Artificial Intelligence Systems in the European Union: Guidelines and Architectures for Compliance-by-Design

Daniele Canavese, Afonso Ferreira, Romain Laborde, Abdelmalek Benzekri

Artificial Intelligence Systems in the European Union: Guidelines and Architectures for Compliance-by-Design

Daniele Canavese[a], Afonso Ferreira[a], Romain Laborde[a], Mohamed Ali Kandi[a]

[a] *IRIT, CNRS, 118 Route de Narbonne, Toulouse, 31062, France*

Abstract

The Artificial Intelligence (AI) Act, published in July 2024, represents a pioneering legislative effort to regulate artificial intelligence within the European Union (EU). Aimed at fostering innovation while ensuring safety and funda- mental rights, the AI Act seeks to create a comprehensive legal framework for developing, deploying, and using AI systems. However, the requirements, guide- lines, and best practices mandated in the AI Act are spread amongst various articles and annexes. They are intermingled with legal and sometimes financial issues, making their reading cumbersome for engineers and computer scientists. In addition, the AI Act is not self-contained, as it delegates several issues to other EU legislation, like the Cyber Resilience Act, including technical details. In this paper, we propose an in-depth analysis of both acts, from which we extracted a clear list of technical and organizational requirements that can be used, primarily by developers, to produce systems compliant with these regulations. The requirements are presented in tables, organized according to their categories, and have a workflow to navigate them quickly. We also provide generic architectures to help developers design fully compliant solutions with the AI Act and Cyber Resilience Act.

Keywords: Artificial Intelligence, Software Engineering, Artificial Intelligence Act, Cyber Resilience Act, Compliance-by-Design, Compliance, European Union Regulation

1. Introduction

In the last few years, we have witnessed an explosion in the adoption of Artificial Intelligence (AI) technologies across various sectors [1]. This surge is driven by significant advancements in machine learning, natural language processing, and data analytics, which have made AI more accessible, efficient, and applicable to real-world problems. The proliferation of AI tools and platforms has democratized access to these technologies since companies of all sizes can now leverage AI to optimize their operations and drive innovation [1].

Despite these advancements, the rapid integration of AI also raises important ethical and regulatory considerations [2, 3]. Data privacy, algorithmic bias, and

the impact on employment require careful attention to ensure that AI benefits are distributed equitably, and risks are mitigated.

In this regard, the AI Act, published by the European Union (EU) in July 2024, represents a pioneering legislative effort to regulate artificial intelligence [4]. Aimed at fostering innovation while ensuring safety and fundamental rights, the AI Act seeks to create a comprehensive legal framework for developing, deploying, and using AI systems. It introduces a risk-based approach and provides a set of obligations to developers, distributors, and also users.

AI is often heralded as the promised land of opportunities due to its transformative potential across various sectors. However, this potential also comes with significant risks. AI can be exploited for various nefarious purposes, such as creating deepfakes that can spread misinformation or conducting cyberattacks with unprecedented levels of sophistication [5]. This situation is exacerbated by the fact that many AI systems are black boxes, making it challenging to ensure transparency and trust in their decisions. However, researchers are actively investigating explainable AI techniques [6] to increase the interpretability of machine-generated outputs.

In this context, the AI Act enforces manufacturers and distributors to deploy AI systems that are secure by design. The regulation does not specify any particular technical detail since this is delegated to other legislation and standards, particularly the Cyber Resilience Act (CRA).

The CRA was designed to address the growing threats and vulnerabilities in an increasingly interconnected digital landscape. By setting common standards for digital products and services, the CRA aims to strengthen the cybersecurity framework within the European Union. By establishing robust security requirements for hardware and software, the CRA seeks to enhance the resilience of products against cyberattacks and reduce the risk of data breaches.

From the viewpoint of AI ecosystems, the AI Act and the CRA are meant to be used together to create better and safer AI products. It should be noted that at the time of writing, the CRA is still in the form of the proposal by the European Commission (EC), and, thus, this is the legal text with which we conform in this paper [7]. Contrarily to the AI Act, though, it is not expected that the EU co-legislators will introduce significant changes to the EU proposal. Both regulations mandate several requirements, guidelines, and best practices. However, these are spread amongst various articles and annexes and intermingled with legal and sometimes financial issues. Therefore, we have analyzed both the AI Act and the CRA and extracted a clear list of technical and organizational requirements that can be used to produce systems compliant with
these regulations.

This paper also aims to be a *call to arms* for developers and manufacturers in producing AI systems and cyber security products with a compliance-by-design or regulation-by-design approach [8, 9]. A systematic methodology integrating regulatory requirements into the production processes of software and hardware products is needed more than ever, especially in today's complex EU legal landscape. By integrating compliance from Day 1 in designing a new product, manufacturers will reduce risks early on, avoiding financial penalties and reputational damage in the long run.

Our main contributions in this paper are twofold:

- We have extracted the technical and organizational requirements from the AI Act and CRA regulations — They are exposed in tables, organized according to their categories. We also offer a workflow to navigate them quickly;
- We provide some generic architectures to help developers design solutions that are fully compliant with the AI Act and the CRA, fostering a compliance-by-design approach.

In addition, even though this paper focuses on AI systems, our analyses of the CRA are valuable in their own right for researchers and practitioners who are solely interested in the cybersecurity aspects of their systems.

1.1. *Related Work*

This paper is timely in that it is the first that proposes a technical - rather than purely legal - analysis of the final legal version of the AI Act. While it is true that several papers have already analyzed the AI Act, providing essential insights from diverse viewpoints [10, 11, 12, 13, 14], at time of writing, barely a month after the publication of the final version of the AI Act on the Official Journal of the EU, all such papers referred to draft versions of the Act, or even to the original proposal by the EC [15]. However, the final version of the Act is very different from the EC's proposal, in many aspects. For instance, the original proposal did not even mention general purpose AIs (e.g., Large Language Models). Therefore, the contributions from the literature on this subject must be taken cautiously, depending on the legal text of reference.

In addition, our work is also tied to the Cyber Resilience Act. Producers (and users) of high-risk AI systems are mandated by the AI Act to consider cyber security, making the bond between these two regulations even stricter. The current literature contains very few analyses of the CRA [16, 17, 18] both from technical and legal perspectives. The technical analysis papers of the CRA tend to focus only on the implications of the Regulation on very narrow Information Technology (IT) fields, usually Internet of Things (IoT). On the other hand, to the best of our knowledge, our work is the first general technical analysis of the Cyber Resilience Act.

1.2. *Structure of the paper*

This paper is structured as follows. We start by providing an overview of the main provisions of both regulations in Sections 2 and 3. Then, in Sections 4 and 5, we offer guidelines and architectures to help develop compliant-by-design AI systems. Such guidelines and architectures are directly extracted from the technical and organizational requirements expressed in the regulations. Finally, in Section 6 we close the paper with finishing remarks and avenues for future research.

2. EU Artificial Intelligence Act

The EU AI Act was published in the Official Journal of the European Union on July 12, 2024. This act, formally known as Regulation (EU) 2024/1689, sets forth a comprehensive regulatory framework for Artificial Intelligence within the EU. It is the first of its kind globally, aiming to harmonize rules for using AI across EU Member States and beyond.

The AI Act is a complex regulation. In the following sections, we will describe its birth and adoption timeline (Section 2.1), how the Act defines and classify the AI systems (Sections 2.2 and 2.3), and the primary obligations and governance for the manufacturers (Sections 2.4 and 2.5).

2.1. *Timeline*

The Act entered into force on August 1, 2024, but its provisions are applied staggered, as follows.

February 2, 2025 Provisions related to banned AI practices, such as those posing unacceptable risks, start to apply.

August 2, 2025 Compliance requirements for General Purpose AI (GPAI), which include sophisticated AI models like Large Language Model (LLM), and the establishment of governance structures, including the AI Office and the European Artificial Intelligence (AI) Board, come into effect.

February 2, 2026 The EC issues implementing acts for high-risk AI providers' post-market monitoring plans.

August 2, 2026 Most of the Act's obligations are enforced, including those for high-risk AI systems. This includes requirements for risk management, data governance, technical documentation, and transparency.

August 2, 2027 Final provisions for high-risk AI systems used in regulated safety components are enforced.

The Act also imposes strict obligations on providers and users of high-risk AI systems, including requirements for conformity assessments, documentation, and transparency. Penalties for non-compliance can be severe, reaching up to €35 million or 7% of annual global turnover, whichever is higher, as stated in Paragraph 3 of Article 99.

2.2. *Definition of AI systems*

In the AI Act, an AI system is defined as follows.

> 'AI system' means a machine-based system that is designed to operate with varying levels of autonomy and that may exhibit adaptive- ness after deployment, and that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments.

This definition of AI systems for the AI Act may as well evolve in order to gain precision for the sake of jurists.

We note that the rather long and convoluted Recital 12 already tries to clarify it, stating that the definition is not intended to cover *simpler traditional software systems or programming approaches and should not cover systems that are based on the rules defined solely by natural persons to automatically execute operations*.

However, the recital is not as straightforward as necessary, as, e.g., it states that approaches that *infer from encoded knowledge or symbolic representation of the task to be solved*, are under its scope, even though it stated just before that software that is based on the rules defined solely by natural persons, are

not in scope. The recital also states that *the objectives of the AI system may be different from the intended purpose of the AI system in a specific context*, which may be challenging to understand by engineers and computer scientists alike.

In any case, the Commission should publish guidelines on applying this definition, and this recital explicitly mentions machine learning and self-learning.

2.3. *Classification of AI systems*

The EC 's goal for the AI Act was to address the risks of AI and position Europe to play a leading role globally in the field of AI. As explained by an EC official in a conference organized by ENISA in June 2023, the framework was needed from a safety viewpoint, with the important proviso that the protection of fundamental rights is to be included in the overall notion of safety. Hence, also, the sister liability directives.

That is why the EC proposed a risk-based approach, where some AI systems are considered as posing unacceptable risks and being therefore forbidden (with some derogations, e.g., for law enforcement), others are heavily regulated because they pose high-risks, and finally the remainder AI systems are considered safe enough to be only mildly regulated.

In addition to the prohibited practices described below, the EU AI Act explicitly classifies AI systems to create differentiated regulations.

Figure 1 illustrates the classification of AI systems, above, according to the AI Act. We will explore these in the following.

2.3.1. *Prohibited AI systems*

The Regulation defines a *prohibited AI system* as an AI system that follows a prohibited practice, forbidding its placement on the market. These practices are eight in total, as shown in Figure1, and described in detail in Article 5, as follows:

a. deceptive subliminal techniques meant to control people's behavior;

b. exploitation of people's vulnerabilities due to their age, disability, or a specific social or economic situation, meant to control their behavior;

c. social scoring;
d. prediction of people's possible criminal activity;
e. untargeted scraping of facial images to create or expand facial recognition databases;

f. inferring people's emotions at the workplace or education institutions, except for medical or safety reasons;

g. biometric classification, except in lawful operations;

h. the use of real-time remote biometric identification systems in publicly accessible spaces for law enforcement, except in three detailed cases (e.g., the threat of a terrorist attack).

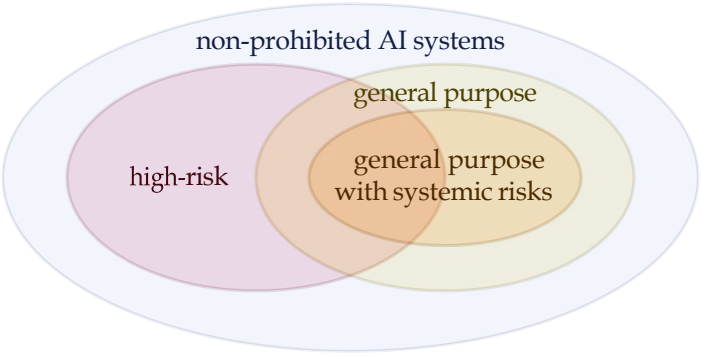

Figure 1: The non-prohibited AI system classes defined in the AI Act.

2.3.2. High-risk AI systems

A *high-risk AI system* is an AI system used in critical environments that can significantly impact people's lives. Some of these high-risk systems are specifically listed in Annex III of the Act and include the following categories of usage:

Biometrics. AI systems are used in people's identification, categorization, and emotion recognition.

Critical Infrastructure AI systems used to manage and operate critical infrastructure, such as electricity, water, and transportation, where failure or malfunction could endanger life or health.

Education and Vocational Training AI systems are used to determine access to education or training or to assess students in educational settings.

Employment, Workers Management, and Access to Self-Employment AI systems used for recruitment, CV sorting, or evaluating candidates in job interviews.

Essential Private and Public Services and benefits AI systems used to determine access to essential services like loan applications, social security, and emergency services.

Law Enforcement AI systems used by law enforcement authorities for predictive policing, crime analytics, and profiling.

Migration, Asylum, and Border Control Management AI systems are used to verify the authenticity of travel documents, assess the eligibility of applications for visas or asylum, and risk assessment of travelers.

Administration of Justice and Democratic Processes AI systems that assist judicial authorities in interpreting laws, legal research, and other judicial functions, but also those that are intended to be used to influence people's voting behavior.

In addition to the aforementioned list, the Act also provides a catalog (in Annex I) of EU harmonization legislations that cover products categorized as high-risk AI systems. In particular, *for products falling in the scope of Annex I, an AI system will be considered to be high-risk if it is intended to be used*

as a safety component of such a product, or it is itself such a product. These products are industrial machinery, toys, lifts, equipment, and protective systems intended for use in potentially explosive atmospheres, radio, pressure and recreational craft equipment, cableway installations, appliances burning gaseous fuels, medical devices, in vitro diagnostic medical devices, automotive, and aviation.

2.3.3. *GPAI systems*

A General Purpose AI (GPAI) model is defined as *An AI model, including where such an AI model is trained with a large amount of data using self- supervision at scale, that displays significant generality and is capable of competently performing a wide range of distinct tasks regardless of the way the model is placed on the market and that can be integrated into a variety of downstream systems or applications, except AI models that are used for research, development, or prototyping activities before they are placed on the market*.
Key characteristics include:

- *significant generality*: the ability to perform a wide range of tasks;
- *data training*: typically involves training with large datasets using self-supervised learning;
- *integration capability*: can be integrated into various downstream systems or applications.

Noticeably, their definition excludes General Purpose AI (AI) models used solely for research, development, or prototyping before market placement.

GPAI with Systemic Risks are defined based on their potential high-impact capabilities: *A GPAI model is classified as a GPAI model with systemic risk if it has high impact capabilities (evaluated on the basis of appropriate technical tools and methodologies, including indicators and benchmarks) or is identified as such by the European Commission. A GPAI model is presumed to have high impact capabilities if its training requires more than* 10^{25} *floating point operations of computational power*.

2.4. *Obligations*

All AI systems intended to interact directly with natural persons are subject to obligations under the AI Act, as stated in Article 50. At a minimum, they must transparently indicate that the interaction is with an AI system. In addition, any content produced or manipulated by an AI system must be clearly labeled as such to natural persons or in machine-readable and machine- detectable format.

In the following paragraphs, we list the primary obligations for all the AI system classes defined in the AI Act. Note that these obligations are *cumulative*, i.e., if an AI system is both a high-risk and GPAI system, it must respect the obligation for both these classes.

2.4.1. *Further obligations for high-risk AI systems*

Article 16 of the EU AI Act mandates stringent requirements for high-risk AI systems to ensure their safe and ethical use. These requirements encompass several key areas: compliance, risk management, data quality, transparency, and

accountability, as listed below (Section 4 will discuss them in more detail):

a. quality Management System to ensure compliance with the regulation;

b. risk management;

c. data and data governance;

d. technical documentation;

e. record-keeping;

f. transparency and provision of information;

g. human oversight;

h. accuracy, robustness, and cybersecurity;

i. assessment of impact on fundamental rights;

j. cooperation with competent authorities.

2.4.2. Further obligations for GPAI model providers

There are several regulatory obligations regarding GPAI models. Article 53 describes the following.

1. technical documentation about training and testing processes and evaluation results;

2. general documentation explaining the model's capabilities and limitations, containing at a minimum:

   (a) a general description of the general-purpose AI model, including eight different information classes, as defined in Annex XII;

   (b) a detailed description of the model's elements and the process for its development;

3. put in place a policy to comply with the EU law on copyright and related rights;

4. publish a sufficiently detailed summary of the content used for training the GPAI model in a template provided by the AI Office.

GPAI models where all model parameters are modifiable and allow for further distribution are not bound by Obligations 1 and 2 above unless such models have systemic risks.

On the other hand, GPAI model providers are subject to at least three further obligations in addition to those listed above. They are:

1. *notification and cooperation*: providers must notify the European Commission if a GPAI model reaches a systemic risk threshold; additionally, providers must cooperate with authorities by providing access to necessary documentation and information;

2. *ethical*[1] *and legal compliance*: GPAI providers must comply with all relevant ethical standards and legal requirements, including non-discrimination, data protection, and consumer rights;

3. if the provider is established in a third country, then it must appoint an authorized representative established in the European Union who will abide by the AI Act.

Compliance with EU law may be demonstrated through the use of codes of practice until standards are published or by other means that need to be approved by the European Commission.

2.4.3. *Further obligations for GPAI model with systemic risks*

Providers of such models must comply with another set of obligations:

1. identify and mitigate systemic risks by performing state-of-the-art model evaluation, including adversarial testing;

2. assess and mitigate systemic risks that may stem from the development, the placing on the market, or the use of general-purpose AI models with systemic risk.
3. notify the appropriate authorities of incidents and possible corrective measures to address them;

4. ensure adequate cybersecurity protection for the model and its physical infrastructure.

2.5. *Governance*

The AI Act devotes a whole chapter (Chapter VII) to governance aspects. They include bodies such as:

AI Office To develop EU expertise and capabilities in the field.

European AI Board Composed of one representative per Member State, to advise and assist the European Commission and the Member States to facilitate the consistent and practical application of the regulation.

Advisory Forum To provide technical expertise and advise the AI Board and the European Commission. We note that the Fundamental Rights Agency, ENISA, CEN/CENELEC, and ETSI are its permanent members.

Scientific Panel of Independent Experts To support the enforcement activities under the AI Act. While Member States may call upon such experts for support, they may be required to pay fees for their advice.

National Competent Authorities Single Points of Contact Which each Member State should have designated to supervise the new rules. In particular, their personnel should have an in-depth understanding of AI technologies, data, and data computing, personal data protection, cybersecurity, fundamental rights,

[1] For instance, the AI Act explicitly mentions the Ethics guidelines for trustworthy AI, available at https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai.

health and safety risks, and knowledge of existing standards and legal requirements.

Voluntary *codes of conduct* will be proposed in a process facilitated by the AI Office, and regulatory sandboxes should be put in place at the national level to facilitate responsible innovation.

In this respect, we note that *standards* will play an essential role in the AI Act's implementation. Accordingly, in May 2023, the EC issued a request to CEN/CENELEC (international non-profit associations that are officially recognized as European Standardization Organizations, alongside European Telecommunications Standards Institute (ETSI) to develop standards in support of the regulatory requirements set out in the AI Act, to have them adopted as harmonized standards.

It should also be noted that the AI Act is not alone in the EC drive to establish a legal framework for AI to address fundamental rights and safety risks specific to AI systems. Interestingly, the AI and CRA are intimately related, even with cross-citations. For instance, the standards mentioned above must be compatible with the CRA. One could see their relationship as the AI Act ruling on procedures and the CRA ruling on cybersecurity content. Furthermore, in addition to the CRA, the AI Act has two sister legal acts, which are less known, but nonetheless important, namely the AI liability directive and the product liability directive — adapting liability rules to the digital age and to AI.

That is why we propose a detailed analysis of the CRA in relation to AI systems in the following.

3. EU Cyber Resilience Act

The EU Cyber Resilience Act (CRA) is officially known as *Horizontal cybersecurity requirements for products with digital elements* [7]. Here, we will explore its key provisions and implications for the software industry. The European Council has adopted the Act on 10 October 2024.

3.1. Timeline

The CRA is a recent proposal and in the next paragraphs we present a short history of its evolution.

September 15, 2022 The EC presents the legislative proposal.

July 19, 2023 The European Parliament (EP) and the Council adopted their individual amended versions.

November 30, 2023 The co-legislators reached a provisional agreement on the Act text.

March 12, 2024 The EP approved the agreed text after the Council and the EP confirmed the agreement.

October 10, 2024 The European Council adopted the CRA.

3.2. Main provisions

The proposed CRA text establishes stricter obligations for software providers

to ensure the security of their products and services. Under this Act, software companies must adopt a risk-based approach to cybersecurity, identifying potential threats and vulnerabilities and implementing appropriate security measures to mitigate them. This proactive approach should promote a culture of security- first software development, safeguarding users from potential cyber-attacks. The CRA aims to safeguard consumers and businesses buying or using products or software with a digital component, particularly those described as being part of the IoT.

The problem addressed by the proposed Regulation is two-fold. Firstly, there is an insufficient level of cybersecurity inherent in many products, and inadequate security updates to such products and software do not help. Secondly, there is our collective inability, as consumers or businesses, to determine which products are currently cyber-secure or to set them up in a way that ensures their cybersecurity is protected. These are very important, as shown by, lest we forget, the Mirai botnet attack in 2016, which mainly targeted consumer devices such as Closed-Circuit TeleVision (CCTV) cameras and home routers connected to the Internet.
The proposed CRA would hence guarantee:

- harmonized rules when bringing products or software with a digital component to market;
- a framework of cybersecurity requirements governing the planning, design, development, and maintenance of such products, with obligations to be met at every stage of the value chain (even if suppliers are based outside the Union);
- an obligation to provide a duty of care for the entire lifecycle of such products.

It was proposed that, when the Regulation enters into force, CRA-compliant software and products connected to the Internet would bear the CE marking to indicate that they are secure by explicit cybersecurity standards.

The proposed Regulation would apply to all products connected directly or indirectly to another device or a digital network, except for specified exclusions, such as open-source software or services already covered by existing rules, such as medical devices, aviation, and cars.

3.3. Main Obligations

Here are a few important new features regulating the IoT domain.

- cybersecurity must be considered in all phases of the product: planning, design, development, production, delivery, and maintenance;
- all cybersecurity risks are documented;
- manufacturers will have to proactively report exploited vulnerabilities and incidents;
- once a product is sold, manufacturers must ensure that vulnerabilities are handled effectively for the expected product lifetime or five years (whichever is shorter);

- products must carry clear and understandable instructions for their use;
- security updates must be made available for at least five years.

3.4. Products with digital elements

The products regulated by the CRA, called *products with digital elements* in the Regulation, are split into categories with increasing risk levels:

Basic products with digital elements are low risk products such as image editors, word processors, electronic spreadsheets, and video games[2].

Important products with digital elements are riskier products and are split into two classes (Annex III gives a comprehensive list of the application categories and their class):

Class I important products are medium-risk products such as identity management systems, browsers, password managers, anti-viruses, and VPN systems.

Class II important products are high-risk products such as hypervisors, firewalls, and tamper-resistant microprocessors.

Critical products with digital elements are very high-risk products and consist of hardware devices with security boxes, smart meter gateways, and smartcards or similar devices (Annex IV of the Act lists the devices belonging to this category).

The main implication of belonging to a specific category involves how the declaration of conformity process is executed. We will discuss this in detail in Section 5.2.

3.5. Revisions agreed by the Council and the European Parliament

According to the EP's website [19], the agreed text simplifies the methodology for classifying products with digital elements. The list of covered devices is expanded with products such as identity management systems software, password managers, biometric readers, smart home assistants, and private security cameras. The support period for manufacturers should be at least five years, with the differentiation between security (automatically installed) and functionality updates. As for the reporting, initial recipients will be competent national authorities, who will notify ENISA to assess the situation and inform other Member States to take the necessary steps if it estimates that the risk is systemic. The application of the Regulation is postponed to three years after it is put into force to give manufacturers sufficient time to adapt. The amended proposal also includes support measures for small and micro enterprises, including specific awareness-raising, education and training programmes, collaboration initiatives, strategies to enhance workforce mobility and support for testing and conformity assessment procedures.

---

[2] The Regulation does not explicitly name them 'basic'; they are just described as non-important and non-critical products. We introduce this name to avoid ambiguity.

3.6. *Commentary*

There are many positive points in the CRA. One of the significant impacts of the proposal on the digital industry is the introduction of mandatory certification and compliance requirements. Providers of digital products would have to undergo rigorous assessments to obtain a cybersecurity certification that attests to the safety and reliability of their products. This certification is meant to enhance the industry's overall credibility, fostering trust among customers and businesses while boosting competitiveness within the EU market.

In addition, the Act mandates that software companies promptly report any cybersecurity incidents to relevant authorities and affected customers. This provision aims at increasing transparency and accountability, enabling authorities to respond swiftly to cyber threats and offering users crucial information about potential risks they may face. While incident reporting can pose challenges for businesses, it ultimately contributes to a collective effort to fortify the EU's cybersecurity infrastructure.

The proposal also emphasizes collaboration among products with digital elements providers, cybersecurity experts, and governmental bodies. Such open information exchanges enable faster threat detection, sharing of best practices, and a coordinated response to emerging cyber threats. To facilitate such information exchanges, the proposal tries to establish clear guidelines to protect proprietary information.

Finally, by setting clear standards and guidelines, the proposal wanted to provide a strong foundation for innovative cybersecurity solutions, encourage companies to invest in cutting-edge technologies that can withstand emerging threats, and ensure that the EU remains at the forefront of global cybersecurity advancements.

Nevertheless, the Act does come with both positive and negative consequences. While users and larger corporations might benefit from increased security measures and improved market credibility, Small and Medium-sized Enterprisess (SMEs) and innovative start-ups may face significant challenges in complying with its costly and complex requirements. Ambiguity in implementation and enforcement, potential impacts on innovation, and possible but little-understood global impact raise concerns about the Act's effectiveness.

4. Architecture Patterns for Compliance with the Artificial Intelligence Act

As introduced in Section 2, the AI Act offers a series of guidelines for creating and using AI systems. The following paragraphs will first discuss the technical and organizational requirements expressed in the Regulation. We will then present an architecture to help providers develop AI systems compliant with the AI Act.

4.1. *Requirements*

The AI Act specifies many different technical and organizational requirements according to the specific role of a person or organization. In particular, two roles are prominent in our analysis:

- *providers*, defined as *a natural or legal person, public authority, agency or other*

body that develops an AI system;

- *deployers*, described in the Act as *a natural or legal person, public authority, agency or other body using an AI system*.

Most obligations are for providers of high-risk AI systems. However, some requirements are also for other AI systems and roles. Figure 2 contains a workflow that can be used as an index to quickly identify the requirements depending on the context of an AI system. The requirements themselves are expressed in a sequence of tables, from Table A.1 to A.12, in which we report a name (given by us, the excerpt from the Regulation, and its location in the AI Act.

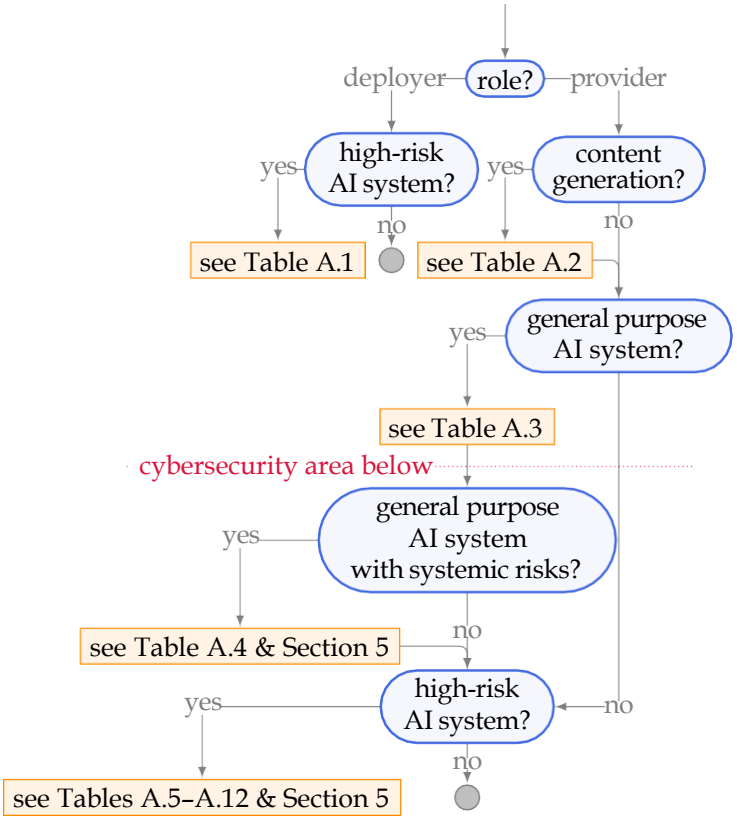

Figure 2: AI Act workflow.

We marked a *cybersecurity area* in the lowest part of the diagram. If any of the two questions here get a *yes* answer, then the AI system must be cyber-secure. It is important to note that the AI Act does not directly legislate or regulate cyber-security; this is demanded by other standards and laws, such as the CRA (see Sections 3 and 5).

The following paragraphs will describe the workflow, question by question, and their related requirements.

Role

The workflow entry question concerns the role of the person involved with an AI system: Is he a deployer or a provider? Most obligations are for providers; however, providers might also have some responsibilities.

High-risk AI system (for deployers) [Table A.1]

The Regulation states that deployers (i.e., users) of high-risk AI systems must follow two requirements: keep the logs and, if they realize or suspect a risk, they must inform the provider, distributor, or competent authorities about it. Interestingly, the AI Act only enforces some responsibilities on high-risk AI system users. Still, no obligation is imposed on deployers of other AI systems (e.g., GPAI systems with systemic risks).

Content generation [Table A.2]

In the workflow, the first question for a provider of an AI system regards the type of data produced. The synthetic output must be marked and detectable as artificially generated if the AI system produces text or multimedia elements (images, video, or audio). This requirement has important implications for all the LLM and Large Vision Model (LVM) developers since, if they want to provide an AI Act-compliant product, they must somehow appropriately tag their outputs.

General Purpose AI systems [Table A.3]

The next choice in the workflow is about dealing with GPAI systems. Although their definition in the AI Act is vague, GPAI systems are essentially LLMs, LVMs, and similar generative systems. Providers of such products have some obligations regarding technical documentation. In particular, they must keep it up-to date, describe the training and testing process, and publicly offer a summary describing the content used for training. Regarding the last point, the Regulation does not force a deployer to disclose the entire training set, only a summary of its content.

GPAI systems with systemic risks [Table A.4]

If a provider of a general purpose AI system is producing a system with systemic risks, then some additional obligations apply too. First, the provider must perform the model evaluation according to some standard[3] and perform some adversarial testing of the model[4]. The second obligation of the provider of such systems regards its security; that is, the provider must ensure an 'adequate' level of protection, as well as for the physical infrastructure hosting the model. As stated before, the AI Act does not provide any specific requirement on cybersecurity; this is delegated to other standards, such as the CRA (see Sections 3 and 5).

---

[3] The list of accepted standards is not mentioned in the Regulation.

[4] Adversarial testing is a family of techniques where testers try to mount attacks to expose weaknesses in a system. In several AI systems, especially Machine Learning (ML) ones, the goal of many adversarial testing techniques is to try to learn how a system behaves, maliciously manipulate it, or reverse engineer it (since many models are fundamentally black boxes, such as neural networks).

High-risk AI system (for providers) [Tables A.5–A.12]

The last choice in the workflow regards high-risk AI systems. This is where the vast majority of technical and organizational obligations are mandatory. For readability purposes, we have split them into multiple tables.

Overarching requirements [Table A.5]. These requirements are about general concepts not explicitly investigated in the Regulation. The AI Act imposes that high-risk AI systems must be accurate, robust to data errors, and also cyber-secure; as stated before, the AI Act does not specify what kind of security or best practice must be adopted since this is out-of-scope.

Risk management [Table A.6]. The Regulation imposes a series of obligations regarding risk management. First, a risk management procedure must be established for high-risk AI systems. This procedure must be continuously iterative throughout the entire system's lifecycle. High-risk AI systems must also be tested to identify appropriate risk management measures using pre- defined metrics and thresholds[5].

Data and data governance requirements [Table A.7]. Data is critical to many AI systems (especially ML-based ones), so the AI Act lists a series of obligations for data management. In particular, data sets should be error-free, complete, and representative. They must also adhere to appropriate data governance and management practices. Furthermore, providers may only process personal and private data to ensure bias detection and correction; this must be performed with some appropriate management techniques to safeguard the fundamental rights and freedoms of natural persons.

System output requirements [Table A.8]. The AI Act lists two critical requirements regarding the output of high-risk AI systems. First, if a continuous learning system is in place, it must be developed to reduce the risk of biased outputs. Second, the output of high-risk AI systems must be developed so that human beings can sufficiently interpret their results.

Technical documentation requirements [Table A.9]. The Regulation enforces a series of obligations on the technical documentation for high-risk AI systems. First, the documentation must be created before a high-risk AI system is placed on the market or put into service. Furthermore, it must be pro- vided in a digital format, be kept up-to-date, and contain the information reported in Annex IV of the AI Act. This information includes, for instance, a description of the system, the required hardware, the training process (if applicable), the accuracy levels, and relevant metrics for the AI system.

Record keeping requirements [Table A.10]. The AI Act strictly imposes that all the high-risk AI systems implement some system for automatic event recording (logs) for the entire product lifetime. The logs should allow trace- ability of the system's functioning and retain specific information, such as usage periods, input data, and personnel involved in result verification. Additionally, providers of high-risk AI systems must store the

[5] The AI Act does not list nor suggest any metrics and thresholds to use.

logs appropriately. This requirement is aligned with the analogous obligation for deployers of high-risk AI systems (see Table A.1).

Reporting requirements [Table A.11]. When a high-risk AI system poses a risk, its providers must immediately investigate the causes, collaborate with the reporting deployer (if applicable) and inform the market surveillance authorities. Noteworthy, the Regulation does not report a time window for the reporting, differently from the ENISA reporting requirement in the CRA (see Section 5 and Table B.17).

Human oversight requirements [Table A.12]. The AI Act obliges providers of high-risk AI systems to build them with some human oversight capabilities so that they can be effectively monitored and supervised by natural persons. The goal of human oversight should focus on preventing or minimizing risks to the health, safety, or fundamental rights of people.

4.2. *Declaration of conformity*

The AI Act defines a relatively liberal procedure for assessing the conformity of high-risk AI systems. This is primarily described in Articles 40 and 41, and the Regulation offers two choices: an internal control procedure or an evaluation involving a third party. The chosen route depends on whether the product adheres to some harmonized standards (described in Annex VI) or not.

Internal control

If a high-risk AI system provider proves that his product fully complies with some harmonized standards, it can opt for a simple internal control conformity procedure.

Third-party assessment

If a provider cannot prove that his product fully complies with these standards, no common specifications are available, or when the provider deems that his product necessitates some external verification, it must perform a third-party assessment procedure. Noteworthy, at least in the initial phase of application of the Act, the AI Act tends to favor internal conformity controls, as written in the Recital 125:

Given the current experience of professional premarket certifiers in the field of product safety and the different nature of risks involved, it is appropriate to limit, at least in an initial phase of application of this Regulation, the scope of application of third-party conformity assessment for high-risk AI systems other than those related to products. Therefore, the conformity assessment of such systems should be carried out as a general rule by the provider under its own responsibility.

4.3. *Architecture*

Building an AI system compliant with the AI Act will pose various challenges. To ease this burden, we propose a generic architecture for AI systems that conform to this Regulation. Our architecture is flexible enough to accommodate a variety of AI models and systems, spanning from modern neural networks to more classical logic-inference-based engines.

Figure 3 shows our proposed architecture and its interactions with a generic

deployer.

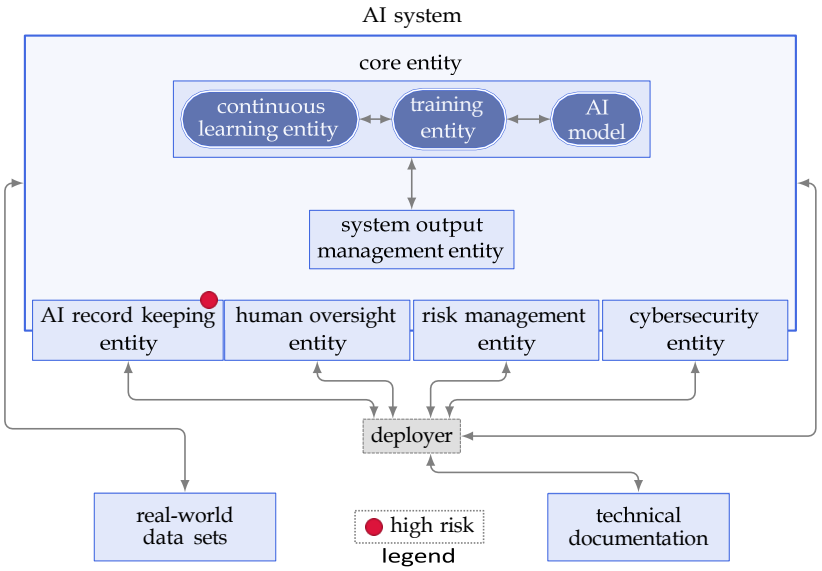

Figure 3: Our AI Act-compliant proposed architecture (from the deployer POV).

The AI system entity contains six sub-entities, which are closely related to how we split the requirements into tables for the AI Act:

Core entity The AI system's heart (or brain). It contains the bare logic and data to make the AI system work. It might include an *AI model*, if applicable, such as a neural network or a decision tree. It might also contain a continuous learning entity that will use a training entity to update the AI model with new observations gathered during its lifetime.

System output management entity This component performs various post-processing activities on the system outputs, such as performing bias detection and correction or tagging it as *artificially generated*.

AI record keeping entity This module consists of logging facilities. It logs both operational and training logs.

Human oversight entity This component, mandatory only for high-risk AI systems, is a human-machine interface that allows a human operator to supervise an AI system.

Risk management entity This entity manages the risks for health, safety, and fundamental rights of the AI system.

Cybersecurity entity This final component is in charge of handling the cybersecurity of the AI system — this module can be modeled using our proposed CRA architecture (see Section 5).

The record keeping, human oversight, and risk management entities are placed on the border of the AI system box to indicate that these modules might

be implemented using a mixture of software/hardware components but also by human beings and that they are connected with all the other entities.

The real-world data sets and technical documentation represent the data exchange by the AI system when it is put into service and the guides that an AI system provider must have written.

The diagram is also marked with a red circle where the obligations of the deployer lay according to the various entities. In this case, as reported in Table A.1, the deployer has only one requirement concerning the record keeping entity (i.e., he must appropriately store the logs).

Figure 4 instead shows the same architecture but from the provider point-of-view.

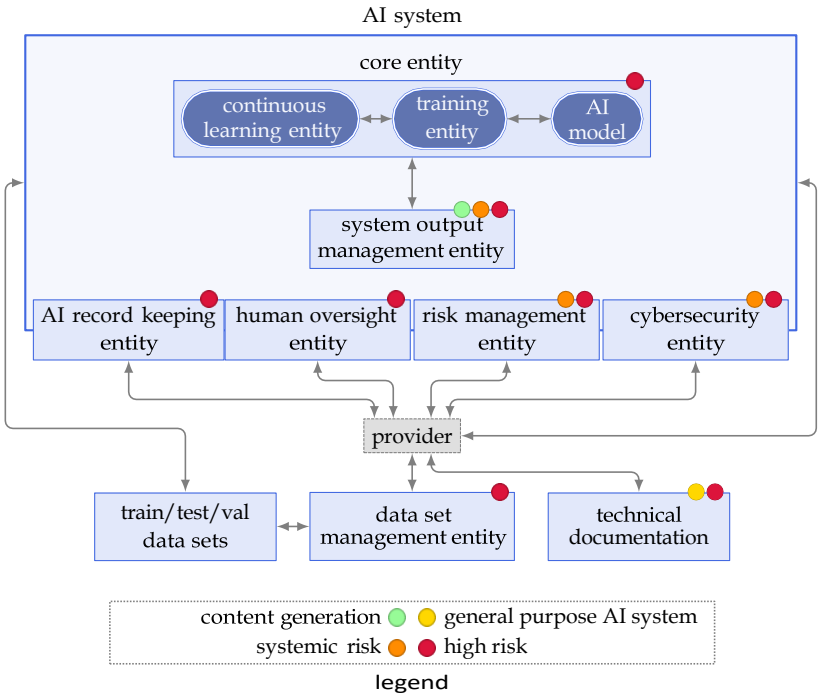

Figure 4: Our AI Act-compliant proposed architecture (from the provider POV).

The architecture differs from the provider version, mainly regarding the data sets at play. First, the provider will potentially use training, test, or validation data sets to build an AI model; this is typical of ML systems, so it might only apply in some scenarios. The *data set management entity* is in charge of performing the data governance and ensuring that the data sets used meet a certain level of quality (i.e., the data is consistent, free of errors, ...).

As in the previous architecture, we have also marked the entities regulated by the AI Act within differently colored circles. As expected, most components must obey specific regulations when dealing with a high-risk AI system.

5. Architecture Patterns for Compliance with the Cyber Resilience Act

The CRA (see also Section 3) aims to strengthen the cybersecurity of

products with digital elements by imposing a secure-by-design approach. Analogously, as we did for the AI Act (see Section 4), we will first analyze the technical and organizational requirements of the regulation. Then, we will provide an architecture for manufacturing CRA-compliant products.

5.1. *Requirements*

The Act places responsibility on a variety of roles (e.g., manufacturers, developers, distributors, and importers), only the manufacturers are involved in the technical requirements so that this section will focus only on them. All the CRA-compliant products with digital elements, regardless of their Class (no Class, Class I, and Class II), must comply with all the requirements listed from Table B.13 to B.17, described in the following paragraphs.

Overarching requirements [Table B.13]

The CRA emphasizes that products with digital elements should be secure by default, meaning they must come with secure configurations and the ability to reset to factory settings. In addition, the legislation requires manufacturers to implement access control mechanisms, protect the availability of essential functions (e.g., against Denial-of-Service (DoS) attacks), and minimize the attack surface[6]. In addition, all CRA-compliant products must appropriately log all security-related events and undergo continuous testing to maintain an appropriate security level during their entire lifecycle.

Data requirements [Table B.14]

The CRA mandates that products with digital elements ensure data confidentiality, protecting the privacy of stored, transmitted, and processed information, both personal and non-personal. Furthermore, data integrity must also be safeguarded, ensuring that data, commands, configurations, and programs are generally protected from unauthorized manipulation or modification.

Vulnerability management requirements [Table B.15]

The Act dictates that regular security patches must be available, allowing vulnerabilities to be promptly addressed through automatic updates and notifying users of these updates[7]. These security updates must be securely available for the product's expected lifetime or up to ten years from its market release, whichever is longer. In addition, manufacturers must disclose patched vulnerabilities to the public.

Technical documentation requirements [Table B.16]

The legislation requires technical documentation to properly document all vulnerabilities and how to handle them, in addition to documenting all the product components. It also mandates providing instructions on how to securely configure the product and use it securely throughout its lifetime. In addition, the documentation must be constantly updated for the product's expected lifetime.

---

[6] The CRA does not explicitly list proper security mechanisms and controls to adopt. It is up to the manufacturer to select the most appropriate ones.

[7] Noteworthy, the legislation also mandates that security patches must be free of charge.

Reporting requirements [Table B.17]

The CRA imposes strict requirements whenever a manufacturer becomes aware of an exploited vulnerability or a security incident. They must report the event to a designated Computer Security Incident Response Team (CSIRT) and European Network and Information Security Agency (ENISA) within 24 hours with an early warning notification, a full notification within 72 hours, and a final report within 14 days for a vulnerability or within one month for an incident.

5.2. *Declaration of conformity*

To fully comply with the CRA, a product with digital elements must undergo a strict declaration of conformity procedure. Similarly to the AI Act, the legislation proposed several control procedures, and which one to undertake depends on the product category, as described in Articles 27 and 32.

Application to a harmonized standard or common specification

This presumption of conformity applies when a manufacturer applies to a harmonized standard (see Regulation (EU) No 1025/2012) or some common specification (see Regulation (EU) 2019/881). This procedure is valid for Basic and Class I Important Products.

Internal control

The manufacturer ensures that the product meets all essential requirements and processes without resorting to any external entity. This conformity procedure can be used by Basic Products, but also by a Class I Important Product when the manufacturer has also applied for harmonized standards or common specifications.

Full quality assurance

This is a stricter conformity assessment procedure, and it involves ensuring not only that the product respects all the requirements but also imposes some controls on its manufacturing process involving a notified body. This procedure suits all Products.

EU-type examination procedure

A manufacturer provides documentation and proof of compliance to a designated notified body (e.g., the ANSSI in France or the ACN in Italy), which then assesses the product and issues an EU-type examination certificate. This procedure suits all Products.

Application to a European cybersecurity certification scheme

This presumption of conformity applies when a manufacturer has obtained a European cybersecurity certification scheme as expressed in the Regulation (EU) 2019/881 with an assurance level of 'substantial' or 'high'. This procedure suits all Products.

5.3. *Architecture*

As we did for the AI Act, we also built an architecture for products with digital elements to conform with the CRA. Figure 5 shows our proposed architecture and its interactions with a generic manufacturer.

A product with digital elements consists of a *cybersecurity entity*, regulated by the CRA, and the system to be protected.

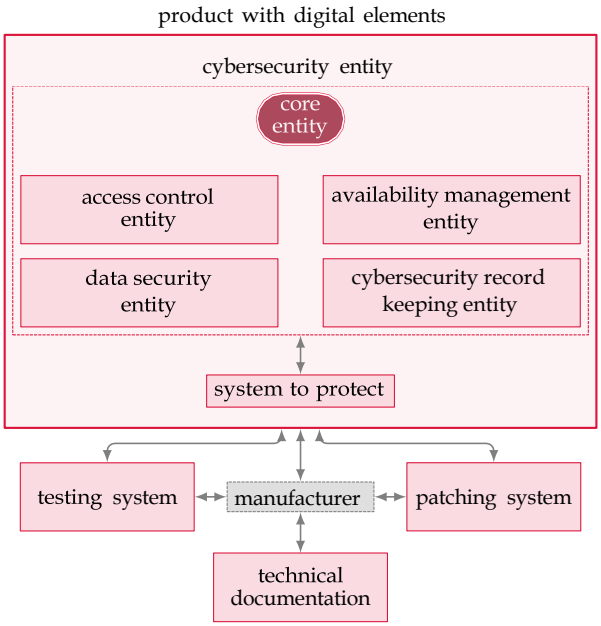

Figure 5: Our Cyber Resilience Act-compliant proposed architecture for the manufacturers.

The cybersecurity entity can be split into five sub-components:

Core entity This is the main security function, such as a firewall or VPN terminator module.

Access control entity This component is in charge of regulation, which can access and alter the functionalities of the system and its core entity.

Availability management entity This module ensures the product's availability, especially of its core entity. For instance, this module can safe- guard the system from denial-of-service attacks.

Data security entity This entity ensures that all the critical data is kept confidential and intact (e.g., by using encryption or message authentication codes).

Cybersecurity record keeping entity This component logs all the security-related events.

In addition to the product with digital elements, several other entities managed by the manufacturer are at play. Notably, the manufacturer must provide and keep up-to date the *technical documentation*, it must put in place a *testing system* to perform continuous testing for the entire duration of the life of the product, and also a *patching system* for securely and timely deliver security patches to the users of the product.

This schema can be used to model the cybersecurity entity in our AI Act architecture (see Section 4), guaranteeing double compliance with the AI Act

and the CRA.

6. Conclusion

In this paper, we addressed the complex issue of software design and development compliance by exploring the impact of a recent duo of EU regulations, namely the Artificial Intelligence and Cyber Resilience Acts.

We proposed a detailed analysis of the requirements imposed on AI systems that fall in the scope of this legislation. In particular, we gave comprehensive tables showing all the technical and organizational requirements that can be extracted from the many articles in both Acts. Such tables and their accompanying diagrams will guide the software industry in ensuring compliance in developing AI applications that are meant to enter the EU market directly or via global supply chains.

The approach followed in this paper opens wide avenues for further research, given the large number of legal acts regulating digital socio-economic eco-systems that the European Union has produced in the past six years (cf, for instance, the overview of EU legislation in the Digital Sector compiled in [20]).

For instance, we addressed the CRA about AI systems. Although the main focus of the CRA is IoT devices, like CCTV, routers, and other connected objects, extracting the requirements for such a category of electronic devices will be very valuable also for other appliances. It will do the same for the Data and the Data Governance Acts, the European Health Data Space, NIS 2, and many others. As such, a detailed analysis is needed to clarify several laws that benefit the digital applications industry.

Acknowledgments

This work was partially supported by the European research projects H2020 LeADS (GA 956562) and Horizon Europe DUCA (GA 101086308), ARN TrustIn- Clouds, and CNRS IRN EU-CHECK.

References

[1] Stanford University, Artificial Intelligence Index Report 2024 (2024).
URL https://aiindex.stanford.edu/report/

[2] A. Casovan, Editorial, Journal of AI Law and Regulation 1 (2) (2024). doi:10.21552/aire/2024/2/3.
URL https://doi.org/10.21552/aire/2024/2/3

[3] S. Schmitz-Berndt, Interfering with judicial independence?, Journal of AI Law and Regulation 1 (2) (2024). doi:10.21552/aire/2024/2/6.
URL https://doi.org/10.21552/aire/2024/2/6

[4] E. Union, 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) (2024).
URL https://eur-lex.europa.eu/eli/reg/2024/1689/oj

[5] N. Kaloudi, J. Li, The ai-based cyber threat landscape: A survey, ACM Comput. Surv. 53 (1) (February 2020). doi:10.1145/3372823.
URL https://doi.org/10.1145/3372823

[6] A. Abusitta, M. Q. Li, B. C. Fung, Survey on explainable ai: Techniques, challenges and open issues, Expert Systems with Applications 255 (2024) 124710. URL https://doi.org/10.1016/j.eswa.2024.124710

[7] European Commission, Proposal for a regulation of the European Parliament and of the Council on horizontal cybersecurity requirements for products with digital elements and amending regulation (EU) 2019/1020 (2022).
URL https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex:52022PC0454

[8] K. Prifti, J. Morley, C. Novelli, L. Floridi, Regulation by design: Features, practices, limitations, and governance implications, Minds and Machines 34 (2) (2024) 13. URL https://doi.org/10.1007/s11023-024-09675-z

[9] E. Ortega, M. Tran, G. Bandeen, Ai digital tool product lifecycle governance framework through ethics and compliance by design†, in: 2023 IEEE Conference on Artificial Intelligence (CAI), 2023, pp. 353–356. URL https://doi.org/10.1109/CAI54212.2023.00155

[10] M. D. Mueck, A. E. O. Bar, S. D. Boispean, Upcoming European Regulations on Artificial Intelligence and Cybersecurity, IEEE Communications Magazine 61 (7) (2023) 98–102. URL https://doi.org/10.1109/MCOM.004.2200612

[11] Burges Salmon, Navigating the EU AI Act: flowchart (2024).
URL https://blog.burges-salmon.com/post/102ixon/ navigating-the-eu-ai-act-flowchart

[12] J. Schuett, Risk Management in the Artificial Intelligence Act, European Journal of Risk Regulation (2023) 1–19 URL https://doi.org/10.1017/err.2023.1

[13] N. Helberger, Futurenewscorp, or how the AI Act changed the future of news, Computer Law & Security Review 52 (105915) (2024). URL https://doi.org/10.1016/j.clsr.2023.105915

[14] E. Drouard, O. Kurochkina, R. Schlich, D. Ozturk, The interplay between the AI Act and the GDPR, Journal of AI Law and Regulation 1 (2) (2024). URL https://doi.org/10.21552/aire/2024/2/4

[15] European Commission, Proposal for a regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (artificial intelligence act) and amending certain union legislative acts (2021).
URL https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex\%3A52021PC0206

[16] M. R. Shaffique, Cyber resilience act 2022: A silver bullet for cybersecurity of IoT devices or a shot in the dark?, Computer Law & Security Review 54 (2024) 106009. URL https://doi.org/10.1016/j.clsr.2024.106009

[17] C. Schmittner, O. Veledar, T. Faschang, G. Macher, E. Brenner, Fostering cyber resilience in Europe: An in-depth exploration of the cyber resilience act, in: M. Yilmaz, P. Clarke, A. Riel, R. Messnarz, C. Greiner, T. Peisl (Eds.), Systems, Software and Services Process Improvement, Springer Nature Switzerland, Cham, 2024, pp. 390–404.

[18] P. Eckhardt, A. Kotovskaia, The EU's cybersecurity framework: the interplay between the cyber resilience act and the nis 2 directive, International Cybersecurity Law Review 4 (2) (2023) 147–164. URL https://doi.org/10.1365/s43439-023- 00084-z

[19] European Parliament (20 June 2024), Horizontal cybersecurity requirements for products with digital elements, URL https://www.europarl.europa.eu/legislative-train/theme-a-europe-fit-for-the-digital-age/file-european-cyber-resilience-act

[20] K. Zenner, J. S. Marcus, K. Sekut, A dataset on EU legislation for the digital world (2024). URL https://www.bruegel.org/dataset/dataset-eu-legislation-digital-world

Appendix A. AI Act requirements

This Appendix contains all the requirements grouped into tables mentioned in Section 4.

name	excerpt	location
deployers' log retention	deployers of high-risk AI systems shall keep the logs	art. 26
deployers' risk reporting	where deployers have reason to consider that the use of the high-risk AI system in accordance with the instructions may result in that AI system presenting a risk […], they shall, without undue delay, inform the provider or distributor and the relevant market surveillance authority, and shall sus- pend the use of that system	art. 26

Table A.1: Requirement for deployers of high-risk AI-systems.

name	excerpt	location
artificially generated output	providers of AI systems […] generating synthetic audio, image, video or text content, shall ensure that the outputs of the AI system are marked in a machine-readable format and detectable as artificially generated or manipulated	art. 50

Table A.2: Requirement for content generation AI-systems.

name	excerpt	location
documented training and testing	draw up and keep up-to-date the technical documentation of the model, including its training and testing process and the results of its evaluation, which shall contain, at a minimum, the elements set out in Annex XI	art. 53
training content summary	draw up and make publicly available a sufficiently detailed summary about the content used for training of the general-purpose AI model	art. 53

Table A.3: Requirements for general purpose AI-systems.

name	excerpt	location
model evaluation	perform model evaluation in accordance with standardised protocols and tools reflecting the state-of-the-art, including conducting and documenting adversarial testing of the model	art. 55
cybersecure	ensure an adequate level of cybersecurity protection for the general-purpose AI model with systemic risk and the physical infrastructure of the model	art. 55

Table A.4: Requirements for general purpose AI-systems with systemic risks.

name	excerpt	location
reliability	high-risk AI systems shall be designed and developed in such a way that they achieve […] an appropriate level of accuracy, robustness and cybersecurity	art. 15
resilience	high-risk AI systems shall be as resilient as possible regarding errors, faults or inconsistencies […], in particular due to their inter- action with natural persons or other systems	art. 15

Table A.5: Overarching requirements for high-risk AI systems.

name	excerpt	location
risk management system	a risk management system shall be established, implemented, documented and maintained in relation to high-risk AI	art. 9
continuous risk management	the risk management system shall be understood as a continuous iterative process planned and run throughout the entire life- cycle of a high-risk AI system	art. 9
acceptable residual risks	the risk management measures […] shall be such that the relevant residual risk […] is judged to be acceptable	art. 9
testing for risks	high-risk AI systems shall be tested for the purpose of identifying the most appropriate and targeted risk management measures	art. 9
testing against metrics	testing shall be carried out against prior defined metrics and probabilistic thresholds that are appropriate to the intended purpose of the high-risk AI system	art. 9

Table A.6: Risk management requirements for high-risk AI systems.

name	excerpt	location
data set governance	training, validation and testing data sets shall be subject to data governance and management practices appropriate for the intended purpose of the high-risk AI system	art. 10
data quality	training, validation and testing data sets shall be relevant, sufficiently representative, and to the best extent possible, free of errors and complete	art. 10
data set localization	data sets shall take into account […] the characteristics or elements that are particular to the specific geographical, contextual, behavioural or functional setting within which the high-risk AI system is intended to be used	art. 10
personal data exception	to the extent that it is strictly necessary for the purpose of ensuring bias detection and correction […], the providers of such systems may exceptionally process special categories of personal data, subject to appropriate safeguards for the fundamental rights and freedoms of natural persons	art. 10

Table A.7: Data and data governance requirements for high-risk AI systems.

name	excerpt	location
biased output mitigation	high-risk AI systems that continue to learn […] shall be developed in such a way as to eliminate or reduce as far as possible the risk of possibly biased outputs	art. 15
transparency	high-risk AI systems shall be designed and developed in such a way as to ensure that their operation is sufficiently transparent to enable deployers to interpret a system's output and use it appropriately	art. 13

Table A.8: System output requirements for high-risk AI systems.

name	excerpt	location
technical documentation prior existence	the technical documentation of a high-risk AI system shall be drawn up before that system is placed on the market or put into service	art. 11
technical documentation up-to date	the technical documentation of a high-risk AI system […] shall be kept up-to date	art. 11
technical documentation content	a high-risk AI system related to a product covered by the Union harmonisation legislation listed in Section A of Annex I is placed on the market or put into service, a single set of technical documentation shall be drawn up containing all the information set out in paragraph 1, as well as the information required under those legal acts	art. 11
technical documentation accessibility	high-risk AI systems shall be accompanied by instructions for use in an appropriate digital format or otherwise that include concise, complete, correct and clear information	art. 13
documented accuracy	the levels of accuracy and the relevant accuracy metrics of high-risk AI systems shall be declared in the accompanying instructions of use	art. 15

Table A.9: Technical documentation requirements for high-risk AI systems.

name	excerpt	location
logging	high-risk AI systems shall technically allow for the automatic recording of events ('logs') over the lifetime of the system	art. 12
traceability by logs	to ensure a level of traceability of the functioning of a high-risk AI system […] logging capabilities shall enable the recording of events relevant for: […] identifying situations that may result in the high-risk AI system presenting a risk […] facilitating the post-market monitoring […] monitoring the operation of high-risk AI systems	art. 12
log content	the logging capabilities shall provide, at a minimum: […] recording of the period of each use of the system […] the reference database against which input data has been checked by the system […] the input data for which the search has led to a match […] the identification of the natural persons involved in the verification of the results	art. 12
providers' log retention	providers of high-risk AI systems shall keep the logs	art. 19

Table A.10: Record keeping requirements for high-risk AI systems.

name	excerpt	location
providers' risk reporting	where the high-risk AI system presents a risk […] and the provider becomes aware of that risk, it shall immediately investigate the causes, in collaboration with the reporting deployer, where applicable, and inform the market surveillance authorities	art. 20

Table A.11: Reporting requirements for high-risk AI systems.

name	excerpt	location
human oversight	high-risk AI systems shall be designed and developed in such a way, including with appropriate human-machine interface tools, that they can be effectively overseen by natural persons	art. 14
risk minimization by human oversight	human oversight shall aim at preventing or minimising the risks to health, safety or fundamental rights	art. 14

Table A.12: Human oversight requirements for high-risk AI systems.

Appendix B. Cyber Resilience Act requirements

This Appendix contains all the requirements grouped into tables mentioned in Section 5.

name	excerpt	location
secure by default	be made available on the market with a secure by default configuration, unless otherwise agreed between manufacturer and business user […], including the possibility to reset the product to its original state	ann. I., sec. 1
access control	ensure protection from unauthorised access by appropriate control mechanisms	ann. I., sec. 1
data minimization	process only data, personal or other, that are adequate, relevant and limited to what is necessary in relation to the intended purpose of the product with digital elements	ann. I., sec. 1
availability of essential functions	protect the availability of essential and basic functions, also after an incident, including through resilience and mitigation measures against denial-of-service attacks	ann. I., sec. 1
attack surface minimization	be designed, developed and produced to limit attack surfaces, including external interfaces	ann. I., sec. 1
logging	provide security related information by recording and monitoring relevant internal activity, including the access to or modification of data, services or functions, with an opt-out mechanism for the user	ann. I., sec. 1
resilience by design	be designed, developed and produced to reduce the impact of an incident using appropriate exploitation mitigation mechanisms and techniques	ann. I., sec. 1
continuous testing	apply effective and regular tests and reviews of the security of the product with digital elements	ann. I., sec. 2

Table B.13: Overarching requirements for products with digital elements.

name	excerpt	location
data confidentiality	protect the confidentiality of stored, transmitted or otherwise processed data, personal or other	ann. I., sec. 1
data integrity	protect the integrity of stored, transmitted or otherwise processed data, personal or other, commands, programs and configuration against any manipulation or modification not authorised by the user, and report on corruptions	ann. I., sec. 1

Table B.14: Data requirements for products with digital elements.

name	excerpt	location
automatic security patching	ensure that vulnerabilities can be addressed through security updates, […] through the notification of available updates to users, and the option to temporarily postpone them	ann. I., sec. 1
coordinated vulnerability disclosure	put in place and enforce a policy on coordinated vulnerability disclosure	ann. I., sec. 2
promptly security patching	distribute updates for products with digital elements to ensure that vulnerabilities are fixed or mitigated in a timely manner	ann. I., sec. 2
security patching disclosure	once a security update has been made available, share and publicly disclose information about fixed vulnerabilities	ann. I., sec. 2
patch secure distribution	provide for mechanisms to securely distribute updates	ann. I., sec. 2
security patching lifetime	ensure that each security update […] remains available after it has been issued for a minimum of 10 years or for the remainder of the support period, whichever is longer	art. 13

Table B.15: Vulnerability management requirements for products with digital elements.

name	excerpt	location
vulnerability documentation	document vulnerabilities and components contained in products with digital elements, including by drawing up a software bill of materials in a commonly used and machine-readable format covering at the very least the top-level dependencies of the products	ann. I., sec. 2
product information and use instructions	[cybersecurity-related information and] the necessary measures during initial commissioning and throughout the lifetime of the product with digital elements to ensure its secure use	ann. II
documentation lifetime	the technical documentation shall be drawn up before the product with digital elements is placed on the market and shall be continuously updated	art. 31

Table B.16: Technical documentation requirements for products with digital elements.

name	excerpt	location
vulnerability reporting	notify any actively exploited vulnerability […] to the CSIRT […] and to ENISA	art. 14
vulnerability reporting deadlines	the manufacturer shall submit an early warning notification of an actively exploited vulnerability […] within 24 hours, […] a vulnerability notification […] within 72 hours, […] a final report, no later than 14 days	art. 14
incident reporting	notify any severe incident having an impact on the security […] to the CSIRT […] and to ENISA	art. 14
incident reporting deadlines	the manufacturer shall submit an early warning notification of a severe incident […] within 24 hours, […] an incident notification […] within 72 hours, […] a final report, within one month	art. 14

Table B.17: Reporting requirements for products with digital elements.