Int. Cybersecur. Law Rev.
https://doi.org/10.1365/s43439-021-00043-6

Cybersecurity certification of Artificial Intelligence: a missed opportunity to coordinate between the Artificial Intelligence Act and the Cybersecurity Act

Federica Casarosa

Received: 16 November 2021 / Accepted: 1 December 2021

Abstract In April 2021, the Commission published a draft proposal for a regulation on artificial intelligence (AI) systems aimed at striking a balance between the market need for a competitive and dynamic ecosystem and the need to minimise risks to the safety and fundamental rights of users and citizens. Among the set of obligations that apply to high-risk AI technologies, the AI Act includes a specific provision addressing the security and robustness of AI systems. This provision overlaps with existing legislation addressing cybersecurity, namely the certification process defined in Regulation 2019/881 on the European Union Agency for Cybersecurity and on information and communication technology cybersecurity certification. Although the AI Act hints at a possible path towards mutual recognition of certifications, a deeper analysis of the provisions and a comparison between the underlying features of the certification mechanisms show that the different approaches adopted in the two acts may undermine the goal of certification mechanisms as trust-enhancing and transparency instruments. As a result, this paper provides evidence of the missed opportunity for the AIA proposal to link and coordinate in a more structured way with the cybersecurity framework.

Keywords Conformity assessment · Stakeholder participation · Transparency · Certification bodies · Information and communication technologies

Federica Casarosa (✉)
Centre for Judicial Cooperation, European University Institute, Florence, Italy
E-Mail: federica.casarosa@eui.eu

Published online: 05 January 2022

1 Introduction

In April 2021, the Commission published a draft proposal for a regulation on artificial intelligence (AI) systems (AI Act or AIA) as the final step in the creation of the first legal framework applicable to AI, with no comparable examples either at the European or global levels.[1] The intervention follows the principles already set out in the European AI strategy presented in 2018[2] and the White Paper on AI delivered in 2020[3], and complements other legislative initiatives in the digital sector, which include the Data Governance Act[4], the Digital Services Act[5] and the Digital Markets Act.[6]

The Commission's ambitions in this area are twofold: to foster the development of AI technologies in the European Union (EU) attracting regional and foreign market investments, and to ensure a human-centric approach is adopted in the development of AI. Accordingly, the most relevant challenge for the AIA proposal is to strike a balance between market needs for a competitive and dynamic ecosystem and the need to minimise risks to the safety and fundamental rights of users and citizens. In order to achieve this result, the Commission adopts a risk-based approach distinguishing between different types of AI technologies, namely low, high and unacceptable risk technologies.[7]

Excluding the case of AI involving unacceptable risks[8], the AIA proposal provides a set of obligations that apply to high-risk AI technologies[9] that address risk prevention activities (such as the adoption of risk assessment and mitigation systems)[10], quality management systems (applicable particularly to data used to train AI)[11], documentation obligations (addressed both to national authorities and users)[12], organisational features (including the need to ensure human oversight)[13] and finally

[1] See the first reactions to the AIA Proposal by academics and civil society: Burri and von Bothmer (2021); Glauner (2021); Greenleaf (2021); Ebers (2021); BEUC (2021); EDRi (2021); ECNL (2021a).

[2] European Commission (2018).

[3] European Commission (2020a).

[4] European Commission (2020b).

[5] European Commission (2020c).

[6] European Commission (2020d).

[7] Note that the distinction is presented in the Explanatory memorandum to AIA, p. 12. See also Vaele and Zuiderveen Borgesius (2021, p. 3).

[8] See Art. 5 AIA.

[9] The definition of high-risk technologies includes a non-exhaustive list provided in Annex III of the AIA. The annex lists inter alia biometric identification and categorisation of natural persons, AI used in law enforcement activities, in migration, asylum and border control management, and in the administration of justice and democratic processes. Moreover, the AIA leaves open the opportunity for further developments of technologies, affirming that the list may include all the AI that may "pose a risk of harm to health and safety, or a risk of adverse impact on fundamental rights" (Art. 7(1) lit. b).

[10] See Art. 9 AIA.

[11] See Art. 10 AIA.

[12] See Art. 11 and 12 AIA.

[13] See Art. 14 AIA.

security and robustness obligations.[14] All these obligations are to be verified through a process of conformity assessment that allows the AI system to be put on the market or put to use.[15]

The provision addressing the limits of conformity assessment in the AIA proposal has already been criticised by several civil society organisations.[16] The main problematic issue emerging from this analysis is the absence of an ex ante evaluation of the impact of AI systems on human rights, rule of law and democracy, which would improve the existing distinction between high-risk and low-risk AI.[17] This paper, however, does not address the fundamental rights perspective. Instead, it focuses on a different and specific feature of the AIA certification mechanism, which shows another case of insufficient coordinating mechanisms between the current AIA proposal and the existing applicable legislation.\mathrm{

In fact, the AIA proposal overlaps with existing legislation addressing cybersecurity, namely the certification process defined in Regulation 2019/881 on the European Union Agency for Cybersecurity (ENISA) and on information and communications technology cybersecurity certification (Cybersecurity Act or CSA).[18] Art. 42(2) AIA acknowledges that high-risk AI systems that have been certified under a cybersecurity scheme created according to the process provided by the Cybersecurity Act "*shall be presumed to be in compliance with the cybersecurity requirements set out in Article 15 of this Regulation*". Therefore, the AIA proposal hints at a possible path towards mutual recognition of certifications. However, a deeper analysis of the provisions in the AIA proposal and a comparative exercise with those provided in the CSA shows that the different approaches adopted in the two acts may undermine the aim of certification mechanisms as trust-enhancing and transparency instruments.

This paper briefly presents the objectives and advantages of the certification mechanisms (). It then highlights the features of the certification mechanism provided in the Cybersecurity Act, addressing not only the actors involved in the drafting and enforcement process but also the substantive factors that may be addressed by the certification. Then, the certification mechanism adopted in the AIA proposal is presented.[19] The following section provides a comparison between the two certification mechanisms, showing that among few similarities some substantial differences emerge. Finally, some tentative conclusions are provided.

As a result, the paper provides evidence of a missed opportunity for the AIA proposal to link and coordinate in a more structured way with the cybersecurity framework.

[14] See Art. 15 AIA.

[15] See Art. 43 AIA.

[16] See ECNL (2021b); EDRi (2021) and Article 19 (2021); Access now (2021).

[17] Note that Art. 29(6) AIA requires a data protection impact assessment, although it does not cover the larger number of fundamental rights that could be affected by AI decisions. See Access now (2021, p. 22); ECNL (2021b).

[18] Regulation (EU) 2019/881 on ENISA and on information and communications technology cybersecurity certification and repealing Regulation (EU) No 526/2013 (Cybersecurity Act).

[19] The paper does not claim to provide a systematic analysis of AIA certification frameworks, which means that some procedural or other aspects are not included in the analysis. In particular, analysis of the market control mechanisms in Arts. 61–68 AIA is not addressed in the paper.

2 Certification mechanisms

A certification scheme is a set of practices that involves at least two phases, a conformity assessment and an attestation of conformity, the latter being a statement that the underlying process, product or person complies with a set of pre-defined requirements that are identified on the basis of the objectives and reach of each certification scheme.[20] The requirements are usually related to safety[21], health and environmental protection[22], and only in a few cases is the human rights dimension taken into account.[23]

There are different types of certification that may involve self-assessment or the involvement of certification authorities. For instance, the definition provided by ISO affirms that the conformity assessment in all their certification schemes is to be carried out by a third party[24] through a formal review that ensures fulfilment of the product/process/person requirements.[25] The request to be certified is usually voluntary, although as soon as the certification process is started compliance with the requirements becomes binding for the company requesting certification.

The final attestation of conformity usually allows the certified producer, service provider or person to apply a seal or a label to their product/process. Although the certification cannot equate to insurance to third parties that the product or process meets the requirements[26], it still signals the quality of the certified item, distinguishing it from those which do not comply with the certification requirements. This is crucial in any sector where third parties (e.g. customers, consumers or users) are not able, due to time constraints, lack of expertise or mere unwillingness, to check in advance the existence of specific requirements that may influence their (bargain-

[20] For an analysis of the different definitions of certification, see Lachaud (2019).

[21] See Directive 2009/48/EC on the safety of toys. The directive provides a set of standards addressing physical and mechanical properties, flammability, chemical properties and electrical properties in order for toys to receive the CE certification mark.

[22] See the case of the Forest Stewardship certification, which provides a voluntary process for verifying responsible forest practices that include not only taking into account the environmental impact of logging activities and the maintenance of the ecological functions and integrity of forests but also includes criteria regarding recognition and respect for indigenous people's rights, respect for worker's rights in compliance with the International Labour Organisation (ILO) convention and equitable use and sharing of benefits derived from forests.

[23] See the voluntary due diligence requirements provided in Regulation (EU) 2017/821 laying down supply chain due diligence obligations for Union importers of tin, tantalum and tungsten, their ores and gold originating from conflict-affected and high-risk areas. The regulation imposes an obligation on EU importers of tin, tantalum, tungsten and gold to verify whether goods purchased from third countries contributed to forced labour or other illicit activities.

[24] Third party assessment is performed by a party different to the organisation that seeks certification (first party) and different to the entity requiring the organisation to be certified (second party). For a detailed description of different types of certification, see Daskalova and Heldeweg (2019).

[25] See ISO/IEC 17000:2004 subclause 5.5. Note that ISO identifies different methods of conformity assessment such as testing, inspection (usually onsite assessments of the conformity of product samples or their production processes), sampling and audit (regarding the conformity of management systems).

[26] Note that the attestation of conformity can have a time limitation, which usually ranges from 5 to 10 years, in order to take into account subsequent sector-specific developments. In some cases, the certification can be renewed if conformity has been maintained for an equal duration.

ing) decisions.[27] Certification thus becomes a competitive advantage for the certified company signalling a recognised level of quality.

In EU law there are several areas where certification mechanisms have been adopted and have flourished. First and foremost, the CE (*Conformité Européene*) mark is one of the best known cases[28], but also in the food production chain organic certification can be taken as an example.[29] Most recently, the Commission fostered the use of certification in the digital market by adopting the above-mentioned Cybersecurity Act, and also including certification schemes in the General Data Protection Regulation.[30]

The following section describes the certification mechanism provided in the Cybersecurity Act and compares it to that provided in the AIA proposal in order to verify whether and how the two mechanisms can be deemed interchangeable in terms of the accuracy, robustness and cybersecurity of AI systems.

3 The cybersecurity certification mechanism

The CSA[31] entered into force in 2019 with the aims of enhancing the level of cybersecurity protection in the EU and strengthening the effective protection of citizens' rights.[32] The regulation addressed two main issues: on the one hand, the role and tasks allocated to ENISA and, on the other hand, the introduction of a cybersecurity certification scheme. The objective of the CSA was to enhance cybersecurity capacity and resilience as a reaction to previous cyberattacks that showed an inability of the legal framework to quickly respond to such threats.[33]

One of the strategies adopted to achieve this objective was to adopt a unified procedure that would allow ENISA to publish European certification schemes that can be adopted by any business in Europe. The definition of the certification scheme pursuant to Art. 2(9) CSA provides that:

[27] See the first formulation of this approach to the market in Viscusi (1978).

[28] See Regulation (EC) No 765/2008 setting out the requirements for accreditation and market surveillance relating to the marketing of products and repealing Regulation (EEC) No. 339/93. Note that Art. 49 AIA provides that the CE marking can be affixed visibly, legibly and indelibly for high-risk AI systems that comply with the requirements provided in the proposed AIA.

[29] See Regulation (EU) 2018/848 on organic production and labelling of organic products and repealing Council Regulation (EC) No 834/2007.

[30] See Art. 42 of Regulation (EU) 2016/679 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation). On the GDPR certification scheme, see Lachaud (2018); Rodrigues et al. (2016); Hornung and Bauer (2019).

[31] See Regulation (EU) 2019/881 on ENISA and on information and communications technology cybersecurity certification and repealing Regulation (EU) No 526/2013.

[32] See also European Commission (2013).

[33] A well-known case was Wannacry, which was a ransomware cyberattack that affected computers running a Microsoft Windows OS in May 2017. For more, see https://en.wikipedia.org/wiki/WannaCry_ransomware_attack, accessed 11 October 2021. The attack was able to infect 200,000 computers in 150 countries, with an economic impact of around $4 billion.

> "European cybersecurity certification scheme" means a comprehensive set of rules, technical requirements, standards and procedures that are established at Union level and that apply to the certification or conformity assessment of specific information and communication technology (ICT) products, ICT services or ICT processes.

The process to define European certification schemes does not discourage national certification schemes defined in Art. 2(10) CSA. However, it strongly pushes towards a shift from existing national and international schemes and mutual recognition systems towards the European cybersecurity certification framework, avoiding "*multiplication of conflicting or overlapping national cybersecurity certification schemes and thus reduc[ing] costs for undertakings operating in the digital single market*"[34].

Although the CSA acknowledges the need for a common approach in the Union and horizontal requirements for European cybersecurity[35], it does not adopt a bottom-up approach for the development cycle of cybersecurity certification. In fact, the procedure is clearly centralised and based on a triggering role of the Commission and ENISA.[36]

According to Arts. 47–49 CSA, the initiator of the procedure to develop a certification scheme is the EU Commission on the basis of the Union rolling work

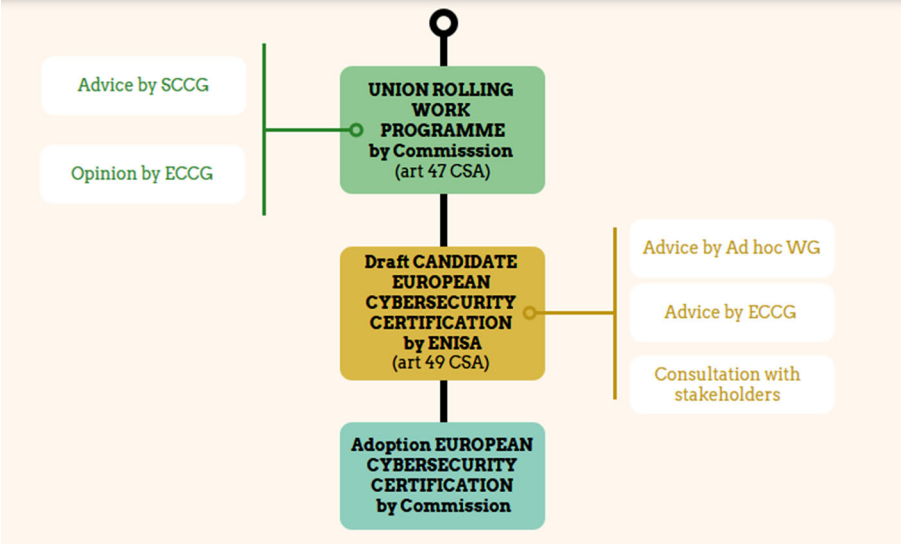

Fig. 1 The drafting procedure for cybersecurity certification. *SCCG* Stakeholder Cybersecurity Certification Group, *ECCG* European Cybersecurity Certification Group, *CSA* Cybersecurity Act, *ENISA* European Union Agency for Cybersecurity, *WG* Working Group

[34] Recital 69 CSA.

[35] Weber and Studer (2016).

[36] Kamara (2020).

programme. This programme is an annual Commission publication outlining strategic priorities for cybersecurity certification schemes and it takes into account *inter alia* the risk of fragmentation emerging from the existence of overlapping national cybersecurity certification schemes, market demand and developments in the cyber threat landscape. The Commission then requests ENISA to draft a candidate scheme, setting very specific goals and requirements such as the subject matter and scope of the scheme, the types or categories of ICT products, systems and services covered, the purpose of the scheme and references to technical standards and specifications etc.[37] These requirements are to be strictly followed by ENISA in order to ensure coherence and uniformity of the certification scheme structure, taking into account differences that may clearly emerge depending on the scope, sector and context of each scheme (see Fig. 1).[38]

As emerges from Fig. 1, although the trigger for the development of a certification scheme is centralised, there are opportunities for stakeholders to contribute to elements that will be included in the scheme. First, the Stakeholder Cybersecurity Certification Group (SCCG)[39] provides advice and comments on the opinion of the Commission's Union Rolling Work Programme, suggesting the most relevant areas where a certification scheme should be provided and highlighting the need for coordination among the proposed certification schemes.[40] Moreover, ENISA sets up an Ad Hoc Working Group (AHWG) to support the preparation of a candidate EU cybersecurity certification scheme involving selected members representing industry, participants from accreditation bodies and EU Member States.[41] The AHWG directly collaborates with ENISA in the development of a draft certification scheme, reviewing existing certification schemes in the same area and helping ENISA to identify the precise scope of the candidate scheme and the pre-qualification of elements that ENISA needs to include in the scheme.

Finally, the draft certification scheme is subject to an open consultation with all stakeholders, with comments and suggestions to improve the draft scheme before its adoption being invited.

[37] A non-exhaustive list of requirements is included in Art. 54 CSA.

[38] See the example of the draft version of the Certification Scheme for Cloud Services, ENISA (2020a).

[39] According to Art. 22 CSA, the SCCG is composed of members selected from recognised experts representing relevant stakeholders (e.g. standardisation bodies, producers, providers, consumer associations, etc.). See the current members at https://digital-strategy.ec.europa.eu/en/library/stakeholder-cybersecurity-certification-group, accessed 11 October 2021.

[40] See the SCCG's opinion on the Draft Union Rolling Work Programme adopted 19-02-2021, available at https://digital-strategy.ec.europa.eu/en/library/stakeholder-cybersecurity-certification-group, accessed 11 October 2021.

[41] According to Rec. 59 CSA, the members of the ad hoc working groups are selected according to the highest standards of expertise, aiming to ensure an appropriate balance according to the specific issues in question between the public administrations of the Member States, Union institutions, bodies, offices and agencies, and the private sector, including industry, users and academic experts in network and information security.

For instance, the Ad Hoc Working Group set up for the preparation of a candidate EU cybersecurity certification scheme on cloud services was composed of 20 industry representatives (e.g. cloud service providers, cloud service customers, conformity assessment bodies) and 12 representatives from accreditation bodies and EU Member States. See ENISA (2020a, p. 4).

From a substantial point of view, an additional element included in each certification scheme is the so-called assurance level, which may be basic, substantial or high.[42] Each assurance level takes into account the resilience of the ICT product, process or service to potential security threats either based on past experience or potential vulnerabilities. For instance, the three levels of assurance take into account different types of cybersecurity attackers: in the recently published European Cybersecurity (candidate) Certification Scheme for Cloud Services, the typical attacker profile for the basic level is a single person with limited skills repeating a known attack with limited resources; for the substantial level it is a small team of persons with hacking abilities and access to a wide range of known hacking techniques, including social engineering, but with limited resources, in particular to launch wide attacks or to discover previously unknown vulnerabilities; and for the high level it is a team of highly skilled persons with access to significant resources to design and carry out attacks, get insider access and discover or buy access to previously unknown vulnerabilities.[43]

Once the certification scheme is adopted by the Commission, the operational cycle starts by involving a set of country-based authorities in charge of accreditation and certification of products, processes and services devised by manufactures and providers.[44] Art. 58 CSA requires each Member State to designate one (or more) financially and institutionally independent authority in charge of enforcing the rules included in European cybersecurity certification schemes and monitoring the compliance of ICT products, services and processes with the requirements of the European cybersecurity certificates. Accordingly, the certification authorities enjoy both investigative and enforcement powers allowing them to carry out investigations (i.e. audits) of conformity assessment bodies, European cybersecurity certificate holders and issuers of EU statements of conformity to verify their compliance[45], and in cases of infringement to impose penalties in accordance with national law.[46]

The validity of certifications depends on the specific features of the scheme, usually with a maximum duration of five years. This is due to the fact that development of technology in this sector is extremely rapid, and revisions of the scheme may be deemed necessary in shorter timeframes in order to avoid obsolescence. In fact, we may distinguish between administrative validity, which is the duration of the certificate provided by the initial certification scheme, and technical validity, which is instead the capability of the underlying procedures, processes, etc. to deliver the underlying objective of cybersecurity, i.e. resistance of the certified product to attack. Accordingly, the CSA acknowledges a need for periodic reviews of certification schemes adopted, and consequently the need for certified products and services to be subject to subsequent scrutiny when they aim to maintain the certification seal.

[42] See Art. 52 (6) and (7) CSA.

[43] See ENISA (2020a, pp. 19–20).

[44] Note that Art. 53(1) allows also for a conformity self-assessment to be carried out by the manufacturer or provider itself. However, this option is available only for the basic assurance level.

[45] See Art. 58 (8) (b) CSA.

[46] See Art. 58 (8) (f) CSA.

It is important to acknowledge that the CSA does not preclude the possibility for producers and service providers to apply for the release of a certificate in any Member State without any territorial limitation due to the jurisdiction where the main establishment is settled. Although this feature may entail the risk of forum shopping, with companies deciding to apply in countries where the certification body is (presumably) more lenient, the fact that the certification scheme is a European one subject to the same requirements and obligations should guarantee that the results of conformity assessments are fair and comparable.[47]

4 The certification mechanisms adopted in the AIA proposal

The AIA proposal sets up a detailed organisational structure requiring Member States to create new national notifying authorities responsible for setting up and carrying out the necessary procedures to assess, designate and notify conformity assessment bodies and to monitor them.[48] Moreover, it provides that high-risk AI systems may receive certificates from notified bodies acknowledging the positive results of conformity assessment procedures[49], and accordingly the AI system may be included in the EU database of stand-alone high-risk AI systems.[50]

It is clear that the requirements defined in Arts. 8–15 provide a set of procedures and obligations that each AI system developer and manufacturer should take into account. The Explanatory Memorandum to the proposed regulation affirms that the legal requirements are a minimum standard based on the state of the art for AI operators.[51] The requirements take into account views and recommendations not only from the High Level Expert Group on AI[52] but also from the stakeholders that were involved in the AI Alliance[53] and from those commenting on the 2019 AI White Paper.[54]

Nonetheless, the wide variety of high-risk AI systems covered by the proposed regulation requires a more detailed analysis of the risks and specific technical obligations which should consider the technological and scientific progress that characterises this field.

[47] In addition, the requirement for a peer review mechanism which subjects all national cybersecurity authorities to assessments by their 'peers', i.e. the competent authorities of other Member States, is a further safeguard against forum shopping.

[48] See Art. 30 AIA.

[49] See Art. 44 AIA. Note that the certificates may be valid for a maximum of five years and can be renewed based on a re-assessment in accordance with the conformity assessment procedures applicable.

[50] See Art. 61 AIA.

[51] See the Explanatory Memorandum, p. 8.

[52] The High-level expert group on artificial intelligence is a task force appointed by the European Commission to provide advice on its artificial intelligence strategy. For more, see https://digital-strategy.ec.europa.eu/en/policies/expert-group-ai, accessed 11 October 2021.

[53] The AI Alliance is a multi-stakeholder forum launched in June 2018. For more, see https://ec.europa.eu/digital-single-market/en/european-ai-alliance, accessed 11 October 2021.

[54] See the detailed list of stakeholder events and activities in Annex 2 of the Regulation Impact Assessment.

It should be mentioned that Art. 40 AIA pinpoints harmonised standards that are adopted according to the procedure for technical standardisation.[55] However, no standard is available at the moment.[56] In the absence of harmonised standards, Art. 41 AIA provides the possibility for the Commission to adopt common (technical) specifications that should clarify the requirements set out in Arts. 8–15 AIA. In this case the procedure is only sketched in the article: the responsibility for defining the common specification is allocated to the Commission through the creation of an internal committee.[57] This should "gather the views of relevant bodies or expert groups established under relevant sectorial Union law"[58]. An advisory role is also allocated to the newly created European Artificial Intelligence Board, which shall issue opinions, recommendations or written contributions on the use of harmonised standards or common specifications.[59]

Throughout the proposed regulation, no additional element is provided regarding the process to adopt common specifications.

Before any AI system is put on the market, the AI system providers should follow a conformity assessment procedure, which can be in the form of self-assessment or with the involvement of a notified body.[60] For the AI system providers that fall into the categories that are listed in the proposed regulation's Annex III, namely high-risk ones, the conformity assessment can be in the form of self-assessment, whereas for those that do not fall into the relevant categories, when harmonised standards or common specifications are lacking the self-assessment procedure cannot be carried out.

Conformity assessments in the form of self-assessment and those carried out by notification bodies are described in the proposed regulation's Annexes VI and VII. Both focus on compliance with a quality management system and technical documentation determined according to Art. 17 and Annex IV respectively of the proposed regulation, which pinpoint the requirements set out in Arts. 8–15 AIA.[61]

The notified bodies can require evidence or ask for tests to assess the conformity of their AI, or they can carry out tests directly.[62] Moreover, notified bodies can access the premises where designing, developing and testing of the AI systems are taking

[55] Rec. 61 AIA points to the importance of standardisation and the application of Regulation (EU) No 1025/2012 on European standardisation.

[56] Note that in March 2021, CEN and CENELEC established a new Joint Technical Committee 21 on 'Artificial Intelligence', which will be in charge of developing standards for AI. See https://www.cencenelec.eu/news/brief_news/Pages/TN-2021-013.aspx, accessed 11 October 2021.

[57] Art. 41(1) AIA clarifies that the Commission will adopt the common specification in accordance with the examination procedure ex Art. 5 of Regulation 182/2011.

[58] See Art. 41(2) AIA.

[59] See Art. 58 AIA. Note that the composition of the Board includes only representatives of the national regulatory authorities, with the possibility of involving external experts and observers only on request. Therefore, the European AI board also does not adopt a more inclusive approach to potential stakeholders.

[60] See Art. 43(1) AIA.

[61] As previously mentioned, Art. 40 AIA acknowledges that compliance with harmonised standards should be presumed to be in conformity with the regulation's requirements.

[62] Note that point 4.5 of Annex IV AIA provides that on a reasoned request the notified authority shall also be granted access to the source code of the AI system.

Table 1 The Artificial Intelligence Act (AIA) proposal on Artificial Intelligence (AI) certification structure

Standard	Assessment mechanism
Harmonised standards	No conformity assessment needed
Common specifications[a]	Only self-assessment
AIA requirements set out in Arts. 8–15	For high-risk AI systems covered by Annex III: only self-assessment
	For all other high-risk AI systems: conformity assessment by notified bodies
Cybersecurity certification	No conformity assessment needed (only for the requirements in Art. 15 AIA)

[a]An interesting element is the fact that the proposed regulation implicitly acknowledges that common specifications are voluntary as it allows providers to justify the adoption of alternative technical solutions that are at least equivalent to that included in the common specifications. See Art. 41(4) AIA

place, and carry out periodic audits to ensure that conformity is maintained over time. After a positive result of the conformity assessment, the AI provider receives a certificate allowing it to put the AI system on the market or to put it into service.

From the above description, it is possible to affirm that the AIA proposal sets a complex structure for AI certification. This is summarised in Table 1.

5 A comparison between the two certification systems

The certification systems presented above are clearly different yet they address overlapping technical issues that AI system providers are obliged to address in the development of their technologies. Anticipating cybersecurity threats is crucial for responsiveness, robustness and resilience of AI systems. Therefore, cybersecurity risks have to be included in the risk management system defined in Art. 9 AIA and in the technical documents defined in Annex III.[63]

Cybersecurity (jointly with accuracy and robustness) requirements are set out in Art. 15 AIA, providing in particular that AI systems should "*be resilient as regards attempts by unauthorised third parties to alter their use or performance by exploiting the system vulnerabilities. [...] The technical solutions to address AI specific vulnerabilities shall include, where appropriate, measures to prevent and control for attacks trying to manipulate the training dataset ('data poisoning'), inputs designed to cause the model to make a mistake ('adversarial examples'), or model flaws*"[64].

This is a very general set of requirements which does not do justice to the wide number of potential threats that can be envisaged in the development of AI

[63] Moreover, AI can play a major role in cybersecurity not only as a subject, i.e. as a technology that should be made responsive, robust and resilient to cyberthreats, but also as a tool able to ensure the responsiveness, robustness and resilience of other technologies. See Taddeo, McCutcheon and Floridi (2019).

[64] See also rec. 51 AIA, affirming "Cyberattacks against AI systems can leverage AI specific assets, such as training data sets (e.g. data poisoning) or trained models (e.g. adversarial attacks), or exploit vulnerabilities in the AI system's digital assets or the underlying ICT infrastructure. To ensure a level of cybersecurity appropriate to the risks, suitable measures should therefore be taken by the providers of high-risk AI systems, also taking into account as appropriate the underlying ICT infrastructure".

systems. A comparison between Art. 15(4) AIA and the recent ENISA study on AI cybersecurity risks[65] clearly shows the level of detail that the risk analysis should achieve. The main risks that the ENISA report identifies are the following:

- Nefarious activity/abuse
- Eavesdropping/interception/hijacking
- Physical attacks
- Unintentional damage
- Failures or malfunctions
- Outages
- Disasters
- Legal[66]

This macro-threat taxonomy is then developed in more detail, distinguishing 74 specific threats that are included in the macro-categories. It is interesting to note that the report provides a set of maps situating threats on the basis of the actors involved, the lifecycle phase of AI development and the potential impact vis-à-vis the properties of AI systems.

Accordingly, it is more than probable that cybersecurity risks will have to be defined in more detail either by the AI system developer or by the Commission in the common specifications to be adopted pursuant to Art. 41(1) AIA. Therefore, it may be that in the near future there will be an option for an AI system provider to decide among different options: (1) internally defined requirements set on the basis of Art. 15 AIA; (2) harmonised standards pursuant to Art. 40 AIA; (3) common specifications defined by the Commission regarding cybersecurity risks pursuant to Art. 41 AIA; (4) CSA-based certification. Apart from the case of harmonised standards that may be developed following the International Organization for Standardization (ISO) or European Committee for Standardization/European Committee for Electrotechnical Standardization (CEN/CENELEC) standardisation process, it is interesting to verify whether the certification mechanisms set out in the CSA and AIA proposals are really comparable in terms of guarantees of a thorough analysis of the most updated technologies and of showing competitive advantages vis-à-vis other non-certified technologies. Table 2 below clarifies the specificities that should be taken into account when comparing the options available.

The certification mechanisms adopted in the AIA proposal and in the CSA show clear convergences regarding the actors which are in charge of implementing the certification schemes. In both cases governance is organised around national supervisory authorities and certification bodies qualified as notified bodies in the AIA proposal framework. The certification authorities have monitoring and supervision power. In both cases this governance structure is aimed at ensuring the competence of conformity assessment bodies, which should have sufficient expertise and knowledge in the sector in order to verify the compliance of AI system developers.

[65] Note that ENISA (2020b), on the other hand, acknowledges the interdependencies between AI and cybersecurity, distinguishing three main dimensions: cybersecurity for AI; AI to support cybersecurity; and malicious use of AI.

[66] See ENISA (2020b, p. 27).

Table 2 A comparison between the Artificial Intelligence Act (AIA) proposal and the Cybersecurity Act (CSA) certification mechanisms

Scheme/ features	AIA certification		CSA certification
	Art. 15 requirements	Common specifications	
Geographical scope	Company only	EU level	Only EU level (no certification at national level)
Voluntary/ mandatory	Mandatory	Voluntary (differences to be justified)	Voluntary in principle, mandatory possible in MS
Minimum scheme content	Only sketched out in Art. 15 AIA	Only sketched out in Art. 15 AIA	Generally provided in CSA
Granularity	None	Not determined. At the discretion of the Commission committee	Three assurance levels (basic—substantial—high) for certification
Subject in charge of drafting	AI system developer	Commission internal committee	ENISA
Stakeholder involvement	None	Requested but not formalised	Advice by Ad hoc WG Consultation with stakeholder
Conformity assessment	For AI systems covered by Annex III: only self-assessment. For all other AI systems: conformity assessment by notified bodies	Only self-assessment	Conformity assessment by national certification authorities. NB self-assessment only for basic level certification
Certification authority	Notified bodies	Notified bodies	Certification authorities
Revision	Not provided	Every 5 years	Every 5 years, evaluation by ENISA

WG Working Group, *MS* Member States, *ENISA* European Union Agency for Cybersecurity

On the other hand, there are several differences between the certification mechanisms, which are partially related to the legal areas where the two regulations are applicable: the CSA case is a regulation that provides for organisational and procedural provisions applicable to any technology that can be developed in the future and have as a common denominator the need to consider and prevent any potential cybersecurity risk or threat. The AIA proposal is a regulation that sets up a wide framework with principles and obligations for AI system developers not only in the cybersecurity area, which should provide a starting point for the criteria to be adopted.

However, the level of detail in the obligations addressing cybersecurity is very limited in the AIA proposal, leaving wide discretion to AI system developers, and consequently gives great responsibility to notified bodies regarding appraisal of the solutions adopted for cybersecurity at the company level. The alternative of common specifications adopted by the Commission in this area should be welcomed as it may provide common ground that would at least cover an even playing field for market

actors. Moreover, companies will have the opportunity to decide whether to follow (and comply with) the common specifications defined by the Commission or to adopt their own specifications on the basis of Art. 15 AIA. In the former case, notified bodies will find in the common specifications a point of reference in order to evaluate the choices of AI system developers.

Nevertheless, there are limits emerging from the common specification procedures, as their development is not well defined. The AIA proposal relies on the work of the Commission's internal committee, and no clear involvement of external experts or civil society stakeholders is defined. This is a step back from the continual dialogue with stakeholders and experts who were directly engaged in the AIA drafting process. Although the CSA procedure for developing any certification scheme is a centralised one[67], the involvement of stakeholders (including not only experts and market actors but also certification and accreditation authorities) represents added value for the capability of the certification scheme to be updated and in line with the needs of the sector.

Another important difference lies in the conformity assessment procedures: the CSA allows for self-assessment only when the certification scheme provides for a basic level, which is the case for low-risk ICT products, services and processes. Otherwise, the certification authorities are in charge of verifying compliance with the requirements set in the certification scheme. The AIA proposal, on the other hand hand, not only relies on the responsibility of AI system developers to identify and deal with cybersecurity risks but it also fully trusts AI system developers to assess whether or not the solutions adopted are sufficient (or comply with common specifications) in order to put the product on the market.

6 Conclusion

Certification mechanisms are instruments aimed at offering transparency and increasing trust in the certified object and organisation, as they may reduce information asymmetry vis-à-vis citizens, organisations and businesses, offer insights for audit purposes and assure the reliability of the AI system developer. Cybersecurity is an area where certification has been deemed crucial in order to improve the robustness, response and resilience of AI systems against third-party attacks.

Analysis of the AIA proposal and the CSA show that the two certification frameworks follow a similar approach regarding the actors in charge of monitoring and supervising the implementation of the certification scheme. However, the similarities end there. In particular, the two regulations do not fully align when looking at the level of autonomy of AI system developers in setting the process and procedures to respond to cybersecurity attacks and the type of conformity assessment to be adopted to verify compliance with the principles and requirements set out in the certification scheme. These elements show that the risk approach adopted by the Commission in drafting the proposed AIA is not fully developed, or at least is not comparable to that adopted in the CSA: a high risk AI system developer in the AIA proposal can

[67] Kamara (2020).

perform self-assessment, whereas this does not apply to CSA certification as this option is only available for the basic level in the certification scheme, i.e. when the (cybersecurity) risk is deemed low. If the overall objective of the proposed AIA is to increase the trustworthiness of high-risk AI systems provision a more stringent certification process is needed.

These differentiations may affect both choices by AI system developers and users of these technologies when certification schemes are available, with the effect of undermining the goal of certification as a trust and harmonisation instrument.

Acknowledgements This research was supported by the ERDF project "CyberSecurity, CyberCrime and Critical Information Infrastructures Center of Excellence" (No. CZ.02.1.01/0.0/0.0/16_019/0000822) in the framework of the visiting fellowship at Mazaryk University. The author would like to thank Dianora Poletti, Francesca Fanucci and Radim Polcak for their comments and suggestions in earlier drafts of the paper. The usual disclaimer applies.

References

Cited literature

Access now (2021) Access now's submission to the European Commission's adoption consultation on the artificial intelligence act. https://www.accessnow.org/cms/assets/uploads/2021/08/Submission-to-the-European-Commissions-Consultation-on-the-Artificial-Intelligence-Act.pdf. Accessed 11 Oct 2021

Article 19 (2021) EU: New proposal on artificial intelligence must protect human rights. https://www.article19.org/resources/eu-artificial-intelligence-and-human-rights/. Accessed 11 Oct 2021

BEUC (2021) Regulating AI to protect consumers—Position paper on the AI act. https://www.beuc.eu/publications/beuc-x-2021-088_regulating_ai_to_protect_the_consumer.pdf. Accessed 11 Oct 2021

Burri T, von Bothmer F (2021) The new EU legislation on artificial intelligence: a primer. https://ssrn.com/abstract=3831424. Accessed 11 Oct 2021

Daskalova V, Heldeweg M (2019) Challenges for responsible certification in institutional context: the case of competition law enforcement in markets with certification. In: Rott P (ed) Certification—Trust, accountability, liability. Springer, Cham, pp 23–71

Ebers M (2021) Standardizing AI—The case of the European Commission's proposal for an artificial intelligence act, in the Cambridge handbook of artificial intelligence: global perspectives on law and ethics. https://ssrn.com/abstract=3900378. Accessed 11 Oct 2021

ECNL (2021a) ECNL position statement on the EU AI Act. https://ecnl.org/news/ecnl-position-statement-eu-ai-act. Accessed 11 Oct 2021

ECNL (2021b) Evaluating the risk of AI systems to human rights from a tier-based approach. https://ecnl.org/news/evaluating-risk-ai-systems-human-rights-tier-based-approach. Accessed 11 Oct 2021

EDRi (2021) European Commission adoption consultation: Artificial intelligence act. https://edri.org/our-work/edri-submits-response-to-the-european-commission-ai-adoption-consultation/. Accessed 11 Oct 2021

ENISA (2020a) European Cybersecurity (candidate) certification scheme for cloud services published in December 2020 subject to public consultation. https://www.enisa.europa.eu/publications/eucs-cloud-service-scheme. Accessed 11 Oct 2021

ENISA (2020b) AI cybersecurity challenges—Threat landscape for artificial intelligence. https://www.enisa.europa.eu/publications/artificial-intelligence-cybersecurity-challenges. Accessed 11 Oct 2021

European Commission (2013) Joint communication to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, cybersecurity strategy of the European Union: an open, safe and secure cyberspace. COM Join/2013/01/final

European Commission (2018) Communication on the strategy for artificial intelligence in Europe. Brussels. COM 2018:237

European Commission (2020a) White paper on AI. Brussels. COM 2020:65

European Commission (2020b) Proposal for a Regulation of the European Parliament and of the Council on European data governance (Data Governance Act). Brussels. COM/2020/767 final

European Commission (2020c) Proposal for a regulation of the European Parliament and of the council on a single market For digital services (digital services act) and amending directive 2000/31/EC. Brussels. COM/2020/825 final
European Commission (2020d) Proposal for a Regulation of the European Parliament and of the council on contestable and fair markets in the digital sector (digital markets act). Brussels. COM 2020:842 (final)
Glauner P (2022) An assessment of the AI regulation proposed by the European commission. In: Ehsani S et al (ed) The future circle of healthcare: aI, 3D printing, longevity, ethics, and uncertainty mitigation. Springer, Cham (Forthcoming)
Greenleaf G (2021) The 'Brussels effect' of the EU's 'AI Act' on data privacy outside Europe. 171 Privacy laws & business international report, vol 1, pp 3–7
Hornung G, Bauer S (2019) Privacy through certification?: The new certification scheme of the general data protection regulation. In: Rott P (ed) Certification—trust, accountability, liability. Studies in European economic law and regulation, vol 16. Springer, Cham, pp 109–131
Kamara I (2020) Misaligned Union laws? A comparative analysis of certification in the cybersecurity Act and the general data protection regulation, TILT Law & Technology Working Paper No. 002/2020. https://ssrn.com/abstract=3732846. Accessed 11 Oct 2021
Lachaud E (2019) What could be the contribution of certification to data protection regulation? Dissertation, Tilburg University
Lachaud E (2018) The general data protection regulation and the rise of certification as a regulatory instrument. Comput Law Secur Rev 34(2):244–256
Rodrigues R, Barnard-Wills D, De Hert P, Papakonstantinou V (2016) The future of privacy certification in Europe: an exploration of options under article 42 of the GDPR. Int Rev Law Comput Technol 30(3):248–270
Taddeo M, McCutcheon T, Floridi L (2019) Trusting artificial intelligence in cybersecurity is a double-edged sword. Nat Mach Intell 1:557–560
Veale M, Zuiderveen Borgesius F (2021) Demystifying the draft EU artificial intelligence act (July 31, 2021). Computer Law Review International (2021) 22(4). https://ssrn.com/abstract=3896852. Accessed 11 Oct 2021
Viscusi K (1978) A note on "Lemons" markets with quality certification. Bell J Econ 9:277
Weber R, Studer E (2016) Cybersecurity in the Internet of things: legal aspects. Comput Law Secur Rev 32(5):715–728

European legislation

Directive 2009/48/EC of the European Parliament and of the Council of 18 June 2009 on the safety of toys, OJ L 170, 30.06.2009, p. 1–37
Regulation (EC) No 765/2008 of the European Parliament and of the Council of 9 July 2008 setting out the requirements for accreditation and market surveillance relating to the marketing of products and repealing Regulation (EEC) No 339/93, OJ L 218, 13.08.2008, p. 30–47
Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), OJ L 119, 04.05.2016, p. 1–88
Regulation (EU) 2017/821 of the European Parliament and of the Council of 17 May 2017 laying down supply chain due diligence obligations for Union importers of tin, tantalum and tungsten, their ores, and gold originating from conflict-affected and high-risk areas, OJ L 130, 19.05.2017, p. 1–20
Regulation (EU) 2018/848 of the European Parliament and of the Council of 30 May 2018on organic production and labelling of organic products and repealing Council Regulation (EC) No 834/2007
Regulation (EU) 2019/881 of the European Parliament and of the Council of 17 April 2019on ENISA (the European Union Agency for Cybersecurity) and on information and communications technology cybersecurity certification and repealing Regulation (EU) No 526/2013 (Cybersecurity Act). OJ L 151, 07.06.2019, p 1
Regulation (EU) No 1025/2012 of the European Parliament and of the Council of 25 October 2012on European standardisation, (OJ L 316, 14.11.2012, p. 12)