Compliance and enforcement in the AIA through AI

Compliance and Enforcement in the AIA through AI

Hans-W. Micklitz and Giovanni Sartor[1]

21-12-24

[1] Both authors received funding from Giovanni Sartor's ERC grant on CompuLaw – Governance of Computational Entities through an Integrated Legal and Technical Framework, https://site.unibo.it/compulaw/en/project

1 Context and Argument

Compliance and enforcement of the EU Artificial Intelligence Act[2] (in the following AIA) is only possible, feasible and affordable if both private parties and public authorities develop appropriate procedures, skills and tools, in isolation and/or jointly, to maximise compliance with a broad range of due diligence obligations. The public and academic debate on the AIA focuses on the material substance of the Union law, the yardstick against which the legality of AI shall be measured, and debates on what exactly counts as high risk or what kind of AI should be forbidden.

What is sometimes overlooked is that the EU has established a vast ensemble of requirements that have to be implemented by the AI companies themselves through appropriate technological and organisational measures.[3] Public authorities will have to survey and monitor the compliance systems established by private companies and establish appropriate discovery mechanisms which render ex-officio investigation and enforcement possible. The present paper aims to dive deep into the intricacies of compliance and enforcement of the AIA, focusing on the interaction between private companies and public authorities in their joint efforts to prevent the marketing of AI systems that are dangerous to people and society. Based on the analysis of the striking complexity of the AIA, we will then advocate the use of AI in compliance and enforcement: in handling of the communication between hundreds of competent national authorities and between the EU and the private compliance bodies, in generating and processing of information on the design and use of AI systems, in preparing

[2] Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act).

[3] Cl. Novelli, Ph. Hacker, J. Morley, J. Trondal, L. Floridi, A Robust Governance for the AI Act: AI Office, AI Board, Scientific Panel, and National Authorities (May 5, 2024). Available at SSRN: https://ssrn.com/abstract=4817755 or http://dx.doi.org/10.2139/ssrn.4817755

regulatory action in the EU as a whole. In this sense, the paper discusses both dimensions of AI systems, the risk they produce for the economy and society, and the opportunities AI systems offer to support new modes of compliance-checking and enforcement. The paper aims to highlight the opportunities AI offers in the compliance and enforcement of complex and multi-level compliance and enforcement legislation such as the AIA. It analyses the three main building blocs of AIA Enforcement The *first* building bloc consists of legally binding obligations imposed on private parties, including 'due diligence' obligations the parties must implement through appropriate measures (Chapter II Sections 2 and 3 AIA). The *second* building bloc is the integration of private parties, such as certification bodies, as de facto and, to some extent, de jure enforcers (Chapter II Section 4 AIA). The *third* building bloc is the implicit repositioning of public enforcement bodies. Outside and beyond their ex officio competence, such bodies are bound to survey and monitor private parties' compliance mechanisms and procedures.

This highly complex mechanism for compliance and enforcement is meant to trigger responsive regulation requiring cooperation and transborder communication.[4] The result is an intermingling of law-making and law enforcement through an ever-closer mutual dependence between private actors and public authorities. In the AIA, the EU combines and elaborates ideas developed in other domains: from the EU health/safety regulation, it takes the imposition of public policy duties on private parties or intermediaries;[5] from EU market surveillance regulation, the blueprint for organising enforcement,[6] from EU financial and accounting regulation, the promotion of due diligence obligations, and the delegation of responsibility for public goods to private parties.[7] The AIA pursues privatising law-making and law enforcement within a broadly defined regulatory frame. Thus, it is inspired by a regulatory model developed in 1985 with the New Approach on Standards and Technical Regulation. It paved the way for adopting the Single European Act in 1986, thereby gaining a quasi-constitutional status, now rediscussed within digital constitutionalism.[8]

1.1 Basic Legal Concepts

The paper is built on legal concepts, for which we rely on previous research on legislative and doctrinal definitions.[9] We use the term AI as defined in the AIA, which aligns with the OECD recommendation.[10] We include in the concept of compliance any activity through which the

[4] R. Baldwin and J. Black. Really Responsive Regulation, The Modern Law Review, vol. 71, no. 1, 2008, pp. 59–94. JSTOR, http://www.jstor.org/stable/25151178. Accessed 8 Sept. 2023.

[5] Ch. Joerges/ J. Falke/ H.-W. Micklitz/ G. Brüggemeier, Sicherheit von Konsumgütern und die Entwicklung der Europäischen Gemeinschaft, 1988 (English Version: European Product Safety, Internal Market Policy and the New Approach to Technical Harmonisation and Standards - Reissued), Hanse Law Review 2010 Special Issue, http://hanselawreview.eu/wp-content/uploads/2016/08/HanseLRVol6No02.pdf,

[6] Regulation (EU) 2019/1020 of the European Parliament and of the Council of 20 June 2019 on market surveillance and compliance of products and amending Directive 2004/42/EC and Regulations (EC) No 765/2008 and (EU) No 305/2011 *OJ L 169, 25.6.2019, p. 1–44*

[7] N. Molony, EU Securities and Financial Markets Regulation. Fourth Edition, 4th edition, OUP 2023.

[8] G. de Gregorio, Digital Constitutionalism in Europe, Reframing Rights and Powers in the Algorithmic Society, CUP 2022, https://doi.org/10.1017/9781009071215

[9] G Sartor/A. Reichmann, Algorithms and Regulation, in H.-W. Micklitz/ O. Pollicino/ A. Reichman/ A. Simoncini/ G. Sartor/ G. De Gregorio (eds.) Constitutional Challenges in the Algorithmic Society, Cambridge University Press, 2021, pp. 131

[10] OECD Recommendation of the Council on Artificial Intelligence, 2019 https://legalinstruments.oecd.org/en/instruments/oecd-legal-

addressees of legal duties —here both private actors and public bodies, in their role of developers and deployers of AI systems — try to abide by EU laws, either directly or through the involvement of third parties, such as certification bodies. In enforcement, we cover third-party activities to survey and monitor the addresses' compliance. Public bodies can exercise enforcement at the national and/or the European level. The public authorities may engage in investigations and take regulatory action, ranging from advice and recommendations to pecuniary and non-economic sanctions at various levels of severity.[11] However, enforcement through business and civil society organisations having been granted legal standing and the role of courts in collective actions are excluded. The emphasis lies on private-administrative law and administrative enforcement.[12]

Compliance is linked to accountability, a concept that has become prominent in EU technology regulation.[13] According to Article 6 (2) GDPR, accountability requires that 'the controller shall be responsible for, and be able to, demonstrate compliance' with the data protection principles. Accountability requires providers and deployers of AI systems to adopt a proactive approach (rather than merely responding to the competent authorities' indications). They must embrace suitable technical, institutional, procedural, and substantive measures and demonstrate compliance upon request to make the information accessible and comprehensible to stakeholders.

1.2 Subject of Analysis – the AIA

The EU digital policy legislation covers many projects[14] that raise similar compliance and enforcement issues. The AIA is among the most prominent and heavily debated by EU institutions, national democratic bodies, and civil society.[15] The European Parliament had proposed 771 – seven hundred seventy-one – amendments to the first proposal of the European Commission,[16] which had to be addressed in the interinstitutional agreements. Such amendments concerned crucial aspects of the AIA: the governance structure, including the role of the European Artificial Intelligence Office (EAIO), the European AI Board (EAIB), the AI Forum (AF) and the Scientific Panel (SP); the regulation of General Purpose AI; and the fundamental rights impact assessment for deployers of high-risk AI systems. Most of them made it into the finally adopted version.

0449#:~:text=The%20Recommendation%20aims%20to%20foster,human%20rights%20and%20democratic%20values

[11] See F. Cafaggi and H.-W. Micklitz, New Frontiers of Consumer Protection – the Interplay between Private and Public Enforcement, Elgar, 2009. In the introduction the authors are trying to systematize how enforcement could be structured along the line of public vs private and collective vs. individual,

[12] R. Vallejo, The Idea of Private Administrative Law, EUI PhD 2021.

[13] See M. Bovens, Analysing and assessing accountability: A conceptual framework. Euro-pean Law Journal, 2008, 13:447–468; N. Bamforth, P. Leyland (eds.) Accountability in the Contemporary Constitution, OUP 2013; P. Balboni, M. Taborda Barata, A. Botsi, A., and K. Francis, Accountability and enforcement aspects of the EU general data protection regulation - methodology for the creation of an effective compliance framework and a review of recent case law. The Indian Journal of Law and Technology, 2019, 102–239.

[14] The references provides for a comprehensive overview https://www.bruegel.org/sites/default/files/2023-07/Tables_Scott_Kai.pdf

[15] https://edri.org/wp-content/uploads/2023/07/Civil-society-AI-Act-trilogues-statement.pdf

[16] P9_TA(2023)0236 Artificial Intelligence Act Amendments adopted by the European Parliament on 14 June 2023 on the proposal for a regulation of the European Parliament and of the Council on laying down harmonized rules on Artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts (COM(2021)0206 – C9-0146/2021 – 2021/0106(COD, https://www.europarl.europa.eu/doceo/document/TA-9-2023-0236_EN.pdf. All references to the EP are taken from this document.

1.3 How the Argument is Built

We will develop our analysis in four steps.

The *first step* (Section 2) is to examine the jungle of applicable legal rules, breaking them down into (a) different types of regulation, from binding legal requirements to voluntary codes of practices, and (b) the range of responsibilities imposed on the various addressees. The *second step* (Section 3) pertains to examining the regulatory mechanisms in place for ensuring compliance: (a) self-certification as well as third-party certification and (b) private and public enforcement mechanisms at the national and European levels. The *third* step dives into the thicket of national and European authorities involved in compliance and enforcement, tied together through a net of information exchange and cooperation duties (Sections 4 and 5). Information exchange presupposes a shared understanding of what kind of information shall be generated, by whom and on what basis. The information generated under the AIA has to circulate within the net of national and European enforcement authorities and be made accessible within the boundaries of confidentiality. It is precisely here —in the overwhelming number of competent national and European authorities tasked with the generation, supply and circulation of information generating,— that we identify the potential for the use of AI, which might help to improve public enforcement through public agencies and to some extent private enforcement through non-governmental organisations.[17] *The fourth step* (Section 6) concerns the potential contribution of AI to compliance and enforcement: (a) how AI tools may facilitate compliance by private parties in reducing and mitigating risks, as well as in monitoring activities and preparing documentation and (b) how AI tools may support investigations and controls by public bodies or by third parties. The conclusions sum up the findings and define a set of recommendations on how the enforcement of the AIA could benefit from the use of AI.

2 Layers of Norms and Responsibilities

The EU AI policy is to be implemented through different layers of norms, which go much beyond the text of the AIA. These various layers of regulation must be considered for determining the responsibilities of all 'operators that develop, import or use AI systems' (recital 3 of AIA), as well as the extent to which regulatory actions of public authorities depend on cooperation with private parties. The AIA targets the scope and reach of responsibilities in light of the role and function of these operators in the AI supply chain. The result is a complex network of rules that is difficult to follow. Additional regulatory layers complementing the AIA further complicate the regulatory framework.

When it comes to compliance and enforcement, the AIA faces problems well-known from product safety regulation: shared responsibility between private parties (companies and notified bodies tasked with certification), the multitude of public enforcement authorities at the Member States and the EU level and their interaction (or lack of interaction) in the management of 'serious' incidents and/or 'risks' to health and safety. [18] The latter two will be analysed in sections 3 and 4.

[17] Private enforcement of AI requires a separate analysis.

[18] H.-W. Micklitz/ Th. Roethe/ St. Weatherill (eds.), Federalism and Responsibility – A study on Product Safety Law and Practice in the European Community, Graham & Trotmann, London, 1994

2.1 Layers of norms in the EU AI policy

The AIA operates with different layers of regulation, which range from binding and semi-binding rules to all sorts of non-binding recommendations, guidelines, and standards. One might distinguish rather crudely five layers: 1) mandatory binding legal requirements, 2) harmonised technical standards, 3) common specifications, 4) guidelines and recommendations from the European Commission, 5) codes of conduct, codes of practices, model terms and non-binding guidance on benchmarking and finally, 6) voluntary industry standards (national, European, international). Layers 2 to 5 come under co-regulation, the interaction between law and harmonised rules, reflected in staggered legal effects.

These effects are united by the philosophy that both the binding rules and the non-binding standards/guidelines/codes, etc., which govern AI systems should reflect the 'generally acknowledged state of the art' (Articles 8 (1), 25 (4), 50 (2), 96 (1)) AIA. For all systems, compliance demands are limited to measures that fit the generally acknowledged state of the art (Recital 64). Thus, the AIA refers to the least stringent level of protection against risks in the established tripartite hierarchy: 'state of the science' on top, 'state of science and technology' in the middle and 'generally acknowledged state of the art' at the bottom.[19] Interestingly, the classification behind the AIA categorisation did not raise resistance from the EP and civil society organisations. Finally, there is the voluntary set of rules: technical standards, codes of conduct, codes of practice, and compliance. Contrary to harmonised technical standards, there are no incentives for adopting codes of practice and codes of conduct.[20]

2.1.1 Binding Requirements

Binding legal requirements are ranked *first* in the AIA. The AIA establishes in Title III Chapter II multiple requirements for AI systems and obligations for providers, deployers and other third parties of AI systems. Such requirements and obligations are different according to the degree and kind of risks presented by each AI system: some AI systems are prohibited (Article 5 AIA); high-risk AI systems are subject to risk assessment and certification procedures (Chapter III); transparency requirements apply to certain kinds of AI systems, (Chapter IV), the remaining non-high-risk systems may be governed by voluntary codes of conduct (Article 95).

The requirements for AI systems may be understood as institutional safeguards on risk management, data governance, record keeping, transparency, human oversight, accuracy and robustness (Chapter III Section 2 Article 8 – 15 for high risks, Chapter IV Article 50 for certain risks, Chapter X Article 95 for minimal risk systems). These institutional requirements, i.e. the rules about the governance of entities providing and deploying AI systems, are translated in Chapter III Section 2 into obligations to be met by the relevant operators, an umbrella term which covers providers, product manufacturers, deployers, authorised representatives, importers or distributors (Article 3 (8)). The distinction between institutional - Section 2 - and personal requirements - Section 3 – applies to high-risk systems. On closer inspection, however, both dimensions are intertwined by directly connecting personal obligations to institutional requirements. The transparency requirements for specific non-high systems merge both dimensions in a single rule (Article 50). The European Commission enjoys far-reaching power to close potential gaps in the binding legal requirements

[19] There is a huge amount of literature on the classification and ranking of risks. Here it suffices to refer to the reference in (fn 5).

[20] P. Pałka, Data Management Law for the 2020s: The Lost Origins and the New Needs, 68 Buff. L. Rev. 559 (2020)

established by the AIA through delegated and implementing acts.[21] There is no leeway for the Member States once these requirements have been adopted due to the full harmonisation principle unless it is explicitly foreseen in the AIA.[22]

The AIA (Chapter III Section 2) contains a whole range of obligations on providers of high-risk AI systems concerning compliance with institutional requirements, quality management systems, technical documentation, logs, conformity assessment, registration duties, corrective actions, notification and information to public authorities and notified bodies (certification bodies) as well as co-operation with public authorities. These obligations are further concretised in the subsequent rules of Section 3 (Article 17-23). The institutional requirements and personal obligations concerning 'non-high-risk' AI systems are limited. Besides the transparency requirement, non-high-risk systems may be governed by voluntary codes of conduct (Article 50 (7)).

2.1.2 Harmonised Standards

The *second* layer is *harmonised European standards* to be elaborated by the European standardisation organisations (ESOs) with the participation of stakeholder organisations, mandated and co-financed by the European Commission (Article 40 AIA). References to harmonised European standards are widely spread in the AIA. Such standards form a constitutive bloc of the AI regulatory framework to shape the broadly worded binding legal requirements. Whilst their use is voluntary, the presumption of conformity in case of compliance (Article 32 AIA) sets a strong incentive to use the available harmonised standards so that they may be de facto-binding.[23] Harmonised technical standards – this is to be recalled – are meant to give shape to 'human-centric, secure, ethical and trustworthy AI' (the standard formula of AI legislation[24]) by concretising the obligations of Chapter III Sections 2 - 5 (conformity assessment, Article 43).

Harmonised standards have a double function: they set the specifications to be respected in developing high-risk AI systems and define the criteria for conformity with such specifications. Therefore, harmonised standards address institutional, procedural and substantive requirements *and* conformity assessment mechanisms to ensure compliance. To understand the co-regulatory nature of requirements imposed on the operators of AI systems, it does not suffice to look at the ambitious language of the binding legal requirements; the future harmonised technical standards must be included. This is not the place to analyse the tension between the harmonised technical standards which will have to be developed by the ESOs and the already existing AI ISO/IEC and IEEE standards. There are solid reasons to assume that the power of the EU is somewhat limited to adapting existing international AI standards to 'European values' – in particular, fundamental rights.[25]

---

[21] M. Chamon, The European Parliament and Delegated Legislation, An Institutional Balance Perspective, 2021

[22] N. Helberger/O. Lynskey/H.-W. Micklitz/P. Rott/M. Sax/J. Strycharz, EU Consumer Protection 2.0: Structural asymmetries in digital consumer markets, A joint report from research conducted under the EUCP2.0 project, BEUC, March 2021, 207 pages; https://www.beuc.eu/sites/default/files/publications/beuc-x-2021-018_eu_consumer_protection_2.0.pdf

[23] ECJ Case C-171/11 Fra.bo ECLI:EU:C:2012:453

[24] H.-W. Micklitz, The Role of Standards in Future AIA: A Consumer Perspective, Report commissioned by BEUC and ANEC, July 2023, 194 pages; https://www.beuc.eu/sites/default/files/publications/BEUC-X-2023-096_The_Role_of_Standards_in_Future_EU_Digital_Policy_Legislation.pdf

[25] For a more detailed analysis, H.-W. Micklitz, The External Dimension of AI Standards in the EU Digital Policy Legislation, Texas Journal for International Law, (59) 2024, forthcoming

2.1.3 Common Specifications

The Commission may enact common specifications to fill gaps left through delayed or insufficient harmonised technical standards (Article 41 AIA). The regulatory power shall serve as a stick behind the door to push the ESOs into action. So far, the European Commission has limited experience using common specifications as a substitute for harmonised standards in market sectors. Where they were used, it turned out that they remained at a rather general level. Their application required additional detailed specifications by the actors involved in the standardisation of products. The European Commission has established the European Centre for Algorithmic Transparency (ECAT) within the JRC to be better prepared. ECAT will probably play a key role in elaborating AI standards.[26]

The ESOs are critical of the residual power of the European Commission. They fear a downgrading of private through political standardisation, as the European Commission might use its power to substitute private standards through legally binding common specifications. Civil society groups, on the contrary, tend to support public regulatory powers, as they believe that public interests are better off in the hands of public authorities, an assumption which seems highly debatable.[27] Knowledge has been relocated from the administration to industry since industrialisation began in the late 19th century.[28] In the 21st century, the relevant AI expert knowledge is only partly represented in the national and European Parliaments or the executive institutions of the EU and the Member States. There is no alternative to co-regulation as the digital industry holds most of the AI expert knowledge. However, stakeholders' roles and functions in elaborating harmonised AI standards are desperately needed. The legal deficits of Reg. 1025/2012 are not and cannot be compensated by the Advisory Forum (AF) and the Scientific Panel (SP).[29]

2.1.4 Non-Binding Guidelines, Benchmarking

The *fourth* layer includes non-binding guidelines, recommendations, and benchmarking the AIA promotes. Under pressure from the EP, the final version of the AIA is more outspoken about the role and function of soft law. However, the EP failed to establish the EAIO as an independent unit. The power to develop non-binding guidelines and recommendations lies in the hands of the Commission alone (Article 96 AIA).

The Commission is granted widely conceived powers to adopt guidelines for the implementation of the institutional requirements of high-risk AI systems (Articles 8-15 and 25 (supply chains)), prohibited practices and the practical implementation of transparency obligations (Article 96 (1) a)-f)), as well as regulating non-high-risk systems through codes of conduct (Article 95).

In cooperation with relevant stakeholders and organisations such as metrology and benchmarking authorities, the Commission shall encourage the development of benchmarks

[26] https://algorithmic-transparency.ec.europa.eu/index_en

[27] H Schepel The Constitution of Private Governance (Oxford: Hart Publishing, 2005), with a recent revival, O Kanevsk AIA-P, The Law and Practice of Global ICT Standardization, CUP 2023, M Gérardy, The 'Standards Effects': The Public Instrumentalisation of technical standards in EU law, PhD University of Luxemburg, 202; see also P. Delimatsis (ed), The Law, Economics and Politics of International Standardisation, CUP 2015

[28] K-H Ladeur, The Evolution of General Administrative Law and the Emergence of Postmodern Administrative Law' (2011) Comparative Research in Law & Political Economy. Research Paper No 16

[29] On stakeholder participation Micklitz, Role of Standards, loc. cit. pp. 182, on the tasks and functions of the AF and the SP, under 4.2. in this paper.

and measurement methodologies (Article 15 (2) AIA). Under the EP proposal, the expected levels of performance of the AI system should be specified by the AI Office, together with national and international metrology and benchmarking authorities, who should address not only the technical aspects of measuring appropriate levels of accuracy and robustness in Article 15 but also provide cost-effective guidance and capabilities to measure and benchmark aspects of AI systems and AI components. For a more detailed and comparative analysis, it is essential to consider Article 15 of the Cybersecurity Act and the role of ENISA (European Agency for Cyber Security).[30]

The AI Office supports the Commission in adopting soft law measures and assists in preparing guidance and guidelines.[31] and developing a template for a questionnaire of the fundamental rights impact assessment through an automated tool (Article 27 (5)). The EAIB may adopt guidelines and opinions upon request of the European Commission or on its own initiative on a whole series of issues listed in Article 66 (e) (i)-(vii). The Advisory Forum (Article 67 (8)) may take similar measures only upon request of the Board or the Commission.

2.1.5 Codes of Conduct, Codes of Practice and Model Terms

The *fifth* layer consists of codes that the operators involved in making, deploying, distributing, importing, or authorising AI systems or authorising their use are developing out of their own motion. The AIA distinguishes between codes of conduct and codes of practice without defining them. Codes of conduct deal with bilateral b2b and eventually b2c relations, whereas codes of practice address professional behaviour.

Under Article 95, the EAIO and the Member States shall 'encourage' and/or 'facilitate' the development of codes of conduct for 'non-high risk' AI systems bound to minimum requirements. The Commission is not directly involved but indirectly participating through the AI Office. Article 95 distinguishes codes for 'other than high-risk systems' para 1 from codes for 'all AI systems' para 2. For 'other types', the AIA lays down requirements. Codes for all AI systems shall include 'some or all of Chapter III Section 2 requirements (Articles 8-15)' – but not Section 3, which concerns the due diligence obligations – 'taking into account best practices.

Does this mean that codes of conduct which do not meet Article 95 (1) requirements are prohibited? The Commission should encourage and promote the voluntary application of the requirements in Section II when meaningful. Still, it has no obligation to encourage and promote the requirements in Section III. Is it a promotion without a compliance check? Similar questions arise concerning codes for all AI systems. Article 95 (2) supports the elaboration of codes of conduct provided they define clear objectives and key performance indicators to measure such objectives, such as ethical guidelines for trustworthy AI, environmental sustainability, AI literacy, and inclusive and diverse design. They have to be elaborated with the participation of stakeholders, and they shall have no negative impact on vulnerable persons. The list is not exhaustive. That is why the consumer acquis could easily be integrated.

[30] F. Casarosa, Correction to: Cybersecurity certification of Artificial Intelligence: a missed opportunity to coordinate between the Artificial Intelligence Act and the Cybersecurity Act. Int. Cybersecur. Law Rev. 3, 245 (2022). https://doi.org/10.1365/s43439-022-00053-y.

[31] COMMISSION DECISION of 24.1.2024 establishing the European Artificial Intelligence Office, Brussels, 24.1.2024 C(2024) 390 final (Article 3 (2) (c).

Article 95 (3) clarifies that neither the provider, the deployer, nor the respective business organisations are obliged to develop such codes of conduct. They '*may*' do so or not. Like Article 6 of the Unfair Commercial Practices Directive 2005/29, the AIA does not set incentives for operators ready to engage in such a demanding exercise. The AIA does not clarify who oversees the compliance of codes encouraged and/or facilitated by public authorities. The European Data Protection Board (EDPB) is composed of the representatives of the Member States.[32] It is empowered to write opinions on the application of codes of conduct out of its own motion (Article 66 e) (ii) AIA). Such exercise, if taken seriously, presupposes surveillance. All in all, the purpose of voluntary codes of conduct remains opaque. They are semi-binding in that the AIA recognises them and even sets minimum standards but does not put precise mechanisms in place to ensure compliance and enforcement.

The AIA lays down a special regulatory regime for codes of practice on general-purpose AI (Article 56) and for transparency of specific AI systems (Article 50 (7)). In general-purpose AI, the AI Office and the AI Board operate in tandem. Para 2 defines a set of requirements which have to be respected; Para 3 obliges the AI Office and AI Board to monitor and evaluate the achievement of the objectives laid down in Para 2 and grants the European Commission the power to approve a code of practice through an implementing act to 'give it general validity in the Union', provided there is no code available by 2nd August 2025, or the proposed code does not respect the requirements para 9. Article 50 (7) leaves the initiative to develop a code of practice for detecting and labelling artificially generated or manipulated content in the hands of the AI Office alone. The Commission may adopt implementing acts to approve those codes or develop and implement 'common rules' if the code is inadequate.

Standard terms, whether elaborated by business organisations or companies, are the classical tool to transpose due diligence obligations, such as those in Chapter III Section 3. Many of the due diligence obligations for providers of high-risk AI systems in Chapter III Section 2 define contractual relations benchmarks.[33] The same is true about the transparency requirements for specific AI systems (Article 50). Its meaning and importance reach far beyond the interpretation in the UCTD (Unfair Contract Terms Directive and the UCPD (Unfair Commercial Practices Directive), if not the GDPR (General Data Protection Regulation).[34] The EP had proposed the introduction of Article 28 (a) into the AIA. Thereunder, unfair terms that an enterprise unilaterally imposes on SMEs or start-ups should not be binding'. Such a regulatory approach would have fitted well into EU regulations on the fairness of standard terms in b2b relations, value chains in the food sector, platform regulation,[35] and most comprehensively in the Data Act.[36] However, the ambitious proposal failed. All that survived is the possibility for the AI Office to develop and recommend 'voluntary' model terms between providers of high-risk AI systems and third parties which supply tools, services or processes that are used for or integrated into high-risk AI systems (Article 25 (4)). Such model terms shall consider possible contractual agreements applicable in specific sectors or business cases. The vague formula is

[32] On a more detailed analysis of the governance structure see under 4.

[33] M. Ebers, M Ebers, Standardizing AI – The Case of the European Commission's Proposal for an Artificial Intelligence Act (August 6, 2021). The Cambridge Handbook of Artificial Intelligence: Global Perspectives on Law and Ethics, Available at SSRN: https://ssrn.com/abstract=3900378 or http://dx.doi.org/10.2139/ssrn.3900378

[34] A. Jablonowska/G. Tagiuri, Rescuing transparency in the digital economy: in search of a common notion in EU consumer and data protection law, Yearbook of European Law, Volume 42, 2023, 347–387.

[35] Regulation (EU) 2019/1150 of the European Parliament and of the Council of 20 June 2019 on promoting fairness and transparency for business users of online intermediation services *OJ L 186, 11.7.2019, p. 57–79.*

[36] Reg. 2023/2854 Article 13.

far from the mandatory minimum standards imposed on elaborating codes of conduct and practice.

2.2 Responsibilities

The second domain to be analysed is the operators' responsibilities, which differ considerably depending on the risks of the AI system. There must be a correlation between the type of risks, the breadth and depth of responsibilities, and the layers of regulation. The higher the risk, the more encompassing the responsibilities and the more detailed the layers of regulation—and vice versa.

2.2.1 High-Risk Systems and General-Purpose AI (GPAI)

The rules on responsibilities in the AIA focus on high-risk systems, to which the most significant section of the Act (Chapter III) is dedicated. The European Commission is empowered to amend the list of high-risk systems by delegated acts (Article 7 (1) and (3) in connection with Annex III). Also, the fundamental rights impact assessment deals only with high-risk systems (Article 27). The provider and deployer of 'certain AI systems' are freed from specific obligations but must respect the transparency requirements (Article 50).

The risk-based approach requires that the risks of an AI system must be assessed in advance to make it available for commercial purposes. This does not fit well with the 'general purpose AI (GPAI)' as defined in Article 3 (65). These AI systems have a wide range of possible uses, both intended and unintended by the developer, and they can be applied to many different tasks in various fields, often without substantial modification and fine-tuning. The primary examples are the LLMs (large language models), such as the GPT family. These systems have been trained on massive datasets of documents. They can produce multiple kinds of new content that fit the requests and context provided by users: abstracts, summaries, essays, stories, articles, letters, translations, memos, interactive responses, computer programs, images, films, etc. The same LLMs can be used for different purposes and applications: to generate helpful information or even disinformation, to enable computer systems to support users meaningfully or rather to mislead and manipulate them, to help students produce better texts, or rather cheat with examiners and teachers, etc.

These examples demonstrate that most risks are related to a particular use of an LLM rather than the LLM itself. Measures meant to prevent and reduce risks have to deal with application-specific risks to be adopted by downstream actors, i.e., either by the providers of the AI system which will embed the LLM into a specific application (e.g., an interactive bot) or by the deployer who will use the LLM in a particular task. The AIA regulates general-purpose AI in Chapter V Articles 51-56, a verbose set of provisions that, besides requiring measures to prevent and mitigate the general risks related to the model (e.g., controlling the quality of the training sets), calls for support to the developers or deployers in preventing and mitigating application-specific risks.

2.2.2 Addressees and Layers of Regulation

The AIA breaks down the potential addresses of high-risk AI systems into fine-grained categories. The key actors are the providers and deployers of an AI system. The final version substituted the term 'user' (occurring in the EC proposal) with the term 'deployer'. Article 3 (4) clarifies this category's commercial dimension, which does not include non-commercial users such as consumers. Whilst providers and deployers seem to be the prime addressee of the AIA, the rules address manufacturers, as the AIA is meant to cover AI components built

into products, providers of general-purpose AI systems, providers of regulatory sandboxes Chapter V Article 57-62, authorised representatives of companies located outside the EU (Article 3 (5)), importers Article 3 (6), distributors Article 3 (7) and other actors in the AI value chains (Article 25.

The different categories are, by and large, taken from product safety and product liability legislation, except for the authorised representative. There is a tendency in the AIA to broaden the scope of obligations beyond providers and deployers to businesses that are not at the centre stage of the regulated activity. However, this does not mean that these 'peripheral' businesses are on an equal footing with the central players. The complexity is further aggravated by the insertion of SMEs and startups, which benefit from a reduction, even when they are providers of high-risk AI systems, e.g., on technical documentation (Article 11 (1)). This exception met resistance from civil society organisations.[37] Further rules meant to favour SMEs concern regulatory sandboxes (Article 57 (9)), start-ups (Article 62), and even the development of codes of conduct (Article 95 (4)).

The degree to which the various operators are submitted to the different types of regulation differs widely. The horizontal categorisation in the table below breaks down the other operators, from providers to deployers, manufacturers, authorised representatives, importers, distributors, operators in the value chain, GPAI providers and providers of sandboxes. The vertical categorisation distinguishes between the different layers of the norms and the degree to which they mainly address one particular or all operators. The table only mentions explicit references in the AIA. An empty cell means that the AIA does not provide a certain kind of guidance to the category of operators. This does not exclude, however, that these operators may become, at some point, the direct addressee of particular regulatory tools such as harmonised standards, typical specifications, guidelines – recommendations – codes or benchmarking. For instance, harmonised standards are only explicitly referred to regarding providers' obligations, although they may address all other operators, depending on the subject matter.

The relationship between different levels of risk and other sets of obligations through various layers of regulation has to be considered. The AIA refers to harmonised standards in Article 40, common specifications in Article 41, codes of conduct in Article 95, guidelines and recommendations in Article 96, and benchmarking in Article 15 (2) without specifying the potential addressees. These tools can address whatever operator as long as the content remains within the scope of due diligence obligations. There are exceptions. Sometimes, the AIA points to tools such as harmonised standards, common specifications, guidelines or benchmarking and addresses particular operators.

The obligations of the providers and the deployers are laid down in Chapter III, Section 3. Article 16 a) obliges the provider to ensure high-risk AI systems comply with Section 2 requirements. Section 2 and 3 requirements are deeply interconnected through binding legal requirements and harmonised standards. Article 40 (2) obliges the European Commission to issue a standardisation request covering all Chapter III Section 2 requirements, a mandate the

[37] BEUC AI AND GENERATIVE AI: TRILOGUE NEGOTIATIONS FOR THE AI ACT, BEUC recommendations, 25.7.2023, https://www.beuc.eu/sites/default/files/publications/BEUC-X-2023-101_AI_and_Generative_AI_trilogue_negotiations_for_the_AI_Act.pdf

Commission had already fulfilled before adopting the AIA.[38] The providers are not obliged to use harmonised standards, but compliance offers access to the Internal Market (Article 43 (1)). Somewhat surprisingly, the AIA contains two explicit references to harmonised standards on quality management systems in Article 17 (1) e) and (4) and in Annex IV (7) on technical documentation in connection with Article 11, references which are not needed due to Article 16 (a) which refers to Section 2. The European Commission is empowered to amend Annex IV, considering technical progress through a delegated act, thereby changing the requirements that harmonised standards must meet. The common specifications (Article 41) do not address particular operators. Benchmarking (Article 15 (2)) has providers of AI systems as prime addresses due to Article 16 a). Guidelines, recommendations and codes of conduct have no specific addressees. However, they implicitly address the provider and the deployer of AI systems. Article 50 (7) explicitly addresses the provider of AI systems carrying certain risks. The headline of Chapter IV deals with 'codes of practice' on the detection and labelling of artificially generated content for providers and deployers.

There are many particularities concerning the addressees of the different layers of the norms. The AIA does not make deployers the direct addressees of harmonised standards, except for potential standards on human oversight and logs (Article 26 (2) and (6)). This raises the question of why harmonised standards on risk management, data governance, technical documentation, recording keeping, accuracy, robustness and cybersecurity are not addressing deployers directly. The concept of a provider is extended, for instance, to cover manufacturers of AI safety components of products covered by the Union harmonisation legislation listed in Section A of Annex I (Article 25 (3)) as well as distributors, importers, deployers or other third-parties who put their name or trademark on a high-risk AI system, modify such a system, or deviate from its intended purpose (Article 25 (1)). Finally, the AIA introduces approved codes of practices as quasi-equivalent to harmonised standards to speed up the development of common technical requirements. Implementing acts by the European Commission pursue the same function for providers of regulatory sandboxes to avoid fragmentation (Article 58 (1) and (2)). GPAI providers are not the sole addressee of model terms (Article 25 (4)). The Commission may adopt an implementing act on the terms and conditions of the participants of regulatory sandboxes.

	Provider Article 16-21	Deployer Article 26	Manufacturer Article 25 (3)	Authorised Representative Article 22	Importer Article 23	Distributor Article 24	Operators in AI Value Chain Article 25	Provider GPAI (Articles 53-55	Provider of Sandbox Article 57-63
Harmonised standards	Article 16 (a) + Article 40 (2) Chapter III Article 17 (1) e) and (4) quality management Article 43 + Annex V (6) Conformity assessment Article 16 a) Annex IV (7) technical documentation	Article 26 (3) without prejudice to other obligations (2) human oversight about Article 14? (6) keep the logs about Article 19	Article 25 (3) of safety components = provider	Article 22 (3) a) EU declaration of conformity	Article 23 (1) a) relevant conformity assessment Article 43	Article 24 (1) EU declaration of conformity	Article 25 (1) = when provider under a) same or trademark +b) substantial modification Articles 16/40	Article 55 (2) approved code of practices (substitute)	Article 58 (1) + (2) a) -i) implementing act on common principles
Benchmarking	Articles 16 (1) + 15 (2) on								

[38] C(2023)3215 – Standardisation request M/593 COMMISSION IMPLEMENTING DECISION of 22.5.2023 on a standardisation request to the European Committee for Standardisation and the European Committee for Electrotechnical Standardisation in support of Union policy on artificial intelligence

	accuracy, robustness and cybersecurity								
Guidelines and Recommendations	Article 6 (5) classification + Article 80 (3) providers of non-high risks Article 96 guidelines (no limitation)	Article 96 guidelines	Article 96 guidelines	Article 96 guidelines	Article 96 guidelines	Article 96 guidelines	Article 96 guidelines		
Codes of conduct	Article 95 (1) other than high risk	Article 95 (2), including deployers	Article 95 (1) other than high risk	Article 95 (1) other than high risk	Article 95 (1) other than high risk	Article 95 (1) other than high risk	Article 95 (1) other than high risk		
Codes of practices	Article 50 (7) detection and labelling of artificially generated content	Article 50 (7) detection and labelling of artificially generated content						Article 56 codes of practices	
Model terms of contract							Article 25 (4)		

2.2.3 Addressees and Range of Obligations

The EC proposal put providers of high-risk AI systems at the centre stage. However, under pressure from the EP, the responsibilities of deployers were upgraded considerably through the extension of particular obligations and, most prominently, by making deployers responsible actors for the fundamental rights impact assessment of specific high-risk AI systems (Article 27). In this sense, one might speak of a double-headed approach to the design of responsibilities.

The content and the scope of the various obligations imposed on the different kinds of operators reflect the position of these operators in the AI value chain, following EU product safety regulations. The degree to which operators other than the provider – here to be equated with the manufacturer – shall be submitted to the same set of obligations as the provider/manufacturer was and has been subject to a battlefield ever since. The finetuning of public law responsibilities of the various operators does not necessarily imply private liabilities. There is a considerable mismatch between EU product safety and risk-based EU digital framework legislation guided by public law responsibilities and product liability regulation under Directive 85/374/EC, as revised by Directive 2024/2853.[39] The recent revision of the Product Liability Directive and the pending AI Liability Act do not address the potential liabilities of the different operators upfront.[40] The scope and reach of operators' private liability remain governed by national tort law alongside EU product liability. This can lead to friction in cases where national product liability law rules are stricter than the EU Directive.[41]

The following two tables aim at demonstrating the different obligations – broken down horizontally along the due diligence obligations under Chapter III Section 3: the obligations

[39] Directive (EU) 2024/2853 of the European Parliament and of the Council of 23 October 2024 on liability for defective products and repealing Council Directive 85/374/EEC (Text with EEA relevance)OJ L, 2024/2853, 18.11.2024, ELI: http://data.europa.eu/eli/dir/2024/2853/oj

[40] Ph. Hacker, The European AI Liability Directives _ Critique of a Half-Hearted Approach and Lessons for the Future https://arxiv.org/abs/2211.13960; Piotr Machnikowski (ed), European Product Liability: An Analysis of the State of the Article in the Era of New Technologies (Intersentia, Cambridge 2016)

[41] N. Reich/H.-W. Micklitz/P. Rott/K. Tonner, European Consumer Law, Intersentia 2014, chapter 6, 239 pp.

laid down in Art 16, which refer to Section 2; those on quality management, documentation keeping, conformity assessment, automatically generated logs, corrective actions and duty to inform each other and the authorities, co-operation obligations and post-market monitoring duties. These sets of obligations are adjusted vertically to the various economic operators. The special treatment of general-purpose AI and regulatory sandboxes justifies separating the two. The two tables amply demonstrate the enormous complexity of the AIA, which results from the attempt to differentiate between the operators and their respective obligations. We do not aim to analyse the due diligence obligations in detail. For our purposes, it is sufficient to highlight the significant differences and the most striking particularities between the various operators.

The overall rationale of the obligations is taken from product liability law. The provider is the key addressee, and as has been noted, other operators taking an active role are assimilated into providers (Article 25). The deployer is the party that obtains an AI system and transforms it into a use case. Deployers will often be SMEs not developing the AI system but getting it from the GAFAs or other big companies with the resources to build it. The AIA puts a heavy burden on deployers in that they are in charge of the fundamental rights impact assessment (Article 27). This makes sense as the risks of an AI system depend on the context in which they are applied. However, it may have the effect of 'kicking the can down the road' so that the often-small deployers bear the burden and the risk of being held liable.[42] The market for AI systems is global. Providers and deployers are not necessarily located where they are selling their products. The AIA reacts through an alignment of 'distributors' and 'importers' obligations and the introduction of the authorised representative. The overall idea is the same. Those distributors, importers or authorised representatives doing business across the border must ensure that the provider standing behind the AI systems has taken the necessary precautionary measures to ensure compliance with the AIA.

Three of the seven obligations in the table deserve to be analysed in more detail. The paper aims to demonstrate AI's potential to promote compliance, conformity and enforcement: technical documentation keeping, logs generation, and post-market monitoring. The AIA seeks to ensure that the technical documentation is available for 10 years after the product has been brought on the market and that the responsibility for its existence lies equally in the hands of the manufacturer, the authorised representative, the importer, the distributor and those operating in the AI value chain. Automatically generated logs have to be kept by the providers, deployers, and authorised representatives, who have to do so provided the log is under their control (Article 26 (6) and Article 22 (3) c)). Risks may occur after the AI systems have been put onto the market. The AIA lays down post-market monitoring duties on the provider (Article 17 h)), and logs have to be designed to facilitate post-market monitoring (Article 12 (2)).

Post-market monitoring duties must be kept separate from information and notification duties. Theoretically, every economic operator may identify or know information on potential risks independent of their role and function as providers, deployers or any other distribution or value chain member. The AIA distinguishes information and notification of risks. The information generated within the distribution systems, within the AI supply chain, or simply

[42] J Laux, S Wachter and B Mittelstadt, 'Three Pathways for Standardisation and Ethical Disclosure by Default under the European Union Artificial Intelligence Act' (February 20, 2023). Available at SSRN: https://ssrn.com/abstract=4365079 or http://dx.doi.org/10.2139/ssrn.4365079.

in contractual relations, first and foremost, remains in the hands of the operators, who have to inform each other. Public authorities may gain access to the information only upon request or reasoned request. The situation changes dramatically once the operators know about a serious incident. The addressee is mainly the provider but eventually also the deployer (Article 26 (5)). The complex interaction within distribution systems and supply chains, as well as between them and the public authorities, let alone the distinction between risks and incidents and between serious and non-serious, will be discussed under five below.

	Obligations Article 16	Quality Management Article 17	Documentation keeping Article 18	Conformity Assessment (no longer a separate article)[43]	Automatically generated logs Article 19	Corrective action and duty of information (Article 20	Co-operation with authorities (Article 21	Post-market monitoring duties Article 72
Provider (Article 16-23, Manufacturer Article 25 (3), treated as providers	Article 16 Compliance with Section 2 (a),	Article 17 h) post-market monitoring Article 72 Article 73 (9) Reporting of serious incidents	Article 18 (1) documentation keeping for a period ending 10 years after being placed on the market (a)-(e)	Article 16 (f) conformity assessment, (k) upon reasoned request demonstrates conformity.	Article 19 (1) kept for at least six months Article 12 (2) facilitates post-market monitoring	Article 20 (1) information on non-conformity to the distributor, deployer authorised representative, importer Article 20 (2) investigates risks and takes corrective action, informs authorities and notified bodies	Article 21 on reasoned request (1) Chapter 2 compliance info and doc + (2) logs Article 12 (1) under control	Article 72 (3) post-market monitoring plan by providers Reporting of serious incidents Article 73
Authorised representative Article 22	specified mandate Article 22 (2), copy to authority (3) – tasks (3) a)-e) + ensuring compliance with AIA Article 22 (3) after e)		Article 22 (3) b) keeps the contract details of the provider for 10 years after it is placed on the market.	Article 22 (3) b) declaration of conformity c) reasoned request demonstrate compliance.	Article 22 (3) c) reasoned request assess to logs Article 12		Article 22 (3) d) on a reasoned request for co-operation	
Importers (Article 23	Ensure conformity (1) a)-d9 (6) Make available upon reasoned request		Article 23 (5) keeps 10 years after placing on the market copy of the certificate (6) reasoned request section 2 requirements	Article 23 (1) a) declaration of conformity (6) reasoned request demonstrates compliance		Article 23 (2) Inform the provider, authorised representative, and authorities.	Article 23 (7) shall cooperate.	
Distributor (Article 24			Article 24 (1) verifies CE marking, compliance with Article 16 b)-c) Article 24 (5), upon a reasoned request for compliance	Article 24 (1) Verify compliance declaration of conformity Article 24 (5), upon a reasoned request for compliance		(Article 24 (4) Information based on possession to provider, importer and authorities	Article 24 (6) shall cooperate.	
Responsibilities along the AI value chain Article 25	Article 23 (1) a)+b) +c) Any distributor, importer, deployer or third party provider, if renamed, if substantial modification or modify purpose		Article 25 (2) former provider cooperates + information available to the new provider (3) provider and third-party written agreement					
Deployers/ users (Article 26	Article 26 (1) ensures compliance with	Article 26 (2) duties on human oversight			Article 26 (6) keeps logs for at least 6 months.	Article 26 (5) informs providers by Article 72, notify serious incidents to	Article 26 (12) shall cooperate	Article 26 (5) notify serious incidents to providers,

[43] In the original EC proposal conformity assessment obligations were concretised in Article 19. In the finally adopted version they are integrated into the type of duties imposed on the various operators.

	Article 26 (3) and (6)					providers, importers, distributors, national authorities (7) inform workers of representation before AI is put into service (9) inform data protection impact assessment (11) Inform natural persons on high-risk AI	with relevant authorities	importers, national authorities Article 72 (2) collects data 'which may be provided by the deployer' Article 25 (5) eventually Article 73

GPAI providers and providers of regulatory sandboxes are regulated in Chapters V and VI. They are legally defined in Articles 3 (63) and (55). Neither fit easily into the co-regulatory approach, the interaction between binding mandatory requirements and harmonised technical standards, nor into the requirements on high-risk AI systems, their classification, their requirements and their obligations (Sections 1 to 3). The regulations on general-purpose AI try to reproduce the dichotomy between high-risk and non-high-risk systems by distinguishing between general-purpose AI systems with systemic risks and non-systemic risks, with differing due diligence obligations, Artt 53-55; for instance, the difference between high-risk AI systems and general-purpose AI systems with systemic risks. A further issue pertains to whether systemic risks may benefit from 'systemic risk' in financial services. General-purpose AI providers may rely on approved codes of practice if no harmonised technical standards exist. They may benefit from model terms elaborated by the AI office.

Regulatory sandboxes shall facilitate innovation. Member States shall ensure the establishment of at least one regulatory sandbox by 2nd August 2026—which means one year after it enters into force. The rules under which the regulatory sandboxes may operate are fully harmonised. They can be concretised through potential implementing acts (Article 58). One might wonder whether such a heavy-handed approach will produce the desired results or whether providers of AI systems may seek a safe harbour through the SME claim, allowing them to downgrade the obligations under Chapter III.

	Obligations Article 16	Quality Management Article 17	Documentation keeping Article 18	Conformity Assessment (no longer a separate article)[44]	Automatically generated logs Article 19	Corrective action and duty of information (Article 20	Co-operation with authorities (Article 21	Post-market monitoring duties Article 72 and reporting serious incidents Article 73
Provider of GPAI Article 51 Article 54 authorised representative	Article 53 (1) obligations for GPAI a) - d) Article 55 Additional obligations for GPAI with systemic risks Both may rely on codes of practice as of Article 56		Article 53 (1) a) +b) draw up and keep up-to-date technical documentation (Annex XI as a minimum)	Article 53 (1) c) compliance with copyrights		Article 53 (1) b) makes available information and documentation to providers of AI systems.	Article 51 (3) shall cooperate with the Commission and the national authorities.	Article 52 (1) notification to the European Commission Article 55 (1) c) report on serious incidents
Provider of Sandboxes Article 57	Article 57 (5) provides for a controlled environment +	Article 59 Protection of	Article 59 Protection of Personal Data	Article 58 implementing acts on procedures of	Article 59 (1) h) logs of processing personal data		Article 57 (13) is designed for cross-	Article 58 Maybe through

[44] In the original proposal conformity assessment obligations were concretised in Article 19. In the finally adopted version they are integrated into the type of duties imposed on the various operators.

Article 62 particular measures for providers, deployers, SMEs and start-ups Article 63 derogations of microenterprises	(9) a)-e) objective Article 58 implementing acts as enlisted	personal data) Article 60 Testing Article 61 informed consent to participate	Article 60 Testing Article 61 informed consent to participate	application, terms and conditions	have to be kept for the duration of the sandbox		border cooperation Article 58 (2) (f) involvement of notified bodies, standardisation bodies, companies, experimentation facilities, centres of excellence, individual researchers	implementing acts At. 60 (7) Report on serious incidents in real-world testing

3 Compliance and Conformity, Self and Third-Party Certification

The AIA does not define compliance but uses the term repeatedly. To understand what is behind it, one must go back to the Market Surveillance Regulation 2019/1020, which defines 'non-compliance' in Article 2 (7) as *'any failure to comply with any requirement under the Union harmonisation legislation or this Regulation'*. Chapter III Section 3 Article 8 (1) links compliance to the legal requirements of high-risk AI systems, laid down and specified in Articles 8-15, taking into account their 'intended purposes' and the 'generally acknowledged state of the art of AI and AI-related technologies', or in the case of products with AI components enlisted in Annex I compliance with 'all the applicable requirements' formulated in 'their respective Union harmonisation legislation'. Whilst Articles 8-15 provide for a long list of requirements ranging from risk management to accuracy, robustness and cyber security, the reference to the 'intended purpose' and the 'generally acknowledged state of the art' set a cross-cutting benchmark independent of the responsibilities of the various operators. Both are meant to limit the legal requirements – intended use excludes foreseeable use, excluding the expectations of end-users (consumers) and the potential use of AI systems by deployers beyond 'intended use'. If deployers go beyond the intended use, they run the risk of being regarded as providers (Article 25 (1) c)). The reference to the generally acknowledged state of the art excludes high standards of knowledge in developing an AI system.[45]

Conformity is another complex category. Article 3 (20) tells us that conformity assessment '*means demonstrating whether the requirements set out in Chapter II, Section 2 relating to a high-risk AI system have been fulfilled'*. Emphasis has to be put on 'demonstrating'. The operators in charge have to take measures to show what they have done to implement the legal requirements, either by themselves through self-certification or by a third party, the conformity assessment bodies (Article 3 (21)), handling testing, certification and inspection.

3.1 Self-certification

Self-organised compliance activities can be broken down into two strands: (1) compliance with substantive law and (2) demonstrating compliance. The *first* is rather obvious: operators have to comply with the AIA. An element of newness results from the co-regulatory approach, where operators must consider the many different layers of regulation. Studying binding legal rules is just a starter. The key benchmarks are found in delegated and implementing acts, harmonised standards and common specifications, EC guidelines and recommendations, codes of conduct and codes of practice that serve explanatory purposes. Ideally, all non-binding rules should find their way into harmonised standards.

[45] On the three levels of risk regulation, see references under 5.

The second innovative *strand* intermingles compliance with accountability. Accountability has a double meaning: operators are responsible for adopting technical and organisational measures that ensure compliance and for demonstrating that such measures are in place. Here, the due diligence obligations of the AIA may be connected to private governance. Therefore, the AIA (Chapter III Sections 2 and 3) requires that the operators of AI systems must adopt appropriate safeguards and *institutional*, *substantive* and *procedural* measures.[46] (on risk management and data governance, accuracy and robustness, human oversight, etc.); and engage in record keeping, transparency, and documentation preparation. The set of obligations is specified in the various Annexes, emphasising technical documentation (IV post-market control included), conformity assessment (Annex V-VII) and registration of high-risk AI systems. These obligations require skills and resources, which challenge SMEs and startups. They do not fit easily into established mechanisms of judicial control in the UCTD, the UCPD and the GDPR. These three instruments deal with a particular form of the *substance* – standard terms, commercial practices and data privacy policies. The CJEU has enlarged unfair terms control to the *procedural* dimension, missing in the UCPD and the GDPR. The institutional dimension is unchartered territory. The legal doctrinal question is whether the consumer and data protection acquis can be extended to include the *institutional* dimension.[47]

Providers obtain the EC label based on the conformity assessment, which gives them access to the Internal Market. The EC label conveys to the outside world that the responsible operator has taken all the necessary institutional, substantive and procedural safeguards. Self-certification is the dominant paradigm in the AIA, even in self-standing high-risk AI systems, despite the strong resistance from civil society organisations.[48] The AIA suggests that harmonised European standards are the appropriate benchmark at both ends: the substance of the obligations imposed by law and concretised through standards and the demonstration of compliance through proper procedures.

3.2 Third-Party Certification

The *second* layer, the involvement of certification bodies, is not new in risk regulation. If products are classified as particularly risky – such as medical products — the producer is no longer allowed to self-certify but must involve a certification body. Regarding the distinction between self-certification and third-party certification, different rules apply to high-risk AI systems, which are safety components of products and standalone AI systems.

In products with AI components, existing third-party certification provided for the types of products enlisted in Annex I is extended beyond product safety to include protection against

[46] The distinction is in line with theories of governance St. Grundmann/ F. Möslein/ K. Riesenhuber (eds.), Contract Governance – Dimensions in Law and Interdisciplinary Research, OUP, 2015,

[47] The CJEU has widened the scope of the UCTD towards procedural requirements, prominently through the ex officio doctrine A. Beka, The Active Role of Courts in Consumer Litigation. Applying EU Law of the National Court's Own Motion, Intersentia, 2018. Including the institutional dimension would include the organization of the internal safeguards, division of labour, organization of responsibilities of the implementation of the due diligence obligations.

[48] BEUC Regulation AI TO Protect The Consumer Position Paper on the AI Act 2021, https://www.beuc.eu/sites/default/files/publications/beuc-x-2021-088_regulating_ai_to_protect_the_consumer.pdf .

physiological, psychological and economic harm. Article 5 (1) a)[49] and respect for fundamental rights (mentioned 104 times in the AIA) (Article 43 (3)). In standalone AI systems with little experience, self-certification is the rule (Article 43 (2) Annex VI). The only exceptions are remote biometric identification systems (Article 43 (1)), where the AI provider may choose between internal control or third-party control, provided the harmonised standards fully cover the Section 2 requirements (Article 43 (1) second subpar a)-d)). Under EU law, third-party certification can be exercised by public or private bodies. The ultimate responsibility for the accreditation of certification bodies lies with public authorities.

The EU legislator's reliance on third-party certification in the New Approach regulation has boosted the development of private certification bodies, such as the German TÜV. The PIP scandal (concerning defective silicone implants used for plastic surgery) disclosed the conceptual deficiencies of the EU third-party certification. Before the medical device directive had been amended, certification bodies were neither obliged to monitor and survey compliance after the certification nor explicitly empowered to conduct on-site inspections without prior notification.[50] The imposition of such obligations on certification bodies (and the granting of corresponding powers) would turn such bodies into a post-market control entity. The AIA contains a chapter on 'surveillance' in Annex VII 5,[51] without specifying what surveillance entails and without reference to Article 72, which addresses the post-market monitoring duties of the provider. Annex VII 5.3 requires the notified body to '*carry out periodic audits to ensure that the provider maintains and applies the quality management*' but does not clarify whether these visits must be notified in advance. The AIA reiterates the uncertainties that led the CJEU to hold that certification bodies were not liable in the PIP scandal.[52] Article 43 remains silent on the exact scope of obligations of certification bodies.

Public authorities have the power of unannounced on-site inspections, according to Article 74 (5).[53] This power has great significance for making controls effective.[54] This power might be extended to private certification bodies through delegated acts complementing the post-market surveillance duties in the AIA (Article 43 (5) and (6))). *Delegated* acts are non-legislative acts of general application which supplement or amend certain non-essential elements of AIA. Unannounced on-site inspections supplement post-market surveillance duties.

Self- and third-party certifications are close to outsourcing public enforcement responsibilities to private parties. Providers who self-certify must exercise post-market

[49] Outspoken and also mentioning societal harm rec. 5 (Article 5 runs as follows: 'The following AI practices shall be prohibited: a) the placing on the market, putting into service or use of an AI system that deploys subliminal techniques beyond a person's consciousness or purposefully manipulative or deceptive techniques, with the objective to or the effect of materially distorting a person's or a group of persons' behaviour by appreciably impairing the person's ability to make an informed decision, thereby causing the person to take a decision that that person would not have otherwise taken in a manner that causes or is likely to cause that person, another person or group of persons significant harm; see P. Pałka, AI, Consumers & Psychological Harm (July 15, 2023). 'AI and Consumers,' Larry DiMatteo, Cristina Poncibò, Martin Hogg, Geraint Howells (Eds.), Cambridge University Press (2023/2024), Available at SSRN: https://ssrn.com/abstract=4564997

[50] P. Rott,Certification – Trust, Accountability, Liability, Springer ; 2019 ; H.-W. Micklitz/ N. Reich/ L. Boucon, L'Action de la victime contre l'assureur du producteur RIDE, 2015, 37-68

[51] AIA Annex VII under 5.

[52] CJEU Case C-219/15 Schmitt ECLI:EU:C: 2017:128.

[53] Under reference to Article 14(4), points (d) and (j) Regulation 2019/1020

[54] R. Schütze, European Union Law, Chapter 9, new edition 2024/2025 on file with the author.

control of high-risk AI systems (Article 72), and private certification bodies have at least 'surveillance duties', Annex VII 5. The legislature recognises that private parties are closer to the business practice, closer to discovering non-compliance, and closer to finding potential infringements. Public authorities are put into a backstage role. However, operators need a particular reason to take corrective actions in their own motion, i.e. a strong motivation that is not self-evident. The reliance on self-control sharply deviates from the Cyber Resilience Act (Article 45 CRA), where the Commission can act directly based on ENISA's annual reports.

The different conformity requirements for AI systems as safety components and standalone systems lead to a paradoxical result: third-party assessment might have a role to play where technology is an 'add-on' to already existing products, whereas third-party assessment has practically no role to play in the world of new AI systems, which are more susceptible to new risks addressed in the act, namely physiological, psychological, societal and economic harm and infringements of fundamental rights. In the AIA, self-assessment is ideally balanced through appropriate public enforcement mechanisms and a Commission-run European 'registry' of high-risk AI systems (Article 49). The explanatory memorandum to the first proposal of the European Commission highlighted the interaction between ex-ante conformity assessment and 'strong' ex-post enforcement.[55] Recital 148 stresses the need for a new governance framework but does not maintain the original wording.

4 Public Enforcement Mechanisms

EU reality confronts us with a thicket of national and European authorities involved in compliance and enforcement, tied together through a net of information exchange and cooperation duties. Information exchange presupposes a shared understanding of what information will be generated. The AIA lays down a bulk of rules which specify the type of information to be generated and how this information has to circulate within the net of national and European enforcement authorities and be made accessible within the boundaries of confidentiality. We will first look deeper into the public enforcement mechanism before moving on to the particularities of reporting, information, circulation and access between public agencies, public agencies, and private conformity assessment bodies (under 5).

4.1 National and European Public and Private Authorities/Entities

Under the Treaty, Member States are the prime enforcers. The original AIA proposal of the EC from 21.4.2021 and the EP amendments from 23.6.2023 show different approaches to the role and responsibilities of the EU in enforcement. The EC proposal, while not attributing enforcement functions directly to the Commission, put the Commission on centre stage as co-ordinator and governor of Member States' enforcement authorities. The AI Board, composed of the Member States authorities, was supposed to provide advice and support to the Commission in clear fields. The EP proposal instead advocated a solution where the tasks of the AI Office and the European Commission should be separated from each other. Thus, the AI Office envisaged by the EP would have had a more independent role and could have served as an entry card to promote a European AI agency with regulatory powers like ESMA.

[55] EC Proposal on AIA COM(2021) 206 final, 21.4.2021 Explanatory Memorandum under 5.2.3. p. 14.

The finally agreed version (Chapter VIII Section1)[56] concentrates expertise in the hands of the Commission, supported by external advisory bodies, whilst the enforcement remains formally in the hands of the Member States, Chapter VI Section 2. The AI Office (Article 64 AIA)[57] is downsized to an integral part of the Commission, meant to develop Union expertise and capabilities and contribute to implementing Union law on AI. The AI Board *(*Article 65 AIA), composed of one representative per Member State, is tasked with facilitating an efficient and consistent application of the AIA. The Scientific Panel of Independent Experts (Article 68 AIA) integrates the scientific community, and the Advisory Forum (Article 67 AIA) takes care of stakeholder input into the implementation at the Union and national levels. The governance structure shall guarantee that the competent national authorities apply similar standards, coordinated and exchanged within the EAIB, the EAIO, the Scientific Panel and the Advisory Forum.

In the final interinstitutional negotiations, the emergence of Large Language Models (in particular GPT) played a crucial role and led to the integration of what has been termed generative AI. The AIA includes definitions in Article 3 (66) on general-purpose AI, classification, and due diligence obligations in Chapter V, modelled after Chapter III. Those on supervision, investigation, enforcement and monitoring are laid down in Chapter IX Section 3. The Commission has exclusive powers to supervise and general-purpose AI models (Article 88 (1) AIA). At the same time, the role and function of the EAIO and the Scientific Panel have been upgraded. The Commission must entrust supervision and enforcement over general-purpose AI to the AI Office, a sub-unit within the Commission. The Scientific Panel *may* provide a qualified alert to the AI Office.[58] After informing the AI Board, the Commission *may* assess the matter and eventually take regulatory action. There is no obligation.

The question remains of how to embed Member State authorities competent for the AIA into the existing institutional framework regarding the enforcement of EU competition, data protection, product safety law, finance, electronic communication, energy, transport, fundamental rights, environmental and consumer protection. The EU legislature is keen to integrate the AIA governance structure into national and European agencies.[59] This might explain the references to market surveillance authorities (Article 74 (1)) and various articles in the Market Surveillance Regulation 2019/1020[60]. The latter required the Member States to designate a competent market surveillance authority with harmonised regulatory powers to monitor and survey directives and regulations adopted under the New Approach/New

[56] Rec. 148, on the governance structure Novelli et al. (fn 3), and H.-W. Micklitz, The Incomplete Interaction of Public/Private Governance in the AIA in R. Brownsword/L. Di Matteo (eds.), 'Governance of Technology: Discontent, Disruption and Doubts', CUP forthcoming 2025.

[57] Already established C(2024) 390 final Commission Decision, 24.1.2024 establishing the European Artificial Intelligence Office, https://digital-strategy.ec.europa.eu/en/library/commission-decision-establishing-european-ai-office#:~:text=A%20European%20Artificial%20Intelligence%20Office,to%20its%20annual%20management%20plan.

[58] Details under 5.3.5.

[59] For an overview of all the European agencies https://european-union.europa.eu/institutions-law-budget/institutions-and-bodies/search-all-eu-institutions-and-bodies_en?f%5B0%5D=oe_organisation_eu_type%3Ahttp%3A//publications.europa.eu/resource/authority/corporate-body-classification/AGENCY_DEC.

[60] Regulation (EU) 2019/1020 of the European Parliament and of the Council of 20 June 2019 on market surveillance and compliance of products and amending Directive 2004/42/EC and Regulations (EC) No 765/2008 and (EU) No 305/2011 *OJ L 169, 25.6.2019, p. 1–44.*

Legislative Framework.[61] The Market Surveillance Regulation applies to all AI systems (Article 74 (1) AIA), regardless of their risks. The AIA puts national market surveillance authorities in a prominent and decisive position. The AIA requires the market surveillance authority to ensure the application and implementation of the AIA (Article 70 (1)). However, Member States may also involve additional authorities in this task. Supervision and monitoring of AI in finance shall remain in the hands of agencies controlling financial markets (Article 74 (6)). The same applies to competition authorities (Article 74 (2)). The AIA does not deal with national and European agencies dealing with electronic communication, energy, environmental and consumer protection, although the AIA covers these policy fields.

The AIA does not request Member States to establish a genuine AI authority, as discussed but not realised in Germany. [62] Currently, Member States oscillate between centralisation and de-centralisation: centralisation implies the establishment of one single authority competent for all AI matters; de-centralisation involves the assignment of supervision and monitoring to existing market surveillance authorities and sectorial agencies. The enforcement structure of the 27 Member States is highly diverse. More than one national authority will often have to deal with AI matters.[63] The AIA assigns the European Data Protection Supervisor as the sole authority on using AI within European institutions, except for the CJEU (Article 74 (9)).

A typical way to facilitate exchange and cooperation between the many different national authorities and the European Commission whilst not touching upon the diversity of competence distribution in the Member States would have been the requirement to designate no more than one authority per country and to make the list publicly available, such as in the Market Surveillance Regulation (Article 10 (3)) and the DSA (Article 49).[64] There was no political agreement on such a valuable and pragmatic solution during the political negotiations on the AIA. The AIA only states that Member States shall establish or designate as national competent authorities *at least one* notifying authority.[65] and *at least one* market surveillance authority (Article 70 (1) 1)), publish a list of national competent authorities, and communicate them to the European Commission, who will make the list publicly available (Article 70 (2)). There will be 54 designated authorities if not many more.

Contrary to energy, finance, telecom, and transport, EU law does not force Member States to establish a consumer agency, at least not under the supervision of national consumer law. Reg. 2017/2394[66] obliges the Member States to entrust the cooperation in cross-border enforcement to one national public agency that is ultimately responsible. The AIA has not

[61] Most of them are listed in Annex I AIA.

[62] On the discussion in Germany, Bundesministerium der Justiz und für Verbraucherschutz (Federal Ministry of Justice and for Consumer Protection) und H. Schulte-Nölke (ed.) Neue Wege zur Durchsetzung des Verbraucherrechts, Springer, 2017, with contributions on the situation in several other Member States.

[63] In Germany there are 7 ministries involved and this is only at the Federal level.

[64] Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market For Digital Services and amending Directive 2000/31/EC (Digital Services Act) (Text with EEA relevance) OJ L 277, 27.10.2022, p. 1–102

[65] Article 3 (19) notifying authority' means the national authority responsible for setting up and carrying out the necessary procedures for the assessment, designation and notification of conformity assessment bodies and for their monitoring.

[66] Regulation (EU) 2017/2394 of the European Parliament and of the Council of 12 December 2017 on cooperation between national authorities responsible for the enforcement of consumer protection laws and repealing Regulation (EC) No 2006/2004 (Text with EEA relevance), OJ L 345, 27.12.2017, p. 1–26

been added to the list of EU laws where Reg. 2017/2394 applies, contrary to the Data Act.[67] Therefore, national consumer agencies in 25 Member States (in Austria and Germany, there is none) competent to enforce inter alia unfair terms and unfair commercial practices are excluded from the enforcement structure of the AIA. The EU legislator opted for a clear division of competencies between the authorities competent for enforcing consumer laws and those qualified for enforcing the AIA. It insinuates that such a clear borderline between the two fields of law can be drawn, a highly debatable assumption, if one dives deeper into the relationship between the EU consumer law acquis and the AIA.[68] The reason for separating the competent bodies might result from the Member States' resistance against indirect interference in law enforcement's internal organisation and structure. In practice, such a separation might be dysfunctional.[69]

4.2 Competent National and European Authorities

A first attempt to structure the jungle of competent authorities in ensuring compliance and enforcement of the AIA looks like this: The European Commission adopts delegated and implementing acts, guidelines and recommendations, deals exclusively with enforcement over generative AI and monitors surveillance and enforcement of the competent national authorities through the AI Board, the AI Office, the Scientific Panel and the Advisory Forum. Guidelines and recommendations do not bind the national authorities.[70] The EU Data Protection Supervisor supervises the EU institutions. European agencies enforce the regulated markets on finance, energy, electronic communication, transport, and fundamental rights. The AIA lacks guidance on the communication, exchange of information and cooperation between these European agencies, the European Commission and the competent Member States authorities of the AIA.[71]

At the national level, the designated market surveillance authorities and notifying authorities function as liaison offices. They should pass any question —from the Commission, Member states, and economic operators to the national enforcement authority, which is locally, regionally or product-wise competent. Thus, the designated liaison offices have distinct tasks that are not to be confused with those of national authorities competent for surveillance, inspection, and enforcement. The liaison is needed since the current list of competent market surveillance authorities distinguishes between 30 different fields, designates the competent authorities per field and is 322 pages long,[72] There are more than 1,000 entities in total. One might wonder what steps the Member States will take to allocate powers under the AIA among the many enforcement authorities.

Notifying authorities have the task of managing the procedures for the assessment, designation and notification of conformity assessment bodies, and performing third-party

[67] Article 47 of the DA, Regulation (EU) 2023/2854 of the European Parliament and of the Council of 13 December 2023 on harmonized rules on fair access to and use of data and amending Regulation (EU) 2017/2394 and Directive (EU) 2020/1828 (Data Act), OJ L, 2023/2854, 22.12.2023, ELI: http://data.europa.eu/eli/reg/2023/2854/oj

[68] I. Domurath/ H.-W. Micklitz, EU Digital Private Law: Tattering or New Beginning? , ECRL 20(4): 263–319

[69] Not least due to the development of common technical instruments in consumer law enforcement, see Ch Riefa, L. Coll, Director, The transformative potential of Enforcement Technology (EnfTech) in Consumer Law, 2024 https://www.enftech.org/report

[70] Case C-28/15 Opta ECLI:EU:C:2016:692, for the saga behind Z. Georgieva Not so flexible? The instrumental usage of soft law in EU telecommunications regulation, 30.11.2023, https://eulawenforcement.com/?p=8714

[71] 4.4.

[72] https://ec.europa.eu/docsroom/documents/59254 , List of national market surveillance authorities by sector

conformity assessment activities, such as testing, certification and inspection. Their number is not yet known. Notified bodies are conformity assessment bodies designated (notified) by the notifying authority. Those who are recognised by the notifying authorities must be made public. Of particular interest are private and public certification bodies. Last but not least, national qualified entities 'consumer agencies and consumer organisations may take collective action against infringements of the AIA provided they affect the collective interests of consumers (Article 110 AIA), currently 49 under the list of qualified entities.[73] However, the list is incomplete as not all Member States have yet notified the Commission of the qualified entities that will have a standing under Directive 2020/1820 on Representative Action.

4.3 Layers of Competent Enforcement Authorities

A first attempt to structure the jungle of competent authorities at the national and European levels involved in ensuring compliance and enforcement of the AIA looks like this:

European Commission,

- Adopts delegated and implementing acts
- Adopts guidelines and recommendations
- Supervises and controls GPAI
- European competition law

AI Office

- Monitors implementation and enforcement of the AIA
- Investigating possible infringements
- Powers to monitor and supervise compliance of GPAI
- May develop and recommend voluntary model terms for GPAI contracts

Scientific Panel

- Advise and support the AI Office
- At their request, supporting the market surveillance authorities

Advisory Forum

- Provide technical expertise
- Upon request of the Board or the Commission, opinions, guidelines and recommendations

European Artificial Intelligence Board

- Coordinates enforcement
- Upon request or ex officio, adopt recommendations and opinions of codes of practice
- Assist AI Office

EU Data Protection Supervisor

- Supervises AI systems run by EU institutions

European Sectorial Agencies

- Enforce sectorial EU laws in finance, energy, electronic communication, transport
- Enforce fundament rights through the European Union Agency for Fundamental Rights

[73] https://representative-actions-collaboration.ec.europa.eu/cross-border-qualified-entities

Designated Authority

- Member States shall establish or designate
- At least one notifying authority and
- At least one market surveillance authority

Notifying Authority

- Assesses, designates and notifies conformity assessment bodies and monitors them

Conformity assessment bodies

- Performs third-party conformity assessment activities, including testing, certification and inspection

Notified body'

- Are conformity assessment bodies being designated (notified) by the notifying authority

National competent authority

- Market Surveillance Authorities with adequate technical, financial and human resources as well as infrastructure to be reported to the European Commission
- Other competent authorities: competition, data protection, finance, fundamental rights

Private entities (business and civil society organisations)

- Model terms in GPAI b2b relations
- Control unfair terms, commercial practices and data privacy policies in B2C relations

One might distinguish between two key actors, the European Commission and the Member States. The European Commission is the spider in the net, even if it has enforcement powers only on GPAI. The integration of the AI Office into its services, the possible reliance on the Scientific Panel and the integration of stakeholders in the Advisory Forum enables the European Commission to generate the necessary knowledge and expertise to guide surveillance and enforcement. The Member States are participating through the EAIB. The limitation to one representative in the EAIB frees the Commission from diving into the complexity of national enforcement authorities. Still, it sets the problem aside that those in charge of a concrete problem are not participating.

The market surveillance authorities play a key role. They have to master a vast field of activities, which cuts across the 30 sectors already listed and requires competence and expertise with AI technology. For the first time, the Member States must provide the European Commission with information on the resources they assign to the competent authorities (Article 70 (3)). Only a few Member States will likely play an active role in enforcing the AIA. In any case, regulatory decisions taken by national competent authorities will result from a mixture of EU law and national laws. The phenomenon and the potential consequences of judicial review of individual administrative actions are widely discussed in migration law but are still somewhat underestimated in market regulation.[74]

Last but not least, private parties, civil society organisations, consumer and environmental organisations will have a role to play in the enforcement of the AIA, not so much via powers granted through the AIA but by an extended interpretation of the existing EU law acquis in

[74] S. Demiková, Automated Decision-Making and Effective Remedies, The New Dynamics in the Protection of EU Fundamental rights in the Area of Freedom, Security and Justice, Elgar 2023

environmental and consumer law. The new European Commission has promoted collective redress on its agenda for 2025. One of the objectives is to bring together the judges involved in litigation and provide appropriate training.

4.4 Co-operation Member States and Commission

Contrary to the DSA, for instance, the AIA does not formulate rules on the cooperation of the Member States—the market surveillance authorities and the European Commission—in the surveillance, inspection, and enforcement of AI systems.[75] The reason is that Chapter IX, Market Monitoring, Information Sharing and Market Surveillance, heavily relies on the Market Surveillance Regulation 2019/1020. Two articles are particularly outspoken, located in Section 3 of Chapter IX on Enforcement. Article 74 deals with market surveillance and control of AI systems in the Union Market. Para 1 starts with 'Regulation EU 2019/1020 shall apply to AI systems covered by this regulation', and para 11 addresses joint activities, including joint investigations 'concerning specific categories of high-risk AI systems that are found to present a serious risk, across two or more Member States by Article 9 of Regulation 2019/1020. The AI Office shall provide co-ordination support for joint investigations.' The Market Surveillance Regulation is equally applicable to general-purpose AI, where the AI office operates as the sole and exclusive surveillance authority (Articles 75 (1) and 75 (3)).

The reference to the MSR raises questions on the interrelationship between the AIA and the MSR. At first-hand sight, it looks as if the MSR is lex generalis and the AIA lex specialis, which adds competencies to the market surveillance authorities specific to the surveillance of AI systems, such as

- full access to the technical documentation (Article 74 (12)) and access to the source code (13)
- definition of AI systems presenting a risk in Article 79 (1) AIA, also defined in Article 3 (1) MSR,
- obligation of market surveillance authorities to inform relevant national public authorities on possible risks to fundamental rights (Article 79 (2))
- a Member State obligation to inform the Commission and the other Member States, in case the risk is not restricted to its territory (Article 79 (3)),
- procedural rights of economic operators against which the market surveillance authority has taken regulatory action (equally applicable for providers of GPAI Article 94),
- monitoring power on AI Systems classified by the provider as non-high-risk in the application of Annex III (Article 80 (8)), power to perform checks on high-risk AI systems (Article 11 MSR),
- Union AI testing support structures about tasks listed under Article 84 AIA in combination with Article 21 (6) MSR,
- obligation of market surveillance authorities to consider complaints of any natural or legal person on infringements of the AIA (Article 85).

The special rules are fundamental as they change the information exchange mechanism, the involvement of external complaints in the risk assessment and the contribution of the EAIB, the EAIO, the Scientific Panel and the Advisory Forum. Of particular relevance is the establishment of the standard sub-group through the AI Board, which acts as the

[75] H.-W. Micklitz, Compliance and Enforcement of the DSA, forthcoming.

Administrative Co-operation Body in the meaning of Article 30 MSR about issues related to market surveillance. This Body is much broader than the AI Board, as also representatives of stakeholders are allowed to participate. The AI Office will have a decisive role, not least through its exclusive competence in controlling GPAI. The Scientific Panel shall support and advise the AI Office and support the market surveillance authorities at their request. The role and function of the Advisory Forum is less clear. In practice, much will depend on the members and their commitment. Interestingly, the Fundamental Rights Agency, ENISA, CEN, CENELEC, and ETSI are permanent members, whereas the other stakeholders 'with recognised expertise' (Article 67 (3)) must be selected by the European Commission.

The two references to the MSR and particular rules of the MSR raise the question of whether the MSR could close gaps in the AI enforcement structure, for instance, concerning mutual information on serious incidents, in contrast to high risks. Or, more generally, are the market surveillance authorities in the Member States allowed and even entitled to use the MSR as the basis for law enforcement, surveillance, monitoring and investigation, and investigate the AIA only about particular issues, such as access to the technical documentation, the source code and cybersecurity?

The uncertainties of empowering market surveillance authorities to initiate joint activities are even more profound. Article 74 (12) presupposes 'serious risks' across 'two or three Member States'. The AIA does not define the category of serious risk, contrary to the MSR, which draws an explicit distinction between 'risks' and 'serious risks' (Article 3 (19) and (20) MSR). Article 79 (1) only refers to 'risks', which should be distinguished from 'serious incidents' (Article 3 (49)). This is not yet all. Article 74 (11) limits joint activities in 'one or more Member States' to 'serious risks' – not serious incidents. Article 3 (61) provides for an explicit definition of widespread infringements, which, however, is not referred to in Article 74 (11) but only in Article 73 (3), which equates widespread and serious incidents considering reporting duties. These uncertainties and partly confusing definitions and interrelationships between the AIA and the MSR will now be discussed in more detail as they are crucial for the operationality of the AIA governance and enforcement structure.

5 Reporting, Information, Circulation, Access

One might distinguish between the sources, types, and methods of information generation, the circulation of risk information, which is crucial for effective enforcement, the modes of access to information, and the barriers to overcome through a request or a reasoned request. The findings allow tentative considerations on the way forward within the existing governance structure, leaving AI as a tool to improve compliance, conformity and enforcement of section 6.

5.1 Sources, Types and Methods

The information relevant to compliance and enforcement comprises five primary sources: 1) technical documentation on compliance, 2) information on possible infringements – in the language of the AIA – serious incidents, 3) information on risks defined under the market surveillance regulation AIA/MSR, 4) information on serious risks AIA/MSR and 5) information on non-compliance and risks despite compliance. Their relevance and meaning can only be fully disclosed by deepening the interaction between the AIA and the Market Surveillance Regulation.

5.1.1 Technical Documentation

The scope and function of the *minimum technical documentation* are laid down in Article 11 (1), supplemented by detailed requirements in Annex IV, subject to amendment by delegated acts. This documentation might very well turn into *the* key source of information for the enforcement authorities on the inner mechanics of the AI system, which is crucial for understanding its functioning, supervision and monitoring. The requirements are broken down into nine categories (Article 11 Annex IV) inter alia:

(1) General description of the AI system,
(2) Detailed description of the elements of the AI system and the process for its development, including (b) the *design specifications of the system*, namely the general logic of the AI system and the *algorithms*; the *key design choices,* including the rationale and assumptions made, also about persons or groups of persons; (c) the description of the *system architecture* explaining how software components build on or feed into each other and integrate into the overall processing; the *computational resources* used to develop, train, test and validate the AI system, (d) *training methodologies*, (e) assessment of *human oversight* (g) *validation and testing procedures*;
(3) Detailed information about the *monitoring, functioning and control of the AI system*,
(4) Description of the appropriateness of the *performance metrics* for the specific AI system,
(5) Description of the *risk management system*,
(6) Description of any changes by the provider to the system through its *life-cycle*,
(7) List of *harmonised standards* applied,
(8) Copy of the EU *Declaration of Conformity* and
(9) Description of the system meant to evaluate the AI system's *post-market* performance, including a post-market *monitoring* plan (emphasis added HWM/GS).

Technical documentation is crucial for understanding the technology used in high-risk AI systems and addressing problems and incidents. Therefore, such documentation is needed by public authorities to initiate an investigation, as well as by potential victims of AI systems to engage in private litigation, individually or collectively. The providers of AI systems will insist on the confidentiality of the technical documentation. At the same time, potential victims and NGOs might call for full transparency and even open access to initiate litigation. Politically, the situation is not different from the conflict on whether credit bureaus should make publicly available credit scores and the criteria on which they are built.[76] To the dismay of BEUC, the EP succeeded in reducing SMEs from this heavy administrative burden, allowing them to provide the technical documentation in a 'simplified manner', a model to be established by the European Commission (Article 11 (1) sentence 5).[77] One might wonder whether an appropriate solution consists of using AI tools to facilitate the production of such technical documentation (see below).

5.1.2 Serious Incidents in High-Risk AI Systems

Article 3 (49) AIA defines a '*serious incident – incident or malfunctioning'* within an AI system - independent of the category of risk - which leads to death, personal injuries, infringes property, environment and fundamental rights or produces a serious and irreversible disruption of the management or operation of the critical infrastructure disruptions of critical

[76] German Advisory Board for Consumer Affairs (Sachverständigenrat für Verbraucherfragen), Fair Credit Scoring (Verbrauchergerechtes Scoring), 2020.
[77] BEUC (fn 37).

infrastructure, the latter being defined in Article 3 (62) AIA under reference to Article 2 (4) Dir. 2022/2557.[78]

The provider must report all serious incidents to the surveillance authorities (Article 73 (1)). This is not a mere information duty. The provider must deliver a report analysing a possible or likely causal link. Interestingly the AIA does not foresee an implementing act on a standardised format. The deployer is involved, and he may submit an initial report if he is the one who discovers the incident (Article 73 (6)). The delay for the delivery of the report differs according to the type of incident (Article 73 (2)-(4)) – widespread infringements and infringements of the infrastructure within two days, death immediately, and all other incidents – including fundamental rights within 15 days. The definition of a serious incident is not adjusted to the broadening of the scope of protection under Article 5 AIA, which includes significant harm. Property and harm are not the same.

The report shall be sent to the competent national market surveillance authorities *where* the incident occurred (Article 73 (1)). Paragraph 3 deals with widespread infringements of the infrastructure (Article 3 (49) (b)). Widespread, however, is not meant to cover infringements concerning two or more Member States, as defined in Article 3 (66) or referred to in Article 74 (11) on market surveillance. One may wonder whether serious incidents are always regarded as widespread and relevant for several Member States or the EU. Such an interpretation, however, does not fit into Article 73 (1), which suggests that the provider has to report serious incidents to the authority of the Member State where it 'occurred.'

In their key role as designated body (Article 70 (1)), the market surveillance authorities will receive reports of all Article 3 (49) categories, health, infrastructure, fundamental rights, property, environment and the critical infrastructure of essential services using high-risk AI systems.[79] They shall immediately notify the European Commission of any serious incident and inform them whether they have taken regulatory action (Article 73 (11)). One should expect that Article 73 also specifies the distribution of the report between the different national authorities and the Member States' authorities. However (Article 73, paras 7 and 9) only deal with reports on infringements of fundamental rights (Article 3 (49) c)). Para 7 obliges the market surveillance authorities to forward the notification of a serious incident received by the provider to fundamental rights bodies (Article 73 (7) together with Article 77 (1)), para 9 deals with equivalent reporting duties on fundamental rights of the providers of a high-risk AI system. There are no rules on property infringements, which could affect competition and the environment. However, all Member States have national agencies in both fields, infringing on the critical infrastructure providing essential services. The European Commission is requested to develop 'dedicated guidance' to facilitate compliance with the tasks of the fundamental rights authorities (Article 77 (1). Competent national authorities are not obligated to automatically inform all other national supervisory authorities of the Member States of serious incidents. An exception to the rule concerns the handling of fundamental rights. Article 77 (1), in the second sentence, obliges the competent national authority to inform the market surveillance authority of another Member State to provide the

[78] AIA Article 3 (62) refers to Article 2 (4) Dr. 2022/2557 Critical Entities: critical infrastructure' means an asset, a facility, equipment, a network or a system, or a part of an asset, a facility, equipment, a network or a system, which is necessary for the provision of an essential service.

[79] Directive (EU) 2022/2557 of the European Parliament and of the Council of 14 December 2022 on the resilience of critical entities and repealing Council Directive 2008/114/EC (Text with EEA relevance) OJ L 333, 27.12.2022, p. 164–198

infringement concerns of other territories; at least, this is how the somewhat opaque rule should be understood. The special status of infringements of fundamental rights might be counterproductive due to the lack of appropriate case law of the CJEU and the ECtHR on the Charter and the European Convention on Human Rights.[80]

5.1.3 Risks vs. Incidents, High-Risks vs Serious Risks

The scope and range of '*serious incidents'* in Article 3 (49) AIA must be distinguished from '*products presenting a risk*' in the meaning of Article 3 (19) MSR. The MSR does not only cover health and safety but also consumer and environmental protection as well as public interests.[81] The broadening of Article 3 (49) AIA has led to approximating the scope of protection against 'serious incidents' and 'products presenting a risk' under the exclusion of consumer protection, though it is not explicitly mentioned.

However, what about potential differences between 'incidents' and 'risks' and '*serious* incidents' and 'risks' versus 'serious risks? The question's relevance results from the overall interaction between the AIA and the MSR. The AIA and the MSR start from the same or nearly the same understanding of 'risks' (Article 3 (2) AIA and Article 3 (18) MSR. The identity is crucial when it comes to the procedure at the national level for dealing with AI systems presenting a 'risk' (Article 79). The competent authorities might use the same benchmark to assess the risk, whether it is a risky product or a risky AI system, but there is one difference. Article 79 (2) puts particular emphasis on the supervision of potential risks to groups of vulnerable persons and their social and economic vulnerabilities(Article 5 (1) (b)).

The AIA demarcates risks in Article 3 (2) AIA as '*the combination of the probability of an occurrence of harm and the severity of that harm*', but not '*incident'*. The Oxford Dictionary defines an incident as 'something that happens, especially something unusual or bad'.[82] Therefore, we may conclude that incidents are materialised risks. Article 3 (49) focuses on a 'serious incident' as a malfunction that leads to serious consequences, such as death or serious harm to health, irreversible disruption of the infrastructure, infringement of fundamental rights, and serious harm to property. Thus, 'certain' incidents causing 'certain' harms are always serious, while infringements of property are only 'serious' when they produce serious (considerable) harm. Contrary to the AIA, the MSR distinguishes between 'risks', 'products presenting a risk' and 'serious risks' (Article 3 (18) - (20) MSR).[83] The

[80] There is little case law so far, concerning consumer protection, see B. Kas, Ensuring Digital Fairness in EU Consumer Law through Fundamental Rights: Is the EU Charter Fit for Purpose? In H.-W. Micklitz/ N. Helberger/ B. Kas/ M. Namyslowska/ L. Naudts/ P. Rott/ M. Sax/ M. Veale, Digital Fairness for Consumers, Study commissioned by BEUC March 2024; https://www.beuc.eu/reports/digital-fairness-consumers, 2024, 146.

[81] Article 3 (19): product presenting a risk' means a product having the potential to affect adversely health and safety of persons in general, health and safety in the workplace, protection of consumers, the environment, public security and other public interests, protected by the applicable Union harmonization legislation, to a degree which goes beyond that considered reasonable and acceptable in relation to its intended purpose or under the normal or reasonably foreseeable conditions of use of the product concerned, including the duration of use and, where applicable, its putting into service, installation and maintenance requirements;

[82] https://www.oxfordlearnersdictionaries.com/definition/american_english/incident#:~:text=1%5Bcountable%5D%20something%20that%20happens,incident%20sticks%20in%20my%20mind.

[83] (20) 'product presenting a serious risk' means a product presenting a risk, for which, based on a risk assessment and taking into account the normal and foreseeable use of the product, the combination of the probability of occurrence of a hazard causing harm and the degree of severity of the harm is considered to require rapid intervention by the market surveillance authorities, including cases where the effects of the risk are not immediate;

definition of risk in Article 3 (18) MSR is identical to Article 3 (2) AIA. Article 3 (19) MSR determines 'products presenting a risk' through the scope of protection: 'health and safety of persons in general, health and safety in the workplace, protection of consumers, the environment, public security and other public interests. Seriousness in Article 3 (20) MSR is bound to a risk assessment which requires 'rapid intervention'. The distinction sounds very much hands-on in light of the long-lasting debate on defining risks and separating risks from danger. The comparison of the AIA and MSR raises difficult questions: what is the relation between serious risks in the MSR and serious incidents in the AIA? One might start from the premise that the two are identical and that 'serious incidents' to reporting in Article 73 are 'serious risks' in Article 79. But what about 'high-risk' AI systems? What is the difference between high risks defined in Article 6 in conjunction with Annex III and 'serious risks' equated with 'serious incidents'?

Further irritation results from the different treatment of powers and regulatory tools on serious risks. Can the rules in the MSR be transferred to the AIA? The MSR provides a whole set of tools in Article 9, MSR joint activities, Article 14, powers of market surveillance; article 19 (2) MSR risk assessment; article 25, MSR controls on entering the Union market, Article 26, MSR suspension and Article 28 MSR refusal. The AIA refers only to Article 9 MSR in Article 74 (11) explicitly, speaking of 'serious risks' equated with widespread reach in two or more Member States. Does it mean that only 'widespread infringements' are considered serious risks, not the other three, as defined in Article 3 (44)?

One may twist and turn the two regulations. Still, considerable uncertainty remains as to whether there are conceptual consequences behind the different terms – risks vs. incidents, serious incidents vs. serious risks -linking the type of risk to the speed of regulatory action. Most of these issues can be addressed through interpretation, whether serious incidents have to be reported to the Commission only Article 73 (11) or also to the market surveillance authorities of the Member States. If *mere* risks in the meaning of Article 3 (19) MSR must be reported to all the other Member States authorities provided the non-compliance is not restricted to the national territory (Article 79 (2)), the same consequence would equally have to apply to *serious* risks in the meaning of serious incidents (Article 3 (44) and Article 73 (11)). One may hope that the different soft-law instruments will provide clarification.

5.1.4 Union Safeguard Procedure

The Union safeguard procedure in Article 81 applies in case of disagreement between the competent national authorities on the scope and reach of a regulatory action of one Member State directed against an economic operator who is not willing to take corrective actions to remedy an identified risk in the meaning of Article 79 (1) and (2). The Union safeguard procedure covers products with an AI component and stand-alone AI systems. The competent authority in charge has to inform the Commission and the other Member States 'without undue delay' on the measures taken (Article 79 (5-6)), in particular the potential reasons for non-compliance with Article 5, with Chapter III Section 2 (obligations of economic operators), shortcomings in harmonised standards and common specifications and transparency obligations Article 50.

The Union safeguard procedure in Article 79 (5-9) and Article 81 aims at coming to a unified regulatory action by the market surveillance authorities of all Member States – in two constellations – agreement and disagreement on the risk assessment and the regulatory action. Mutual information between all market surveillance authorities and the European

Commission is the starting point. The other market surveillance authorities and the European Commission have three months or 30 days in case of non-compliance with the prohibition of the AI practices under Article 5 to start their own investigation on the notified measures. If neither the other market surveillance authorities nor the Commission raises objections, the decision becomes final, and the other market surveillance authorities have to take 'appropriate restrictive measures' (Article 79 (7-9). Suppose there is an objection from the market surveillance authorities and/or the Commission. In that case, the Commission decides whether the original national measure is justified (Article 81 (1)). If, according to the evaluation of the European Commission, the national measure is justified, all other Member States have to take 'appropriate restrictive measures' (Article 81 (2)); if it is not justified, the market surveillance authority of the respective Member State has to withdraw the unjustified measure and inform the Commission.

The Union safeguard procedure respects the competence of the Member States in law enforcement, but this leads to strange results. The Commission decides on the existence or non-existence of risk in the meaning of Article 79 (1) – with all the uncertainties on the interaction with serious incidents in the meaning of the AIA and serious risks of the MSR. In contrast, the Member States choose the appropriate regulatory tools and impose them on the respective economic operators of the AI system. In case the operator does not accept and goes to court, the national court must decide on the risk assessment made by the European Commission in the litigation between the operator and the competent national authority.[84] The Union safeguard procedure does not foresee the national market surveillance authorities making identical decisions; there is a certain leeway in deciding on the 'appropriate restrictive measures. Therefore, the exact risk may lead to different regulatory actions in the Member States concerned.

5.1.5 Compliance and Formal Non-Compliance

Last but not least, there are '*risks despite compliance'* and merely '*formal non-compliance*' with Chapter III. 'The first case presupposes an in-depth investigation on risk to health and safety, fundamental rights and public interest protection, despite compliance with the obligations imposed on the economic operators in managing high-risk AI systems. The second case defines an irreversible assumption that non-compliance with particularly relevant obligations suffices to justify regulatory action. For each of the two cases, the AIA provides for notification and information requirements in two constellations: compliant AI systems presenting risks (Article 82) and merely formal non-compliance (Article 83).

Article 82 (1) deals with 'risks' to health, safety, fundamental rights and public interests. The procedure is the following: 1) the market surveillance authority of one Member State discovering such a risk must immediately require the provider to take adequate measures, (2) the Member State immediately informs the other Member States and the Commission, 3) the European Commission makes the final decision on the justifiability of the initial measures after consultation with the Member States and the relevant operators and 4) the European Commission immediately informs them both on its decision (Article 82 (3-5)).

Article 83 enlists seven (a-g) findings of non-compliance, including, as requested by the EP, non-compliance with the EU database, non-compliance with the duty to appoint an authorised representative and non-availability of the technical documentation. The

[84] See again Demiková (fn 74).

competent market surveillance authority has to take action against the 'relevant provider' first to put an end to non-compliance and, if the non-compliance persists, to take 'appropriate and proportionate measures. There is no formal obligation to inform the other Member States and the European Commission of non-compliance on such merely formal matters, nor is there an obligation of the different market surveillance authorities to take 'appropriate measures. The gap will have to be closed via Article 9 MSR, which paves the way for joint activities in case of non-compliance.

5.2 Circulation of Risk Information

Different types of risk information circulate in two different though overlapping circles – the inner circle of exchange between the various economic operators in the AI supply chain – in the private business environment – and the outer circle between the market surveillance authorities in the public enforcement environment - other enforcement authorities, notifying authorities, eventually notified bodies, the European Commission, the AI Office and the AI Board. In the inner circle, the operators will try to control risk information and remedy possible deficiencies without involving public authorities via self and third-party compliance. The outer circle becomes involved when operators are formally obliged to report or when the competent public authorities discover 'risks' from their own motion.

The circulation of risk information in the outer circle, i.e., between market surveillance authorities, differs according to the type of risk and information. Before taking regulatory action, the information stays within the investigating authorities. To some extent, providers and deployers must inform the authorities in the Member States whose territory is affected by serious incidents and malfunctioning. These market surveillance authorities and the AI Office about GPAI may also generate information through their ex officio investigations. The overall rationale is that the information reported by the provider and the deployer of the AI system should only circulate between the 'affected' Member States' market surveillance authorities and the European Commission. Once the Member States' market surveillance authority in charge has taken action, *all* other authorities and the European Commission must be informed to allow joint EU-wide regulatory action.

In the *inner* circle, one might distinguish between contractually agreed reporting obligations within the AI supply chain and mandatory risk information established by the AIA. Contractual reporting enables risk management through the provider of the AI system or any other operator. Mandatory information duties in the AIA are somewhat limited and follow the distinction between risk information – before taking action — and risk information on missing corrective actions. There was much debate between the EC and the EP on the conditions under which operators are legally obliged to report to the public authorities. The first issue concerned the limitation of notification duties to 'serious incidents'; the second pertained to the concretisation of risk information, which Article 79 (1) limits to 'risks' while the EP proposed to include 'potential risks'. The difference matters and is related to the scope and reach of post-market monitoring duties (Articles 72-73) for high-risk AI systems. The post-market monitoring plan to be implemented by the European Commission provides room for leeway. Much will depend on the template, which does not yet exist—the third issue concerns who must inform whom within the AI supply chain, on what and when.[85]

[85] Table under 2.2.3.

The notified bodies are located between the inner and outer circle, being widely disconnected from the economic operators and the public enforcement authorities. Like national market surveillance authorities, they can access the technical documentation Article 11 (1) to meet their responsibilities. Chapter III Section 4 deals with notifying authorities and notified bodies more particularly. The primary addressee for the notified bodies is the national notifying authority. The AIA does not assist in identifying who the notifying authorities could be, whether this role could be covered by the national market surveillance authorities, which would facilitate access to information, or by a ministry or another public authority, which would complicate the flow of information. The notified bodies could be either private or public. Article 33 defines the requirements to be fulfilled by the notified body. The body shall have inter alia legal personality (1), be independent and impartial (4-6) and have liability insurance, 'unless liability is assumed by the Member State in which they are established by national law or that Member State is itself directly responsible for the conformity assessment' (9). The formula is vague enough to release the Member States from introducing mandatory liability insurance.

In the conformity assessment, the notified bodies generate information on the type of AI system, including its *potential* risks. Seen through the lenses of the provider of the AI system, the notified bodies belong to the inner circle, on which they rely to be informed not only on existing risks but also on potential risks. The AIA does not oblige the notified bodies to inform the market surveillance authorities at the earliest possible stage, i.e., on potential risk. Upon request only, the notified body has to provide all the documentation and information to the notifying authorities – not the market surveillance authorities— so that the notifying authorities can fulfil their tasks in supervising and monitoring the notified bodies (Article 34 (3)). The notifying authorities are not obliged to provide reasons for such requests, which enables them to be constantly updated on the ongoing conformity assessments. Article 37 (1) grants the Commission the power to investigate the competence, the continued fulfilment or the applicable responsibilities of the notified authorities on all the requirements laid down in Article 31. These requirements are concretised through harmonised standards, compliance with which triggers a presumption of conformity according to Article 32. The investigative powers might become highly relevant if Member States neglect their monitoring and surveillance duties over notified bodies but do not compensate for the lack of knowledge on *potential* risks being disclosed in the conformity assessment. The notified bodies have only to inform the notifying authorities – again not the market surveillance authorities — of any refusal, restriction, suspension or withdrawal of an EU certificate (Article 45 (1) b)) as well as the other notified bodies (Article 45 (2) a)). On the other hand, the national market surveillance authorities must inform notified bodies of the call for corrective actions and respective regulatory actions (Article 79 (2)).

The wide variety of notifying authorities and notified bodies – public and private – leads to further complexities in the circulation of risk information. The competent notifying authorities and the designated notified bodies will be listed and made public (Articles 35 (2)and 70 (2)). Article 38 obliges the Commission to ensure cooperation and coordination between the notified bodies. In the Medical Device Directive, 43 notified bodies are notified; there will probably be many more in the AIA. Notified bodies may have subsidiaries, for which they must take full responsibility and should be notified equally (Article 33 (2)). One may wonder whether Article 33 (2) interferes with company law, provided the notified body is private and the mother company tries to give subsidiaries legal personality to make them economically

independent. Each Member State will designate at least one notifying authority, usually the competent national ministry.

5.3 Access to Internal Information, Confidentiality and Regulatory action

The national enforcement authorities, i.e., the market surveillance authorities, the European Commission, the EAIO, the Scientific Panel and the Advisory Forum, need access to information on AI systems to fulfil their responsibilities effectively. One may distinguish between three different types of information: (1) publicly available information, (2) information which providers and deployers have to report to the enforcement authorities and the Commission and (3) information which lies in the hands of the operators, the provider, deployer, manufacturer, importer, distributor and authorised representative and which is neither publicly available nor automatically available to the national enforcement authorities.

5.3.1 Publicly Available Reports

The overview begins with the publicly available and relatively scarce information on high-risk AI in the EU database, which must be user-friendly (Article 71 (4)). Then there are the public reports from the providers of general-purpose AI (Article 53 (1) d)), (2) and, with limitations, the report concerning regulatory sandboxes (Article 57 (8)) (if the provider and the national authority agree) and the planned and existing sandboxes (Article 57 (15)). These reports are equivalent to the information on high-risk AI systems in the AI systems database. Last but not least, there is the annual report from the Advisory Forum Article 67 (10) and the information generated within the Scientific Panel, where only the declaration of potential conflicts of interests must be public, but not the debates in the Panel (Article 68 (4)).

Key information will result from the reporting of 'serious incidents' by providers and deployers. Setting aside possible investigations that market surveillance authorities will take on their own motion, the question remains: when and under what conditions, the authorities have access to the internal information, first and foremost to the technical documentation (Article 11) and secondly to the record keeping (logs) and eventually the source code (Article 12) generated by the providers and other economic operators or notified bodies in the performance of their mandate.

5.3.2 Reasoned Request on Chapter III Sections 2 and 3

The AIA distinguishes between information to be made available to the market surveillance authorities '*without request*', '*on request*' and on '*reasoned request*'. Whatever public authorities would like to know from all other operators depends on whether they have 'reasons', except the obligation of the authorised representative to hand over the mandate of the provider to market surveillance authorities, Art. 22 (3). The AIA does not clarify the potential difference between a mere request and a reasoned request, let alone how specific the arguments by the competent authorities have to be to underpin their requests.

Article 16 ties the disclosure of the conformity assessment to all Section 2 requirements for high risks to the existence of a *reasoned* request: risk management, data governance, technical documentation, record keeping, transparency of information to the deployers, human oversight, accuracy and robustness is bound to a reasoned request (Article 16 (1) k)), just as the obligation of the provider to cooperate with the market surveillance authorities, to grant access to automated logs (Article 21 (1)-(2)). The other economic operators have narrower obligations to fulfil on reasoned request again: the authorised representative on information, documentation and cooperation with the authority (Article 22 (3) c)-d)),

importers on information and documentation (Article 23 (6)), just as distributors (Article 24 (5)). However, a distributor, importer, deployer or other third party may become a provider and be submitted to fully fledged disclosure if they put their name on the AI system or modify it (Article 25 (1) a)-c).

Whether or not providers shall grant access to the source code of an AI system was particularly conflictual. In the end, the proposal of the EP went through to tie access to the source code to two additional conditions: the necessity to assess the conformity of a high-risk AI system with the requirements set out in Chapter III, Section 2 and (evidence that) testing or auditing procedures and verifications based on the data and documentation provided by the provider have been exhausted or proved insufficient (Article 74 (13)). The burden of the provider will be somewhat eased through the obligation of the Commission to designate one or more testing support structures to perform the tasks listed in Article 21 (6) Reg. 2019/102 in the area of AI (Article 84 (1-2)). However, at the time of writing, it is unclear what the testing support structures will look like and whether they can meet the requirements of sound AI testing as requested and proposed by the Oxford research team.[86]

5.3.3 Request/reasoned Request on GPAI and FR Agencies

The AIA contains special information rules on the provider of general-purpose AI, where the European Commission has exclusive power to obtain information. On *request*, the provider has to disclose technical requirements to the AI Office and the market surveillance authorities of the Member States (Article 54 (1) a)); on *request,* the authorised representative has to lay bare the mandate (Article 54 (3)) and on *reasoned request* he has to provide additional information under Article 54 (3) b), and be willing to cooperate with the AI Office and the competent authority (Article 54 (3) c)-d)). The provider himself may launch a *reasoned request* to the Commission for the reassessment of general-purpose AI with systemic risks (Article 52 (5)). Where a market surveillance authority is unable to conclude its investigation of the high-risk AI system because of its inability to access certain information related to the general-purpose AI model despite having made all appropriate efforts to obtain that information, it may submit a *reasoned request* to the AI Office, by which access to that information shall be enforced (Article 75 (3)). These are not yet all the particularities.

National public authorities or bodies which supervise or enforce the respect of obligations under Union law protecting fundamental rights *about the use of high-risk AI systems* shall have the power to request and access any documentation created or maintained in accessible language and formation (Article 77 (1))**.** Where the documentation referred to in paragraph 1 is insufficient to ascertain whether an infringement of obligations under Union law protecting fundamental rights has occurred, the public authority or body may make a *reasoned* request to the market surveillance authority to organise testing of the high-risk AI system through technical means (Article 77 (3)). The context insinuates that testing means a fundamental rights impact assessment based on Article 27.

5.3.4 Request/reasoned Request of Notified Bodies/Notifying Authorities

Last but not least, notified bodies are tied to a combination of requests and reasoned requests. Notified bodies shall make available and submit upon *request* all relevant

[86] J Laux, S Wachter and B Mittelstadt, 'Three Pathways for Standardisation and Ethical Disclosure by Default under the European Union Artificial Intelligence Act' (February 20, 2023). Available at SSRN: https://ssrn.com/abstract=4365079 or http://dx.doi.org/10.2139/ssrn.4365079.

documentation, including the providers' documentation, to the notifying authority (Article 34 (3)). In the event of the restriction, suspension or withdrawal of a designation, the notifying authority shall take appropriate steps to ensure that the files of the notified body concerned are kept and made available to notifying authorities in other Member States and market surveillance authorities at their *request* Article 36 (6). The notifying authority shall provide the Commission, on *request*, with all relevant information related to the notification or the maintenance of the competence of the notified body concerned (Article 37 (2)). Where the Commission ascertains that a notified body does not meet or no longer meets the requirements for its notification, it shall inform the notifying Member State accordingly and request it to take the necessary corrective measures, including the suspension or withdrawal of the notification if required (Article 37(4)).

Notified bodies shall provide the notifying authority various information items, inter alia: any request for information which they have received from market surveillance authorities regarding conformity assessment activities (Article 45 (a)); on *request*, conformity assessment activities performed within the scope of their notification and any other activity performed, including cross-border activities and subcontracting (e). Moreover, each notified body shall inform the other notified bodies of (a) quality management system approvals which it has refused, suspended or withdrawn, and, upon *request*, of quality system approvals which it has issued; (b) Union technical documentation assessment certificates or any supplements thereto which it has refused, withdrawn, suspended or otherwise restricted, and, upon *request*, of the certificates and/or supplements thereto which it has issued (Article 46 (2)). Each notified body shall provide the other notified bodies carrying out similar conformity assessment activities covering the same types of AI systems with relevant information on issues relating to negative and, on *request*, positive conformity assessment results (Article 47 (1)). A copy of the EU declaration of conformity shall be submitted to the relevant national competent authorities upon request.

The rationale is obvious: decisions on refusal, suspension or withdrawal have to be shared immediately, and more sophisticated and sensible information has to be provided on *request* or even on reasoned request when more detailed information on the conformity assessment is needed. Upon *reasoned request*, the notified body shall also be granted access to the training and trained models of the AI system, including its relevant parameters, where this is necessary to assess the conformity of the high-risk AI system with the requirements set out in Chapter III, Section 2, and after all other reasonable means to verify conformity have been exhausted and have proven to be insufficient. The AIA treats the notified bodies as quasi-enforcement authorities in their own right.

5.3.5 RAPEX (Rapid Exchange of Information Systems), Safety Gate and Alerts

There is a regulatory gap in the AIA, which is amazing as the AIA follows the risk-based approach developed in product safety regulation. The latter regulation provides for an electronic notification system of *potentially* unsafe products. All Member States have to notify the European Commission of such products, and the notifications are integrated and governed under a standardised format, formerly called RAPEX, since the adoption of the Product Safety Regulation 2023/988, renamed into Safety Gate Rapid Alert System (Article 25). Recital 68 sums up the purpose: The Safety Gate Rapid Alert System is the internal system through which authorities and the Commission exchange information on measures concerning dangerous products and which can contain confidential information. An extract of alerts should be published on the Safety Gate Portal to inform the public about dangerous

products. The Safety Business Gateway is the web portal through which businesses inform market surveillance authorities of the Member States about dangerous products and accidents. To the best of our knowledge, there is no single word in the EC proposal nor in the EP document which could pave the way for a similar alert mechanism on 'serious incidents' and/or 'risks'.

It is currently being discussed whether risky products with an AI component should be included in the existing Safety Gate system, but no agreement has been reached. A significant gap concerning stand-alone AI systems would remain even if this inclusion were made. Notifications on serious incidents are certainly not a substitute for an alert system. This could only be a solid information exchange system based on missing potential risks. If any, the alert system in Article 90 could serve as a substitute. The AIA puts the responsibility and the power to launch an alert into the hands of the Scientific Panel. Article 90 suggests that the EU legislature trusts independent scientific expertise more than public officials in market surveillance authorities. However, the alert is limited to general-purpose AI, a limitation which is challenging to understand.

5.4 Consequences

Passing the analysis of the reporting duties, notification duties, and information exchange between the economic actors and between the competent authorities and the European Commission in review, one cannot but be astonished by the complexity of the 'governance structure', the sheer number of authorities and bodies involved, sophisticated distinctions between information which has to be shared on request and reasoned request, and the available information on known and potential risks which are neatly distinguished from serious incidents.

The relationship between the Market Surveillance Regulation and the AIA is far from clear and sometimes even contradictory, for instance, when it comes to distinguishing between risks and serious risks. The AIA puts an enormous burden on the Member States authorities, whoever they are. One may wonder whether the mutual trust between Member States authorities has reached a level to solve problems 'pragmatically' – in a grey legal zone – instead of relying on over-complex rules.[87] This is particularly true regarding information exchange on risks, potential risks, incidents and serious incidents. The EU Treaty emphasises the sovereignty of each Member State's authority to transform shared knowledge into regulatory action. The dissent is not so much on the factual side but on the degree to which the facts justify regulatory action and, if yes, in what form. The AIA inserts two tools: a kind of lead country principle in Article 73 (1) and a 'safety net' in Article 74 (11). Both could heavily reduce legal complexity and advance administrative trust and flexibility.

The country where the infringement '*occurred'* shall remain in charge, even if other Member States might potentially be affected (Article 73 (1). 'Occurred' might well mean 'first discovered', but 'first discovered' may not be the country where the flaw has started. Insofar as it might not be easy to determine the lead country in case of disagreement on 'occurrence'. That is why the rather broadly worded safety net in Article 74 (11) might play a decisive role. Market surveillance authorities and the Commission shall be able to propose joint activities, including joint investigations, to be conducted by either market surveillance authorities or market surveillance authorities jointly with the Commission that have the aim of promoting

[87] Micklitz/Roethe/Weatherill loc. cit. (fn 18).

compliance, identifying non-compliance, raising awareness or providing guidance concerning specific categories of high-risk AI systems that are found to present a serious risk across two or more Member States by Article 9 Reg. 2019/1020. The EAIO shall provide coordination support for joint investigations.

6 Compliance and enforcement through AI

The AIA relies on 'strong' public enforcement mechanisms. However, the proposed interaction between privatised ex-ante conformity assessment and public ex-post enforcement seems inspired more by wishful thinking than complex realities, considering the existing deficiencies in product safety regulation and AI application's complexity, pervasiveness and multiplicity. Whilst little empirical knowledge is available, there is enough evidence to condense the state of affairs in the following way: an overtly complex net of national and European authorities broken down into sectors and legal fields,[88] an EU regulatory framework consisting of partially overlapping legislative instruments and enforcement procedures,[89] understaffed national authorities often lacking competent expertise, tremendous differences in the human and financial resources between the Member States (and consequently, the de facto leading role of a few Member States authorities due to their superior capacities),[90] Underdeveloped spirit of cooperation between Member States' authorities, disagreement on the threshold a risk has to reach to trigger regulatory action.[91]

True, many identified problems result from the division of enforcement competencies in the Treaty, under which the Member States remain primarily responsible for the organisation and the equipment of national enforcement authorities. A delegation of enforcement to the European level or the intensification of cooperation between the Member States (with or without the European Commission's involvement) may exceed the boundaries established in the Treaty. In environmental law, the European Commission has used the infringement procedure to promote the enforcement of EU law by Member States. Such a move is hard for the AIA to imagine, and it does not need to be implemented, as the regulation is directly effective. Therefore, enforcement depends on the willingness of the Member States' authorities. The European Commission may exercise pressure on Member States only by requiring compliance with their obligation to provide the resources necessary to implement the AIA. It is hard to imagine, though, that the European Commission will bring Member States to the CJEU that are not ensuring 'adequate financial and human resources', Art. 70 (3), let alone the lack of a definition of what the requirements imply.

The philosophy or ideology behind the EU Treaty assumes that Member States have equal enforcement capacities and must accomplish the same tasks independently of their different resources and attitudes. All the EU can do is repeatedly claim that the enforcement

[88] A. Ottow, Market and Competition Authorities: Good Agency Principles, OUP 2015.

[89] M. Cantero Gamito/H.-W. Micklitz, Too much or too little? Assessing the Consumer Protection Cooperation (CPC) Network in the protection of consumers and children on TikTok (BEUC, 17-02-23) https://www.beuc.eu/sites/default/files/publications/BEUC-X-2023-018_Assessing_CPC_Network_in_the_protection_of_consumers_and_children_on_TikTok-Report.pdf.

[90] Hard evidence is limited, prior to Brexit the national supervisory authorities competent for the financial supervision looked at the UK authorities, statement by BaFin in interviews, see M. Cantero Gamito/H.-W. Micklitz with regard to TIKTOK (Fn. 89).

[91] Evidence in H.-W. Micklitz/ Th. Roethe/ St. Weatherill (eds.), Federalism and Responsibility (fn. 18) from the early 1990s.

authorities must be adequately equipped, that cooperation should be enhanced, etc. The differences between the Member States are difficult to overcome. A push towards effectively enforcing the AIA may result from integrating AI and other technologies into the enforcement process. The availability of adequate technological support across the EU may increase the effectiveness of enforcement and, on the other hand, contribute to equalising enforcement capacities (to the extent that such tools will be available to all enforcement agencies).

6.1 Hints in the AIA

The AIA views AI as the object to be regulated, not as a tool to enhance compliance, conformity and enforcement: it does not address the opportunity to improve compliance, conformity and enforcement with the help of AI and other information technologies. Many of the tasks to comply with the AIA may benefit from using advanced techniques to support monitoring, risk management, communication, and documentation without the support of AI techniques, with compliance with the AIA risks being impossible or at least very burdensome and costly.

For instance, the risk-management measures described in Article 9 require advanced methods and tools for monitoring, testing and validating AI systems. Similarly, the data and data governance measures in Article 10 call for state-of-the-art techniques, such as identifying biases and shortcomings in the available data. AI techniques, such as generative technologies, could help prepare the documentation required by Article 11 in combination with Annex IV, referring to 'design choice' and 'architecture'. Analysing the logs collected according to Article 12 could profit from AI techniques. Transparency and interpretability Article 13 needs the deployment of AI methods and tools.

AI technologies can contribute to ensuring that AI systems are subject to human oversight: the system itself should be based on technological solutions that provide adequate information to human operators (Article 14). Only in this way can we effectively oversee systems operating at a speed that vastly exceeds our capacities and presents a complexity that makes it impossible for us to make sense of their internal processing. Technological solutions, most likely relying on AI, are required to protect against cyberattacks (Article 15): this is needed to detect threats and target them with appropriate protections timely.

The deployment of advanced solutions is envisaged for quality management (Article 17), such as the following: techniques for designing the system, controlling and verifying the design (b); techniques for quality control and assurance (c); test and validation procedures; (d) systems and procedures for data management, including data acquisition, data collection, data analysis, data labelling, data storage, data filtration, data mining, data aggregation, data retention and any other operation (f).

Documentation keeping is addressed by Articles 18 and 19, which do not mention technological solutions but list a vast amount of documentation to be produced and kept. Technologies are required to support the generation of such documentation, controlling its adequacy, summarising it and extracting answers to specific queries. The documentation requirements in the AIA will generate a vast set of materials. The competent authorities need appropriate NLP (neuro-linguistic programming) and other technologies to effectively analyse such materials and identify and assess the relevant information.

Unique technical solutions are required, according to Article 50, to ensure that users can distinguish content produced by humans and by machines: such technical solutions should be

'effective, interoperable, robust and reliable'. This applies to texts, images and multimedia content.

Technological tools for implementing the AIA could assist providers and deployers, certification and benchmarking bodies and market surveillance authorities: (1) the tools support providers and deployers to comply accurately and cost-effectively with the requirements of the AIA; (2) the tools are helpful to private certification and benchmarking bodies, to competently exercise the controls and tests delegated to them; (3) the tools are needed by competent authorities (the national market surveillance authorities and the other authorities competent at the national level as well as the Commission, the AI office, the AI board, at the EU level) to exercise control, identify violations and provide guidance. Such authorities should not be able to become dependent on their regulators for access to technological support. In the next section, we will further consider what technological tools, particularly AI-based ones, might help enforce the AIA.

6.2 Public – and private - enforcement in the search for support

A superficial reading of the AIA may lead to the idea that the Regulation assumes that a single national enforcement authority oversees supervising AI systems of whatever type and builds alliances with other Member States' authorities to increase the efficiency and effectiveness of market surveillance. However, when looking at the various enforcement mechanisms in the AIA, we have discovered a multitude of actors involved in monitoring and enforcement. At the national level, they include, for each Member state, market surveillance authorities, data protection authorities, sectorial agencies, competition authorities, notifying authorities, and fundamental rights authorities. At the EU level, they include the competent directorates of the European Commission, the EU Data Protection Officer, the sector-related European Agencies in charge of AI, the European Agency of Fundamental Rights, the AI Office, the AI Board, and last but not least the various boards set up in secondary EU law to co-ordinate supervision about particular fields of law relevant for AI, such as the General Data Protection Regulation, the Audiovisual and Media Services Directive and the Regulation of Transborder Enforcement of Consumer Law (in case in of consumer harm). As noted above, several private bodies, such as certification and benchmarking bodies, are also involved. Due to the accountability principle, providers and deployers are at the forefront of the application of the AIA. Finally, there is an opportunity for private enforcement, not only by non-governmental organisations enjoying legal standing but by other societal initiatives working pro bono for public interests.

In such a complex scenario, supervision of the implementation of the AIA involves multiple activities, in particular: (1) supervision of compliance, self-certification, and third-party certification, (2) managing notifications on 'serious incidents' and 'risks' and (3) ex officio investigations. In each area, the 'should be' needs to be distinguished from the 'it is'. The 'should be' refers to the set of rules which establish (1) the requirements of AI systems and the corresponding obligations of providers, deployers, and certification bodies, which, according to the principle of accountability, concerns both substantive compliance (risk prevention) and its demonstration; (2) the preconditions and consequences of notifications; and (3) the authorities' power to investigate ex officio.

In applying such rules, it is necessary to consider the five different layers of rules that determine and concretise AI regulation. These normative requirements are to be matched to the factual reality, which includes, on the one hand, the substantive facts about the

functioning and use of AI systems and, on the other hand, the documentation of such facts that private actors prepare. Both aspects are to be considered by the competent authorities, which usually start by examining the documentation and will move to collect further information only if the documentation is deemed insufficient or unreliable. For instance, assessing the performance of AI systems will start with documents delivered by providers (technical documentation) or certification bodies (compliance reports) and then possibly move to further enquiries on the systems and their usage. Thus, the supervisory activity will first match the 'law' (the five layers) with standardised reports and other documents by private actors. If these reports and documents are unsatisfactory, further facts on the field will be directly considered. As the technical documentation and the compliance reports are available 'on request' (or 'reasoned request'), access to such documentation for public authorities should, in principle, be granted. A different question is whether this documentation is truthful and exhaustive. Similar considerations apply to the detection and assessment of incidents and risks, where the activity of the supervisory authority is generally triggered by the notification provided by providers or deployers.

In this scenario, we can consider various uses of digital technologies, particularly AI, to enhance compliance and enforce the AIA.

The *first* aspect pertains to assessing the functioning of AI systems and determining their risks. Providers/deployers, certification bodies, and public authorities can use various AI techniques at all levels. Techniques are available to verify the functioning of AI systems, particularly to determine accuracy and consistency (calibration) over different data classes.[92] Data analytics technologies can be deployed to monitor current practices and elicit problematic cases. For instance, methods and tools have been developed to assess AI systems' fairness and mitigate any observed unfairness issues. Several diverse metrics for individual and collective fairness have been defined, and multiple methods have been proposed for their implementation so that incompatible outcomes can be obtained according to different metrics and methods. This shows that fairness cannot be reduced to any statistical properties, which can only provide proxies for the essentially contested socio-political idea of fairness (equity) and the contextual applications of this idea. However, only through advanced statistical AI methods is it possible to detect and understand instances of unfairness in AI applications and find ways to remedy or mitigate them by modifying the functioning of such applications.[93] Thus, the legal-ethical assessment of AI systems critically includes technologies deployed to identify and remedy possible unfairness.

A crucial issue in high-risk systems is determining the importance of each risk about their probability and severity, to the extent that uncertainty on both can be overcome, to assess whether adequate mitigation measures have been adopted, and to compare the residual risks and the benefits to be provided by a system to determine whether such risks are acceptable, taking into account the economic viability of the proposed solutions. Such an evaluation cannot be reduced to a merely rhetorical exercise or to simply filling a checklist; it would be

[92] S.A. Seshia, D. Sadigh and S.S. Sastry (2022). Toward verified Artificial intelligence. Commun. ACM, 65(7):46–55.

[93] S. Barocas, M. Hardt and A. Narayanan, A. (2023). Fairness in Machine Learning. Limitations and Opportunities. Mit Press; R. Calegari, G.G. Castané, M. Milano, M., and B. O'Sullivan, B. (2023). Assessing and enforcing fairness in the AI lifecycle. In 32nd International Joint Conference on Artificial Intelligence (IJCAI 2023), pages 6554–6562, Macau, China. IJCAI; F. Lagioia, R. Rovatti and G. Sartor (2023). Fairness through group parties? The case of COMPAS-SAPMOC. AI and Society, pages 1–20.

necessary to build computational tools to support such a complex assessment, inspired by legal theory, multicriteria decision-making, computer simulation and optimisation. Similar tools may also be developed to support the fundamental proper assessment by deployers.[94]

A *second* field pertains to preparing and evaluating the documentation describing the functioning of AI systems and extracting relevant information from such documentation. In this regard, systems for processing natural languages can be helpful, on the one hand, to facilitate the generation of the relevant documentation and, on the other hand, to enable controllers to extract from such documentation the information that is relevant to assess both the feature of the system at stake (according to its description) and the adequacy of the information being provided. A real breakthrough in this area has been provided by LLMs, which are amazingly good at generating information based on specifications (prompts) and extracting information from textual documents.[95] Several contributions already exist, showing how LLMs in the legal domain can be used to summarise cases, prepare drafts of court documents, or extract relevant information by querying textual documents. As these systems provide tentative answers with limited reliability, careful human control over the obtained outcomes is indispensable to ensure their truthfulness and accuracy. The speed of technical improvements in LLMs justifies the hope that better tools will be available soon.

A *third* field pertains to supporting communication and interaction between the parties involved by supporting workflows, suggesting activities and supporting the preparation of questions and responses. In this regard, integrating technologies for workflow management and NLP tools for document analysis and generation may provide helpful support. LLMs, integrated with a chatbot interface, could provide information on the functioning of AI systems. A key issue to be addressed pertains to generating explanations of individual decisions to comply with the obligation established by Article 86. Several techniques can be used for this purpose —from the construction of transparent models to substitute and complement opaque ones to the provision of contrastive explanation to the use of argumentation-based approaches, to self-explanations by generative model— as explainable AI is a key domain not only for research but also for commercial activities.[96] Explanation facilities could be helpful not only for deployers and users of AI systems but also for certification bodies and public authorities engaged in controlling the functioning of such AI systems.

A *fourth* field matches the information obtained by examining the documentation, the system and the context of its usage to legal requirements to determine whether all such requirements are being fulfilled. In this regard, different approaches are to be considered. Some criteria to be respected may be given a computable representation using symbolic knowledge representation techniques (formal models of legal knowledge and reasoning) to detect compliance with them automatically. Consider, for instance, the case in which quantitative benchmarks for fairness or accuracy are violated, some cases in which personal data are unlawfully collected, or some requirements on documentation are not complied with. Implementing this task requires providing formal representations of the normative knowledge to be applied (typically a set of rules plus an ontology of concepts or a set of cases

[94] A. Mantelero, (2022). Beyond Data: Human Rights, Ethical and Social Impact Assessment in AI. Springer.

[95] A. Galassi, F. Lagioia, R. Liepina, P. Palka, G. Sartor, Reconciling Comprehensiveness and Comprehensibility of Privacy Policies: Rethinking the Law and Harnessing the Power of LLMs. In publication.

[96] R. Guidotti, A. Monreale, F. Turini, D. Pedreschi and F. Giannotti, F. (2018). A survey of methods for explaining black box models. ACM Computing Surveys, 51(Article 93):1–42.

represented through their features) and a reasoning method to apply such knowledge to new cases. Datasets of systems cases and features that have been considered unlawful by authorities of legal experts can be used to identify cases that may deserve a similar assessment. For these purposes, techniques based on machine learning can be deployed, possibly in combination with techniques from case-based reasoning. A large dataset of cases must be available to apply machine learning techniques, possibly by merging the cases provided by different authorities. Relevant information can be extracted from such cases by available techniques, from advanced (conceptual) information retrieval to automated summarisation and keyword extraction, identifying trends and emerging patterns, and predicting possible outcomes. While AI tools have had limited success in dealing with legal cases (except for information retrieval), new exciting results have been recently obtained, mainly through LLMs. Techniques for symbolic representation can be integrated with approaches to machine learning so that the first operates based on the knowledge extracted by the latter. For instance, symbolic normative rules are applied to factual features automatically extracted from the available documentation.

The fifth field concerns special AI techniques needed to verify whether violations may be committed through AI systems. Consider copyright violations, which would require comparing outcomes automatically generated with pre-existing works or determining whether copyrighted works have been used to train a system against the will of the right holders.

As an example of a system that performs legal assessments based on textual data, we can mention Claudette, which is used for the automated detection of unfair standard terms and unfair data privacy policies.[97] The legal benchmark has been taken from the respective EU laws, according to which we have produced a tripartite classification of commercial practices –unfair, potentially unfair, and fair terms and practices. The categorisation is inspired by the regulatory approach to the control of unfair standard terms, which operates with blacklisted terms being different from potentially unfair terms. The system has been trained based on classifications provided by legal experts (who built the system's training set). It has proved successful in detecting and classifying unfair clauses in new documents. The experience gained in using Claudette-like techniques in other fields of the law, such as insurance.[98] and tax law,[99] as well as in the handling of different languages[100] May contribute to supporting compliance in the AI domain.

We should not, however, underestimate the significant differences between Claudette's approach to contract terms and the requirements of the assessment of AI systems. First, Claudette only deals with language: it determines whether a particular clause is lawful, but it does not consider whether the actual practice fits the commitment expressed in the clause (whether, for instance, a company processed personal data in ways that go beyond what it declared in its privacy policy). On the contrary, the supervision of AI systems cannot stop at the level of checking the documentation provided by businesses. Supervision must include

[97] F. Lagioia, A. Jablonowska, R. Liepina, K. Drazweski, AI in Search of Unfairness in Consumer Contracts: The Terms of Service Landscape, Journal of consumer policy, 2022, Vol. 45, pp. 481–536

[98] A. Jablonowska, F. Lagioia, R. Liepina, H.-W. Micklitz and G., Automating Compliance Assessment of Insurance Contracts Using AI: Promises and Challenges, C. Poncibo & P. Tereszkiewicz (eds), European Insurance Law in the Age of Digitalisation, Springer AIDA Europe Research Series (to be published 2024).

[99] Th. Dal Pont, F. Galli, A. Loreggia, G. Pisano, R. Rovatti, G. Sartor, Legal Summarisation through LLMs: The PRODIGIT Project, arXiv preprint arXiv:2308.04416.

[100] A. Galassi, F. Lagioia, A. Jablonowska, M. Lippi, Unfair Clause Detection in Terms of Service across Multiple Languages, AI&Law (under review).

verifying whether the documentation is trustworthy. Secondly, Claudette examines standard terms and data privacy policies, which are publicly accessible, as they are part of the overall business strategy. This is not the case for technical documentation delivered by providers, which, on the one hand, is not accessible to the public and, on the other hand, maybe reticent about internal practices of dubious legality. Thus, there is a restriction on the actors that can control the documentation (excluding private actors such as NGOs). On the other hand, there is a limitation in what the documentation tells about the AI application at stake. In applying the Claudette approach to privacy policies, a recent example pertains to extracting information from privacy policies through LLMs using a set of detailed queries. A similar approach could also extract information from documentation on high-risk AI systems and check whether the documentation content complies with the law.

Notwithstanding the limitations, we consider Claudette a promising example of how technologies can enable authorities, NGOs and individual citizens to exercise some control over technology-driven commercial practices. We urge public institutions to support the creation of such technological enforcement tools and to engage in making them available to public and private actors.

7. Conclusion

The AIA is a very complex piece of legislation: the process leading to its adoption, including more than 700 amendment proposals by the European Parliament, led to a hypertrophic document of 144 pages, which merges political considerations, statements of legal principles and detailed procedural and technical prescriptions in a complex entanglement. Compliance by its addressees (developers and deployers of AI systems) and application by EU and national authorities and other bodies (standard and benchmark setting and certification bodies) will not be easy and will require cooperation by everyone. Interpretation and implementation must fit the elusive pervasive nature of AI, its distributed development, its duplicitous and unpredictable dynamics, and its global development and user base.

The key challenge and critical aspect for the success of the Act consists indeed in developing compliance and enforcement processes that ensure the achievement of the goals of the AIA — preventing evitable risks while enabling beneficial uses of AI— and provide an acceptable level of legal certainty, without unduly burdening providers and deployers. We hope that the analysis we have developed here may contribute not only to the early detection of inconsistencies, complexities and indeterminacies that may interfere with the achievement of these goals but also to the identification of possible solutions based on communication and cooperation between the multiple actors having a stake in the implementation of the AIA. To ensure that the activity of this network of public and private actors, with overlapping tasks and competencies, produces consistent outcomes, efficient processes must be put in place, on the one hand, for the generation and verification of information and on the other hand for interaction and communication between all parties. We have argued that AI technologies can make a decisive contribution to this goal, providing much-needed support to human experts.