Computers & Security 158 (2025) 104628

Contents lists available at ScienceDirect

Computers & Security

journal homepage: www.elsevier.com/locate/cose

Full length article

AI algorithms under scrutiny: GDPR, DSA, AI Act and CRA as pillars for algorithmic security and privacy in the European Union

Marta Beltrán

Agencia Española de Protección de Datos (AEPD), Scientific Area, Innovation and Technology Division, C/Jorge Juan 6, 28001 Madrid, Spain

ARTICLE INFO

Keywords:
Algorithms
Artificial Intelligence
Cybersecurity
Data protection
Privacy
Regulation
Sociotechnical systems
Trust

ABSTRACT

The General Data Protection Regulation (GDPR), Digital Services Act (DSA), Artificial Intelligence Act (AI Act) and Cyber Resilience Act (CRA) are essential pillars for algorithmic security and privacy in the European Union. Each of these regulations addresses specific aspects of technology, such as personal data protection, trustworthy online services, safe AI systems, and secure digital products while fostering trust in algorithm-based systems. Together, they can establish a robust framework for ensuring the security and privacy of AI algorithms in the EU by addressing critical concerns through a risk-based approach. This paper proposes a multi-layered approach to algorithmic security and privacy, based on these four instruments, considering organisational risk, risks to rights and freedoms, systemic risks and risks to national security. An illustrative example demonstrates how the EU can establish a global standard for trustworthy innovation and the protection of fundamental rights by leveraging the direct and indirect synergies of these laws.

1. Introduction

The interconnected nature of the digital age implies that algorithms are deeply embedded within complex socio-technical systems, where technology, people and processes interact to shape our daily lives (Willson, 2019; Ulbricht and Yeung, 2022). Specifically, AI algorithms are broadly applied in automated decision-making, regulatory oversight, predictive analysis, and the automation of tasks ranging from workplace management to content recommendation or moderation.

Ensuring algorithmic security and privacy in the European Union, therefore, requires a holistic approach that considers the technical aspects of algorithms and their broader social and organisational contexts. As AI systems become increasingly important in crucial issues, there is a growing need for robust algorithmic regulation (Brkan, 2019; Seyfert, 2022). By effectively addressing security and privacy concerns (Fischer-Hübner et al., 2021), the EU aims to protect consumers from potential harm and hold algorithms accountable, thereby mitigating risks. The goal is to boost the EU's economy by promoting innovation and the uptake of trustworthy AI that can be used "for good" (Michael et al., 2019; Busuioc, 2021; Riefa).

The EU's regulatory framework, particularly the General Data Protection Regulation (GDPR), the Digital Services Act (DSA), the AI Act (EU AI Act and AI Act are interchangeable terms and refer to the same legislation in this paper, the European Union's Artificial Intelligence Act), and the Cyber Resilience Act (CRA), serve as essential and interconnected pillars in establishing algorithmic security and privacy. Therefore, a holistic understanding and application of this interconnected regulatory framework is crucial to effectively establish trust within the EU (Tamò-Larrieux et al., 2024).

These legislative instruments are not designed to operate in isolation but as mutually reinforcing components of a comprehensive approach (Pollak; Ferreira and Goldman, 2025). The effectiveness of these legal frameworks relies on the interplay among technological capabilities, regulatory requirements, and the human element in socio-technical landscapes (Abbas et al., 2023). The main contributions of this paper are (1) Analyse each of these regulations, examining their individual and collective contributions to AI algorithms security and privacy; (2) Discuss their multi-layered approach to risk management, considering both direct and indirect synergies and organisational risks, risks to rights and freedoms, systemic risks and national security risks and (3) Provide an illustrative example of this multi-layered risk-based approach and the challenges and limitations it implies.

The rest of this paper is organised as follows. Section 2 summarises the related work and outlines the motivation for this research. Section 3 provides a foundational understanding of the essential concepts and the four regulatory pieces: GDPR, DSA, AI Act and CRA. Section 4 presents our analysis of the multi-layered approach for risk management that can be deployed considering these four instruments and their synergies. Section 5 illustrates this approach with an example that highlights the challenges and limitations of the current regulatory framework. Section 6 compiles our findings and discusses the key elements of the

E-mail address: mbeltran@aepd.es.

https://doi.org/10.1016/j.cose.2025.104628
Received 30 April 2025; Received in revised form 1 August 2025; Accepted 5 August 2025
Available online 19 August 2025
0167-4048/© 2025 The Author. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

proposed approach. Finally, Section 7 presents our conclusions and some lines for future work.

2. Related work and motivation

2.1. Algorithmic security and privacy

The EU has established a regulatory framework for security and privacy, focusing on several key pillars to ensure robust protection and compliance. Previous research has explored the intertwined relationship between security and privacy. It has been highlighted their focus on a risk-based approach (assessing and managing risks, ensuring that measures are proportionate to the level of risk involved), the "by-design" concept (these measures must be integrated into the design and development of systems and processes from the outset, rather than being added as an afterthought), reporting obligations and certification schemes (Fischer-Hübner et al., 2021; Mantelero et al., 2020; Olukoya, 2022; Amoo et al., 2024; Chiara, 2024).

On the other hand, different studies have shown that while they share common approaches and concepts, tensions can arise in areas such as logging, data retention, or incident handling (Allahrakha, 2023). The increasing adoption of AI further complicates this balance, introducing new threats and challenges concerning security and privacy (Yanamala et al., 2024; Novelli et al., 2024).

Several research works, mainly focused on AI, have analysed the European approach to algorithmic regulation. The EU is identified as a key player systematically addressing the regulation and legal limits of new digital technologies such as AI (Baldini et al., 2025; Bygrave, 2025). The concept of an evolving *EU digital acquis*, a comprehensive body of law regulating digital aspects, has emerged in the last decade (Baldini et al., 2025).

Scientific papers and studies have compared the GDPR and the AI Act, focusing on their intersections, divergences and implications for governance, innovation, and compliance (Wolff et al., 2023; Musch et al., 2024; Schwartmann et al., 2024). Previous research work highlights their complementary roles in ensuring trustworthy AI while respecting fundamental rights and fostering responsible technology design and use. However, different papers underscore the need for harmonisation to ensure effective compliance and innovation. It is worth mentioning that various institutions and authorities have explored the interplay between the GDPR and the AI Act, examining how these two major regulations overlap, complement, and sometimes conflict when AI systems process personal data. Most published guidance and analysis documents focus on regulatory requirements and enforcement (Data Protection Authority of Belgium, 2024; EDPB, 2024; CNIL, 2025).

Previous initiatives have provided less explicit conclusions on the DSA and the CRA directly related to the security and privacy of AI algorithms. However, their role in regulating online platforms and digital products, which involves algorithmic processes, has been acknowledged (Ulbricht and Yeung, 2022; Junklewitz et al., 2023; Rampášek, 2023).

Finally, several papers have identified the challenges and limitations of the current EU approach and propose lines for future research. The reliance on criteria such as "state of the art" in current legislation, without sufficient judicial guidance, is seen as a limitation (Bygrave, 2025). The AI Act links the concept of a "generally acknowledged state of the art" to fundamental technical requirements for high-risk AI systems. While driven by the rapid pace of technological advancement, it refers to established, feasible practices and performance levels rather than pure novelty. Obligations to update or change systems are primarily triggered by initial market placement, substantial modifications, or as part of continuous risk management, particularly to address evolving threats and maintain compliance with standards that reflect the state of the art as it develops. However, the inherent speed and complexity of AI development, coupled with ambiguity in the regulation, make defining and continuously adhering to the state of the art a significant challenge for regulated entities. The European Data Protection Board, for example, has been required to guide the "state of the art" concept within data protection and GDPR in the past due to similar practical challenges (EDPB, 2020).

In addition, the effectiveness of the GDPR's "right to explanation" and transparency requirements for complex algorithms remains questionable (Seizov and Wulf). Some studies explore the challenges in achieving algorithmic transparency due to different complexities (Grochowski et al., 2021; Lazcoz and De Hert, 2023; Busuioc et al., 2023; Wulf and Seizov, 2024) and other obstacles such as the requirement to use sensitive data to prevent discrimination (Van Bekkum and Borgesius, 2023). The fact that regulatory efforts have significantly fallen behind the practice of AI deployment is noted (Busuioc, 2022). It has to be considered that the concept of algorithmic regulation itself is maturing and requires a cross-disciplinary analysis (Ulbricht and Yeung, 2022; Fortes et al., 2022).

2.2. Motivation

As mentioned before, previous research has shown that governments and stakeholders need to understand the global regulatory landscape to minimise the risks posed by AI algorithms in different contexts. The existing literature emphasises the need for improved coherence among existing regulations to ensure robust governance while fostering technological advancement responsibly. Differences in implementing different laws should not create inconsistencies or unnecessary burdens. On the contrary, an integrated policy landscape that works in synergy with the existing regulatory framework should enable the shared goal of protecting individual rights. Harmonising GDPR, DSA, AI Act, and CRA is crucial to trustworthiness. All these regulations complement each other, but overlapping areas and potential ambiguities must be interpreted and navigated effectively to ensure seamless integration.

This paper will advance this global goal by proposing a multi-layered risk-based approach that considers the direct and indirect synergies in these four regulations, leveraging them to manage organisational risks, risks to rights and freedoms, systemic risks, and risks to national security. This new approach covers a research gap, transforming uncertainty into structured action and governance, enabling organisations to deliver secure, privacy-preserving, and compliant systems.

3. European Union regulatory framework

This section establishes the foundations for understanding the relationship between the four analysed regulatory instruments and the algorithms' trustworthiness.

3.1. Preliminary definitions and concepts

The term trustworthiness is widely used in the context of European policy regarding digital systems, particularly concerning AI and digital identity, as outlined in EU regulations, guidelines, and frameworks. However, it is not defined as a standalone legal term.

Definitions and guidelines for understanding trustworthiness in digital systems (NIST, 2018; Belanger et al., 2002; Zhang and Gupta, 2018) or specifically in AI (Liu et al., 2022; Li et al., 2023) can be found in the literature on system design and risk management. Trustworthiness is typically related to the degree to which a system maintains confidentiality, integrity, availability, and privacy or the capability of a system to operate within defined risk levels despite disruptions and attacks.

Considering available definitions, this research primarily relates trustworthiness, as a technical term, to (cyber)security and privacy as defined within the Cybersecurity Act and the Charter of Fundamental Rights of the European Union. Therefore, cybersecurity is related to "activities necessary to protect network and information systems, the users of such systems, and other persons affected by cyber

threats" (Regulation (EU) 2019/881, 2019) while privacy, considered a fundamental right, can be formulated as "Everyone has the right to respect for his or her private and family life, home and communications" (Charter of Fundamental Rights of the European Union, 2000).

Table 1 considers other possible components of AI trustworthiness, such as fairness, transparency, explainability, accountability, or safety, that could be explored in the future. Although these terms have different and specific meanings in the context of each considered law, in general, we refer to entities (data processing, services and platforms, systems and models, products) avoiding unexpected functioning, including producing harmful or discriminatory outcomes across different groups (fairness), entities understandable to different stakeholders with clear documentation of their processes and decisions (transparency/explainability), entities demonstrating compliance and providing precise mechanisms to hold stakeholders accountable for outcomes (accountability) or entities proactively preventing harm or danger caused by accidental hazards, failures or malfunctioning (safety).

At this point, it is required to provide some additional definitions because, although there is an intended alignment and relationship between these concepts in the considered regulations, there are also some gaps and areas of conceptual vagueness (Reichman and Sartor, 2021; Bagni and Seferi, 2025). While the AI Act adopts a risk-based approach, which implies addressing threats, harms, risks and vulnerabilities, the specific application and interpretation of these terms have not been tailored to the AI context in the law, potentially differing from how they are understood in broader cybersecurity or privacy contexts:

- Threat is any circumstance or event with the potential to adversely impact an AI system, organisational operations (including mission, functions, image, or reputation), organisational assets, individuals (users or affected persons), other organisations, the environment or the Nation including but not limited to cyber threats and AI-specific threats (such as model flaws or data poisoning). This definition is consistent with the one provided by the Cybersecurity Act in the specific case of cyber threat "potential circumstance, event or action that could damage, disrupt or otherwise adversely impact network and information systems, the users of such systems and other persons".
- Harm is an adverse or detrimental effect on natural persons, groups of persons, or society, including any material or immaterial harm, such as physical injury, psychological impact, societal disruption, or economic loss. Harm can arise from the use of an AI system following its intended purpose or under conditions of reasonably foreseeable misuse. The impact is the magnitude of harm that can be expected to result from the consequences of an occurrence or materialisation of a threat.
- Risk is the extent to which an entity is affected by a threat and typically a function of (i) the impacts that would arise if the threat materialises and (ii) the likelihood of this occurrence. This definition is consistent with the one provided by the AI Act, "the combination of the probability of an occurrence of harm and the severity of that harm", and with the one provided by the CRA for cybersecurity risk, "the potential for loss or disruption caused by an incident and is to be expressed as a combination of the magnitude of such loss or disruption and the likelihood of occurrence of the incident".
- Vulnerability is a condition, weakness, susceptibility, or flaw in an AI system, its underlying data, development process, or infrastructure that could enable a threat to materialise. This is consistent with the definition provided by the CRA "weakness, susceptibility or flaw of a product with digital elements that can be exploited by a cyber threat".

3.2. The General Data Protection Regulation (GDPR): Establishing foundational principles for personal data protection

The General Data Protection Regulation (GDPR) applies to the processing of personal data, defined broadly as any information relating to an identified or identifiable natural person. Processing is defined as any operation or set of operations performed on personal data, whether or not by automated means. This includes activities commonly performed when developing or using AI algorithms, such as collection, recording, organisation, structuring, storage, alteration, retrieval, consultation, use, disclosure, dissemination, restriction, erasure or destruction.

The GDPR's primary goal is to protect the fundamental rights and freedoms of natural persons (data subjects) concerning the processing of personal data. It contains several core principles (Article 5): lawfulness, fairness, transparency, purpose limitation, data minimisation, accuracy, storage limitation, integrity, confidentiality and accountability (Regulation (EU) 2016/679, 2016). These principles collectively guide the use of algorithms to process personal data in a manner that is legal, responsible, and ethical.

AI algorithms often process personal data for various purposes, including automated decision-making and profiling, which directly impact the rights and freedoms of individuals. The rights of the data subject, as defined in Articles 12 to 23, must be respected during the design, development, deployment, and use of algorithms. Particularly, Article 22 of the GDPR grants individuals the right not to be subject to decisions based solely on automated processing, including profiling, that produce legal or similarly significant effects. Exceptions exist for contracts, law, or explicit consent, but even then, safeguards such as the right to human intervention, to express their view, and to contest the decision must be in place.

Article 24 establishes that the data controller shall implement appropriate technical and organisational measures to ensure and be able to demonstrate that processing is performed by the GDPR. Those measures shall be reviewed and updated where necessary. Furthermore, Article 25 introduces the principles of data protection by design and by default. These principles are crucial for developing and deploying secure and privacy-respecting AI algorithms that proactively minimise personal data processing and embed safeguards from early design stages.

Article 32 mandates the implementation of appropriate technical and organisational measures to ensure a level of security appropriate to the risk. These measures are vital for maintaining the integrity and confidentiality of the data processing, protecting it from unauthorised access, alteration, or loss, and contributing to the overall security and reliability of algorithmic processes. Finally, Article 35 settles that organisations must conduct a Data Protection Impact Assessment (DPIA) for high-risk data processing activities, ensuring risks to individuals' rights are identified and mitigated.

Supervising the implementation and enforcement of the GDPR is primarily entrusted to the supervisory authorities established in each Member State, known as Data Protection Authorities (DPAs).

3.3. The Digital Services Act (DSA): Enhancing fairness, transparency and accountability of online platforms' algorithms

The Digital Services Act (DSA) regulates the obligations of digital services that act as intermediaries in their role of connecting consumers with goods, services, and content. This includes a range of online services and platforms. The DSA has particular provisions for very large online platforms (VLOPs) and very large online search engines (VLOSEs), which are subject to enhanced obligations due to their significant societal impact.

The DSA aims to create a safer, more predictable and trusted online environment by establishing harmonised rules for protecting users and promoting fairness, addressing illegal content and harmful activities (Regulation (EU) 2022/2065, 2022). Its main goal is, again, to

Table 1
Regulatory framework overview concerning algorithmic trustworthiness (main articles related to these terms excluding recitals).

Aspect	GDPR	DSA	AI Act	CRA
Security/ cybersecurity	Art. 32 Security of the processing	Art. 28 Online protection of minors, Art. 42 Transparency reporting obligations and Art.48 Crisis protocols	Art.9. Risk management system, Art. 15 Accuracy, robustness and cybersecurity and Art. 55 Obligations of providers of general-purpose AI models with systemic risk	Globally
Data protection/privacy	Globally	Art. 28 Online protection of minors and Art. 46 Codes of conduct for online advertising	Art. 10 Data and data governance, Art. 26 Obligations of deployers of high-risk AI systems and Art. 59 Further processing of personal data for developing certain AI systems in the public interest in the AI regulatory sandbox	–
Fairness	Art. 5 Principles relating to processing of personal data	Globally (fairness by design)	Globally (prohibited AI practices, fundamental rights impact assessment)	–
Transparency/ explainability	Art. 5 Principles relating to processing of personal data and Art.12 to 23 Rights of the data subject	Globally	Art. 13 Transparency and provision of information to deployers, Art.50 Transparency obligations for providers and deployers of certain AI systems and Art. 86 Right to explanation of individual decision-making	–
Accountability	Art. 5 Principles relating to processing of personal data, Art. 24 Responsibility of the controller, Art.25 Data protection by design and by default and Art.35 Data protection impact assessment	Globally (annual report)	Art. 17 Quality management system, Art. 18 Documentation keeping and Art.40 to Art.49 Standards, conformity assessment, certificates, registration	Globally (conformity of the product, market surveillance)
Safety	–	Art. 28 Online protection of minors and Art. 31 Compliance by design	Globally (intertwined with health, robustness and cybersecurity)	Globally (cyber-safe)

protect the fundamental rights and freedoms of natural persons (mainly users) with a special focus on vulnerable groups such as minors (Article 28, Online Protection of Minors).

Providers of intermediary services have transparency reporting obligations regarding any content moderation they engage in, including the use of automated means for this purpose (Article 15). This involves the publication of reports, including a qualitative description, the precise purposes, indicators of accuracy and error rates of automated means, and any safeguards applied. The DSA establishes additional reporting obligations for providers of online platforms (Article 24).

The DSA also impacts targeted advertising practices by requiring greater transparency of advertising algorithms (Article 26). Additionally, providers must ensure that recipients of their services are adequately informed about how their recommender systems influence the display of information (Article 27). They need to present the main parameters of these systems in an easily comprehensible manner, including how information is prioritised, potentially based on profiling and online behaviour.

VLOPs and VLOSEs are subject to enhanced obligations. For example, to ensure accountability, the DSA requires VLOPs and VLOSEs to diligently identify, analyse, and assess systemic risks stemming from the design, functioning, and use of their services (and related algorithmic systems, Article 34). This includes risks associated with the dissemination of content, fundamental rights, public discourse, and public safety. They are then obligated to implement appropriate and proportionate mitigation measures for these identified risks (Article 35). Furthermore, VLOPs and VLOSEs are subject to independent auditing to ensure compliance with their obligations under the DSA (Article 37). The DSA also encourages developing codes of conduct to address specific types of illegal content and systemic risks, such as disinformation.

Additionally, they must provide at least one option for each recommender system that is not based on profiling (Article 38). Providers of VLOPs and VLOSEs that present advertisements must compile and make publicly available a repository of advertisements with information on the content, advertiser, and targeting parameters (Article 39). This aims to facilitate supervision and research into the risks associated with online advertising. Recipients of the service should also have access to information directly from the online interface about the main parameters used for determining which advertisements are presented to them.

Upon a reasoned request, providers of VLOPs or VLOSEs must explain the design, logic, functioning, and testing of their algorithmic systems, including recommender systems (Article 40). Upon a reasoned request, these providers must also provide access to data to vetted researchers for research that contributes to the detection, identification, and understanding of systemic risks in the Union and the assessment of risk mitigation measures. Furthermore, these large providers have additional transparency reporting obligations (Article 42).

Competent authorities in the Member States (Digital Services Coordinators) and the European Commission share the responsibility for supervising DSA implementation and enforcement.

3.4. The AI Act: A risk-based framework for ensuring safety and fundamental rights in AI systems

The AI Act aims to foster the development, placement on the market, putting into service, and use of AI systems and models in the EU by its values, promoting human-centric and trustworthy AI while ensuring a high level of protection of health, safety, and fundamental rights (Regulation (EU) 2024/1689, 2024b). An AI system is defined broadly as a machine-based system designed to operate with varying levels of autonomy, exhibiting adaptiveness after deployment and inferring from input to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments. This definition does not cover systems based solely on rules defined by natural persons to execute operations automatically. The AI Act also includes harmonised rules for the placing on the market of general-purpose AI models, defining them by their generality and capability to perform a wide range of distinct tasks competently.

The AI Act adopts a risk-based approach, categorising AI systems based on their risk level. This approach includes prohibited AI practices

that pose an unacceptable risk (Article 5), high-risk AI systems subject to specific requirements and obligations (Article 6), limited-risk AI systems with certain transparency obligations, and minimal-risk AI systems.

The AI Act establishes several crucial requirements for high-risk AI systems directly relevant to algorithmic security and privacy (Articles 8 to 15). These include:

- Risk management system: Planning and running a continuous iterative risk management process throughout the entire lifecycle of the high-risk AI system (including cybersecurity risks), requiring regular, systematic review and updating.
- Data governance: Ensuring that training, validation, and testing data sets are relevant, representative, and as free of errors and complete as possible.
- Technical documentation: Requiring the creation and maintenance of detailed documentation that allows for the assessment of the AI system's compliance with the Act. This contributes to the transparency and auditability of algorithms.
- Transparency and provision of information to users: Mandating that deployers are provided with sufficient information to understand and use the AI system correctly, including its capabilities and limitations. For AI systems intended to interact with natural persons, individuals should be informed that they are interacting with an AI system.
- Human oversight: Establishing the need for human oversight mechanisms to ensure that high-risk AI systems are used appropriately and that interventions can be made when necessary.
- Accuracy and robustness: Requiring high-risk AI systems to achieve an appropriate level of accuracy and to be robust against manipulation, ensuring they perform consistently and reliably throughout their lifecycle.
- Cybersecurity: Obligating providers to implement measures to ensure adequate cybersecurity protection for high-risk AI systems, safeguarding them against threats.

The AI Act also establishes obligations for various actors involved in the development, distribution, use, and other aspects of high-risk AI systems (Articles 16 to 27). These may include conducting a Fundamental Rights Impact Assessment (FRIA) before deploying a high-risk AI system. This regulation also establishes conformity assessment procedures to ensure compliance (Article 43). Providers of high-risk AI systems must follow these procedures before placing them on the market. Finally, the AI Act establishes requirements and obligations concerning general-purpose AI (GPAI) models in Articles 51–55.

The responsibility for supervising the implementation and enforcement of the AI Act is primarily shared between the Member States (notifying authorities and market surveillance authorities) and the European Union level, explicitly involving the European Commission (acting through the AI Office) and the AI Board.

3.5. The Cyber Resilience Act (CRA): Strengthening the cybersecurity of digital products

The Cyber Resilience Act (CRA) aims to improve the overall cybersecurity of digital products placed on the European Union market: products with digital elements (any software or hardware product and its remote data processing solutions, including software or hardware components to be placed on the market separately) whose intended or reasonably foreseeable use includes a direct or indirect logical or physical data connection to a device or network.

The CRA aims to establish a horizontal regulatory framework at the European level to ensure a higher level of cybersecurity for consumers and businesses (Regulation (EU) 2024/2847, 2024). This initiative aims to mitigate vulnerabilities in digital products and ensure that manufacturers prioritise security throughout a product's lifecycle. The CRA has a broad personal scope, addressing all market participants involved in the supply chain, from manufacturers to distributors and importers. Depending on their classification, economic operators are subject to different obligations. The CRA does not grant rights to individuals, and they are not directly addressed in the regulation. In this sense, it is a regulation similar to the AI Act.

The scope of the CRA is notably broad (Article 2), ranging from consumer internet-connected devices (such as smart toys and speakers) to mobile applications (including health-monitoring apps). Importantly, this inherently includes software and devices that rely on AI algorithms, which are integral digital components. Products with digital elements are classified based on their level of cybersecurity risk under the CRA, distinguishing between default non-critical products, important products (Article 7) and critical products (Article 8).

The CRA mandates several essential cybersecurity requirements for designing, developing, and producing products with digital elements. These include the principles of security by design, ensuring that products are created with security considerations from the outset and are delivered in a secure configuration by default. Manufacturers must assess the cybersecurity risks associated with their products throughout their lifecycle to ensure these requirements are met. Additionally, manufacturers are required to ensure that their products are not delivered with known exploitable vulnerabilities. Furthermore, they must establish processes for effectively addressing vulnerabilities throughout the product's expected lifetime or for five years, whichever is shorter. This includes the provision of necessary security updates.

By mandating these requirements and defining conformity assessment procedures, the CRA significantly contributes to algorithmic security and privacy. Since products employing algorithms and AI are considered products with digital elements, the CRA ensures that the underlying digital infrastructure and software components are designed and maintained to be resilient against cyberattacks. The obligations related to vulnerability handling and security updates directly enhance the security of the algorithms and AI systems embedded within these products, making them less susceptible to data breaches or unauthorised manipulation, for example.

The responsibility for supervising the implementation and enforcement of the regulation is entrusted to national market surveillance authorities designated by each Member State, with the European Commission playing a coordinating role and ENISA (the European Union Agency for Cybersecurity) providing support.

4. A multi-layered approach to manage algorithmic risks

The four analysed regulations collectively address different facets of AI algorithms' security and privacy. Fig. 1 represents this coverage in an AI algorithm-based digital product. This figure is a simplification because the boundaries between layers are sometimes blurred, as will be analysed in the rest of this section. However, it is a good starting point for identifying the direct and indirect synergies that can be found between the four regulations in the context of this research.

4.1. Direct synergies

There are direct synergies where compliance with one regulation can facilitate compliance with others. These direct synergies can be clearly identified in pairs, between the GDPR and the DSA (with a strong focus on protecting the rights and freedoms of natural persons) and between the AI Act and the CRA (with a strong emphasis on product quality and conformity assessments).

The DSA aims to create a safe, predictable and trusted online environment, aligning with the GDPR's objective of protecting individuals in relation to the processing of personal data. Article 26(3) of the DSA explicitly prohibits presenting advertisements based on profiling using special categories of personal data as defined in Article 9(1) of the GDPR. This DSA provision directly reinforces the GDPR's protection of

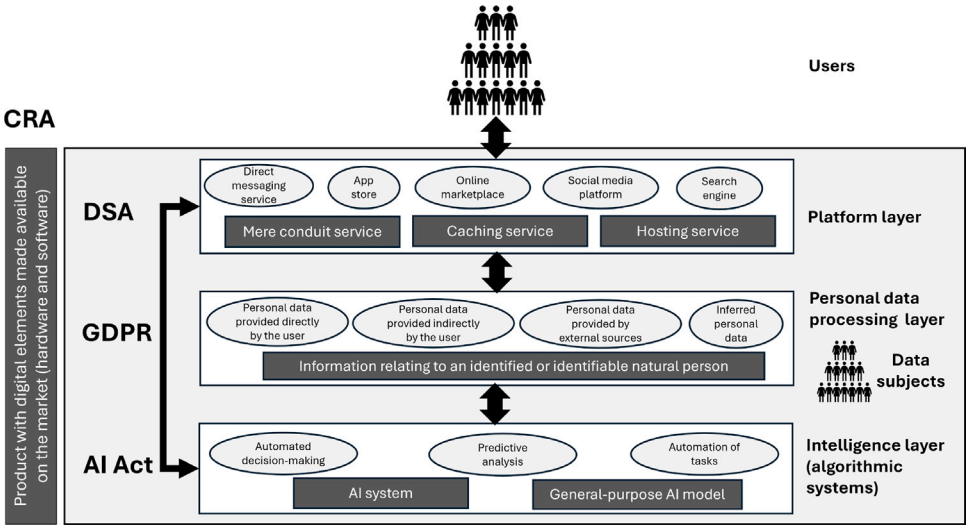

Fig. 1. Simplified relationship of regulations and AI algorithms security and privacy.

sensitive personal data in the context of online advertising. In addition, the DSA mandates content moderation and real-time reporting for online platforms. Implementing systems to support these processes might require a better understanding of the data being processed, including personal data, which is also a requirement under the GDPR for ensuring data accuracy and implementing data subject rights, for example. In general terms, the DSA's focus on fairness, transparency and accountability can support GDPR compliance by encouraging organisations to be more mindful of their data processing activities and the rights of natural persons.

On the contrary, complying with the GDPR's principles and requirements, providers are better equipped to meet DSA's expectations for protecting minors from content, conduct, contact, consumer or cross-cutting risks under Article 28, for example, OECD (2021).

There is a significant synergy where high-risk AI systems falling under both the AI Act and the scope of the CRA can demonstrate compliance with the AI Act's cybersecurity requirements by fulfilling the essential cybersecurity requirements of the CRA. When a high-risk AI system meets the CRA's requirements (Annex I), it is "deemed compliant with the cybersecurity requirements of the AI Act". The EU declaration of conformity under the CRA can serve as evidence of this compliance.

For products with digital elements classified as high-risk AI systems, the relevant conformity assessment procedure required by Article 43 of the AI Act applies. Notified bodies entitled to control the conformity of high-risk AI systems under the AI Act are also entitled to control their conformity with the CRA's requirements. This indicates a streamlined approach to conformity assessment for systems that fall under both regulations.

Furthermore, regulatory sandboxes, mandatory at the national level under the AI Act, can serve as a direct link between the AI Act and the CRA in terms of conformity assessment of products. Both regulations recognise cybersecurity as a priority within their respective sandboxes. The CRA establishes cyber resilience regulatory sandboxes, while the AI Act stipulates that AI sandboxes must facilitate the development of cybersecurity profiles for AI systems. Successful completion in a CRA sandbox could potentially demonstrate the cybersecurity aspects required by the AI Act for a high-risk AI system that is also a product with digital elements.

But there are also cross-synergies. For example, the AI Act's provisions on AI regulatory sandboxes, particularly Article 59 governing the re-use of personal data, are intrinsically linked to the GDPR. Compliance within the sandbox needs to adhere to GDPR principles regarding the processing of personal data. For example, the AI Act requires "effective monitoring mechanisms" for high risks to data subjects' rights as defined in Article 35 of the GDPR and response mechanisms. Furthermore, Data Protection Impact Assessments (DPIAs) conducted under Article 35(7)(d) of the GDPR, which include security measures for personal data, are relevant in the context of AI system development within or outside regulatory sandboxes.

The essential cybersecurity requirements outlined in the CRA are designed to enhance the protection of personal data processing and the privacy of individuals, as mandated by the GDPR. By ensuring the security of products with digital elements, the CRA may directly support the GDPR's goals, specifically compliance with Article 32. Additionally, there are potential synergies between the CRA and the GDPR in areas such as standardisation and certification of cybersecurity aspects, as well as in market surveillance and enforcement. Cooperation between national market surveillance authorities under the CRA and authorities supervising Union data protection law could be very beneficial in this sense.

4.2. *Indirect synergies*

These four key EU regulations, while focusing on specific domains, are indirectly linked by their fundamental reliance on a risk-based approach to regulation. By managing risk effectively, these regulations aim to protect individuals' rights, promote societal well-being, enhance the security of organisations, and contribute to national security by fostering a more resilient digital environment. Their ultimate aim is to mitigate or prevent harm.

Therefore, risk acts as the fundamental lens through which each regulation views the digital landscape and formulates its principles and requirements. It is the common language and the underlying principle that connects these seemingly distinct pieces of legislation. They each focus on a specific domain but share the common thread of using risk as the central concept. The GDPR requires controllers to consider the risks to the rights and freedoms of natural persons posed by processing personal data. The DSA requires providers of very large online platforms and search engines to assess systemic risks stemming from the design, functioning, and use of their services, including those related to algorithmic systems and the dissemination of illegal and harmful content. The AI Act adopts a risk-based approach, identifying AI systems and models with a high risk of harm to health, safety, and fundamental rights. The CRA requires manufacturers to assess the cybersecurity risks associated with products with digital elements throughout their lifecycle. This includes risks to the availability, integrity, and confidentiality of these products.

Furthermore, all these regulations require the assessment of the likelihood and severity of potential harms, as understanding the nature, probability, and impact of risks is crucial for determining the appropriate response. They also impose obligations and requirements proportionate to the risk: the level of regulatory scrutiny and the stringency of obligations increase with the level of risk identified. And they mandate risk management processes. Organisations are generally required to implement ongoing processes for identifying, assessing, and mitigating the risks related to their activities within the scope of these regulations.

This shared emphasis on a risk-based approach and multi-layered risk management creates an indirect synergy by establishing a common regulatory mindset across different domains of the digital landscape. Organisations operating within these domains will develop expertise and implement frameworks for identifying, assessing, and mitigating risks, which can be adapted and leveraged across their compliance efforts for each of these regulations. Considering that there may be other regulations that apply (digital but also related to the internal market, consumer protection, labour, or sectoral) and also require the assessment and management of other different risks.

For example, experience in conducting DPIAs under the GDPR can inform the risk assessments required under the DSA or the AI Act. Similarly, implementing robust cybersecurity measures as mandated by the CRA may help mitigate risks related to personal data security (GDPR) and the robustness of AI systems (AI Act). This interconnectedness, even without direct overlaps in specific obligations, fosters a more coherent and resilient digital ecosystem. The following sections develop these concepts concerning the four most important risk categories from the perspective of these regulations.

4.2.1. Organisational risks

This type of internal risk is not the focus of the four regulations analysed in this work. However, all four regulations can ultimately reduce various organisational risks, such as those producing legal liabilities, financial losses, reputational damage, or business disruptions.

The CRA directly addresses cybersecurity risks in products with digital elements, mandating essential cybersecurity requirements that organisations placing these products on the EU market must adhere to. This reduces the risk of vulnerabilities that could be exploited to harm organisations and their customers. The GDPR mandates strong data protection measures, including security of processing. Compliance with GDPR can reduce the risk of data breaches and associated organisational costs, although it is not the objective of this regulation. The DSA imposes obligations on online platforms, especially very large ones (VLOPs and VLOSEs), to manage systemic risks, including those that can affect the platform's integrity and operations, such as the spread of illegal content and disinformation. This can protect organisations that rely on these platforms. Finally, the AI Act establishes requirements for high-risk AI systems, including cybersecurity aspects, which may reduce the risk of these systems malfunctioning or being exploited, potentially leading to organisational harm.

While aiming for harmonisation, differing scopes and specific requirements across the regulations can lead to confusion for organisations operating across different sectors or offering various types of digital products and services. For example, the definition of a ''product with digital elements'' under the CRA might overlap but not perfectly align with ''high-risk AI systems'' under the AI Act or ''information society services'' under the GDPR and DSA. Enforcement and interpretation differences between national authorities for each regulation may lead to inconsistencies in how organisational risks are assessed and managed.

In addition, compliance burdens may be significant, especially for smaller organisations navigating the complexities of multiple overlapping regulations. The need to comply with specific requirements in each regulation (for example data breach notification under the GDPR and potentially the CRA) may create administrative overhead and increase compliance costs.

4.2.2. Risks to rights and freedoms

The GDPR is primarily focused on safeguarding the rights and freedoms of individuals concerning their personal data. However, the DSA also includes provisions in this area. Specifically to protect freedom of expression and information while also aiming to counter illegal content and disinformation that can harm individuals. It also requires transparency regarding content moderation decisions. The AI Act places a strong emphasis on protecting fundamental rights from the potential harms of high-risk AI systems, including requirements for transparency, human oversight, and non-discrimination. It prohibits certain AI practices deemed to pose unacceptable risks to these rights. The CRA indirectly protects individuals by reducing the risk of their devices and data being compromised due to security vulnerabilities, thereby safeguarding their privacy and fundamental rights (integrity, etc.).

There may be some tensions between security measures and data protection. For example, extensive data collection for cybersecurity purposes (intrusion detection, anti-spoofing safeguards) might conflict with the GDPR's data minimisation principle. The balance between freedom of expression and the need to moderate harmful content under the DSA may also be challenging. Overly broad content moderation could infringe on free speech, while insufficient moderation could allow the spread of harmful content that violates others' rights. The use of AI for law enforcement and border control, while potentially enhancing security, raises significant concerns about potential infringements of fundamental rights such as non-discrimination, the right to a fair trial, and the presumption of innocence, as highlighted in the AI Act. The need for transparency and human oversight in these contexts is crucial to mitigate all these risks. Finally, exceptions for journalistic, academic, artistic, or literary expression under the GDPR (Article 85) must be carefully balanced with obligations under the DSA regarding illegal content and disinformation.

4.2.3. Risks to society (systemic risks)

The DSA directly addresses systemic risks posed by VLOPs and VLOSEs, including the spread of illegal content, disinformation, manipulation of public opinion, and negative impacts on democratic processes and public health. It mandates risk assessments and mitigation measures for these platforms. The AI Act aims to prevent systemic risks arising from the deployment of high-risk AI systems in areas like critical infrastructure, education, employment, and access to essential services, which could have wide-ranging societal impacts.

Furthermore, the concept of systemic risk is explicitly mentioned in the AI Act in relation to general-purpose AI models. Providers of GPAI models presenting systemic risks are subject to additional obligations beyond those for general-purpose AI models. For example, they are required to perform model evaluation using standardised protocols and tools that reflect the state of the art. This includes conducting and documenting adversarial testing to identify and mitigate systemic risks. These providers must assess and mitigate possible systemic risks at the Union level, including the sources of these risks, which may stem from the development, placement on the market, or use of their models. They are obligated to keep track of, document, and report, without undue delay, to the AI Office and, where appropriate, to national competent authorities relevant information about serious incidents and possible corrective measures to address them. If the development or use of a model causes a serious incident despite efforts to identify and prevent risks, the provider must track the incident and report relevant information. All these obligations are closely aligned with the DSA's strategy for managing systemic risk.

The CRA enhances the security of digital products that are integral to the digital ecosystem, reducing the risk of widespread vulnerabilities that could be exploited to cause systemic disruptions across different sectors. For example, AI systems intended to be used as safety components in the management and operation of energy supply systems (such as electricity grids, gas, and heating) are classified as high-risk. A systemic disruption risk involves the failure or malfunctioning of these

AI safety components. The resulting harm can be significant, putting the lives and health of persons at large-scale risk and leading to appreciable disruptions in the ordinary conduct of social and economic activities. The same applies when considering AI systems intended to be used as safety components in the management and operation of road traffic, AI systems designed to be used as safety components in the management and operation of water supply, or AI systems intended to be used as safety components in the management and operation of critical digital infrastructure (like cloud computing centres), to mention additional examples.

Ultimately, the GDPR's emphasis on the security and accountability of personal data processing enhances the overall resilience of data-driven systems that support numerous societal functions.

Defining and assessing systemic risks can be complex and may require coordination across different regulatory frameworks. Identifying the interplay of risks arising from AI systems deployed on online platforms, for example, may require a holistic view that encompasses both the AI Act and the DSA. International cooperation is also crucial for addressing global systemic risks, but differing regulatory approaches across jurisdictions can create challenges for effective risk management. The fragmentation of regulatory agencies and authorities globally can be a significant issue.

4.2.4. Risks to national security

All four regulations can indirectly contribute to national security by enhancing the overall security and resilience of the digital environment.

The CRA's focus on the cybersecurity of products with digital elements can reduce vulnerabilities in critical infrastructure and other systems vital for national security. AI systems are involved in the digital activities of governments and organisations. Cyberattacks targeting these systems can compromise intelligence sources, military strategies, and diplomatic negotiations, leading to serious consequences for national security, including damage to physical objects or cyber–physical systems and harm to institutional processes or purposes, such as electoral processes.

The DSA's efforts to combat disinformation and manipulation can help counter information operations by hostile actors that pose a threat to national security. Disinformation, propaganda, and hate speech, facilitated by AI systems in content moderation or recommender systems, for example, can undermine norms of evidence-based truth-telling. Specific disinformation campaigns are now considered threats to national security.

The AI Act's requirements for cybersecurity in high-risk AI systems relevant to national security applications (law enforcement, border control, etc.) are crucial. The GDPR's emphasis on the security of personal data processing helps protect sensitive information that could be relevant to national security, although exceptions exist for national security purposes.

The fundamental tension between privacy and national security imperatives remains a key challenge. National security agencies may require access to data and communications that are protected under GDPR. Strong encryption (beneficial for security and privacy) can hinder intelligence gathering for national security. Therefore, the debate around encryption is very active in different contexts. Furthermore, data retention requirements for national security purposes may conflict with the GDPR's principles in specific scenarios.

The scope and application of these regulations to activities directly related to national security (intelligence operations, military cyber activities) might be unclear or subject to specific exemptions at the national level, potentially creating inconsistencies. International cooperation on national security matters may be complicated by differing data protection and cybersecurity regulations across countries.

Table 2
Examples of Organisational risks posed by the AI-powered content moderation system in the illustrative example.

#	Risk	Examples of harm
1	Insufficient resources allocated to content moderation processes and enforcement of terms and conditions	Widespread dissemination of illegal content (such as child sexual abuse material or sale of products or services prohibited by Union or national law), disinformation or misleading content (for example, negatively impacting responses to emergencies) or amplification of hate speech and propaganda may cause financial harm, reputational damage, market losses or non-compliance
2	Lack of appropriate measures for human oversight of the content moderation system	Unjustified or disproportionate detrimental treatment of individuals or groups or replication and automation of bias and discrimination may cause financial harm, reputational damage, market losses or non-compliance
3	Failure to ensure that personnel assigned to implement instructions for use and human oversight have the necessary competence and training	Inaccurate or biased decision-making in critical areas, amplified discrimination or increased administrative burdens may cause financial harm, reputational damage, market losses or non-compliance
4	Absence of a post-market monitoring system to assess the content moderation system's performance after deployment and to report serious incidents to relevant authorities	Inability to improve system design and development, lack of accountability or failure to report serious incidents may cause financial harm, reputational damage, market losses or non-compliance

5. Illustrative example

A major social media network designated as a VLOP (subject to the DSA) relies on a third-party AI-powered system to automatically detect and remove illegal and harmful content (subject to the AI Act and considered high-risk). This AI system processes vast amounts of user-generated content, often including personal data (subject to the GDPR). The entire platform, including the AI system and its underlying infrastructure, is a digital product with digital elements (subject to the CRA). In this example, the risks identified in Tables 2 (Organisational risks), 3 (Risks to rights and freedoms), 4 (Risks to society or systemic risks) and 5 (Risks to national security) can be identified. This is not an exhaustive list, and these tables only present some significant risks in a non-hierarchical, unordered way. How can a multi-layered risk-based approach be applied to assess and manage this kind of risk?

The social media network provider must first identify the elements shown in Fig. 1 (AI systems, personal data processing, offered services, digital products placed on the market, users, and data subjects). Furthermore, this provider must establish its role under each of the considered regulations. That is, the data controller of specific personal data processing activities under the GDPR, the provider of an online platform under the DSA, the deployer of a high-risk AI system under the AI Act, and the manufacturer of a digital product under the CRA. After this, the provider can understand its compliance obligations concerning the four regulations and start the risk assessment and management processes.

The application of the considered regulations is not sequential in a "first GDPR, then AI Act" manner. Instead, they are designed to be simultaneously applicable and complementary, starting at the design phase and continuing throughout the platform lifecycle. The application of these regulations is not a simple "if this, then this" linear process but rather a concurrent, continuous, overlapping framework that addresses different facets of the platform and its underlying technologies. The four considered regulations emphasise a shift from punctual or periodic risk assessments to dynamically operationalising compliance in an ongoing risk assessment and management process that should

Table 3
Examples of Risks to rights and freedoms posed by the AI-powered content moderation system in the illustrative example.

#	Risk	Examples of harm
5	Processing more personal data (IP addresses, interaction history, geolocation) than necessary for moderation	Creation of profiles and identity traces or surveillance infringe on the Right to respect for private and family life and the Protection of personal data
6	User profiling, using non-content data (user activity patterns, geolocation) to infer intent, leading to unfair account restrictions	Creation of "algorithmic groups" that share characteristics or automated discrimination and bias infringe on the Right to respect for private and family life, the Protection of personal data and the Non-discrimination and Human dignity rights
7	Lack of transparency and explainability in the decision-making process, making it difficult for users to understand why their content was flagged or removed and to seek redress	Difficulty in challenging decisions or infringement of due process and effective remedy violate the Protection of personal data, the Right to an effective remedy and to a fair trial, the right to Human dignity and the Freedom of expression and information
8	Legal content is mistakenly flagged and removed due to inaccuracies or biases	Chilling effects on civic discourse undermine Freedom of expression and information and Freedom of assembly and association (as restrictions on expression can indirectly impact assembly)
9	Unfair removal or prioritisation of content from certain groups	Perpetuating historical patterns of discrimination or associating content with algorithmic group characteristics (translating it to unfair treatment) infringes on the Non-discrimination and Human dignity rights
10	Potential manipulation of information users are exposed to (function creep)	Inhibiting decisional autonomy or subverting free choice disregard the Right to respect for private and family life, the Protection of personal data and the Human dignity and Consumer protection rights

Table 4
Examples of Risks to society (systemic risks) posed by the AI-powered content moderation system in the illustrative example.

#	Risk	Examples of harm
11	Potential for large-scale data breaches	Massive violation of privacy and data protection rights, increased likelihood of misuse and discrimination or loss of public trust impact on fundamental rights and on public security
12	Dissemination of illegal or harmful content on a large scale because the content moderation system is not fully effective	Direct harm to individuals including minors, the erosion of the integrity of the information ecosystem or incitement to violence and crime have negative effects on democratic processes, civic discourse, and electoral processes, as well as public security and negative effects on public health, minors, and serious negative consequences to a person's physical and mental well-being
13	Widespread deployment of a high-risk AI system for content moderation exhibiting biases or inaccuracies and affecting a large user base	Detrimental or unfavourable treatment of certain natural persons or whole groups, compromised justice and due process or massive automated discrimination impact on fundamental rights and have negative effects on democratic processes, civic discourse, electoral processes and the rule of law
14	Potential for the content moderation system to contribute to the systemic spread of disinformation or harmful content because it is not sufficiently robust against manipulation or adversarial attacks	Direct harm to individuals including minors, the erosion of the integrity of the information ecosystem or incitement to violence and crime have negative effects on democratic processes, civic discourse, and electoral processes, as well as public security and negative effects on public health, minors, and serious negative consequences to a person's physical and mental well-being

Table 5
Examples of Risks to national security posed by the AI-powered content moderation system in the illustrative example.

#	Risk	Examples of harm
15	Foreign interference in democratic processes through the spread of disinformation and propaganda	Cognitive warfare or identity-based disinformation campaigns cause political instability and polarisation, erosion of public trust in institutions, legitimacy crises, strategic weakening, violence and undermine democratic sovereignty
16	Facilitate illegal activities or the spread of extremist content	Terrorist recruitment and operations, incitement to violence or financial scams lead to societal discord, political instability and polarisation, erosion of public trust in institutions and violence
17	State-sponsored actors or cybercriminals may disrupt services or conduct espionage exploiting cybersecurity vulnerabilities in a major social media platform which can be considered critical infrastructure in terms of public communication and information dissemination	Intelligence gathering and sensitive data theft or hindering emergency communications, crisis management, and the dissemination of vital information to the public cause erosion of public trust in institutions, legitimacy crises, strategic weakening and undermine democratic sovereignty

be continually conducted, updated, and improved. Compliance is a continuous and dynamic process rather than a one-time event.

Fig. 2 shows a simplified 3-phase procedure for this illustrative example. Phase 1, Design and pre-deployment, involves the initial acquisition or procurement of the AI system, its integration into the VLOP's operations, and all preparatory assessments and setup before the system is actively used for content moderation.

The GDPR mandates that the social media network provider, as data controller, must have a lawful basis for processing users' personal data for content moderation. The GDPR also requires compliance with the data protection by design and by default principle in the integration of the AI system.

This specific personal data processing is likely to result in a high risk to the rights and freedoms of natural persons: a social media network provider using AI for content moderation involving vast amounts of user-generated content, potentially including special categories of personal data. The controller shall, before the processing, assess the impact of the envisaged processing operations on the protection of personal data (Data Protection Impact Assessment, DPIA). Additionally, it must be considered that a controller wishing to reuse personal data in the context of developing or using high-risk AI systems typically triggers the obligation to carry out a DPIA. Effective monitoring mechanisms and response mechanisms to mitigate risks and halt processing, when necessary, should be identified within the relevant DPIA, in addition to its usual content regarding risks to rights and freedoms.

The primary goal of a DPIA is to assess the particular likelihood and severity of the high risk to the rights and freedoms of natural persons arising from the processing of personal data. It helps controllers identify and mitigate these risks before processing takes place. The DPIA will help identify and mitigate risks ranging from 5 to 10, as listed in Table 3. However, this also applies from 1 to 4 (Table 2) or 11 to 14 (Table 4) from a data protection and rights and freedoms perspective. Organisational risks can have direct implications for individuals' rights and freedoms, and systemic risks in digital ecosystems can easily manifest as data protection risks, for example. This interconnectedness makes categorising risks into distinct, isolated categories almost impossible.

As a VLOP, the DSA imposes specific obligations to the social media network provider regarding the algorithmic systems it uses, including content moderation. The social media network provider must conduct

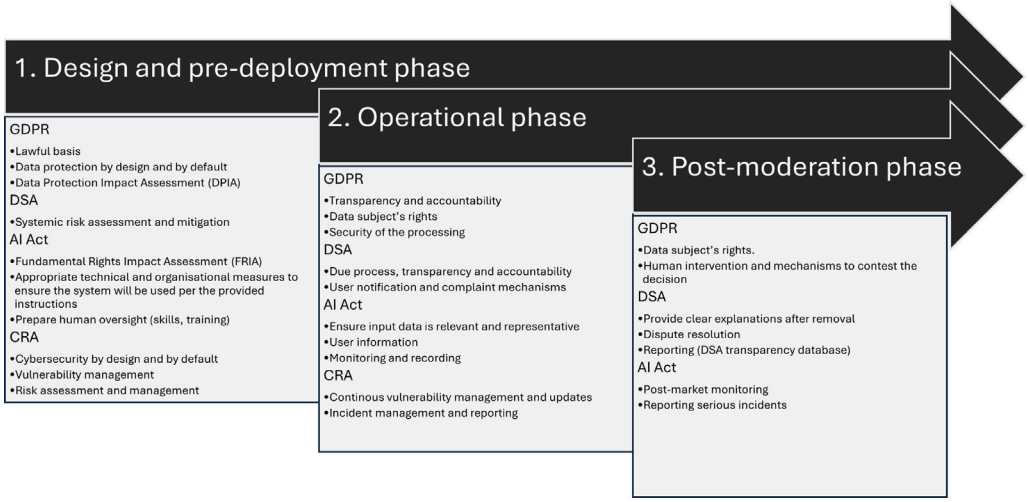

Fig. 2. Simplified procedure applied in the illustrative example by the social media network provider.

a specific risk assessment proportionate to systemic risks, considering their severity and probability. It must include risks related to the dissemination of illegal content, any actual or foreseeable adverse effects on the exercise of fundamental rights (including freedom of expression, data protection, non-discrimination, and the rights of the child), risks arising from the misuse of their service (weaponised disinformation, public opinion manipulation and other risks to national security), any actual or foreseeable adverse effects on civic discourse or electoral processes, or any actual or foreseeable adverse effects concerning gender-based violence, the protection of public health and minors and severe negative consequences to the person's physical and mental well-being. When conducting this risk assessment, the VLOP must consider, in particular, whether and how the following factors influence systemic risks: the design of their AI content moderation system, the applicable terms and conditions and their enforcement, and the data-related practices.

The social media network provider must diligently deploy the necessary means to mitigate the systemic risks identified in this risk assessment, respecting fundamental rights and ensuring reasonable, effective, and proportionate measures. The provider should ensure that its approach to risk assessment and mitigation is based on the best available information and scientific insights. Furthermore, that assumptions are tested with impacted groups, involving representatives of recipients, potentially affected groups, independent experts, and civil society organisations where appropriate. The obligations on risk assessment and mitigation may trigger the need to assess and adjust their content-moderation system's design, deployment or configuration to prevent or minimise biases that lead to discrimination, for example. This process will help identify and mitigate risks from 11 to 14 in Table 4. But also from 15 to 17 or 5 to 10 (from a fairness, transparency, accountability and societal impacts point of view). Therefore, the results of the DPIA and this assessment may be complementary in some respects.

The AI-powered content moderation system is classified as high-risk under the AI Act because it makes decisions that significantly impact freedom of expression. It is subject to detailed obligations for the content moderation system providers (the developers) and the social media network provider (the deployer).

Since using AI for content moderation may have a significant societal impact, the provider is likely to be required to conduct a Fundamental Rights Impact Assessment (FRIA) before deploying the system. This assessment aims to identify specific risks to individuals' rights and to propose measures for mitigating them. The FRIA must identify the deployer's relevant processes where the high-risk AI system will be used, describe the period and frequency of use and the specific categories of individuals likely to be affected, identify particular risks of harm to the fundamental rights of those persons or groups, and determine measures to be taken in case of materialisation of those risks, including governance arrangements and redress procedures. This process will help identify and mitigate risks from 5 to 10 in Table 2. But also from 1 to 3, always from the point of view of an AI system (not data processing, not service). If any of the FRIA obligations are already met through the DPIA, the FRIA should complement the DPIA for the rest of them. But they are different assessments, and this should be clear.

As the deployer of this system, the social media network provider has other specific obligations, including taking appropriate technical and organisational measures to ensure that they use the system according to the provided instructions. Furthermore, to prepare human oversight measures.

The CRA aims for cybersecurity by design and by default principles and imposes a duty of care for the lifecycle of products, including a vulnerability management process. The social media network provider, as the manufacturer, must address cybersecurity risks associated with the product category throughout its lifecycle (planning, design, development, manufacturing, distribution, and maintenance), meeting the essential cybersecurity requirements outlined in Annex I of the CRA. Remember that for products with digital elements classified as high-risk AI systems (such as content moderation systems), the CRA introduces a presumption of conformity: fulfilling the requirements of the CRA's Annex I is deemed to be compliant with the cybersecurity requirements of the AI Act.

Furthermore, risk assessment is a central component of the CRA. As a manufacturer of a product, the social media network provider must assess the cybersecurity risks associated with this product. This process will help identify and mitigate almost all risks in Tables 2, 3, 4 and 5 from a product cybersecurity perspective: protection against unauthorised access, data integrity, availability, and minimisation of impact from vulnerabilities. The GDPR's requirement for the security of personal data processing and the DPIA are linked to some of the results of this cybersecurity assessment. This risk assessment under the CRA, when considering the content moderation system, should consider risks to the cyber resilience of the AI system regarding attempts by unauthorised third parties to alter its use, behaviour, or performance, including AI-specific vulnerabilities like data poisoning and adversarial attacks, as well as relevant risks to fundamental rights as required by the AI Act.

Suppose that the VLOP performs substantial modifications to the third-party AI system. In that case, it is considered a manufacturer under the CRA and becomes subject to all manufacturer obligations,

including cybersecurity risk assessment and conformity assessment. During real-world testing conditions of the AI system, market surveillance authorities may require modifications to testing aspects if the conditions for testing are not met, implying the deployer's involvement in ensuring security compliance during this design and pre-deployment phase.

Phase 2, the Operational phase, covers the ongoing use of the AI system for content moderation, including processing content, making decisions, and managing its performance.

Data protection principles under the GDPR, such as fairness and transparency, require the platform to be clear with users about how the content moderation system uses their data. Additionally, data subjects must be able to exercise their rights to access, rectify, and erase their personal data, as well as other relevant rights, concerning the data used by the content moderation system. If content moderation makes decisions that significantly affect users, Article 22 of the GDPR, concerning automated individual decision-making (including profiling), would be relevant, granting users the right to human intervention and information about the logic involved.

In this phase, appropriate technical and organisational measures should be implemented to ensure the security of the personal data processed by the AI system. As mentioned before, this directly supports the security objectives of the CRA and the AI Act.

The DSA requires transparency regarding how these algorithms operate, including the primary parameters that determine content moderation decisions. Users must be informed about the use of automated tools and have mechanisms to report complaints or contest decisions. In this case, the DSA's focus on algorithmic transparency and accountability complements some of the GDPR and AI Act's requirements for transparency and human oversight.

Deployers must ensure that input data is relevant and sufficiently representative and inform users that they are subject to a high-risk AI content moderation system, including the system's purpose and the types of decisions it makes. The social media network provider must monitor the functioning of the AI system and keep relevant records. Furthermore, it must cooperate with the competent authorities in any actions they take related to the content moderation system. The AI Act also mandates cybersecurity measures for high-risk AI systems, which aligns with and reinforces the objectives of the CRA, focused on continuous risk, vulnerability and incident management.

Finally Phase 3, Post-moderation, refers to actions taken by the VLOP and related parties after the AI system has made or assisted in making content moderation decisions.

As a data controller under the GDPR, the provider must facilitate data subjects' rights about decisions made or assisted by the AI system, including the right to an explanation of automated decisions or human intervention.

The DSA requires providers to provide clear and sufficiently detailed information to recipients when content is removed or access is restricted, including the reasons and means of redress. There must be clear and accessible mechanisms for redressal and appeal of decisions made regarding content moderation. The provider may be required to provide regulators with access to or reporting specific data, including data on the accuracy, functioning, and testing of algorithmic systems for content moderation.

The AI Act establishes the obligation to implement post-market monitoring to analyse the system's performance, especially for AI systems that "learn" after deployment, and to address emerging risks. In addition, to report serious incidents resulting from the use of their AI systems to the relevant authorities. The CRA does not establish explicit post-moderation obligations.

6. Discussion

The example in the previous section shows how the social media network provider faces a complex web of risk management obligations spanning data protection, platform regulation, AI governance, and cybersecurity. Compliance requires a holistic approach, considering the interconnectedness of these regulations and the roles the provider plays as a data controller, online platform provider, AI system deployer and manufacturer of a product with digital elements. Similar reasoning could be applied in other scenarios, such as algorithmic recommendations or algorithms to target advertisements for users within VLOPs, AI-based age estimation solutions, algorithmic management in the workplace, AI-powered smart home devices, or AI wearables used for monitoring and diagnosing purposes.

This paper has proposed a strategy for enhancing the effectiveness of risk management by moving beyond isolated processes to a multi-layered approach that actively leverages efforts and fosters synergies. It is essential to consider that these processes differ significantly in their focus, the entities responsible for conducting them, and the scope of the assessments. Furthermore, in their triggering events: the DPIA is triggered by the likelihood of high risk to individuals from data processing (before the processing), the DSA systemic risk assessment is mandatory for VLOPs/VLOSEs annually and before deploying critical functionalities, the FRIA is mandatory for specific deployers of certain high-risk AI systems before use, and the cybersecurity risk assessment is an ongoing obligation for manufacturers throughout the lifecycle of products with digital elements.

Therefore, it is not about "reducing efforts" or "reusing"; the proposal is about "connecting". By intentionally connecting these risk assessment and management processes (methods, expertise, teams, resources), organisations can better understand potential risks, identify overlaps and interdependencies that might otherwise be missed, and ultimately develop more robust and targeted mitigation strategies. These strategies differ substantially in nature and scope, but they are complementary and must work together. At the same time, there may be overlaps and synergies. The safeguards and mitigation measures under the GDPR, DSA, AI Act and CRA address a different range of risks. This integrated method aims to ensure that the combined outcome of the four processes is significantly more complete, richer, and extensive than the sum of their individual parts would be if conducted independently. In addition, it mirrors the trend in EU regulation towards increasingly integrated approaches to governance, recognising the limitations of simple, isolated measures.

A multi-layered approach to risk management, while offering numerous benefits, also presents several challenges that should be considered:

- Potential for information asymmetries and opaqueness: When assessing risks associated with complex technologies such as AI algorithms, inherent opaqueness and information asymmetries can make it difficult to fully understand how different risks might interact and manifest across layers.
- Difficulty in defining scope and interdependencies: Defining the precise scope of each assessment and how they influence each other requires careful consideration given potential vague legal definitions or formulations and complex interconnections between some risks. As previously discussed, some terms are not explicitly defined in the legal framework, or they have different meanings in each regulation. In addition, risks are often multifaceted, with cybersecurity risks impacting rights and freedoms or systemic risks that have implications for data protection. While the aim is to leverage synergies, the lack of a common vocabulary or poorly coordinated efforts in a multi-layered approach could lead to overlapping scope and redundant analysis, therefore, to an inefficient use of resources.

- Balancing competing objectives and potential conflicts: Different risk assessments might highlight competing objectives. For example, robust cybersecurity measures involve extensive data collection for threat detection, which could conflict with the data minimisation principle. Balancing these competing imperatives requires careful consideration and may involve difficult trade-offs.
- Methodological inconsistencies and lack of harmonisation: Different types of risk assessment may employ distinct methodologies and frameworks. For instance, a DPIA, as outlined in GDPR, focuses on risks to individuals' rights and freedoms related to data processing. In contrast, a cybersecurity risk assessment, as discussed in the context of the AI Act or the CRA, might prioritise threats to the confidentiality, integrity, and availability of products. Bridging these methodological gaps and establishing a harmonised approach across layers can be challenging.
- Evolving nature of risks and regulations: Both the threat landscape and the regulatory environment are constantly evolving. Regulatory uncertainty, or lack of coordination or clarity, might discourage innovation and delay the market entry of new systems. Therefore, maintaining the relevance and effectiveness of a multi-layered risk assessment approach is essential, even if it requires continuous reassessment and adaptation to new threats and regulatory changes.
- Risk of focusing on compliance over genuine risk management: If the multi-layered approach is not well designed, there is a risk of it becoming a bureaucratic exercise concentrated on compliance with multiple requirements ("checkbox compliance") rather than a genuine effort to identify and mitigate interconnected risks effectively. This could lead to a "compliance before trustworthiness" scenario, undermining the intended purpose.

7. Conclusions and future work

The European Union has established a multi-faceted regulatory landscape with the General Data Protection Regulation (GDPR), the Digital Services Act (DSA), the Artificial Intelligence Act (AI Act) and the Cyber Resilience Act (CRA). These instruments have been crucial in establishing data protection, safety, and security standards, emphasising principles such as fairness, transparency or accountability and guaranteeing fundamental rights. Furthermore, introducing risk-based frameworks, setting mandatory requirements and establishing conformity assessment procedures.

This multi-layered regulatory approach is crucial for facing the challenges presented by the increasing deployment of AI algorithms across different sectors. This paper has demonstrated that the combination of these four regulations, considering their direct and indirect synergies, enables holistic management of the different risks associated with AI systems, ranging from privacy breaches and discriminatory outcomes to security vulnerabilities.

An exciting line for future work is proposing a unified risk management framework for AI algorithmic systems that helps to overcome the challenges discussed in the previous section. This framework could encompass standardised vocabulary to ensure clarity and consistency across different regulatory domains, as well as standard methods for impact assessment that integrate both security and privacy considerations throughout the algorithm lifecycle. The result of these unified tools would be a new approach to performing AI algorithmic audits that evaluate the security and privacy of algorithmic systems. Such a unified framework would contribute to a more streamlined and practical approach to algorithmic governance in the EU, enabling better risk identification, mitigation, and, ultimately, greater trust.

Additionally, other aspects of trust, not technical but human, could be investigated with current or future legislation concerning AI algorithms, such as diverse participation, education or socio-technical design.

Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this work, the author used Grammarly to improve language and readability, only for editing her own texts.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This research was funded by the Agencia Española de Protección de Datos (AEPD), the Spanish Data Protection Authority. The author thanks the entire AEPD team and specifically the staff of the Innovation and Technology Division for their involvement in this project and all their generous comments and contributions.

Data availability

No data was used for the research described in the article.

References

Abbas, R., Michael, K., Pitt, J., Vogel, K.M., Zafeirakopoulos, M., Artificial Intelligence (AI) in Cybersecurity: A Socio-Technical Research Roadmap, The Alan Turing Insitute. 2023.

Allahrakha, N., 2023. Balancing cyber-security and privacy: legal and ethical considerations in the digital age. Leg. Issues Age (2), 78–121.

Amoo, O.O., Atadoga, A., Osasona, F., Abrahams, T.O., Ayinla, B.S., Farayola, O.A., et al., 2024. GDPR's impact on cybersecurity: A review focusing on USA and European practices. Int. J. Sci. Res. Arch. 11 (1), 1338–1347.

Bagni, F., Seferi, F. (Eds.), 2025. Regulatory sandboxes for AI and cybersecurity, questions and answers for stakeholders. In: SERICS–Eraclito and CybeRights programs.

Baldini, D., et al., 2025. Legislative intersection perspectives on regulatory sandboxes: Navigating the interplay between the AI Act and the GDPR. In: White Paper on Regulatory Sandboxes for AI and Cybersecurity. Cybersecurity National Lab.

Belanger, F., Hiller, J.S., Smith, W.J., 2002. Trustworthiness in electronic commerce: the role of privacy, security, and site attributes. J. Strat. Inf. Syst. 11 (3–4), 245–270.

Artificial Intelligence systems and the GDPR: A data protection perspective, https://www.autoriteprotectiondonnees.be/publications/artificial-intelligence-systems-and-the-gdpr—a-data-protection-perspective.pdf (2024).

Brkan, M., 2019. Do algorithms rule the world? Algorithmic decision-making and data protection in the framework of the GDPR and beyond. Int. J. Law Inf. Technol. 27 (2), 91–121.

Busuioc, M., 2021. Accountable Artificial Intelligence: Holding algorithms to account. Public Adm. Rev. 81 (5), 825–836.

Busuioc, M., 2022. AI algorithmic oversight: new frontiers in regulation. In: Handbook of Regulatory Authorities. Edward Elgar Publishing, pp. 470–486.

Busuioc, M., Curtin, D., Almada, M., 2023. Reclaiming transparency: contesting the logics of secrecy within the AI Act. Eur. Law Open 2 (1), 79–105. http://dx.doi.org/10.1017/elo.2022.47.

Bygrave, L.A., 2025. The emergence of EU cybersecurity law: A tale of lemons, angst, turf, surf and grey boxes. Comput. Law Secur. Rev. 56, 106071.

Charter of Fundamental Rights of the European Union, 2000. https://www.europarl.europa.eu/charter/pdf/text_en.pdf.

Chiara, P.G., 2024. Towards a right to cybersecurity in EU law? The challenges ahead. Comput. Law Secur. Rev. 53, 105961.

CNIL, 2025. AI and GDPR: new recommendations to support responsible innovation. https://www.cnil.fr/en/ai-and-gdpr-cnil-publishes-new-recommendations-support-responsible-innovation.

EDPB, 2020. Guidelines 4/2019 on Article 25 data protection by design and by default, version 2.0. https://www.edpb.europa.eu/sites/default/files/files/file1/edpb_guidelines_201904_dataprotection_by_design_and_by_default_v2.0_en.pdf.

EDPB, 2024. Opinion 28/2024 on certain data protection aspects related to the processing of personal data in the context of AI models. https://www.edpb.europa.eu/system/files/2024-12/edpb_opinion_202428_ai-models_en.pdf.

Ferreira, A., Goldman, A., 2025. ComplOps research: Navigating the digital regulation revolution. hal-04930047.

Fischer-Hübner, S., Alcaraz, C., Ferreira, A., Fernandez-Gago, C., Lopez, J., Markatos, E., Islami, L., Akil, M., 2021. Stakeholder perspectives and requirements on cybersecurity in Europe. J. Inf. Secur. Appl. 61, 102916.

Fortes, P.R.B., Baquero, P.M., Amariles, D.R., 2022. Artificial Intelligence risks and algorithmic regulation. Eur. J. Risk Regul. 13 (3), 357–372.

Grochowski, M., Jablonowska, A., Lagioia, F., Sartor, G., 2021. Algorithmic transparency and explainability for EU consumer protection: unwrapping the regulatory premises. Crit. Anal. L. 8, 43.

Junklewitz, H., Hamon, R., André, A., Evas, T., Soler Garrido, J., Sanchez Martin, J., 2023. Cybersecurity of Artificial Intelligence in the AI Act. Luxemb.: Publ. Off. Eur. Union. 10, 271009.

Lazcoz, G., De Hert, P., 2023. Humans in the GDPR and AI governance of automated and algorithmic systems. essential pre-requisites against abdicating responsibilities. Comput. Law Secur. Rev. 50, 105833.

Li, B., Qi, P., Liu, B., Di, S., Liu, J., Pei, J., Yi, J., Zhou, B., 2023. Trustworthy AI: From principles to practices. ACM Comput. Surv. 55 (9), 1–46.

Liu, H., Wang, Y., Fan, W., Liu, X., Li, Y., Jain, S., Liu, Y., Jain, A., Tang, J., 2022. Trustworthy AI: A computational perspective. ACM Trans. Intell. Syst. Technol. 14 (1), 1–59.

Mantelero, A., Vaciago, G., Samantha Esposito, M., Monte, N., 2020. The common EU approach to personal data and cybersecurity regulation. Int. J. Law Inf. Technol. 28 (4), 297–328.

Michael, K., Kobran, S., Abbas, R., Hamdoun, S., 2019. Privacy, data rights and cybersecurity: Technology for good in the achievement of sustainable development goals. In: Proceedings of the IEEE International Symposium on Technology and Society. ISTAS, pp. 1–13.

Musch, S., Borrelli, M.C., Kerrigan, C., 2024. Bridging compliance and innovation: A comparative analysis of the EU AI Act and GDPR for enhanced organisational strategy. J. Data Prot. Priv. 7 (1), 14–40.

NIST, 2018. NIST special publication 800-37, revision 2: Risk management framework for information systems and organizations, a system life cycle approach for security and privacy. https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-37r2.pdf.

Novelli, C., Casolari, F., Hacker, P., Spedicato, G., Floridi, L., 2024. Generative AI in EU law: Liability, privacy, intellectual property, and cybersecurity. Comput. Law Secur. Rev. 55, 106066.

OECD, 2021. Children in the digital environment. revised typology of risks. https://www.oecd.org/content/dam/oecd/en/publications/reports/2021/01/children-in-the-digital-environment_9d454872/9b8f222e-en.pdf.

Olukoya, O., 2022. Assessing frameworks for eliciting privacy & security requirements from laws and regulations. Comput. Secur. 117, 102697.

Pollak, D.G., Chapter III. The European Union Digital Strategy: GDPR, DSA, DMA and AI Act, Smart Cities, Artificial Intelligence and Digital Transformation Law, 47.

Rampášek, M., 2023. AI cybersecurity standardisation and its overlap with DSA and CRA. Acta Fac. Iuridicae Univ. Comen. 42 (2), 14–14.

Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 april 2016 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), https://eur-lex.europa.eu/eli/reg/2016/679/oj/eng (2016).

Regulation (EU) 2019/881 of the European Parliament and of the Council of 17 April 2019 on ENISA (the European Union Agency for Cybersecurity) and on information and communications technology cybersecurity certification and repealing Regulation (EU) no 526/2013 (Cybersecurity Act), https://eur-lex.europa.eu/eli/reg/2019/881/oj/eng (2019).

Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 october 2022 on a single market for digital services and amending directive 2000/31/ec (Digital Services Act), https://eur-lex.europa.eu/eli/reg/2022/2065/oj/eng (2022).

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), https://eur-lex.europa.eu/eli/reg/2024/1689/oj/eng (2024).

Regulation (EU) 2024/2847 of the European Parliament and of the Council of 23 october 2024 on horizontal cybersecurity requirements for products with digital elements and amending regulations (EU) no 168/2013 and (EU) 2019/1020 and Directive (EU) 2020/1828 (Cyber Resilience Act), https://eur-lex.europa.eu/eli/reg/2024/2847/oj/eng (2024).

Reichman, A., Sartor, G., 2021. Algorithms and Regulation. Cambridge University Press, pp. 131–181.

Riefa, C., Protecting vulnerable consumers in the digital single market. Eur. Bus. Law Rev. 33 (4).

Schwartmann, R., Keber, T., Zenner, K., Kurth, S., 2024. Data protection aspects of the use of Artificial Intelligence—initial overview of the intersection between GDPR and AI Act. Comput. Law Rev. Int. 25 (5), 145–150.

Seizov, O., Wulf, A.J., Artificial Intelligence and transparency: a blueprint for improving the regulation of AI applications in the EU. Eur. Bus. Law Rev. 31 (4).

Seyfert, R., 2022. Algorithms as regulatory objects. Inf. Commun. Soc. 25 (11), 1542–1558.

Tamò-Larrieux, A., Guitton, C., Mayer, S., Lutz, C., 2024. Regulating for trust: Can law establish trust in Artificial Intelligence? Regul. Gov. 18 (3), 780–801.

Ulbricht, L., Yeung, K., 2022. Algorithmic regulation: A maturing concept for investigating regulation of and through algorithms. Regul. Gov. 16 (1), 3–22.

Van Bekkum, M., Borgesius, F.Z., 2023. Using sensitive data to prevent discrimination by Artificial Intelligence: Does the GDPR need a new exception? Comput. Law Secur. Rev. 48, 105770.

Willson, M., 2019. Algorithms (and the) everyday. In: The Social Power of Algorithms. Routledge, pp. 137–150.

Wolff, J., Lehr, W., Yoo, C.S., 2023. Lessons from GDPR for AI policymaking. Va. J. Law Technol. 27, 1.

Wulf, A.J., Seizov, O., 2024. Please understand we cannot provide further information: evaluating content and transparency of GDPR-mandated AI disclosures. AI SOCIETY 39 (1), 235–256.

Yanamala, A.K.Y., Suryadevara, S., Kalli, V.D.R., 2024. Balancing innovation and privacy: The intersection of data protection and Artificial Intelligence. Int. J. Mach. Learn. Res. Cybersecur. Artif. Intell. 15 (1), 1–43.

Zhang, Z., Gupta, B.B., 2018. Social media security and trustworthiness: overview and new direction. Future Gener. Comput. Syst. 86, 914–925.