The Artificial Intelligence Act: A Jurisprudential Perspective

Michał Araszkiewicz [1], Grzegorz J. Nalepa [2] and Radosław Pałosz [1]

[1] *Jagiellonian University, Department of Legal Theory*
[2] *Jagiellonian University, Department of Physics, Astronomy and Applied Informatics*

Abstract

The aim of this paper is to analyze the Artificial Intelligence Act (AIA) from the point of view of theories from the field of analytical jurisprudence—namely, the Hofeldian theory of fundamental legal concepts and the theory of rules and principles developed by Dworkin and Alexy. This enables us to formulate a research question concerning the reconstruction of the complete set of the normative positions deriving from the AIA and to indicate what values should be taken into account in the process of balancing of principles related to the enforcement of this act.

Keywords

Artificial Intelligence Act, legal concepts, normative positions, principles, values

1. Introduction

This paper provides preliminary insights into the problems related to the 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 (hereinafter AIA) from a general legal–theoretical perspective. The AIA has already drawn comments from the perspectives of European Union law and information technology law. However, a general jurisprudential perspective on the problems arising from this act is still lacking. This paper's purpose is to take the first step toward filling this gap. We focus on two interrelated perspectives: the theory of fundamental legal concepts and the theory of legal norms. Accordingly, we first refer to research based on the paradigm initiated by Wesley Hohfeld, focusing on the normative positions created by legal norms. Normative positions, such as rights, duties, and powers, are interconnected and thus enable important legal inference patterns. This approach indicates how the set of normative positions deriving from the AIA should be determined on an abstract level. However, some normative positions may be determined only on the level of the application of norms in specific situations. Here, the second perspective plays a key role. We focus on the distinction between rules and principles introduced to mainstream legal theory by Ronald Dworkin and then modified and extended by Robert Alexy. According to this distinction, there are important differences in the modes of application of the two types of norms and different procedures for the resolution of incompatibilities between them. This perspective is of crucial importance because the EU norms protecting fundamental rights, which provide both the foundation for the AIA and constraints on its interpretation, often have the structure of principles in the sense of Alexy—that is, optimization commands. These general considerations enable us to anticipate the categories of problems related to the interpretation and enforcement of the AIA.

2. Structure of the Artificial Intelligence Act

In its current form, the AIA consists of 85 articles divided into 12 titles, which are preceded by a preamble. There are also nine annexes that complement the regulation's provisions.

Title I contains the definitions of notions used in the regulation and determines its scope. It should be read in conjunction with Annex I, which provides information on how "artificial intelligence techniques and approaches" should be understood. This is an important part of the definition of AI.

Title II enumerates and describes prohibited AI practices. It includes the prohibition of AI systems that have the potential to manipulate actions of natural persons through subliminal techniques or exploit the vulnerabilities of certain groups to "materially distort their behaviour in a manner that is likely to cause them or another person psychological or physical harm". It also prohibits social scoring systems, like the one used in the People's Republic of China, and (with some exceptions) real-time biometric identification systems in public spaces.

Title III is probably the most important part of the AIA. It regulates high-risk AI systems—a widely diverse group that includes systems that are "intended to be used as safety component of products that are subject to third party ex-ante conformity assessment" and other standalone AI systems with fundamental rights implications, listed in Annex III . Taking a risk-based approach to ensure EU citizens' health and safety, the AIA provides for the imposition of multiple obligations on AI system providers and distributors. Title III consists of five chapters. Chapter 1 determines the scope of high-risk AI systems as described above. Chapter 2 describes the requirements that AI systems should meet. These concern the proper documentation and functioning of AI systems to ensure transparency, robustness, accuracy, and security, which should be controllable during their operation. The legal norms reconstructed from Title III are relatively complex and multilayered, and, to a large extent, their content needs to be supplemented by other sources, such as particular decisions or soft law.

Title IV contains rules on the operational transparency of certain AI systems that may entail a risk of manipulation. The project identifies them as "systems that (i) interact with humans, (ii) are used to detect emotions or determine association with (social) categories based on biometric data, or (iii) generate or manipulate content ('deep fakes')."[1] The aim of these regulations is to ensure that users are aware that they are interacting with AI and not with a living person.

Title V aims to promote innovation in the development of AI systems. Its articles grant competences to national organs to create "regulative sandboxes" to facilitate the creation of innovative AI systems. The norms of this title can be described as classical delegation of power.

Titles VI–VIII provide for the development of a multilevel governance system. They are intended to form the basis for the creation of a European Artificial Intelligence Board and the designation of competent organs of the member states for the enforcement of the AIA and its supervision. To this end, specialized databases will be created, and AI system providers will be obligated to monitor system operations and report to appropriate authorities.Title IX presents the framework for non-high-risk AI system providers to prepare codes of conduct for the implementation of some solutions obligatory for high-risk AI systems.

Titles X–XII contain the final provisions. They have an organizational character, aiming to give the Commission the competence to implement proportional measures to ensure efficiency in the enforcement of the regulation.

3. Rights, duties, and powers

Wesley Hohfeld is considered the first to perform a systematic analysis of normative positions (to which he referred as "fundamental legal conceptions" [1, 2] and the relations between them. Such relations can be seen in the two Hohfeldian squares shown in Figure 1.
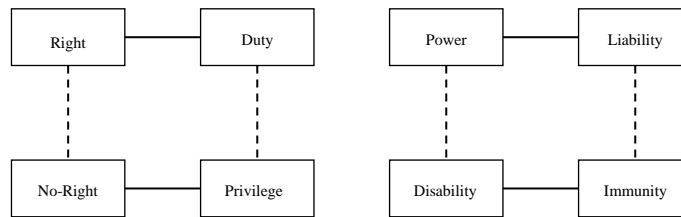
Figure 1: Hohfeldian squares. On the left, deontic concepts and their relations; on the right, potestative concepts and their relations. The solid lines indicate correlatives, and the dashed lines indicate opposites.

It is not our aim to discuss Hohfeld's understanding of these concepts in detail. Suffice it to note that he created a paradigm for the analysis of normative positions using formal tools. Thanks to the development of logic in the 20th century, this led to important contributions by Kanger [3] and Lindahl [4], which resulted in the identification of 255 distinct relations and demonstrated the possibility of implementing them computationally [5]. Research on normative positions continues in different conceptual settings [6, 7]. It is important to note the five theses deriving from this stream of research:

- Both deontic and potestative concepts are interconnected by relations that enable inference (particularly about the normative positions implied by the explicit regulative text).
- Legal provisions are typically ambiguous as to which type of normative concept they express. The types of normative positions may be identified through formal analysis.
- Normative concepts may be analyzed on different levels of granularity. More fine-grained models lead to further distinctions.
- The concept of power (generally speaking, creating new states of affairs in the world of law) leads to the distinction of particularly complex normative positions.
- An understanding of a given normative position may be gained through the identification of its relations with other normative positions.

These considerations lead to the formulation of the following research question: *What is the set of normative positions, both explicit and implicit, that can be extracted from the text of the AIA on an abstract level?* The realization of this task requires the reconstruction of an outline of the AIA domain ontology that encompasses, *inter alia*, the set of entities involved in relations, the objects of these relations, and the different modes of action that these agents can perform and their effects.[2] The conceptual schemes employed in ontologies built for modeling the General Data Protection Regulation domain may serve as a useful benchmark [8].

In such an outline of a domain ontology, the following categories of normative positions should be indicated and elaborated upon:

- Obligations, especially those related to the requirements for high-risk AI systems and the remaining obligations of the providers of such systems
- Powers, particularly those assigned to member states, notifying authorities, notified bodies, national competent authorities, or market surveillance authorities
- Rights, particularly of recipients of decisions made using intelligent systems

It should be noted that these normative positions require fine-grained modeling due to the relatively specific language used in the AIA text. For instance, let us consider the wording used in Art. 15.2: "High-risk AI systems shall be accompanied by instructions for use in an appropriate digital format or otherwise that include concise, complete, correct and clear information that is relevant, accessible and comprehensible to users." The provision involves four features of instructions—conciseness, completeness, correctness, and clearness—and three more relational predicates—relevance, accessibility, and comprehensibility. This example illustrates the relative complexity of the content of the planned provisions even in terms of their structure.

Moreover, the use of open-textured concepts in the formulation of the AIA provisions means that it will often be necessary to conduct a contextual assessment of the scope of these positions in particular (classes of) cases. Moreover, numerous positions listed in the AIA, including obligations, have a

---

[2] Outlines of ontologies for standardizing trustworthy AI are already available; see [9].

meta-level character in the sense that they require conformity with standards that have yet to be developed and adopted and that must themselves comply with certain requirements. One of the important aims of this investigation would be to indicate the subset of normative positions deriving from the AIA so that they may be determined in abstraction (and extent) and a subset thereof that requires further regulatory activities or evaluation stemming from the balancing of values in the context of specific administrative or judicial proceedings.

4. Principles and values

It is possible to identify the values protected by the AIA by applying the theory of principles that Robert Alexy proposed (and presented in their fullest form in [10]) drawing on Ronald Dworkin's ideas. The core concept of Dworkin's theory is that there are two types of legal norms: rules applied on an 'all-or-nothing' basis and principles that serve interpretation purposes in cases of conflict between rules [11]. In developing this theory, Alexy argues that principles are meant to optimize the application of rules. The difference between rules and principles lies in their quality and does not depend on their hierarchical position in a legal system. While a conflict of rules results in the invalidation of one of them, a conflict of principles does not lead to the invalidation of any of them. Different principles have different weights in varying legal contexts. However, due to their optimizing character, in each case, a conflict of principles is solved by the creation of a special rule that indicates the principle that should be applied. Such a conflict should be resolved by answering the question of how a particular legal aim can be achieved by selecting from a given set of physical and legal possibilities. Principles, in the way described above, make it possible to choose from existing options and find the best solution to a legal problem.

According to Alexy's theory, the identification of principles for optimizing the enforcement of the AIA makes it possible to identify values that should provide backing for its interpretation. Moreover, it can lead to delineating possible conflicts between principles. This can complete the picture started by setting the normative positions expressed by the AIA, as some normative positions can be identified only after resolving conflicts between principles.

Most conflicts will probably arise between the principle of protecting EU citizens' fundamental rights and the principle of promoting AI innovation. In many places, the text stresses that due to their autonomous character and extraordinary data-processing capabilities, AI systems could pose a threat to such values as health and safety (motive 1 of the AIA) or fundamental human rights, such as non-discrimination, data protection and privacy, and the rights of the child, through the exploitation of the vulnerabilities of certain groups (motive 15 of the AIA). The published text establishes multiple principles to meet these goals, using general terms to set the standards required for providers of AI systems to ensure their safety and appropriate oversight. It is worth highlighting some provisions concerning high-risk AI systems, such as Art. 9.3 ("The risk management measures referred to in paragraph 2, point (d) shall give due consideration to the effects and possible interactions resulting from the combined application of the requirements set out in this Chapter 2"), Art. 10.2 ("Training, validation and testing data sets shall be subject to appropriate data governance and management practices"), and Art. 13.1 ("High-risk AI systems shall be designed and developed in such a way to ensure that their operation is sufficiently transparent to enable users to interpret the system's output and use it appropriately").

On the other hand, the text also contains regulations that set the boundaries of the abovementioned constraints to ensure that they will not hamper the development of innovative AI systems—for example, Art. 9.6 ("Testing procedures shall be suitable to achieve the intended purpose of the AI system and do not need to go beyond what is necessary to achieve that purpose"). The frequent use of terms such as "reasonable" and "appropriate" also indicates that risk control–related principles are always limited by principles that ensure the market competitiveness of AI system providers and innovation in the development of new types of systems (which can be inferred from motive 1 of the AIA project).

5. Related work

The extant literature on this subject lacks the analysis proposed in this paper. Moreover, there is little work on the values enshrined in the AIA and the relations between them. The most relevant work in this area is probably [12], which criticizes the AIA for failing to ensure an appropriate level of protection of fundamental rights. The authors argue that the project does not meet all the requirements of legally trustworthy AI, as it fails to recognize fundamental rights as claims with "enhanced moral and legal status" and to reflect all the responsibilities associated with different types of AI systems. An outline of the aims and values of the project is presented in [13], which systematizes its provisions. A thorough legal analysis of the AIA can be found in [14], which indicates possible problems with its enforceability and the dangers of far-reaching harmonization that may make it difficult for member states to adapt it to their conditions. Another critical approach to the regulation can be found in [15]. Examining how to best meet the requirement of standardization of AI systems, mentioned in motive 61 of the AIA, Ebers [16] argues that the text is insufficient to appropriately standardize AI systems and that ethical considerations are needed in connection with further political decisions.

Taking different approaches to the AIA, all these works underline the role that a proper understanding and application of the AIA's values plays in achieving its goals. However, as none of these works takes a general jurisprudential approach, the analysis presented herein may contribute to better addressing the highlighted problems. A clear separation of normative positions can undoubtedly enhance the operability of future analyses.

6. Conclusions and further research

In this paper, we provide an overview of the AIA drawing on the jurisprudential theories of normative positions and of rules and principles. In our view, only the realization of the research program outlined in this paper can lead to a general answer to the question of what the interpretation of the AIA and its practical application will look like. The realization of this research program involves the determination of a possibly complete set of normative positions deriving from the content of the AIA and the indication of which of these normative positions can be completely reconstructed on an abstract level. It also involves the indication of patterns of values balancing where the content of those normative positions will need specification regarding the application of provisions in particular cases. This research should serve not only descriptive but also evaluative purposes, as it should lead to the identification of semantic overlaps or indeterminateness in the content of the AIA. It should also provoke normative questions regarding whether the content of the AIA actually determines the resulting set of states of affairs intended by European lawmakers and, if not, what instruments may be considered to achieve a better outcome [17]. These questions are also related to the procedural context of the competences of law-enforcing bodies in applying the AIA provisions, given the dynamically and rapidly changing landscape of AI technology standards and specifications. Last but certainly not least, the realization of this research program may provide a basis for building computational ontologies of the AIA and eventually enabling computational support and automated predictions of decisions. This would lead to an interesting paradox, as the systems used for this purpose would themselves be subject to the requirements stipulated by the AIA and would therefore reason about the legality of their own operation.

7. Acknowledgements

8. References

[1] W. Hohfeld, Some fundamental legal conceptions as applied in judicial reasoning, The Yale Law Journal 23 (1913) 16–59.

[2] W. Hohfeld, Some fundamental legal conceptions as applied in judicial reasoning, The Yale Law Journal 26 (1917) 710–770.

[3] S. Kanger, Law and logic, Theoria 38 (1972) 105–132.

[4] L. Lindahl, Position and Change: A Study in Law and Logic, Reidel, Dordrecht, 1977.

[5] M. Sergot, Normative positions, in: D. Gabbay, J. Horty, X. Parent, R. van der Meyden, L. van der Torre (Eds.), Handbook of Deontic Logic and Normative Systems, College Publications, Rickmansworth, 2013, pp. 353–406.

[6] G. Sartor, Fundamental legal concepts: A formal and teleological characterization, Artificial Intelligence and Law 14 (2006) 101–142.

[7] J. Hage, Foundations and Building Blocks of Law, Eleven International Publishing, Den Haag, 2018.

[8] M. Palmirani, M. Martoni, A. Rossi, C. Bartolini, L. Robaldo, Legal ontology for modelling GDPR concepts and norms, in: M. Palmirani (Ed.), Legal Knowledge and Information Systems: JURIX 2018: The Thirty-first Annual Conference, Groningen, Netherlands, Frontiers in Artificial Intelligence and Applications 313, IOS Press, Amsterdam 2018, pp. 101–110.

[9] D. Lewis, D. Filip, H. J. Pandit, An ontology for standardising trustworthy AI, in: A. G. Hessami, P. Shaw (eds.), Factoring Ethics in Technology, Policy Making, Regulation and AI. doi:10.5772/intechopen.97478.

[10] R. Alexy, A Theory of Constitutional Rights, Oxford University Press, Oxford, 2002.

[11] R. Dworkin, The model of rules, University of Chicago Law Review 1 (1967), 14–46.

[12] N. Smuha, E. Ahmed-Rengers, A. Harkens, W. Li, J. MacLaren, R. Piselli, K. Yeung, , How the EU Can Achieve Legally Trustworthy AI: A Response to the European Commission's Proposal for an Artificial Intelligence Act (2021). doi:10.2139/ssrn.3899991.

[13] M. Kop, EU Artificial Intelligence Act: The European approach to AI. Stanford - Vienna Transatlantic Technology Law Forum, Transatlantic Antitrust and IPR Developments, Stanford University, Issue No. 2/2021, 2021. URL: https://law.stanford.edu/publications/eu-artificial-intelligence-act-the-european-approach-to-ai/

[14] M. Veale, F. Zuiderveen Borgesius, Demystifying the draft EU Artificial Intelligence Act — Analysing the good, the bad, and the unclear elements of the proposed approach, Computer Law Review International 4 (2021), 97–112.

[15] M. Ebers, V. R. S. Hoch, F. Rosenkranz, H. Ruschemeier, B. Steinrötter, The European Commission's proposal for an Artificial Intelligence Act—A critical assessment by members of the Robotics and AI Law Society (RAILS), J 4 (2021), 589–603.

[16] M. Ebers, Standardizing AI - The case of the European Commission's proposal for an Artificial Intelligence Act, in: L. DiMatteo, C. Poncibò, M. Cannarsa (Eds.), The Cambridge Handbook of Artificial Intelligence: Global Perspectives on Law and Ethics, Cambridge University Press, Cambridge, 2022. Published at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3900378.

[17] M. Araszkiewicz, K. Płeszka, The concept of normative consequence and legislative discourse, in: M. Araszkiewicz, K. Płeszka (Eds.), Logic in the Theory and Practice of Lawmaking, Springer, Cham, 2015, pp. 253–297.