J Multidisciplinary Scientific Journal

MDPI

Article

Artificial Intelligence *ante portas:* Reactions of Law

Rolf H. Weber

Faculty of Law, University of Zürich, Rämistrasse 74/38, 8001 Zurich, Switzerland; rolf.weber@rwi.uzh.ch; Tel.: +41-44-634-48-84

Abstract: Artificial intelligence and algorithmic decision-making causes new (technological) challenges for the normative environment around the globe. Fundamental legal principles (such as non-discrimination, human rights, transparency) need to be strengthened by regulatory interventions. The contribution pleads for a combination of regulatory models (hard law and soft law); based on this assessment, the recent European legislative initiatives are analyzed.

Keywords: AI Principles; CAHAI-Report; EU Proposal; hard law; legal framework; soft law

Citation: Weber, R.H. Artificial Intelligence *ante portas:* Reactions of Law. *J* **2021**, *4*, 486–499. https://doi.org/10.3390/j4030037

Academic Editors: Ugo Pagallo and Massimo Durante

Received: 15 July 2021
Accepted: 31 August 2021
Published: 6 September 2021

1. Introduction

Artificial intelligence (AI) offers new informational interaction and data processing opportunities enabling the implementation of innovative communications and business models [1] (pp. 1–2). The use of intelligent "devices" and the availability of algorithms include the potential to replace human activities by software and/or machines. Instead of a human intervention, the programming of the code, which executes the tasks, becomes important.

Artificial intelligence allows one to establish a "regime" of automated decision-making, being conducted in a very timely and effective manner. Automation of this kind is mainly feasible in situations not requiring a specific human input, for example, in the case of an algorithm-driven search or in case of a standardized exchange platform. However, the automated decision-making can also cause many socio-ethical and legal challenges.

The following contribution addresses the potential normative framework for a trust-oriented environment with regulatory tools being suitable to minimize the occurrence of technological risks. Therefore, different elements of a suitable normative framework are considered, in particular the relevant international and national instruments as well as soft law sources. The notion that legalization requires a specific form of discourse no longer reflects the needs of society; moreover, the inclusion of all concerned stakeholders is justified (leading to co-regulation as a new model). Thereafter, those normative key principles that are particularly exposed to specific challenges in the AI context are discussed. Based on these general considerations, the ongoing regulatory debates in the Council of Europe and the recently published proposal of the European Commission for an AI Regulation are assessed in light of the developed normative environment. The contribution closes with a future-oriented outlook.

Since the contribution is a theoretical analysis, a case study will not be presented by purpose, but a future-oriented outlook offers ideas about how the practical developments could be tackled.

2. Need for a Comprehensive Legal Framework

2.1. Legitimacy for Regulatory Intervention

Law constitutes a structural system that is composed of an organized or connected group of objects (terms, units, or categories) forming a complex structure. The function of law is crystallized in a system of rules and institutions that underpin civil society, facilitate orderly interaction, and resolve conflicts and disputes arising in spite of the rules [2] (p. 13

with further references). Law can be created through different processes, for example by negotiations among the concerned norm addressees, by imposition of legal rules through the governing body, or by evolution of self-regulatory mechanisms [2] (p. 13). The legal system is not a predetermined construct but is embedded by way of rulemaking in other socially relevant systems.

In this context, some key questions are to be asked and answered [3] (p. 105 et seqq.): (i) Who is entitled to set the rules? (ii) In whose interest? (iii) By which mechanisms? (iv) For which purposes? The need is given to develop overarching networks and negotiation systems between the different stakeholders, thus forming a cooperative approach to rule-making that includes the whole society, hence dividing responsibilities between public and private actors [3] (pp. 112–113).

Applying a theoretical perspective, rule-making issues can be addressed from the angle of different disciplines [1] (pp. 3–4 and pp. 4–5); nevertheless, also in private matters encompassing AI methods, the discussions must concentrate on the appropriate allocation of duties and responsibilities as well as the adequate structuring of the concerned "organization" that is developing AI systems. In other words, rulemaking, at whatever level of social organization it may take place, refers to setting norms for the conduct of the business in an appropriate way.

The traditional rule-making approach in international matters, namely the conclusion of multilateral treaties, does not fit the objectives of a regulatory framework setting guidelines for AI systems. Moreover, other mechanisms have to play a more important role. A new rule-making model that has been developed (and partly also applied) in the Internet governance (as well as in the climate change and sustainability) context is the multistakeholder participation model [4]. If all concerned persons and organizations of the public and the private sphere are involved in the discussions and negotiations of the regulatory framework for AI processes, the chances increase that the developments are in the interest and to the benefit of the whole society [5] (pp. 249 et seqq.).

Practical experience conveys the lesson that some basic challenges need to be addressed in order to make the multistakeholder concept successful, particularly if several forms of cooperation based on a variety of actors involved are considered. Thereby, four fundamental questions must be tackled [5] (p. 249) [6] (pp. 1 et seqq.): (i) How do concerned groups best match challenges with the organizations and networks? (ii) How can governing bodies/entities be most able to help develop legitimate, effective, and efficient solutions? (iii) How should the flow of information and knowledge necessary for successful legal regime be structured? (iv) How can different groups approach the coordination between the available normative networks in order to avoid conflicting interest? Answers to these questions need a differentiated thinking depending on the given environment. As items of shortcomings, the identification of adverse effects of automated decision-making in the relevant policy field, the facilitation of networking opportunities, and the public support of civil society interventions are to be considered.

2.2. Implementation of a Co-Regulatory Approach

International legal instruments (Conventions, Treaties) usually qualify as hard law. From a theoretical perspective, hard law constitutes an institutional normative order [7] (pp. 42–43). In contrast, technological standards as well as social and ethical guidelines (only) build an informal normative order. During the last years, however, the borderlines have blurred, and the regulatory bodies are becoming intertwined, i.e., the hard/soft law dichotomy does not anymore correspond to the reality [8].

The often-expressed assumption that hard law is qualitatively better than soft law does not hold in the present normative environment. Soft law can entail functions previously tied to hard law, for example the notion of coordinating the concerned actors and of improving the transactional efficiency [9] (p. 11) [10] (pp. 188 et seqq.). So far, the term soft law has not found a generally acknowledged definition. Partly, soft law is seen as a normative model entailing opinions and recommendations of authorities (e.g., the European Union

Aviation Safety Agency (EASA) in European Union (EU) civil aviation law or the European Data Protection Board (EDPB) in EU data protection law). However, soft law is also used as a term for rules created by the participants of a specific community; these rules are usually efficient since they respond to real needs, mirror the technology, and provide the opportunity to flexibly adapt the legal framework to the changing environment. Since such rules are negotiated by the involved communities, the likelihood is high that they enjoy broad acceptance [7] (p. 27). Furthermore, the concerned persons/entities are incentivized to conduct permanent consultation processes to develop and implement the rules [10] (pp. 179 et seqq.). A private or inter-agency process laying down a set of rules might also more easily realize new solutions than hard law [11] (p. 897). In addition, this regulatory model develops and establishes rules independently of the principle of territoriality, i.e., soft law has a global reach at least in principle.

Soft law also carries risks since the private rules are not always transparent and not every relevant group concerned is always involved. Furthermore, the active participants spending human and/or financial resources are also confronted with a "free-rider problem" [7] (p. 28). In addition, self-regulation has a lower democratic legitimization and, particularly, non-compliance with private rules might be confronted with problems of enforcement and does not mandatorily cause sanctions [12] (pp. 282–284). Nevertheless, the role of the governmental regulation could be limited to situations in which the private actors do not find suitable solutions themselves [9] (p. 12) [13] (p. 589).

In a nutshell, soft law must fit into the overall legal environment with the objective of realizing a non-discriminatory and socially acceptable legal framework [9] (p. 13). As experience shows, by overcoming the traditional dichotomy, the concepts of hard law and soft law have moved closer to each other or even intermingled (see also Table 1).

Table 1. Intermingling of the concepts of hard law and soft law.

Hard Law				**Soft Law**
	Legitimized	←	Partly legitimized	
	Top-down bodies	←	Bottom-up bodies	
	Stable/inflexible	←	Adaptable/flexible	
	Enforceable	←	Partly enforceable	

As expressed by the arrows, soft law is moving towards hard law, i.e., notwithstanding the term "soft", the respective rules and guidelines are becoming (or have already become) quite robust and are relatively narrow to the characteristics of the traditional hard law. It is to be expected that this movement will continue during the coming years.

Rules considered by the "governed" persons as adequate guidelines appear to be legitimate since private incentives lead to a need-driven rule-setting process [7] (pp. 22 et seqq.). Insofar, soft law contributes to a more extensive capacity building and its enabling functionality serves to easing the coordination process while at the same time providing directionality to the provision of cross-border standards. Soft law quality allows regulators to enter into arrangements by varying scope and specificity and then to clarify (or change) the expectations of the involved actors [9] (p. 12).

If soft law does not cover the whole scope of a socially desirable legal framework, it must be backed by a governmental "regime", for example in the case that some basic rules should not be left to the private actors. In such a situation, legal doctrine refers to the concept of "co-regulation" (also called regulated self-regulation) [14] (pp. 43, 139–148, 230) [15]. In some circumstances, it makes sense to have soft law approved by a public authority as experiences in certain regulatory ecosystems (financial markets, media markets, Internet markets) have evidenced [16] (pp. 9 et seqq.). Nevertheless, in view of the rapid growth, the complex inter-relationships, and the dynamic changes having taken place in the new technological environment, it appears to be broadly accepted that a more flexible and innovation-friendly model of regulation is required [17] (pp. 32 et seqq.).

Hereinafter, the described theoretical framework encompassing hard law and soft law will be discussed with a special focus on AI communications and business models. AI causes normative challenges in respect of some key legal principles such as non-discrimination or transparency. Therefore, these principles merit a general analysis before specific regulatory approaches are discussed.

2.3. Identification of Substantive AI Principles

As shown, the AI era needs a broader and more complex consideration of values exceeding a narrow perception of rights contained in the traditional legal instruments, particularly the multilateral treaties. Therefore, a values-oriented approach, not restricted to aspects related to the "autonomy" of automated systems and robot devices, respectively [18] (pp. 163 et seqq.) [19] (pp. 154–155), or to the legal personhood of self-learning machines [18] (p. 162) [19] (p. 164) [20] (pp. 819 et seqq.), is needed to tackle the challenges in the digital world. Acting in compliance with ethical principles improves the moral reputation, which in turn helps to gain the trust of users and citizens [21] (p. 4). Hereinafter, the impact caused by AI on the legal framework will be discussed at the hand of three prominent principles, namely (i) non-discrimination, (ii) transparency, and (iii) informational self-determination (i.e., data protection). The subsequent analysis will show how high-value standards in a democratic society can serve the interests of all concerned individuals.

Apart from the three mentioned principles, trust is a key element of human interaction [22] since it also relates to accountability and to good governance; publicly assessable accounts are a pre-condition for a sustainable society [1] (p. 7). In order to improve to foreseeability of AI systems' effects, it is important to implement standards that design the behavioral requirements in a concise manner. Thereby, the mirroring of compliance in reality means to also rely on soft law instruments and/or provide for cooperative models, which are appropriate and fair without replicating an asymmetrical private ordering [21] (pp. 4–5).

2.3.1. Non-Discrimination

Artificial intelligence systems cause a risk of discrimination. Many examples already exist that show a certain degree of structural inequality or problems of foreseeable or unforeseeable biases, amongst others, in respect of gender considerations or the treatment of colored people [23] (pp. 7 et seqq.) [24] (nos. 21 et seqq.). Technologically, self-learning algorithmic programs do exercise the code implemented in the software; however, the design of the code is influenced by the programming person/device.

Most international legal instruments (such as the Declaration of Human Rights of the United Nations) as well as practically all national constitutions around the globe contain a provision prohibiting undue discrimination. Usually, without going into the details, the wording of the respective provisions is rather broad and encompasses direct and indirect discrimination [25] (p. 39 with further references). However, the problem of indirect discrimination consists in the fact that specific criteria seem to be "neutral" even if in their application some persons are more affected than others [24] (no. 27) [26] (pp. 10–13) [27] (pp. 680–687). In such case, the discrimination is solely based on the relation of a person to the protected group [28] (pp. 394–412).

Another practically important aspect concerns the legal situation that—at least in principle—the non-discrimination rules may only be invoked against States as users of algorithmic systems, not directly against private enterprises and individuals, except if the so-called horizontal effect of human rights against private persons is acknowledged in a given legal order. Therefore, by further developing the fundamental rights, it should be generally accepted that the right of non-discrimination constitutes an obligation to analyze and mitigate, throughout AI systems' life cycle, the risk of unjust biases [29] (p. 23).

Furthermore, the non-discrimination principle should also encompass aspects of "fairness" playing an important role in the AI context. Fairness is an aspect showing compliance

with behavioral rules of a conscious human being. The constitutional provisions available usually do not literally refer to this term; nevertheless, it appears to be adequate to apply a broad understanding of "discrimination" [24] (nos. 25–26) [30] (pp. 22–27).

In the meantime, some legislators have become active by implementing secondary legal instruments that contain specific provisions forbidding the discrimination based on gender, race, or ethnic origin in connection with work, social support, or insurance coverage. Respective laws are particularly known in the Western hemisphere but not around the globe [23] (p. 56). Justification reasons for an "unequal" treatment are usually also foreseen but their scope is mostly restricted.

The non-discrimination principle can equally be based on the personality rights being widely recognized by State legislators. The personality rights are expressions of the individual self-determination principle and of human dignity. The exact scope depends on the wording of the stated personality rights [24] (no. 29) [25] (p. 40). Irrespective of the concrete design, such rights' position is always linked to the already generally mentioned concepts of trust and accountability.

The protection against discrimination is also foreseen in soft law guidelines developed and published by international and regional organizations; such guidelines (as for example drafted by an expert group of the EU) often state the principle that AI systems must be based on transparent analyses and enshrine equal treatment principles [31] (nos. 80 et seqq.). According to the Organisation for Economic Co-operation and Development (OECD) Principles, AI actors playing an active role in the lifecycle of AI systems have to see to it that human rights and equal treatment principles are observed in the technological design [32] (no. 1.2). Similarly, private enterprises such as Google [33] and Microsoft [34] state the principle of non-discrimination.

The implementation of the non-discrimination principle can cause difficulties in practice since AI systems are not tailored to the specificities of the available legal instruments. Even if an actual legal vacuum does not exist, experience and academic research studies have shown that the prevailing fundamental rights principles as well as secondary normative provisions do have certain limits and are not in a position to secure a comprehensive non-discrimination protection [29] (pp. 21–22).

2.3.2. Transparency and Accountability

"Transparent" means clear, evident, and obvious. In the AI context, a culture of transparency enshrining the disclosure of the used AI applications, a description of their logic, as well as access to the structure of algorithms and to the introduced datasets should be provided [1] (p. 6). Transparency requires robust and general rules, not necessarily more regulation. Therefore, the improvement of transparency does not mean to have a quantitative increase of information, but "more" in terms of higher information quality [1] (p. 7 with further references).

The transparency principle is only partly contained in international legal instruments and national constitutions. Indirectly, however, transparency can be seen as an expression of the rule of law on a rather abstract level. In addition, access rights to data collections of governmental bodies and—with a more limited scope—to the data processing of private enterprises are also a reflection of the transparency principle [25] (p. 36). In practice, transparency plays a particularly important role in the health sector ("early warning scores") [24] (no. 8).

In real life, transparency means explainability and interpretability of specific AI processes. Only if the concerned persons in their given environment are able to understand how the processes work and which objectives are to be envisaged can an agreement to the automated or algorithmic decision-making be assumed [35] (p. 663) [36] (pp. 41–47 and 189–207). Apart from the technical challenges in making neural networks explainable, the legal issue of existing business secrets and know how protection should not be underestimated and require a diligent balancing of interests [24] (no. 10).

In the meantime, some valuable examples can be found in the existing legislation:

- Article 5 of the EU Regulation 2019/1150 on the Promotion of Fairness and Transparency for the Commercial Users of Online Intermediary Services [37] requires the offerors of intermediary services to make transparent the ranking of the main parameters of the search, the reasons for the relative impact of such parameters, and the potential possibility to influence the ranking by paying a certain sum of money. However, the providers do not have to disclose the algorithms of the programs.
- Similar transparency requirements in favor of the consumers are contained in the EU Directive 2015/29 about illegal business practices in its newly revised version [38,39]. Further specific transparency requirements are found in legal instruments regulating the health sector.
- Many transparency requirements are usually contained in data protection laws, particularly in the EU Data Protection Regulation (GDPR) 2016/679; a person can only give the consent to the data processing if its scope and purpose is known (see below Section 2.3.3 and [24] (nos. 11, 13 and 15)).
- The transparency requirement is now also foreseen in Art. 13 of the new EU proposal for an AI Regulation (see below Section 3.2.2).

Nevertheless, the transparency principle seems to be more clearly addressed in soft-law instruments. Most declarations and guidelines of international and regional organizations having been developed and published in connection with the more widespread use of algorithms contain specific transparency requirements. Examples can be found in the guidelines of the High-Level Group of the EU [31] (nos. 76 et seqq.) as well as in the OECD Principles [32] (no. 1.3). Google also promised to develop AI systems in a way that an adequate feedback mechanism is implemented, which discloses the relevant information to the concerned persons [33] (no. 4). Microsoft equally declares in its AI Principles that the algorithmic systems must be transparent [34] (principle "transparency").

However, some deficits in respect of the transparency requirements related to AI systems cannot be overlooked, particularly since—as mentioned—more transparency does not always have an added value for the concerned person and does not necessarily reach the intended purpose [25] (p. 38). A reason for this assessment has to be seen in the fact that the addressees of information often do not fully understand its contents or do not take proper note of it; furthermore, an information overload leads to confusion [40] (pp. 77–78, 79–81). This problem can even become bigger in case of AI systems due to the high complexity of machine learning systems. Therefore, special efforts of soft law should be guided in the direction of increasing the comprehensibility of information [25] (pp. 38–39).

An often-discussed topic in connection with transparency is the matter of interpretability, which can be divided into local interpretability and global interpretability. While local interpretability tries to understand decisions made by AI for just one instance or person, global interpretability addresses the understanding of general rules that are essential for the functionality of machine learning models. The latter is also connected to the problem of non-discrimination discussed in Section 2.3.1.

As an additional issue being closely related to transparency [41] (p. 320 et seqq. with further references), the topic of accountability must be briefly addressed. Accountability as a pervasive concept includes political, legal, philosophical, and other aspects with each of them casting a different shade on the meaning of the term. As a fundamental principle, accountability can be outlined with three elements: (i) Providing information in a timely manner, (ii) introducing standards that hold governing bodies accountable, and (iii) implementing mechanisms of sanctions [2] (p. 70 et seqq. for new discussions on accountability, with further references).

2.3.3. Data Protection

If AI systems are processing personal data, the applicability of data protection laws is obvious. Even if some applications of AI, such as certain forms of predictive policing do not process personal data, it must be assumed that in many cases, the data protection principles of the applicable legal instrument need to be observed [42,43]. In fact, data

protection is probably the most discussed issue in connection with AI systems [26] (pp. 9–10) [42–44] (pp. 18 et seqq. and p. 44 et seqq.). Thereby, several data protection issues play a role.

The most prominent topic is the prohibition of automated decisions (with certain exceptions). According to Article 22 GDPR, the technological and socio-technical design of each automated decision-making system has to be performed in accordance with the data subject's rights and freedoms as well as based on legitimate interests; therefore, a full assessment of the positive and negative impacts of AI systems is required [42] (p. 848).

In addition, most modern data protection laws contain (partly far-reaching) information duties of the processor of data as well as access rights of the concerned person to his/her data, partly similar to existing consumer law. These general principles apply in all cases of processing personal data, but particularly so in case of AI systems [24] (nos. 16 et seqq.) [25] (p. 35 with further references). Therefore, each AI project does have to include a data privacy analysis even if the risk cannot be overlooked that an efficient and broad realization of the respective principles meets technological difficulties in the AI environment.

Recent data protection laws have introduced the obligation of data processors to conduct a so-called data protection impact assessment if the processing of data can cause substantial risks to the concerned person. Such a data protection impact assessment is usually a complex undertaking and must be planned in view of the given data security environment of the data processor [25] (p. 35); the details of protective measures in the case of AI systems depend on the given circumstances.

Nevertheless, certain challenges of the existing data protection laws with respect to AI systems and algorithmic programs cannot be overlooked: AI technologies are constantly improving their capacity to solve tasks causing the processing of more and more data. Partly, an AI system seemingly performs the impossible by doing tasks thought to be the exclusive purview of human intelligence [42] (p. 849). In addition, a problem in assessing AI concerns the fact that it denotes the whole set of technologies [45] and that those technologies have a general-purpose nature. Therefore, the performance of a socio-technical impact assessment, which is able to weigh and balance all aspects, is necessary; in other words, a law-by-design approach is required [42] (p. 868).

3. Assessment of Regulatory Initiatives in Europe

During the last two years, organizations in Europe have been particularly active in the development of AI guidelines and are in the process of concretizing the need for legal action in normative instruments. The key drivers are the Council of Europe and the Commission of the European Union.

3.1. Council of Europe: Project for an AI Legal Instrument

3.1.1. Analysis of the Normative Environment

The Council of Europe (CoE), incorporated in 1949, with its 47 Members (extending across Europe to the Western part of Asia) has the objective to ensure human rights, democracy, and the rule of law. The basic legal instrument is the European Convention on Human Rights (ECHR) being applicable in all areas of life, including online and offline as well as regardless of the technology.

On 11 September 2019, an Ad hoc Committee on Artificial Intelligence (CAHAI) has been mandated to analyze in a report, on the basis of broad multi-stakeholder consultations, the feasibility and potential elements of a normative framework for the development, design, and application of artificial intelligence. On 17 December 2020, the CAHAI published a feasibility study not only addressing opportunities and risks arising from AI on the values of the ECHR, but also analyzing the possible legal instruments for a normative framework for governing AI processes (CAHAI-Report) [29] (p. 2).

After a discussion of the opportunities and risks arising from the design, development, and application of AI on human rights, the rule of law, and democracy, the CAHAI Report

assesses the impact of AI on the different fundamental rights, which are contained in the ECHR, namely [29] (pp. 7 et seqq.):

- Liberty and security; fair trials; no punishment without law; effective remedy (Articles 5, 6, 7, 13 ECHR);
- Private and family life; physical, psycho-social and moral integrity (Article 8 ECHR);
- Freedom of expression; freedom of assembly and association (Articles 10, 11 ECHR);
- Equality and non-discrimination (Article 14 ECHR, Protocol 12);
- Social and economic rights (Articles 2, 3, 5, 11, 12, 13 and 20 of the European Social Charter).

Furthermore, the CAHAI-Report analyzes the impact of AI on democracy; obviously, a functional democracy relies on open social and political discourse, as well as on the absence of improper voter influence or manipulation [29] (p. 10 et seqq.). In general, the concentration of power in the hands of a few private platforms influencing the public sphere can amplify risks if no regulation is in place.

Furthermore, AI systems might also affect the rule of law [29] (p. 12). In the mission of the CoE, the rule of law exercises a particularly important role; it prescribes that all public authorities act within the constraints set out by law. Therefore, the rule of law requires respect for principles such as legality, transparency, accountability, legal certainty, and effective judicial protection.

3.1.2. Discussions about the Form of the Legal Instrument

Since the Council of Europe is very much engaged in efforts that are suitable to realize normative principles to the benefit of the whole society, it is not surprising that a special and long chapter of the CAHAI-Report is devoted to the "Mapping of Instruments Applicable to Artificial Intelligence" [29] (p. 18 et seqq.). At the outset, the (few) international legal instruments applicable to artificial intelligence are shortly described; from this analysis, the conclusion is drawn that the existing instruments do not always provide adequate safeguards to the challenges raised by AI systems. The frequently available ethics guidelines on AI systems are also described [29] (p. 20, The CAHAI experts have reviewed not less than 116 documents on "ethical AI"); the CAHAI-Report notes that ethics guidelines are useful tools to exert some influence on the public decision-making over AI and to steer its developments towards social good; however, these guidelines cannot substitute mandatory governance [29] (p. 20). Therefore, the CAHAI-Report arrives at the result that, while there is no legal vacuum as regards AI regulation, a number of substantive and procedural legal gaps nevertheless exist and these gaps should be filled [29] (p. 22 et seqq.).

Since a number of essential principles that are relevant for the protection of human rights, democracy, and the rule of law in the context of AI are currently not clearly legally assured, the CAHAI-Report outlines the main elements of a legal framework for the design, development, and application of AI by addressing the key values, rights, and principles derived—in a bottom-up perspective—from sectoral approaches and ethics guidelines as well as—in a top-down perspective—from the requirements of basic fundamental principles as follows [29] (p. 27 et seqq.):

- Human dignity;
- Prevention of harm to human rights, democracy, and the rule of law;
- Human freedom and human autonomy;
- Non-discrimination, gender equality, fairness, and diversity;
- Principles of transparency and explainability of AI systems;
- Data protection and the right to privacy;
- Accountability and responsibility;
- Democracy;
- Rule of law.

For each of the mentioned basic values, the CAHAI-Report contains a list of the "key-substantive rights" and of the "key obligations". The analysis gives a valuable overview of

the concerned legal challenges and can serve as a foundation for further discussions about the design and contents of the normative framework.

The mentioned basic values and their concretization in the form of substantive rights and obligations appear to be a good guideline for the subsequent legislative activities. In this respect, the CAHAI-Report describes a couple of policy options that have been submitted to the multistakeholder environment for comments. Possible legal instruments are the adoption of a new binding legal treaty (Convention or Framework Convention) or, alternatively, different forms of non-binding legal guidelines [29] (p. 46 et seqq.). Partly, a certain complementarity is diagnosed between the horizontal and cross-cutting elements that could form part of a conventional-type instrument and the vertical and sectoral work that could give rise to specific instruments of a different nature [29] (p. 50).

The CAHAI-Report is an excellent feasibility study, which not only addresses the major AI challenges in light of the fundamental normative principles of the ECHR, but also assesses the possible legal instruments that could be developed in order to have governed the AI systems by an appropriate legal framework. Particularly, the CAHAI-Report acknowledges that hard law and soft law instruments, preferably in a co-regulatory framework, should be considered in the design of an appropriate normative environment.

3.2. European Union: Proposal for a New AI Regulation

On 21 April 2021, the European Commission published a proposal for a Regulation of the European Parliament and of the Council "Laying Down Harmonised Rules on Artificial Intelligence" [46]. With this Regulation, the European Commission intends to implement a dense framework that should prevent or mitigate the risks caused by AI systems.

3.2.1. Background of the Regulation

As mentioned, in 2018, the European Commission appointed a High-Level Expert Group that had the task to assess the chances and risks of AI systems in an interdisciplinary way. The purpose of the efforts consisted of the attempt to issue AI guidelines that would particularly address ethics issues; the respective guidelines were published in spring 2019 [31].

In 2020, the Commission published a White Paper on artificial intelligence, setting out policy options on how to achieve the twin objective of promoting the update of AI and of addressing the risks associated with certain uses of such technology [47]. On 21 October 2020, the Council of the European Union published so-called "Conclusions" addressing the opacity, complexity, bias, unpredictability, and the partially autonomous behavior of certain AI systems with the objective of ensuring their compatibility with fundamental rights [48].

The European Parliament has also undertaken a considerable amount of work in the area of artificial intelligence. Between fall 2020 and spring 2021, several resolutions concerning AI topics (for example on ethics, liability, copyright, criminal matters, education, culture, etc.) were discussed and adopted [46] (p. 2 for further details).

The EU Commission's proposal for a new Regulation is designed with the aim to realize the following objectives [46] (p. 3):

- Ensure that AI systems placed on the Internal Market and used are safe and respect existing law on fundamental rights and Union values;
- Ensure legal certainty to facilitate investment and innovation in AI;
- Enhance governance and effective enforcement of existing law on fundamental rights and safety requirements applicable to AI systems;
- Facilitate the development of a single market for lawful, safe, and trustworthy AI applications and prevent market fragmentation.

The proposed AI Regulation states quite detailed harmonized rules for the development, placement on the market, and use of the AI systems in the EU following a risk-based approach; the European Commission is of the opinion that such an approach would be most appropriate and proportional. The legislative project seems to be the first general

and broad attempt worldwide of introducing a legislative instrument that is designed to regulate the AI phenomenon [49].

The form of a Regulation, not of a Directive, has been chosen in order to avoid fragmentation regarding essential normative elements of the Internal Market in the EU and to increase legal certainty for both providers and users of AI systems [46] (p. 6). Between different potential policy options, the quite dense approach of a horizontal legislative instrument with mandatory requirements for high-risk AI systems and codes of conduct for non-high-risk AI systems has been chosen [46] (p. 9).

3.2.2. Contents of the Regulation

After a long Preamble of more than twenty pages, Title I of the Regulation contains rules governing its scope and outlining definitions. The important subject matter is described in Article 1; the contents are (i) harmonized rules for the placing on the market, the putting into service, and the use of AI systems; (ii) prohibitions of certain AI practices; (iii) specific requirements for high-risk AI systems and obligations for their operators; (iv) harmonized transparency rules for AI systems, intended to interact with natural persons, emotion recognition systems, and biometric categorization systems, and AI systems used to generate or manipulate image, audio, or video content, as well as (v) rules on market monitoring and surveillance. Surprisingly, however, the non-discrimination principle is not expressly laid down (see above Section 2.3.1 and [24] (no. 30)). By purpose, liability issues are not addressed, and a proposal for regulating AI liability will be published in the second half of 2021 [49] (p. 362 no. 4).

An "artificial intelligence system" is defined as "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" (Article 3 para. 1). In particular, deep learning and machine learning are covered. The choice for a wide definition and its concretization in an Annex that can be adapted more easily should serve the purpose of having a technology-neutral AI description. However, whether this intention will be fully realized remains doubtful, not at least due to the fact that the proposed AI Regulation distinguishes between "embedded" AI systems and "stand-alone" AI systems being submitted to different regulatory regimes [49] (p. 363 no. 13). The data protection rules as stated in the GDPR are not affected by the new AI Regulation.

Since the AI Regulation is based on a risk-oriented approach, the first provisions deal with the completely prohibited activities. Specifically, Title II of the Regulation contains rules that state which practices shall be prohibited; the respective Article 5 is quite long and lists those AI practices that cannot be justified at all (for example social scoring, biometric recognition under certain circumstances) [49] (p. 365 nos. 25–28 for further details). Insofar, however, the long-term problem should not be underestimated that some AI systems are qualified as "forbidden" even if future technological developments would be able to minimize the risks.

The long Title III of the Regulation addresses high-risk AI systems. The detailed Article 6 distinguishes between two main groups of AI systems, namely the (i) embedded systems that have to comply with the requirements contained in several legal instruments providing for a product safety framework and the (ii) stand-alone systems to be implemented in accordance with Annex II of the Regulation. In addition, the EU Commission will be entitled to amend and modify the specific technological requirements listed in Annex III in order to diminish potential risks of AI systems (Article 7).

A concerned AI provider must fulfil different requirements, for example establish a risk management system (Article 9), implement appropriate data and data governance measures (Article 10), draw up the necessary technical documentation (Article 11), and safeguard the record-keeping (Article 12). The transparency obligations and the duties to provide information to the users are particularly important (Article 13). Furthermore, high-risk AI systems shall be designed and developed in such a way (including appropriate

human–machine interface tools) that they can be effectively overseen by natural persons during the period in which the AI system is in use (Article 14) [49] (pp. 366–368 nos. 33–44, with references to German publications).

In view of this long list of obligations imposed on AI systems providers, the compliance of AI systems with the legal framework will have to be done on the basis of quality management systems in the future (Article 17). Additional provisions address the accuracy, robustness, and cybersecurity elements, the conformity assessments, and the specific obligations of AI systems' providers. Further transparency obligations are foreseen for certain AI systems that (i) interact with humans, (ii) are used to detect emotions or determined association with (social) categories based on biometric data, or (iii) generate or manipulate content (deep fakes).

Title V of the Regulation contributes to the objective of creating an innovation-friendly legal framework by allowing the setup of regulatory sandboxes that are resilient to disruption. The instrument of regulatory sandboxes is already known from financial market regulations and has shown its merits in dealing with newly developing technologies [50]. AI regulatory sandboxes establish a controlled environment allowing for the testing of innovative technologies for a limited time [51].

In case of an AI use other than a high-risk AI system, the implemented risk minimization measures can be based on codes of conduct. Such codes may be drawn up by individual providers of AI systems with the support of the concerned stakeholders or (more likely) by organizations representing the providers or both the providers and users (Article 69). However, in contrast to the legal situation in the data protection context (Art. 40 GDPR), the compliance with such codes of conduct does not relieve the providers of AI systems from the observation of the general compliance and risk management measures [49] (p. 371 no. 70).

The Regulation is complemented with many provisions on governance and implementation issues (including the establishment of a new "European Artificial Intelligence Board") as well as with provisions on criminal sanctions. In the case that an enterprise is violating the proposed AI Regulation, a fine amounting to a maximum 6% of the yearly turnover of the concerned enterprise can be levied [49] (pp. 372–373 nos. 72–82).

3.2.3. Assessment of the Regulation

The proposed AI Regulation is a very detailed legal instrument being directly applicable in the Member States of the EU; the "softer" instrument of a Directive that must be transformed into national law is no longer chosen by the European Commission in the context of the Internal Market regulations (incl. digital services) in order to avoid a normative fragmentation within the EU. Such an approach has certain merits; however, it deprives nation States more and more from their legislative competences in civil law matters.

The proposal for an AI Regulation encompassing a wide variety of AI systems follows a strict risk-based approach. Such a policy option can be chosen; however, other options should also be analyzed in the preparation of the new provisions, for example a rights-oriented model (similar to the GDPR), a damages-based approach, or a cost–benefit analysis (looking at the cheapest cost avoider). The pure risk-orientation is exposed to the problem of over-regulation as the length of the proposed AI Regulation clearly shows [49] (p. 374 no. 86; slightly critical).

As far as the theoretical regulatory model is concerned, obviously, if in place, the AI Regulation will constitute hard law. Even if some instruments foreseen therein—similarly as in the case of the GDPR—belong to the category of soft law (such as the codes of conduct), the question can be raised whether it is appropriate to limit the room for co-regulatory or self-regulatory instruments to such a far-reaching extent. As outlined throughout the above general observations to the fundamental legal principles (see above Sections 2.3.1–2.3.3), a combination of regulatory models (including co-regulation) appears to be the most fruitful way to go forward [52] (pp. 97−112).

4. Outlook

Artificial intelligence confronts the legal environment with substantial but not unsurmountable challenges. The key tension consists in the fact that the use of AI systems impacts fundamental normative principles (such as non-discrimination, human rights, transparency, etc.) and not solely certain sectoral law areas (see above Section 2.3 (introduction)). Therefore, an appropriate normative concept must address the trust and ethics expectations of civil society and include its interests in the decision-making processes. As a result, AI systems can only be implemented in a normative framework securing compliance with the generally accepted fundamental rights.

The most recent initiatives presented in Europe, particularly by the Commission of the European Union, tend to implement a hard law instrument. This approach, with a detailed and dense AI Regulation harmonizing the rules on the Internal Market of the EU, is confronted with the risk that the legal provisions limit innovation by the private actors in overly tight clothes, not leaving much room for maneuver. Often, soft law instruments appear to be more suitable. If the "governed" persons are involved in the rulemaking, the chances increase that the respective legal provisions are timely as well as more business-oriented, more need-driven, and also more likely to be observed.

In the new technological environment, a flexible and innovation-friendly model of regulation is required. Insofar, the model of co-regulation appears to be particularly apt to design and frame the regulatory environment as the experience in the Internet governance context has shown. Co-regulation as a newly developed normative model between soft law and hard law allows regulators to enter into arrangements by varying scope and specificity and then—with the purpose of improving the normative quality of the rules—to clarify (or change) the expectations of the involved actors.

The ongoing discussions within the Council of Europe have concluded that several policy options should be submitted to the stakeholders for further deliberations. This approach allows for weighing and balancing the advantages and disadvantages of the different regulatory models. As outlined in this contribution, a combination of normative frameworks merits attracting more attention. The ongoing debates in Europe appear to be an appropriate laboratory for other countries around the world being confronted with the task of how the design of the legal environment for AI systems should be determined.

Funding: This research received no external funding.

Conflicts of Interest: The author declares no conflict of interest.

References

1. Weber, R.H. Socio-ethical values and legal rules on automated platforms: The quest for a symbiotic relationship. *CLSR* **2020**, *36*, 105380. [CrossRef]
2. Weber, R.H. *Internet Governance at the Point of No Return*; EIZ Publishing: Zurich, Switzerland, 2021. [CrossRef]
3. Weber, R.H. *Shaping Internet Governance: Regulatory Challenges*; Schulthess: Zurich, Switzerland, 2009.
4. Report of the Working Group on Internet Governance, June 2005. Available online: www.wgig.org/docs/WGIGREPORT.pdf (accessed on 30 August 2021).
5. Weber, R.H. Legal foundations of multistakeholder decision-making. *Z. Für Schweiz. Recht* **2016**, *135*, 247–267.
6. Gasser, U.; Budish, R.; West, S.M. *Multistakeholder as Governance Groups: Observations from Case Studies*; Berkman Center for Internet & Society Research Publications: Cambridge, MA, USA, 2015. [CrossRef]
7. Weber, R.H. *Realizing a New Global Cyberspace Framework. Normative Foundations and Guiding Principles*; Schulthess: Zurich, Switzerland, 2014.
8. Weber, R.H. Sectoral Self-Regulation as Viable Tool. In *Law and Economics of Regulation*; Klaus, M., Tor, A., Eds.; Springer: Cham, Switzerland, 2021; forthcoming.
9. Weber, R.H. Overcoming the Hard Law/Soft Law Dichotomy in Times of (Financial) Crises. *J. Gov. Regul.* **2012**, *1*, 8–14. [CrossRef]
10. Guzman, A.T.; Meyer, T.L. International Soft Law. *J. Leg. Anal.* **2010**, *2*, 171–225. [CrossRef]
11. Meyer, T. Soft Law as Delegation. *Int. Law J.* **2008**, *32*, 888–942.
12. Tambini, D.; Leonardi, D.; Marsden, C. *Codifying Cyberspace: Communications Self-Regulation in the Age of Internet Convergence*; Routledge: London, England, 2008.
13. Gersen, J.E.; Posner, E.A. Soft Law: Lessons from Congressional Practice. *Stan. L. Rev.* **2008**, *61*, 573–628.
14. Senn, M. *Non-State Regulatory Regimes, Understanding Institutional Transformation*; Springer: Berlin/Heidelberg, Germany, 2011.

15. Marsden, C.T. Internet Co-Regulation and Constitutionalism: Towards a More Nuanced View, SSRN 2011. Available online: http://dx.doi.org/10.2139/ssrn.1973328 (accessed on 30 August 2021).
16. Marsden, C.T.; Meyer, T.; Brown, I. Platform values and democratic elections: How can the law regulate digital disinformation? *CLSR* **2020**, *36*, 105373. [CrossRef]
17. Black, J. Constitutionalizing Self-Regulation. *Mod. L. Rev.* **1996**, *59*, 24–55. [CrossRef]
18. Teubner, G. Digitale Rechtssubjekte? Zum privatrechtlichen Status autonomer Softwareagenten. *Arch. Für Civ. Prax.* **2018**, *218*, 155–205. [CrossRef]
19. Grinzinger, J. Der Einsatz Künstlischer Intelligenz in Vertragsverhältnissen. In *Privatrecht 2050—Blick in Die Digitale Zukunft, Jahrbuch Junge Zivilrechtswissenschaft 2019*; Beyer, E., Erler, K., Hartmann, C., Kramme, M., Müller, M.F., Pertot, T., Tuna, E., Wilke, F.M., Eds.; Nomos: Baden-Baden, Germany, 2020; pp. 151–180.
20. Chesterman, S. Artificial Intelligence and the Limits of Legal Personality. *ICLQ* **2020**, *69*, 819–844. [CrossRef]
21. Weber, R.H. Ethics in the Internet Environment. In *Global Commission on Internet Governance*; Paper Series No. 39; Centre for International Governance Innovation: Waterloo, CA, USA, 2016.
22. O'Neill, O. *A Question of Trust*; Cambridge University Press: Cambridge, England, 2002.
23. Castelluccia, C.; Le Métayer, D. *Understanding Algorithmic Decision-Making: Opportunities and Challenges*; Study for the European Parliament: Brussels, Belgium, 2019.
24. Braun Binder, N.; Burri, T.; Lohmann, M.F.; Simmler, M.; Thouvenin, F.; Vokinger, K.N. Künstliche Intelligenz: Handlungsbedarf im Schweizer Recht. *Jusletter* **2011**, 1–25. [CrossRef]
25. Weber, R.H.; Henseler, S. Regulierung von Algorithmen in der EU und in der Schweiz. *EUZ* **2020**, *22*, 28–42.
26. Borgerius, F.Z. Discrimination, Artifical Intelligence, and Algorithmic Decision-Making, Strasbourg 2018. Available online: https://rm.coe.int/discrimination-artificial-intelligence-and-algorithmic-decision-making/1680925d73 (accessed on 30 August 2021).
27. Barocas, S.; Selbst, A.D. Big Data's Disparate Impact. *Cal. L. Rev.* **2016**, *104*, 671–732. [CrossRef]
28. Wachter, S. Affinity Profiling and Discrimination by Association in Online Behavioural Advertising. *BTLJ* **2020**, *30*, 367–430. [CrossRef]
29. Council of Europe, Ad hoc Committee on Artificial Intelligence, Feasibility Study, CAHAI(2020)23, Strasbourg, 17 December 2020. Available online: https://www.coe.int/en/web/artificial-intelligence/cahai (accessed on 30 August 2021).
30. Kleinberg, J.; Ludwig, J.; Mullainathan, S.; Rambachan, A. Algorithmic Fairness. *AEA Pap. Proc.* **2018**, *108*, 22–27. [CrossRef]
31. High-Level Expert Group of the European Union, Ethics Guidelines for Trustworthy AI, April 2019. Available online: https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai (accessed on 30 August 2021).
32. OECD. Principles on Artificial Intelligence, 22 May 2019. Available online: https://www.oecd.org/going-digital/ai/principles/ (accessed on 30 August 2021).
33. Google. AI at Google: Our Principles, 7 June 2018. Available online: https://www.blog.google/technology/ai/ai-principles/ (accessed on 30 August 2021).
34. Microsoft. Microsoft AI Principles. Available online: https://www.microsoft.com/en-us/ai/our-approach-to-ai (accessed on 30 August 2021).
35. Zerilli, J.; Knott, A.; Maclaurin, J.; Gavaghan, C. Transparency in Algorithmic an Human Decision-Making: Is There a Double Standard? *Philos. Technol.* **2019**, *32*, 661–683. [CrossRef]
36. Martini, M. *Blackbox Algorithmus—Grundfragen Einer Regulierung Künstlerischer Intelliigenz*; Springer: Berlin/Heidelberg, Germany, 2019.
37. Official Journal of the European Union, L 186/57 of 11 July 2019. Available online: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A62011TN0238&qid=1630891276279 (accessed on 30 August 2021).
38. Official Journal of the European Union, L 149/22 of 11 June 2005. Available online: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3AC2000%2F149%2F22&qid=1630891511912(accessed on 30 August 2021).
39. Official Journal of the European Union, L 328/7 of 18 December 2019.
40. Weber, R.H. From disclosure to transparency in consumer law. In *Consumer Law and Economics*; Mathis, K., Tor, A., Eds.; Springer: Cham, Switzerland, 2020; pp. 73–87.
41. McGregor, L.; Murray, D.; Ng, V. International Human Rights Law as a Framework for Algorithmic Accountability. *Int. Comp. Law Q.* **2019**, *68*, 309–343. [CrossRef]
42. Djeffal, C. The Normative Potential of the European Rule on Automated Decisions: A New Recording for Art. 22 GDPR. *ZaöRV.* **2020**, *80*, 847–879.
43. Mendoza, I.; Lee, A.B. The Right Not. to be Subject to Automated Decisions Based on Profiling. In *EU Internet Law—Regulation and Enforcement*; Synodinou, T.E., Jougleux, P., Markou, C., Prastitou, T., Eds.; Springer: Cham, Switzerland, 2017; pp. 77–98.
44. Edwards, L.; Veale, M. Slave to Algorithm? Why a Right to an Explanation is Probably Not the Remedy You Are Looking For. *Duke L. Tech. Rev.* **2017**, *16*, 18–84.
45. Gasser, U.; Virgilio Almeida, A.F. A Layered Model for AI Governance. *IEEE Internet Comput.* **2017**, *21*, 58–62. [CrossRef]

46. European Commission. Proposal for a Regulation of the European Parliament and of the Council, Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts, SEC(2021) 167 final, COM(2021) 2006 final. Brussels, Belgium, 21 April 2021. Available online: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52021PC0206 (accessed on 30 August 2021).
47. European Commission. White Paper on Artificial Intelligence—A European approach to excellence and trust. *COM* **2020**, *65*, 1–26.
48. Council of the European Union. Presidency Conclusions—The Charter of Fundamental Rights in the Context of Artificial Intelligence and Digital Change. Brussels, Belgium, 21 October 2020. Available online: https://www.consilium.europa.eu/media/46496/st11481-en20.pdf (accessed on 30 August 2021).
49. Spindler, G. Der Vorschlag der EU-Kommission für eine Verordnung zur Regulierung der Künstlichen Intelligenz (KI-VO-E). *Comput. und. Recht.* **2021**, *6*, 361–374.
50. Zetzsche, D.A.; Woxholth, J. *The DLT Sandbox under the Pilot-Regulation*; EBI Working Paper Series 2021 No. 92; EBI: Frankfurt am Main, Germany, 2021.
51. Krönke, C. Sandkastenspiele—«Regulatory Sandboxes» aus der Perspektive des Allgemeinen Verwaltungsrechts. *Juristen-Zeitung* **2021**, *76*, 434–443. [CrossRef]
52. Veale, M.; Zuidereveen Borgesisu, F. Demystifying the Draft EU Artificial Intelligence Act. *Comp. L. Rev. Int.* **2021**, *4*, 97–112.