: Ethics as an escape from regulation, in: Bayamlioglu/Baraliuc/Janssens/Hildebrandt, Being profiled: cogitas ergo sum, Amsterdam University Press 2018, 84-88.

Yampolskiy, Roman V. / Spellchecker, M.S.: Artificial Intelligence Safety and Cybersecurity: a Timeline of AI Failures, arXiv, 2016.