In the Era of AI

Exploring New Frontiers in Cybercrime and Safeguarding Personal and Health Data*

Bianca Nicla Romano**

ABSTRACT:
The rapid evolution of Artificial Intelligence (AI) and its widespread application across various human endeavours pose significant challenges. On one hand, AI brings notable advantages, particularly in the healthcare sector, which have already reshaped and may further transform individuals' lives. On the other hand, concerns arise regarding its impact on cybersecurity, with implications for the protection of data, especially healthcare data used to train AI systems (AIs). This paper explores the complex relationship between AI, cybersecurity and the safeguarding of healthcare data through an analysis of EU and national regulations. It underscores the pressing need for more robust protective measures, particularly given the challenges associated with AIs and large language models.

1. Introduction

The subject matter of this paper has been extensively debated recently, particularly given its pertinence to the necessity of aligning the protection of personal data[1] – notably health

* Contributo sottoposto a revisione tra pari in doppio cieco.

** Researcher in Administrative Law at "Parthenope" University of Naples; biancanicla.romano@uniparthenope.it.

ISSN 2611-8882

data – with the evolving landscape shaped by the virtual world and Artificial Intelligence (hereafter referred to as AI), of which, gradually, greater use is being made.
The evolutionary prospects have rapidly become tangible realities within a very short period. Consequently, we are now experiencing a remarkable cultural evolution characterized by increased intelligence and speed, achieved through the integration of life and human relationships into virtual reality. In this dimension, technology, having become an essential part of daily life, has the merit of improving its quality through the significant increase in services dependent on technological tools, while also providing opportunities for distraction, such as gaming[2], which serves as an indispensable form of entertainment, particularly but not exclusively for young people, and which has become even more widespread during the COVID-19 pandemic lockdown period.
As a consequence of the rapid evolution of the *Internet of Things* (IoT), many common goods – ranging from washing machines and watches to automobiles, credit cards, and, more recently, identity cards – have transitioned into digital tools. These digital tools, previously exclusive to engineers and computer scientists, are now accessible and usable by ordinary individuals, who, driven by both necessity and the allure of their potential, are now able to integrate them effortlessly into their daily lives.
The expansion of the Internet, the Web, and the IoT has led to a substantial increase in data production. This, in turn, has facilitated data collection and predictive analysis, which has contributed to the expansion of AI. Notably, in recent years, AI has experienced a significant surge, attributed to the spread of its technologies. The introduction of natural language processing has further augmented AI functionalities, fostering widespread adoption and accessibility for a diverse user base attracted by its promising capabilities.
Where and why does the debate arise, then, if AI contributes to improving human life and has become easily accessible? It arises, on one hand, from the need to ensure strong and adequate protection to guarantee the same rights recognized to individuals in real life within the virtual realm, and on the other hand, from the awareness that, nevertheless, it is not easy to establish suitable rules given the continuous and rapid evolution of AI. By facilitating cyber-attacks, AI unfortunately has a significant impact on cybercrime, endangering the security of the cyberspace in which data is circulated and, consequently, their protection itself.
Indeed, AI's automatic languages (Large Language Models, LLMs) are no longer solely intended, as they once were, to carry out mundane tasks, as they have been refined to the

[1] The first legislation regarding the protection of individuals' personal data traces back to Directive 95/46/EC on personal data protection. This directive aimed to harmonize national legislation concerning data protection due to the growing transboundary data flows among public and private entities following the establishment of the Single Market. Subsequently, Directive 95/46/EC was superseded by the General Data Protection Regulation (GDPR) (Regulation 2016/679), which addresses personal data processing and the free movement of such data. For further insights, refer to M. Maurino, Cybersecurity*, sicurezza nazionale e trattamento dei dati personali,* in *amministrativ@mente*, 2/2023, 939-972.

[2] This term corresponds to the extensive spread of video games, including online.

point of being capable of complex creative work. It is therefore easy to exploit them to generate more convincing phishing emails, facilitate automated social engineering attacks, create spam scripts, and spread online disinformation[3], to the detriment of cybersecurity.
This describes what has been observed to occur with one of the most renowned Large Language Model currently in use, specifically ChatGPT[4]. ChatGPT is an LLM capable of generating textual responses based on user queries by using a database consisting of internet pages from which it extracts and processes texts to produce coherent responses[5]. Although ChatGPT was not originally designed for malicious purposes and lacks direct ties to cybercrime, it inadvertently facilitates the commission of illicit activities. This is due to its capacity to expedite the identification of vulnerabilities within systems, which are subsequently exploited by hackers to execute more efficient and tailored automated attacks. These activities may result in potential human rights violations.
The repercussions of these developments in the healthcare sector are substantial, given the extensive use of AI across various domains (from disease diagnosis to drug development and drug interaction assessment), thereby accessing a vast amount of data. This data becomes a potential and easy target for cyberattacks like *ransomware* and *supply chain breaches*[6]. Additionally, there are frequent concerns regarding the security of electromedical devices, which are particularly exposed and vulnerable as they are designed solely for specific purposes (such as monitoring heart rate or administering medications), without simultaneous assurance of the security of the data they collect.
Balancing the protection of personal data with the evolution driven by the digital world and AI represents an increasingly pressing need in today's context, particularly as the use and reliance on these technologies expand[7]. The greatest difficulties appear to lie in applying regulations to the learning capacity of machines, which, being based on algorithmic systems, can only mechanically apply rules to particular cases, relying solely on the data that convey their content. This leads to legal uncertainty, which fuels the category of risks

[3] Regarding the so-called 'vulnerabilities' in computer systems, reference may be made to B. N. Romano, *Il rischio di "attacchi" ai sistemi informatici tra fattispecie penalmente rilevanti, tutela dei dati ed esigenze di "buona amministrazione",* in *amministrativ@mente, Rivista scientifica trimestrale di diritto amministrativo*, 3/2021, pp. 545-594.

[4] Chat Generative Pretrained Transformer ('ChatGPT') was developed by the OpenAI Artificial Intelligence research lab in November 2022 and represents a prototype chatbot based on AI capable of providing responses to user inputs (i.e., queries). Specifically, the chatbot is a software designed to simulate a conversation with a human being. See I. Coppola, *Intelligenza artificiale generativa: GPT o chatGPT verso ipotesi di ragionamento automatici e linguaggio artificiale. Tra l'argomentazione giuridica di Bobbio ed il principio di precauzione*, in *Diritto di internet*, March 30, 2023.

[5] ChatGPT differs from Google in that it does not respond to the request by suggesting internet pages where the specific requested topics are discussed, but rather extracts parts of texts from various websites and harmonizes them into a single text according to linguistic models set by the developers.

[6] I.e., hacking the integrated system management across various healthcare facilities connected to a single centralized cloud.

[7] See also R. Trezza, *La tutela della persona umana nell'era dell'intelligenza artificiale: rilievi critici*, in *federalismi.it,* 16/2022, pp. 276-305.

associated with the use of AI systems (AIs), ultimately diminishing the positive and advantageous prospects they can determine.
At the same time, AI can also enhance security and data protection through advanced systems (such as facial recognition and fingerprinting) and cryptographic algorithms, which streamline the detection and prevention of cyber threats. In practical terms, it can ensure secure access to personal data, verification of privacy compliance, and identification of potential breaches for protection against unauthorized access.
However, setting boundaries for AIs is crucial to mitigate highlighted risks and identify cybersecurity measures for safeguarding personal data from diverse cyber threats. *Cybersecurity*[8], being an inherent individual right, must be safeguarded alongside data protection, as the security of the cyberspace directly enhances data security.
Hence, securing the structure through which personal data are processed is imperative. Failure to protect this structure poses risks of political manipulation and commercial persuasion[9], emphasizing the need for robust safeguards.
This paper aims to shed light on the current landscape of AI and cybersecurity, highlighting both its opportunities and challenges. Despite potentially representing one of the new frontiers of cybercrime, with significant impact especially on the healthcare sector, AI also holds promise for enhancing the protection of personal data. This is attributed to the unique characteristic of AIs, which can be programmed based on external inputs, enabling them to be trained to detect criminal activities and strengthen cybersecurity measures.
It will also be highlighted that the proactive cybersecurity strategies driven by AI, essential for tackling emerging threats, often lack adequate regulatory support. Existing norms frequently fall short in anticipating the wide spectrum of risks associated with fast-paced technological advancements, rendering them outdated and insufficient.
Against this background, the article suggests a human-centred regulatory approach, thereby the individual is not just as a recipient of protection but an active participant. Through collaboration with institutions and guided by principles like proportionality, responsibility, and transparency, individuals should play a conscious and proactive role in this process of evolution and risk management[10].

[8] The topic of cybersecurity is now the subject of considerable investigations and studies; reference is made, among others, to R. Ursi (edited by), *La sicurezza nel Cyberspazio*, Milano, *Scritti di Diritto Pubblico*, 2023; but also to S. Rossa, *Cybersicurezza e pubblica amministrazione*, Napoli, 2023.

[9] M. Maurino, *Cybersecurity, sicurezza nazionale e trattamento dei dati personali*, in *amministrativ@mente*, 2/2023, pp. 939-972.

[10] A. Mantelero, *Artificial Intelligence, dati e diritti: spunti di riflessione per i regolatori*, in P. Bertoli, F. Ferrari, G. Ripamonti, G. Tiberi (a cura di). Data protection *tra Unione europea, Italia e Svizzera*, Torino, 2019.

2. AI in light of the evolution of the EU and national regulatory framework

The beginning of this evolutionary process is relatively recent[11]; indeed, it was in the 1970s and 1980s that concrete attempts to simulate human thought activity began, culminating in the developments of today. Currently, AI is fully capable of producing socially and environmentally beneficial outcomes, impacting significant sectors ranging from climate change to healthcare and the public sector, among others.

The term *Artificial Intelligence*[12] generally refers to the evolving capacity of machines to perform tasks that require a form of intelligence. However, before the definitive approval of the Regulation known as the *Artificial Intelligence Act*[13] in March 2024 (hereafter referred to as the 'AIA'), there was no legislative definition of it. The AIA defines AI as a system based on machines designed to operate with various levels of autonomy and capable of generating outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments based on the inputs it receives[14].

As highlighted in the regulatory text of the AIA[15], this notion should be clearly defined and closely aligned with the work of international organizations dealing with Artificial Intelligence, so as to provide the necessary flexibility to adapt to rapid technological developments in this field while ensuring legal certainty. Furthermore, it can be inferred from the text that this notion should be based on the key functional characteristics of AIs and is not intended to cover simpler traditional software systems or programming approaches, which rely on rules defined exclusively by individuals to automatically perform operations.

[11] See A. Simoncini, *Il linguaggio dell'intelligenza artificiale e la tutela costituzionale dei diritti,* in *Rivista AIC,* n. 2/2023, p. 17, which highlights that the roots of this innovative process actually date back to 1955 when, for the first time, the possibility of a form of Artificial Intelligence was envisioned, imagining that a machine, based on precise descriptions, could simulate any characteristic of human intelligence.

[12] The literature on AI is extensive. For example, recent writings include: M. Luciani, *La sfida dell'intelligenza artificiale,* in *Lettera AIC 12/2023 - Libertà di ricerca e intelligenza artificiale*; D. U. Galetta, *Digitalizzazione, Intelligenza artificiale e Pubbliche Amministrazioni: il nuovo Codice dei contratti pubblici e le sfide che ci attendono,* in *Federalismi.it*, 12/2023, pp. 4-14; S. Zorzetto, *La metafora della IA: una giungla lessicale e foresta simbolica,* in *Notizie di Politeia*, 151/2023, pp. 179-185; in the same journal, S. Salardi - M. Saporiti, *Risposte ai commenti e nuove riflessioni,* in *Notizie di Politeia,* 151/2023, pp. 186-189; A. Alaimo, *Il Regolamento sull'Intelligenza Artificiale: dalla proposta della Commissione al testo approvato dal Parlamento. Ha ancora senso il pensiero pessimistico?,* in *Federalismi.it*, 25/2023, pp. 132-149; F. Pizzetti, *Con AI Verso la Società digitale,* in *Federalismi.it,* 23/2023, pp. 4-9; D. Chiappini, *Intelligenza Artificiale e responsabilità civile: nuovi orizzonti di regolamentazione alla luce dell'Artificial Intelligence Act dell'Unione europea,* in *Rivista Italiana di Informatica e diritto,* 2/2022, pp. 95-108; M. Corti, *L'intelligenza artificiale nel decreto trasparenza e nella legge tedesca sull'ordinamento aziendale,* in *Federalismi.it*, 29/2023, pp. 162-170; L. Imberti, *Intelligenza artificiale e sindacato. Chi controlla i controllori artificiali?,* in *Federalismi.it*, 29/2023, pp. 191-201; D. Reiners et al., *The Combination of Artificial Intelligence and Extended Reality: A Systematic Review*, in *Frontiers in Virtual Reality*, 2, 2021.

[13] Proposal for a Regulation of the European Parliament and of the Council laying down harmonized rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts, 21 April 2021, COM(2021) 206.

[14] Art. 3 (1) AIA.

[15] Recital n. 6.

It is deemed appropriate, thus, for the European Commission to develop guidelines that standardize the use and application of Artificial Intelligence systems, which differ from traditional computer programs. Indeed, compared to these, they have the ability to learn to perform tasks better on their own, without being bound by rules and not simply limited to following instructions. They are algorithmic and generally rely on the synthesis and processing of inferences from large amounts of data, managing to provide information to be used in the human decision-making process. Therefore, AI derives from the ability of an algorithm to perform tasks and processes usually reserved for human intelligence, learning from the data it encounters each time and determining reactions that vary according to experiences and different external stimuli.

Based on this functioning, albeit described in a simplistic manner, attempts have been made since 2016 to formulate a clear definition of AI, while simultaneously paying attention to a regulation that has become increasingly urgent and necessary due to its growing presence in every aspect of human life[16].

In scholarly discourse, it has been highlighted that AI technologies have *"nothing intelligent," essentially involving the processing of pre-existing information to derive new and more useful insights*[17]. However, it cannot be denied that they represent a significant advancement in Data Analysis techniques, as the machine, based on its own experience, can learn through the selection of information it can derive from what it already possesses and has access to.

On one hand, therefore, its speed and, on the other, its self-learning capabilities contribute to AI bringing incredible benefits, such as accelerating the fight against diseases and mitigating the impact of disabilities in the healthcare sector or optimizing efficiency in agriculture. Correctly managed, it can swiftly facilitate the achievement of the *United Nations' Sustainable Development Goals* (SDGs) by 2030[18] and economic, social, and cultural rights worldwide, supporting improvements in various aspects of human life[19].

[16] Even smartphone virtual assistants, such as Siri or Google Assistant, are based on Artificial Intelligence; they use natural language processing technologies, Machine Learning, and other AI techniques to understand and respond to user requests naturally, using voice. Moreover, many platforms, including Amazon, Netflix, and Spotify, leverage AI extensively, particularly *Deep Learning*, to analyse user listening behaviour and provide personalized product recommendations that may be of interest to them. *Deep Learning* and *Machine Learning* are the modes through which AI operates; both aim to enable autonomous data evaluation, generating increasingly abstract conclusions. However, they differ in that while *Deep Learning* requires a large amount of data, it operates more automatically and requires much less human intervention compared to *Machine Learning*. The term "matryoshka" is used to refer to them; see A. E. Tozzi, F. Gesualdo, C. Rizzo, *Introduzione all'intelligenza artificiale in medicina per il personale sanitario,* cit. 9; ma anche Y. LeCun, Y. Bengio, G. Hinton, *Deep learning*, in *Nature*, 512, 2015, pp. 436-444.

[17] F. Pizzetti, *Con AI Verso la Società digitale*, in *Federalismi.it*, 23/2023, p. 6.

[18] R. Vinuesa, O. Azizpour, I. Leite, M. Balaam, V. Dignum, S. Domisch, A. Felländer, S. D. Langhans, M. Tegmark & F. Fuso Nerini, *The role of artificial intelligence in achieving the Sustainable Development Goals*, in *Nature Communications*, 233/2020, on https://doi.org/10.1038/s41467-019-14108-y.

[19] K. Jones, *AI governance and human rights. Resetting the relationship, Research Paper, International Law Programme*, Chatam House, January 2023, p. 11.

For this reason, the adoption of active measures has always been considered necessary, including non-legislative tools such as guidelines and codes of conduct, aimed at ensuring that its benefits are distributed equitably to avoid reinforcing and exacerbating social disparities[20], while simultaneously avoiding the risk that the use of its systems may result in harm to affected individuals or the community[21].

To prevent such an occurrence, AI must therefore be "reliable"; it must, in practice, possess three essential components – *legality*, *ethicality*, and *robustness* – which must exist throughout the entire lifecycle of the system and ensure compliance with all applicable laws and regulations, adherence to ethical principles and values, and the use of systems that do not cause harm[22]. To this end, precise guidelines have been developed from the outset aimed at all stakeholders involved in this field[23] and necessary to establish a horizontal foundation that ensures the reliability of AI.

The European legislator has intervened on this delicate issue with more or less annual frequency, aiming primarily to implement an anthropocentric, ethical, sustainable, and respectful approach to AI that upholds fundamental values and rights. This approach necessitates an appropriate European regulatory framework to avoid fragmentation of the internal market while establishing European networks and centres to enhance research, training, and innovation, without compromising trust and legal certainty objectives due to the risks associated with the unpredictability of AI pathways[24].

Indeed, with the *2020 White Paper on Artificial Intelligence*[25], strategic options were defined on how to achieve the dual objective of promoting AI adoption and addressing the risks associated with certain uses of this technology, which inevitably give rise to concerns. Among these concerns is the fear of being deprived of the means to defend one's rights

[20] K. Jones, *AI governance and human rights. Resetting the relationship, Research Paper*, *International Law Programme,* cit., p. 7.

[21] See D. Chiappini, *Intelligenza Artificiale e responsabilità civile: nuovi orizzonti di regolamentazione alla luce dell'Artificial Intelligence Act dell'Unione europea,* cit., p. 97.

[22] Those components were developed in 2018 by the "Independent High-Level Expert Group" appointed by the European Commission to draft the document entitled "The Ethical Guidelines for Trustworthy AI".

[23] From designers, to developers, to distributors, to implementers, and finally to users of AIs.

[24] See the Communication of April 8, 2019, COM(2019) 168, "Building Trust in Human-Centric Artificial Intelligence."

[25] White Paper on Artificial Intelligence - A European approach to excellence and trust COM(2020) 65 final. It is accompanied by the "Report on the implications of Artificial Intelligence, the Internet of Things, and robotics for security and liability," COM(2020) 64, of February 19, 2020. See also, *Sviluppi recenti in tema di Intelligenza Artificiale e diritto: una rassegna di legislazione, giurisprudenza e dottrina,* in *Rivista italiana di informatica e diritto*, *Osservatorio su Intelligenza Artificiale e diritto,* 2/2022, pp. 123-140, which is referred to for an in-depth reconstruction of the regulatory discipline, both at the European and national levels, regarding AI. But also see U. Salanitro, *Intelligenza artificiale e responsabilità: la strategia della Commissione europea*, in *Riv. dir. civile*, n. 6/2020, p. 1246 ff.; A. Fusaro, *Quale modello di responsabilità per la robotica avanzata? Riflessioni a margine del percorso europeo*, in *NGCC*, n.6/2020, p. 1344 ff.; P. Serrao D'Aquino, *Responsabilità civile per l'uso di sistemi di intelligenza nella Risoluzione del Parlamento europeo 20 ottobre 2020: "Raccomandazioni alla Commissione sul regime di responsabilità civile e intelligenza artificiale*," in DPER online, 1/2021, pp. 248-262.

and security in the face of the informational asymmetries of algorithmic decision-making processes[26].

The 2020 White Paper was followed in 2021 by the AI Act, which represents the first Proposal for a Regulation on Artificial Intelligence, presented by the European Parliament and Council, and will be discussed more extensively in the following paragraph. It is anticipated here that this Proposal establishes harmonized rules on AI, particularly regarding the development, placing on the EU market, and use of products and services that rely on it. Furthermore, it identifies categories of risk arising from the use of AIs, taking into account specific parameters and indications outlined in the 2020 White Paper.

To complement the framework outlined by this Proposal for a Regulation, in 2022, the European Commission published the Proposal for a Directive of the European Parliament and of the Council on adapting non-contractual civil liability rules to Artificial Intelligence[27]. The objective sought was to envisage a new discipline to be applied solely in civil judgments brought before national courts in cases of non-contractual liability. Such liability arises from the failure to observe due diligence by anyone (suppliers, developers, users), for the compensation of any type of damage foreseen by national law (life, health, property, privacy, etc.) and for any type of injured party (individuals, legal entities)[28].

In accordance with these European proposals, national legislation has also been established in Italy, consisting of guidelines and recommendations. Specifically, the Agency for Digital Italy (AgID) issued the *White Paper on AI in Service of the Citizen*; it was presented on March 21, 2018, and emerged from the consultation and in-depth analysis conducted by both public and private entities on how AI tools can impact increasingly useful and efficient public services. It has been described as the "first piece" within the debate concerning the sustainable and responsible use of AI in Public Administration, benefiting citizens; indeed, it provides guidance on optimizing the opportunities offered by AI while mitigating criticalities and problematic aspects to develop public services that are increasingly citizen-centric[29].

[26] See paragraph 5 of the White Paper.

[27] AI Liability Directive, September 28, 2022, COM(2022) 496, issued following the European Parliament Resolution of October 20, 2020, regarding damages caused by AIs of any kind (high or low risk). The discipline concerning AI liability serves a dual role: to ensure the right to compensation for the victim of harm and, simultaneously, to incentivize individuals and legal entities to avoid causing harm or prejudice from the outset; furthermore, it quantifies the compensation due for their behaviours. See P. Serrao D'Aquino, *Responsabilità civile per l'uso di sistemi di intelligenza nella Risoluzione del Parlamento europeo 20 ottobre 2020: "Raccomandazioni alla Commissione sul regime di responsabilità civile e intelligenza artificiale"*, cited above.

[28] The legal framework outlined in this proposal refers back to the definitional framework of the AIA, leaving it to the legal systems of the respective Member States to establish the notion of "fault" or "damage," while providing the definition of the "duty of care" (Article 2, No. 9).

[29] In particular, the White Paper addresses the challenges arising from the implementation of AI in Public Administration, adopting a multidisciplinary and systemic approach throughout. Additionally, it includes the "Italian Digital Strategy," developed within the framework of the European Digital Agenda, taking into account IoT, Big Data Analytics, AI, and Blockchain. Concerning these topics, the Three-Year Plan for Informatics in Public Administration was approved in 2017,

Aligned with the European Strategy, the Strategic Program for AI 2022-2024[30] has finally been approved, setting the conditions for its development in Italy, focusing on cooperation, data, and IT infrastructure, as well as researcher training, the importance of research investments, and the adoption of AI and its applications in Public Administration and productive sectors.

3. AI Act and Risk Categories

The AIA represents the first regulation on AI and confirms the goal pursued over the years by the European legislator to promote the adoption of anthropocentric and reliable AI; it must be used to support innovation while ensuring a high level of protection against the harmful effects that its systems can have on health, safety, fundamental rights, democracy, and the environment. Its final version is expected to come into force in May 2024 and be applied two years from that date[31].

This Regulation is part of the so-called *A Europe fit for the digital age strategy* outlined by the European Commission and defines the levels of risk associated with the impact of different AIs on people's lives and their rights, paying specific attention to generative AI models, such as OpenAI's *ChatGPT* and Google's *Gemini*. These are allowed to operate provided that their outputs are clearly labelled as generated by AI[32].

In the recently approved version, the Regulation provides specific rules[33] for general-purpose AI models and those posing systemic risks, which should also apply when such models are integrated or part of an AI system[34]. "Models" for general purposes differ from "systems," as a model should be defined based on its ability to competently perform a

containing operational guidelines aimed at guiding the country's digital transformation, becoming a reference for both central and local administrations in the development of their information systems, and establishing not only fundamental principles but also rules for usability and interoperability.

[30] Approved on November 24, 2021, and developed through collaboration between the Ministry of University and Research, the Ministry of Economic Development, and the Ministry for Technological Innovation and Digital Transition.

[31] The COM(2021) 206 final 2021/0106 (COD) of 21.4.2021, after being voted on by the European Parliament on June 14, 2023, became the subject of a political agreement between the Council and the Parliament in December of the same year. Its final version was approved by the European Parliament on March 13, 2024. The law will be published in the Official Journal of the European Union by May 2024 and will enter into force twenty days later, beginning to apply 24 months after its entry into force.

[32] *ChatGPT-4* is considered a high-impact AI system for which EU law requires *ex ante* application of rules on cybersecurity, transparency of training processes, and sharing of technical documentation before it enters the market. For all other foundational models, of lesser impact, the provisions contained in the AIA will apply when developers commercialize their products.

[33] Recital 60(a).

[34] See E. Millstone et al., *Science in Trade Disputes Related to Potential Risk: Comparative Case Studies*, Siviglia, European Commission, 2004; M. E. Gonçalves, *The risk-based approach under the new EU data protection regulation: a critical perspective*, in *Journal of Risk Research*, n. 23, fasc. 2, 2020, pp. 139-52.

wide range of distinct tasks that constitute its fundamental functional characteristic for general purposes.
Both general-purpose AIs and the models they are based on must comply with transparency requirements as well as EU copyright rules during the training phases of the various models. The most powerful among these and those that could pose systemic risks must also meet additional obligations, such as assessing and mitigating such risks and reporting on incidents.
Although not dealing with aspects related to civil liability, the AIA sets specific objectives that reflect on it, such as ensuring that AIs placed on the Union market and used are safe and comply with current legislation on fundamental rights. It provides that the achievement of these objectives must occur, as mentioned, through a horizontal regulatory approach to AI, balanced and proportionate, *"which is limited to the minimum requirements necessary to address the risks and problems associated with it*", without unduly hindering technological development. Furthermore, the cost of bringing AI solutions to the market should not be disproportionately increased[35], nor should unnecessary trade restrictions be imposed[36].
The aim is to create a discipline consisting of flexible mechanisms that can easily adapt to the evolution of technology and the emergence of new areas of concern. Moreover, the safety of AIs is regulated from a perspective *ex ante*, through a multilevel *risk-based approach*, whereby compliance obligations, more or less stringent, vary depending on the risk that software and AI may pose to fundamental rights.
To this end, four levels of risk are identified and defined, with precise allocations of responsibility among the various parties involved (from algorithm writing to user utilization). Each level corresponds to distinct categories, and each of these is subject to regulation necessary to ensure the safety, transparency, traceability, and non-discriminatory nature of the operations conducted. Given the somewhat general nature of the AIA (necessary precisely to allow flexible harmonization among Member States), risk assessments based on concrete application scenarios are not provided for; in some cases, they are hinted at, but without providing a general methodology for calculating risk, a circumstance that could undermine the effectiveness of the AIA[37].
The systems related to the four risk categories are as follows:

[35] "Proposal for a Regulation of the European Parliament and of the Council laying down harmonized rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts," p. 4.

[36] This clarification is in line with the legal basis of the proposal, primarily constituted by Article 114 of the Treaty on the Functioning of the European Union (TFEU), which provides for the adoption of measures aimed at ensuring the establishment and functioning of the internal market.

[37] See, in this regard, C. Novelli, *L'Artificial Intelligence Act Europeo: alcune questioni di implementazione,* in *federalismi.it,* 2/2024, pp. 94-113. The author highlights three weaknesses identified in the AIA, namely the pre-determination of risk levels, the judgment of risk significance during review, and the assessment of impact on fundamental rights (Fundamental Rights Impact Assessment - FRIA).

a) Unacceptable risk systems

Such systems pose a threat to individuals, their safety, livelihoods, and rights; therefore, they are prohibited. The prohibited practices include, under Article 5, cognitive behavioural manipulation of specific vulnerable individuals or groups, social scoring, and real-time and remote biometric identification systems, such as facial recognition through *non-targeted web scraping*. The latter aims to extract data from a website and then collect them in databases or local tables for analysis to infer race or political opinions, union membership, religious or philosophical beliefs, and sexual orientation[38].

The rule specifies that the use of *real-time* remote biometric identification systems in publicly accessible spaces is only permitted to the extent strictly necessary[39]: for example, in cases of targeted search for specific kidnapping victims or for the prevention of a specific, substantial, and imminent threat to the life or safety of individuals[40]. Furthermore, such use is permissible for the location or identification of a person suspected of committing a crime for the purpose of conducting a criminal investigation; in these cases, AI can be used without prejudice to provisions[41] for the processing of biometric data for purposes other than law enforcement[42].

However, these activities must be duly and pre-authorized by a judicial authority or an independent administrative authority, whose decision is binding on the Member State in which the use is to take place. Authorization is issued based on a reasoned request and in compliance with national law[43]; it can be waived in justified cases of urgency, provided that it is subsequently requested, without undue delay, within twenty-four hours at the latest. If authorization is refused, the use of the system must be immediately discontinued, with the deletion not only of all data but also of the results and outputs of the use itself. Since misuse of such systems could result in discrimination against the individuals involved, the legislator requires authorization only if there is objective evidence and clear indications that convince the authority responsible for granting it, strictly observing the principle of proportionality applied to achieving the objectives contained in the request to be authorized.

[38] Pursuant to letter b *bis)* of Article 5, this prohibition does not concern the labelling or filtering of sets of biometric data legally acquired, such as images, based on biometric data or the categorization of biometric data within the scope of law enforcement. Instead, the use of AIs to assess or classify natural persons or groups of them for a certain period based on their social behaviour or personal characteristics to avoid prejudicial or unfavourable treatment in social contexts unrelated to those in which the data were originally generated or collected is prohibited.

[39] Lett. d), par. 1, art. 5.

[40] Art. 5, lett. d), points (i) e (ii).

[41] Art. 9 GDPR.

[42] Art. 5, lett. d), point (iii).

[43] The aforementioned request must comply with the detailed provisions of national law. Member States may choose to introduce, in accordance with Union law, stricter laws on the use of remote biometric identification systems.

The regulation on the use of *real-time* remote biometric identification systems in publicly accessible spaces is therefore very precise. It provides for a necessary form of preventive protection to ensure the protection of individuals, avoiding uncontrolled use that could lead to serious violations with consequences that may also be difficult to predict. Indeed, it also envisages careful monitoring by national market surveillance and data protection authorities of the Member States. All uses must be notified to the latter so that they can prepare annual reports to be submitted to the European Commission, which will then publish them. Risk assessment tools based on profiling are also prohibited.

b) High-risk systems

These systems do not engage in prohibited behaviours but still pose a high risk; they negatively impact security, health, or fundamental rights and represent a "significant risk," understood as the result of the combination of its severity, intensity, likelihood of occurrence, duration of its effects, on the one hand, and its capacity to affect an individual, a plurality of persons, or a particular group of individuals, on the other[44].

Included in this category are uses related to critical infrastructures, including healthcare[45]; the Regulation provides that, in sensitive contexts like this, such systems must have an obligation to assess and mitigate risks, maintain usage logs, be transparent, accurate, and ensure human oversight. Citizens can file complaints and receive explanations about decisions based on such systems that affect their rights[46].

A system that profiles natural persons is always high-risk[47], as are systems used to dispatch first aid emergency services or to prioritize the dispatch of such services, as decisions are made in situations critical to the life and health of individuals and their property[48]. Conversely, systems that do not present a significant risk of harm to the health, safety, or fundamental rights of natural persons are not high-risk[49].

This category of systems is not prohibited, but the impact they can have, especially on sensitive sectors such as healthcare, is significant; therefore, the legislator requires compliance with very strict obligations, not only when the system is high-risk but also when the pro-

[44] The regulation is contained in Article 6.

[45] Education, vocational training, employment, basic public and private services, as well as certain systems related to migration control, and border management, justice, and democratic processes (such as systems used to influence elections) are also considered high-risk.

[46] The legislation considers high risk the system intended to be used as a security component of a product, as well as one that is itself a product, regardless of whether it is placed on the market or put into service.

[47] According to recital 34, systems intended to be used as safety components for road traffic management as well as for the supply of water, gas, heating, and electricity are also considered high-risk. This is because a failure or malfunction of these systems can endanger the lives and health of a large number of people, causing significant disruptions to normal social and economic activities.

[48] Recital 37.

[49] As occurs, for example, in cases where the system is intended to perform a limited procedural task or is aimed at enhancing the outcome of a human activity previously completed.

vider believes it is not. In such cases, the provider must ensure and certify its evaluation before introducing the system to the market or putting it into service, providing the user with clear, detailed, and adequate documentation containing the necessary information about the system and its purpose so that authorities can assess its compliance and track the results[50].

Specifically, regarding health data, it is envisaged that, to ensure better protection, these systems are developed through algorithm training carried out precisely on sets of health data, securely, timely, transparently, and reliably. To this end, the European space for such data must facilitate non-discriminatory access to them, ensuring that adequate institutional governance guarantees their privacy protection[51].

c) Limited-risk systems

Limited-risk systems, on the other hand, are those that interact with natural persons (e.g., *chatbots*) and create or manipulate sounds, images, and videos (such as *deepfakes*). There is no detailed regulation on the resulting risk; only specific transparency obligations are identified[52], according to which users of these systems must be informed that they are interacting with an AI system. This obligation does not apply to systems authorized by law to detect, prevent, investigate, and prosecute crimes, subject to appropriate guarantees for the rights and freedoms of third parties, unless these systems are also available for reporting a crime by the public. Both system providers, including those from GPAI[53], and users of a system that generates or manipulates image, audio, or video content that constitutes a *deep fake* are therefore subject to transparency obligations. System providers must ensure that the results of the systems are marked in a readable and detectable format as artificially generated or manipulated[54]; users must declare that the content has been artificially generated or manipulated, unless, once again, its use is authorized by law to detect, prevent, investigate, and prosecute crimes.

d) Minimal-risk systems

Minimal-risk systems include applications such as AI-enabled video games or anti-spam filters; for these, the AIA provides codes of conduct[55] aimed at promoting voluntary appli-

[50] Art. 6, § 2b.

[51] Recital 45.

[52] Regulated under Title IV.

[53] GPAI is the *Global Partnership on Artificial Intelligence*.

[54] This obligation does not apply when AIs serve an assistive function for standard editing or do not substantially alter the input data provided by the implementer, or if authorized by law to detect, prevent, investigate, and prosecute criminal offenses.

[55] Regulated under Title IX.

cation to AIs other than those with high-risk, taking into account industry technical solutions and best practices that enable the implementation of such requirements.
The rationale behind the classification and definition of risk in a context such as that of AI is to be found in the need to prevent its occurrence. Indeed, given its high complexity and especially the limited comprehensibility of the systems used, it is appropriate to intervene *ex ante*, through forms of preliminary risk analysis, involving both stakeholders (i.e., developers) and all interested parties. They are called upon to contribute also regarding the risk assessment related to the processing of personal data; the risk assessment is therefore not only a tool to prevent potential prejudices to the rights and freedoms of the data subjects but also plays a fundamental role *"in the dynamics focused on the trust of users that have always characterized the development of technologies"*[56].
The preventive risk assessment contributes to creating a horizontal framework for reliable AI, which, as mentioned, is one of the priority objectives pursued by the AIA proposal. It is significant, therefore, that in the recently approved version, specific provisions on the evaluation of the impact on fundamental rights[57] (*Fundamental Rights Impact Assessment - FRIA*) have been introduced for high-risk AIs for this purpose. This introduction should be positively interpreted, as it makes the preventive rationale pursued by the AIA particularly effective in terms of fundamental rights. Due to this rationale, it is necessary not only to be able to anticipate the evaluation of the detrimental effects of these systems but also to ensure that it is not limited to a mere assessment of compliance with technical requirements. In this regard, the aforementioned provision obliges both public authorities and all distributors of high-risk AIs to analyse, on the one hand, the intended use of the system and, on the other hand, its scope over time and space, describing the processes of the implementer in which the system will be used, in line with its intended purpose. It is also necessary to analyse the period of time and frequency in which each of these systems is intended to be used, in relation to the categories of natural persons and groups that may be affected by such use.
The provision also requires a description of human surveillance measures, in accordance with the instructions for use, and those to be adopted in the event of risk occurrence, including provisions on internal governance and complaint mechanisms. This assessment is carried out in the initial phase and involves stakeholders[58]; it allows distributors to develop plans that, based on the results of the assessment, can reduce or at least mitigate the negative impacts of systems on fundamental rights. If it is not possible to formulate an adequate plan, the distribution of such systems must be stopped after informing both suppliers and national authorities.

[56] A. Mantelero, *Artificial Intelligence, dati e diritti: spunti di riflessione per i regolatori,* cit. 32.

[57] Article 29a, entitled "Assessment of the Impact on Fundamental Rights for High-Risk Artificial Intelligence Systems."

[58] Consumer protection agencies and data protection authorities have six weeks to provide their input for the assessment.

The preventive assessment thus allows to address some shortcomings of the *static* risk model of the AIA: the impact on the values it deals with and on which AI can have an effect cannot be predetermined, and the consequent risk analysis must lead to balanced measures, which are neither too rigid nor too flexible, always keeping in mind that the priority objective is the protection of rights, to which AI must bring an added value in terms of advantage, not disadvantage. This is why it is envisaged that information be continuously updated and that the results of the impact assessment be notified, by the installer, to the market surveillance authority.
However, it must be noted that, at present, there is no single and clear method on the basis of which to develop plans aimed at mitigating the negative impacts on fundamental rights, and unfortunately, each distributor follows their own method, with a result - frankly not desirable - of poor uniformity and consequent loss of functionality of the plan itself compared to a correct and balanced risk analysis.

4. Critical issues concerning ChatGPT in the Europol Report and the intervention of the Italian Data Protection Authority

Regarding the preliminary risk assessment related to AIs, which inevitably impacts cybersecurity and data processing, it is necessary to briefly mention some critical issues concerning *Large Language Models* (LLMs), analysing the case involving the most widely used one, ChatGPT, developed by OpenAI LLC[59] (OpenAI). Indeed, over the past year, it has not only been the subject of a report by Europol[60], but also the recipient of a provision from the Italian Data Protection Authority (DPA) that temporarily restricted its use. Through these two acts, the negative impacts that such an LLM can have on cybersecurity, as well as its failure to comply with principles of responsibility and transparency, have been highlighted.
Indeed, the Europol report from March 2023, resulting from a specific investigation, revealed that ChatGPT, while allowing for the acceleration and enhancement of many "legitimate" workflows, such as research or content translation, can also be used for criminal purposes. It easily reduces barriers to entry into the cybercrime market. This means that even those without particular computer skills can carry out attacks based on *social engi-*

[59] OpenAI is the American company that developed, launched, and manages the ChatGPT AI platform.

[60] Europol is an agency based in The Hague that supports Member States in preventing and combating all forms of serious organized and international crime, cybercrime, and terrorism.

neering, enabling sophisticated and large-scale *phishing* campaigns, as well as activities carried out through *prompt engineering*[61].

What has alarmed researchers the most, however, is the discovery of a lack of transparency regarding the processing of personal data in this technology, caused not only by the fact that much of the datasets used to "train it" were not up to date[62], but also by the prevalent tendency to provide plausible, albeit not always correct, responses to users' questions.

Regarding the same lack of transparency, the Italian DPA notified OpenAI of a violation of data protection regulations. Therefore, in March 2023, it temporarily restricted the processing of such data[63] by OpenAI, the company that owns it, due to infringements of the GDPR. Among these, the lack of information for users and stakeholders explaining how OpenAI collected and processed data within the platform's operations; the absence of a suitable legal basis for both the collection of personal data and their processing for the purpose of training the algorithms underlying the operation of ChatGPT; and finally, the absence of any age verification for users of the service[64].

Due to these infringements, the DPA ordered OpenAI to implement a series of measures and prescriptions in accordance with Article 58, § 2, *sub d)*, of the GDPR[65]; moreover, it temporarily limited the ChatGPT services in Italy.

These services were subsequently restored following the implementation of the technical measures required by the DPA. In fact, the platform has not only articulated a procedure to allow users easier access and to exercise their right to object to data processing, but has also published detailed privacy information on the website regarding the data processed to train the algorithm; finally, it has developed specific mechanisms to allow users to correct or even delete any inaccuracies in the information processed.

The case leads to both a negative and a positive judgment; the negative one refers to the *modus operandi* of ChatGPT, as Open AI did not adopt, from the outset, a risk-based approach and did not respect the accountability and transparency principles set by the legislator, seriously endangering the processing of users' data. The positive one, on the other hand, relates to the intervention of the DPA, which has demonstrated great attention

[61] Prompt engineering is a relatively new type of engineering in the field of natural language processing. It makes it possible to bypass the security mechanisms provided by OpenAI through the "Do Anything Now" (DAN) command, inducing the system to respond to any subsequent input. Recently, the first "GPT" models created by cybercriminals have also been detected, such as *FraudGPT* (https://lnkd.in/dahapTmF) or *WormGPT* (https://lnkd.in/dHMJkDUp). Based on open-source technologies, these GPT models lack all the protections found in market products, although with an acceptable quality of results.

[62] They date back to September 2021.

[63] Decision of March 30, 2023 (web document no. 9870832).

[64] On this point, the DPA requested the mandatory adoption of a plan containing age verification tools, suitable for excluding access to the service by users under eighteen and minors in the absence of an express manifestation of will by those exercising parental responsibility over them.

[65] Decision of April 11, 2023, (web document no. 9874702).

in indicating the method to prevent such systems from compromising other fundamental rights of the individual.
Overall, the ChatGPT case demonstrates that the question of whether and how data can be used to train LLMs remains an open issue. Adequate strategies such as differential privacy[66] do not seem to provide sufficient assistance when applied to such language models, as they offer limited privacy protection[67]. Data is not merely information; it identifies us, characterizes us, and above all, belongs to us. It cannot be acquired without consent for purposes unknown to the data subjects. Risks further escalate when the data itself is used to train AI system algorithms. In such cases, privacy breaches can only be minimized if a generative AI model is trained to meticulously process requests containing personal data, allowing users to be aware of the processing activities conducted on their data and enabling them to autonomously exercise their rights. In practice, users must confidently rely on tested applications and technologies that are both ethically compliant with the regulatory framework and robust in terms of cybersecurity.

5. Synergy between AI and Cybersecurity

The highlighted need to act preventively regarding the risk assessment of AIs naturally also concerns the fight against cybercrime, a topic that unfortunately is still predominantly known only to "insiders", despite being now regulated both at the European[68] and national

[66] *Differential privacy* is a technique introduced in 2006 as an integral part of so-called PETs (Privacy Enhancing Technologies), capable of protecting personal data by masking individual information within a dataset. Regarding the issue of the relationship between AI usage and data collection, it has regained some emphasis following the recent dissemination by NIST (National Institute of Standards and Technology) of a draft guideline (NIST SP 800-226) on the application of differential privacy in AI, providing a detailed and rather technical analysis and highlighting the opportunities for practical implementation. The reason why differential privacy is potentially so valuable is that it offers a balanced solution between data access and analysis and the protection of the privacy of the individuals involved. Indeed, it does not rely on a single algorithm or mathematical method, but on multiple possible mathematical tools that are applicable and more reliable, in an era where attention to data analytics and the associated risks is high.

[67] The use of ChatGPT can have numerous negative or risky legal implications, sometimes not entirely clear even among professionals. These implications should be overcome by the new generation of AI (ChatGPT-5), especially regarding current models in reasoning abilities, data contextualization, and multimodality. The upcoming version of GPT, ChatGPT-5, should make AI safer by introducing new tools for information retrieval from the web and new features for source verification. Regarding privacy, a more granular control is expected, providing users with new tools to better adapt AI responses to the scenario in which it is commonly used.

[68] Reference is made, in particular, to the two NIS Directives: the Network and Information Security Directive 2016/1148/EU (*NIS Directive*) on the security of networks and information systems - transposed into the Italian legal system by Legislative Decree no. 65/2018 (*NIS Decree*) - followed by Legislative Decree no. 82/2021, converted into Law no. 109/2021, which established the National Cybersecurity Agency (*ACN*); and the so-called *NIS2*, EU Directive 2022/2555 of the European Parliament and of the Council of 14 December 2022, concerning measures for a high common level of cybersecurity in the Union, amending Regulation (EU) No. 910/2014 and Directive (EU) 2018/1972 and repealing Directive (EU) 2016/1148. This latter Directive must be transposed by the Member States by 17 October 2024. Also, the TFEU, at Article 83, deals with *cybercrime*, including it among the particularly serious and transnational criminal phenomena, on which the EU has criminal jurisdiction. However, from a regulatory point of view, the first act that addressed the fight

levels, and despite the establishment of an *ad hoc* agency with the specific objective of protecting national interests, security, and resilience in cyberspace[69].
Because cybersecurity must be achieved through all necessary activities «*to protect the network and information systems, the users of such systems, and other individuals affected by cyber threats*»[70], the participation of individuals/users in this context should be more aware and therefore more collaborative with the institutions so that the community as a whole can control its own digital destiny and thus achieve its *digital sovereignty*[71].
To this end, starting from the NIS Directive of 2016, followed by NIS2 in 2022, the EU has intensified the awareness of Member States towards cybersecurity against the commission of cybercrimes, through the adoption of measures aimed at protecting the security of EU networks and information systems. The goal is to ensure freedom of expression, protection of personal data and privacy, overcoming all forms of digital illiteracy, expanding the number of competent entities to address and manage cyber crises[72], and the number of subjects to be protected. Among these, the healthcare sector is particularly included.

against cybercrime was the *Budapest International Convention*, issued by the Council of Europe and ratified by Italy through Law no. 48 of 18 March 2008. It represents the only international treaty on cybercrime and aims to establish a common policy among the Member States and to combat cybercrime effectively. Regarding the so-called 'vulnerabilities' in computer systems, reference may be made to B. N. Romano, *Il rischio di "attacchi" ai sistemi informatici tra fattispecie penalmente rilevanti, tutela dei dati ed esigenze di "buona amministrazione",* in *amministrativ@mente, Rivista scientifica trimestrale di diritto amministrativo*, 3/2021, pp. 545-594; but also V. S. Z. Bonamini Pepoli, *Profili di contrasto al cybercrime* in iure condito *e* de iure condendo, in *Rivista italiana di informatica e diritto,* 2/2022, 109-121.

[69] This concerns the Italian National Cybersecurity Agency (ACN), established by Decree Law No. 82 of June 14, 2021, converted with amendments by Law No. 109 of August 4, 2021, which redefined the national cybersecurity architecture. The aim was to rationalize and simplify the existing system of national competencies to prevent and mitigate the highest number of attacks, promoting the achievement of technological autonomy and thereby safeguarding national security in cyberspace. On this point, reference is made to I. Forgione, *Il ruolo strategico dell'Agenzia Nazionale per la Cybersecurity nel contesto del Sistema di sicurezza nazionale: organizzazione e funzioni, tra regolazione europea e interna*, to R. Ursi, (a cura di), *La sicurezza nel* cyber*spazio*, Franco Angeli, Scritti di Diritto Pubblico, 2023, pp. 95-121; and also to G. G. Cusenza, *I poteri dell'Agenzia per la Cybersicurezza Nazionale: una nuova regolazione del mercato cibernetico*, in R. Ursi, (edited by), *La sicurezza nel* cyber*spazio*, Franco Angeli, Scritti di Diritto Pubblico, 2023, pp. 123-138.

[70] Article 2, paragraph 1), Regulation (EU) 2019/881, concerning ENISA, the European Union Agency for Cybersecurity, and the certification of cybersecurity for information and communication technologies. This regulation is situated between the two NIS Directives.

[71] *Digital sovereignty* refers to the way in which a state regulates and exercises governance over technology and services used in various ways within its national boundaries, thereby addressing the protection of sensitive data, enabling companies, organizations, and individuals to benefit from all opportunities related to the digitization of information, while at the same time maintaining control over where the data resides, flows, and who has control over it. See, among others, R. Baldoni, *Il* cyber-*spazio, un dominio fin troppo umano*, in *AirPress, Mensile sulle politiche per l'aerospazio e la difesa*, fasc. n. 127, 2021, pp. 4 ss.; G. Caggiano, *Sul trasferimento internazionale dei dati personali degli utenti del Mercato unico digitale all'indomani della sentenza Schrems II della Corte di giustizia*, in *Studi sull'integrazione europea*, 2020, pp. 563 ss., especially pp. 564-565.

[72] The NIS 2 Directive indeed envisages the establishment of competent national authorities, crisis management authorities for cyber incidents, "single points of contact," and Computer Security Incident Response Teams (CSIRTs). See, among others, M. Santarelli, *Verso la NIS 2, c'è l'accordo in Europa: ecco le novità su soggetti coinvolti e obiettivi,* in *CyberSecurity.it*, 2022; L. Tosoni, *Direttiva NIS, così è l'attuazione italiana (dopo il recepimento): i punti principali del decreto*, in *AgendaDigitale.eu*, 2021.

The AIA contemplates the cybersecurity profile regarding the impact that high-risk AIs have on it, requiring a design and development aimed at achieving, throughout their lifecycle and in light of their purpose, an adequate level of accuracy, robustness, and cybersecurity[73].

The Regulation also requires a set of requirements for such systems to be used without creating privacy violations or cyber-attacks; they must be *resilient*, meaning they must have the ability to withstand errors, failures, or inconsistencies that may occur within the system or the environment in which the system operates, particularly due to their interaction with natural persons or other systems. The technical and organizational measures to be adopted must therefore be able to eliminate/reduce the possibility that the outcomes achieved in high-risk AIs influence inputs for future operations ("*feedback loop*") due to their ability to continue learning even after market deployment or commissioning.

The cybersecurity requirement of the AIA applies to the AI system as a whole and not directly to its internal components; to ensure compliance with the cybersecurity requirements set out in the Regulation, it is therefore necessary to conduct a security risk assessment taking into account the system's design in order to identify risks and implement necessary mitigation measures. However, it is noted that the regulations are not yet exhaustive; therefore, such compliance requires an integrated and ongoing approach, using proven cybersecurity practices and procedures combined with specific controls for AI that have not yet been introduced. This gap is mainly due to the fact that AI cybersecurity is still an emerging field of study, gathering and combining knowledge and approaches from different fields, such as AI research, adversarial machine learning, and cybersecurity.

It follows that currently consolidated cybersecurity procedures are those used to protect traditional software-based (and hardware-based) systems; unfortunately, they are not able to address the wider range of cybersecurity risks of AIs, characterized by variables that are too specific and partly unknown.

The task of AI cybersecurity should then be to research and address system vulnerabilities[74] in order to develop risk management tools capable of responding to the initial request for AIA standardization, providing a solid foundation in terms of technical controls available to achieve and measure not only the general horizontal cybersecurity requirements but also those specific to AI.

Stakeholders aiming to make systems using emerging AI technologies compliant with the Regulation's cybersecurity requirements should be adequately supported in addressing some issues, namely, first, those of an organizational nature, related to security processes and controls, so that they can manage the AI lifecycle security by adapting existing controls for the respective software. Secondly, to research and develop techniques necessary

[73] Art. 15 AIA.

[74] For example, *adversarial machine learning attacks*, *data poisoning*, or embedded *backdoors* (known as "*porte di servizio*" in Italian, which allow remote access to a system) in AI models.

to address the impacts of AI on cybersecurity, such as assessing attacks on machine learning models, as well as developing specific AI security measures and *hardening* models to strengthen the most advanced methodologies.
European regulation is now oriented towards creating a connection between cybersecurity and the advancement of AI, which, by penetrating the world and market of cybersecurity, should push large companies to increase investments in the sector. This aims to counteract cyber warfare and develop new defence methods against cyber warfare attacks, with cutting-edge detection systems capable of identifying and blocking malicious activities even before they cause damage. Specifically, reference is made to the proposed regulation known as the *Cyber Solidarity Act* (CSA), recently approved, which aims to strengthen EU solidarity and capability in detecting cybersecurity threats and incidents, whether significant or large-scale. It seeks to achieve coordinated crisis management, enhancing response capabilities in each Member State and contributing to ensuring a secure digital landscape for citizens and businesses, to protect critical entities and essential services, such as healthcare[75].
The mechanism the proposed regulation aims to create should support preparedness actions, including conducting checks on entities operating therein, to detect potential vulnerabilities based on common risk scenarios and methodologies. Additionally, it should enable a review of significant or large-scale cybersecurity incidents that have occurred, drawing lessons and formulating recommendations to improve the EU's cyber deterrence position[76].
Pursuing the same goal of making AI the tool through which to ensure cybersecurity, the very recent Regulation 2023/2841[77] also aims to establish measures for a high common level of cybersecurity in EU institutions, bodies, and agencies. It requires each Member State to define a high degree of internal management, governance, and control framework for

[75] Recently (March 2024), an agreement has been reached on the establishment of a European Cybersecurity Shield ('European Cybersecurity Alert System'). This aims to create a pan-European infrastructure consisting of Security Operations Centres (SOCs), both national and cross-border, throughout the EU. Equipped with cutting-edge and highly secure tools, equipment, and infrastructure, these centres are intended to facilitate the exchange of threat intelligence data from various sources on a large scale and in a trusted environment.

[76] Through the three pillars on which it is based - namely the "European Cybersecurity Alert System", the "Cybersecurity Emergency Mechanism," and the "Cybersecurity Incident Review Mechanism" - the CSA, if approved within the expected timeframe, will effectively enable AI to play a significant role in countering cyber threats, potentially impacting significant cybersecurity incidents by allowing and ensuring data analysis for sharing with the CSIRT network. Within a year and a half, the discipline should be enriched with important components in addition to the AIA, including the Digital Operational Resilience Act (DORA), which establishes standards and requirements for managing and mitigating computer and security risks for the financial sector (such as risk management).

[77] Regulation (EU, Euratom) 2023/2841 of the European Parliament and of the Council of 13 December 2023, entered into force on 7 January 2024.

cyber risks, through mature risk management and reporting capabilities, using information sharing, thanks to inter-institutional organization, functioning, and operation[78].
Regarding the processing of personal data, this Regulation[79] provides that it should only occur to the extent necessary[80]; therefore, prohibited processing with respect to data such as those covered by Article 10 of Regulation (EU) 2018/1725[81] will only be allowed if necessary for reasons of significant public interest, in a manner proportionate to the pursued purpose. In cases considered necessary, data controllers[82] may allow it provided that the essence of the right to data protection is respected and appropriate and specific measures are envisaged to protect the fundamental rights of the data subject[83].
Finally, another proposed Regulation that has recently reached an agreement and is awaiting approval should be mentioned. This is the *Cyber Resilience Act* (CRA)[84], containing a series of rules aimed at increasing the security and resilience to cyber threats of all *products with digital elements* (PDEs), from smartphones to toys. Indeed, although the existing legislation in the internal market applies to some products with digital elements, most hardware and software products are currently not regulated by any EU legislation regarding their cybersecurity[85]. These products suffer, therefore, on one hand, from a low level of cybersecurity and, on the other hand, from a poor understanding of the information to which users also have limited access and therefore cannot choose products with adequate cybersecurity properties or use them safely. This means that a cybersecurity incident in a product within a connected environment can harm an entire organization or supply chain,

[78] In a logic of balance among the often conflicting needs of cybersecurity, technological innovation, and respect for citizens' rights, the Regulation provides, in Article 5, that by 8 September 2024, the Interinstitutional Cybersecurity Board (IICB) established pursuant to Article 10 of Regulation 2023/2841, after consulting with the ENISA Agency and receiving guidance from the CERT-EU (Computer Emergency Response Team of the EU, i.e., the EU's computer emergency response team for European institutions, bodies, and agencies), shall issue guidelines to Union entities for carrying out an initial cybersecurity review and establishing an internal framework for risk management, governance, and control. The IICB must also adopt consequential risk management measures as well as a cybersecurity plan.

[79] Art. 4.

[80] It must be carried out by the CERT-EU, the Interinstitutional Cybersecurity Board, and Union entities in accordance with Article 10 of Regulation (EU) 2018/1725. The latter is the Regulation on the protection of individuals with regard to the processing of personal data by the Union institutions, bodies, offices, and agencies and on the free movement of such data, repealing Regulation (EC) No 45/2001 and Decision No 1247/2002/EC.

[81] Namely, the processing of personal data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, or trade union membership, as well as genetic data, biometric data intended to uniquely identify a natural person, data concerning health or sex life, or sexual orientation of the individual.

[82] Specifically, the entities of the Union and the CERT-EU when acting in that capacity.

[83] In accordance with Article 10, paragraph 2, letter g), of this regulation.

[84] This is the COM(2022) 454 final 2022/0272 (COD) of 15.9.2022, proposing a Regulation of the European Parliament and of the Council on horizontal cybersecurity requirements for products with digital elements and amending Regulation (EU) 2019/1020. Negotiations related to it were concluded by the European institutions in December 2023.

[85] The Regulation proposal highlights, in particular, that the current legal framework of the EU does not address the issue of cybersecurity for non-embedded software, even though cybersecurity attacks increasingly target vulnerabilities in such products, causing significant social and economic costs.

quickly spreading to the internal market, disrupting economic and social activities and even becoming a lethal threat.
The underlying rationale of this Regulation is therefore to introduce, through a gradual approach to PDEs security, cybersecurity requirements that must be mandatory for designing, developing, producing, and distributing hardware and software products, to ensure standardization of norms among the various Member States.
But above all, this Regulation introduces a form of liability for manufacturers that must accompany the product throughout its life cycle, consisting of ensuring adequate support and tools to identify and address identified vulnerabilities, enabling prompt handling of emerging issues, conducting regular security tests, or, for products considered "important PDEs", undergoing mandatory conformity assessments[86].
The CRA, like the AIA, also does not explicitly address generative AI; both regulations should therefore provide, through additional implementing acts, technical safeguards proportionate to the attack vectors of a specific LLM, regardless of the risk levels governed by the AIA, bearing in mind that EU legislation lacks specific provisions for disinformation created by generative AI.

6. Data protection in the healthcare system: risks and measures to be implemented

At this point, it is necessary to dwell on the impacts of what has been highlighted so far in the sensitive healthcare sector and, above all, on how to defend against cyber-attacks aimed at it.
At European level, contributions to the digitalisation of the healthcare sector and the sharing of related data have been constant[87]: from the interoperability of information systems to the implementation of digital service infrastructure for *e-health*, to cross-border exchange of health data.
Progressive digitalisation has gradually led to extensive and advantageous use of AI; in fact, the latter, along with IoT, Big Data, Cloud Computing, and Machine Learning, has given rise to an ecosystem that is increasingly interconnected. It is called *Connected Care*,

[86] Art. 7.

[87] On the topic of digitalisation in healthcare and data management, see, among others, D. Morana, T. Balduzzi, F. Morganti, *La salute "intelligente":* eHealth*, consenso informato e principio di non-discriminazione,* in *federalismi.it,* 34/2022, pp.126-151; P. Melpignano, *L'intelligenza artificiale in sanità. Limiti, sfide e opportunità derivanti dall'utilizzo di sistemi che stanno rivoluzionando le modalità di diagnosi e cura dei pazienti*, in *Rassegna di diritto farmaceutico e della salute*, 3/2022, pp. 528-534; G. Lofaro, *Dati sanitari e* e-Health *europea: tra trattamento dei dati personali e decisione amministrativa algoritmica,* su *Astridonline,* 2-2023, pp. 179-208; C. Silvano, *La digitalizzazione dei servizi sanitari alla luce del riparto di competenze tra Stato e Regioni. Il caso del Fascicolo Sanitario Elettronico*, in *Federalismi.it,* 26/2023, pp. 227-249; F. C. Rampulla, G. C. Ricciardi, A. Venturi, *Digitalizzazione delle amministrazioni e accesso ai dati e ai documenti informatici sanitari,* in *Federalismi.it ,* n. 2/2024, pp. 132-175.

and within it, health information is shared with all parties involved in the care process (nurses, doctors, healthcare workers in hospitals and in the community), using diagnostic and medical devices directly at the patient's home.
AI is used in prevention, rehabilitation, but also in telemedicine, robotic surgery, and the development of new drugs. It is particularly advantageous in the analysis of diagnostic images because it allows the physician to save time in formulating the diagnosis by presenting the result of comparing huge amounts of images; but also in the administrative sector, as it enables the completion of documents and the creation of patient records with *synthetic* data (which mimic real data) for research purposes on patient samples.
The profiles we will focus on pertain, therefore, on one hand, to the use of generative AI like ChatGPT in the healthcare sector, and on the other hand, on how to defend healthcare data from the cyber threats they are inevitably exposed to.
Regarding the first profile, it must be highlighted that the use of ChatGPT constitutes an undeniable and valuable support for all administrative processes, benefiting from a reduction in operational times, with precise and accurate results, and consequently, better management of information flows. For example, when it comes to activities such as associating information for administrative purposes, compiling medical records, or prescribing medication, a doctor has the power of immediate review over the document generated by AI, and no healthcare service is provided nor are therapies or diagnoses suggested.
A different discourse, however, arises concerning decision support for diagnostics or research. AI uses data acquired not only from medical records and images but also from diagnostic devices and clinical or population studies. By operating on the data they come into contact with, processing them, and correlating them in a different and more comprehensive manner than humans would, algorithms are able to identify patterns and make decisions through machine learning systems and neural networks that mimic the functioning of the human brain[88]. Indeed, the system can generate new concepts and associations, finding connections between seemingly unrelated or numerous pieces of information that would otherwise be too vast to be analysed differently.
This clinical use, pertaining to intellectual performance, perhaps warrants a higher degree of caution regarding the responses provided by AI.
In the view of the author, indeed, there should be a greater emphasis on acquiring more appropriate skills tailored specifically to the user's profile, taking into account the adequate processing of data and aiming to avoid any potential distortions in system usage. Indeed, erroneous training of operational algorithms could have consequences that impact both individual health and the respect for fundamental rights.

[88] P. Melpignano, *L'intelligenza artificiale in sanità. Limiti, sfide e opportunità derivanti dall'utilizzo di sistemi che stanno rivoluzionando le modalità di diagnosi e cura dei pazienti*, cit.

These aspects inevitably lead to the second aspect to be addressed, namely the defence against cyber threats[89]. Fragmentation in the digitalisation process of healthcare, as well as the absence of a comprehensive security plan and the heterogeneity that unfortunately characterized the digitalisation of public administration in Italy, have a significant impact on this matter. These deficiencies, which not only affect cybersecurity but are even more dangerous in the healthcare sector, may lead to compromised data security and systems, potentially causing malpractice. Conversely, safeguarding personal data and systems is a crucial factor in healthcare efficiency.

In this regard, the new European legal framework provides important assurances, requiring transparency and contestability of the algorithmic process, specific precautions for the outsourcing of treatment, and, more broadly, an overall approach based on risk prevention. This includes the provision of precautionary measures and the adoption of a comprehensive strategy aimed at data protection and accountability of the parties involved in the processing.

However, unfortunately, there are still various gaps that need to be addressed.

Specifically, in the AIA, in addition to proposing a risk-based classification for determining the potential impact of a given AI system on health, safety, and fundamental rights, there is not adequate space dedicated to the healthcare sector or healthcare research. It is true that there is no clarification, for example, regarding the predefined intended use of AI-based devices that would allow identifying the risk class they fall into[90]; these can only be classified as medical devices under the MDR, i.e., the *Medical Device Regulation*[91].

Better protection might be expected, perhaps, from the CRA which, in accordance with the provisions of the NIS2 Directive, requires that measures and technical specifications similar to the essential cybersecurity requirements are also implemented for the design, development, and management of software vulnerabilities provided as a service for systems such as electronic health records, even if developed within healthcare institutions. This provision must also be in line with another proposed regulation, that of the *European Health Data Space* (EHDS) [92], which was recently approved.

[89] The collection, storage, and analysis of healthcare data are the responsibility of healthcare companies (Data Controllers), which are therefore required to adopt measures adequate to ensure their privacy and security in compliance with the regulations provided by the GDPR and the "Digital Administration Code" (CAD).

[90] Reference is made to sophisticated diagnostic systems and to robots useful for assisting and caring for individuals, as mentioned in recital 28.

[91] This is Regulation (EU) No 2017/745 concerning Medical Devices, which replaced Directive 93/42/EEC on medical devices (MDD), introducing new requirements and responsibilities for all Economic Operators.

[92] This is the Proposal "The European Health Data Space," COM(2022) 197/2, presented by the European Commission in 2022 to establish the European Health Data Space, aimed at supporting its use to improve healthcare delivery, research, innovation, and policy-making. It will allow and regulate the secondary use of health data, including those from electronic health records, registries, and medical devices, as well as data relating to individuals' lifestyles. The European Health Data Space is based on regulations such as the GDPR, Regulation (EU) 2017/745 on medical devices, Regulation (EU) 2017/746 on in vitro diagnostic medical devices, the proposed legislation on artificial intelligence, the proposed Data

Indeed, on March 15, 2024, an agreement was reached between the European Parliament and the Council of the European Union on the text of the proposal for the new Regulation on the EHDS. This Proposal aims to establish the European health data space, enabling individuals to access, share, and control their electronic health data. These data will be managed reliably and securely, safeguarding privacy and overcoming the inconsistencies in the implementation and interpretation of the GDPR by Member States, which create significant legal uncertainties and obstacles to the secondary use of electronic health data[93].
The creation of a common European health data space must therefore be based on respect for the principles of transparency and protection of patients' personal data, as well as the strengthening of *data portability*, under Article 20 of the GDPR. In accordance with the NIS2 Directive on cybersecurity, this space aims to enhance security and trust in the technical framework designed to facilitate the exchange of electronic health data for both primary and secondary use. Based on the CRA, more specific security provisions are envisaged in certain sectors; therefore, all those manufacturing products with digital elements classified as electronic health record systems, falling within the scope of the Regulation on the EHDS, are obliged to demonstrate compliance with its essential requirements.
Therefore, the latter could effectively overcome the current shortcomings that are still evident; for example, those related to medical devices using high-risk AIs[94].
Finally, a high level of cybersecurity is foreseen for data flow, considering the increasing risks of attacks on healthcare systems. In this regard, it is worth noting that, according to the recent *Clusit Report* of 2024[95], in Italy the healthcare sector ranked fourth in successful and publicly known cyber-attacks in 2023[96]. Moreover, there was a doubling in the number of cyber-attacks globally compared to the previous year[97], with a strongly increasing trend, demonstrating the healthcare sector's growing exposure to cyber threats. Nearly all incidents[98] had a cybercriminal motivation, while only 5 cyber-attacks were attributed to hacktivism.

Governance Act, the proposed Data Act, Directive (EU) 2016/1148 on the security of network and information systems (NIS Directive), and the Directive on cross-border healthcare.

[93] Already in its work programs for 2021 and 2022, the EU4Health program supports the development and establishment of the European health data space based on the existing infrastructure for primary uses of electronic health data (MyHealth@EU) and the secondary use of electronic health data (HealthData@EU).

[94] See Annex II, Section 2.

[95] The Report of the Italian Association for Cybersecurity *Clusit* is the document prepared by a panel of experts providing an overview of the most significant security incidents that have occurred globally (therefore also in Italy), referring to the four years preceding the reference year. The report for 2024 was presented on March 19, 2024, during the opening session of the Security Summit 2024. The data analysed pertains to the year 2023, compared with those collected in the previous four years.

[96] On this matter, please refer to the in-depth analysis "Cybersecurity in Healthcare: Between Increased Attacks and Regulatory and Technological Innovations," edited by di S. Montegiove, M. Santini, S. Scozzari, *Women for Security*, in the "Clusit Report" 2024, available at https://clusit.it/rapporto-clusit/ p. 151.

[97] In 2022 cyber-attacks were 304, in 2023, 624.

[98] 99%.

Compared to the past, these attacks have significantly increased as they are specifically targeted at exploiting the immense value of healthcare data on the *dark web*, where they are resold[99]. But what is most concerning is the impact they have unfortunately had on the affected healthcare facilities. The most commonly used techniques include *malware* (especially *ransomware*), *exploitation of vulnerabilities*, *compromised accounts*, while *phishing* or *social engineering* and unknown techniques – mainly *data breaches* – have slightly decreased.

The tripling of attacks recorded in Italy over the last four years demonstrates that it should evidently equip itself better, including through a punctual assessment of system vulnerabilities, which, after being attacked, require service restoration activities, unfortunately neither simple nor quick.

Since the use of technology, networks, and digital tools is now widespread, it is of paramount importance to ensure the security of the entire system through, first and foremost, the conscious use that each operator/user must make, thereby also having a basic knowledge of cybersecurity risks and, above all, countermeasures. These latter can indeed be easily invalidated by unaware users, through imprudent or erroneous behaviours, such as, for example, leaving the computer connected to a system requiring authentication, or leaving it without having logged out. Unfortunately, lack of preparedness is a problem that concerns not only healthcare facility employees but also top management of healthcare companies, who, despite having specific skills in many sectors, do not demonstrate the same level of expertise in cybersecurity, despite interacting with and using computer and digital tools on a daily basis[100].

It becomes evident, then, that the first step to take in cybersecurity is precisely that of *training*, which must lead to awareness in the use of digital technologies, to operate securely and not compromise the countermeasures implemented, especially in a sensitive sector such as healthcare, where a "simple" phishing trap can potentially compromise the health of many users, in terms of denying access to treatments, scheduled interventions, and already scheduled visits.

In compliance with the provisions of the NIS2 Directive, Member States are obliged to define a national cybersecurity strategy and to designate competent authorities, SPOCs (Single Point of Contact), and CSIRTs (Computer Security Incident Response Teams), as well as ENISA at the European level. This Directive, as mentioned, extends the scope to a greater number of sectors and types of companies compared to those included in the

[99] As reported by *Il Sole 24 Ore*, in May 2022, a person's medical record can cost up to 2000 dollars.

[100] Indeed, from a survey conducted during the first semester of 2023 by NetConsulting cube, it emerges that in 46% of cases there is a lack of a person entirely dedicated to cybersecurity, with percentages worsening in public health (52%). Where a responsible figure is present, the structure is either small or with competencies not entirely adequate. (cf. https://www.sanita24.ilsole24ore.com/art/aziende-e-regioni/2023-10-03/la-cybersecurity-come-presupposto-necessario-sviluppo-sanita-digitale-italia-101012.php?uuid=AF3w3K5).

previous NIS Directive, including those in the healthcare sector. Unfortunately, however, precisely in this sector, despite significant funding being provided (also in the PNRR[101]) for the strengthening of digital tools, infrastructures, and health records, it is observed that not as much has been invested in specific personnel training. The latter should contribute to the adoption of appropriate security policies and procedures to protect health data and prevent cyber-attacks, thereby saving negative consequences not only of an economic and organizational nature but also specifically related to people's health, patients, and citizens in general.

To protect data and prevent security breaches, healthcare companies should therefore adopt effective security measures, also leveraging AI. For example, by implementing a data security policy that establishes procedures and guidelines for their management, communicated to all company employees and regularly reviewed and updated to keep pace with new security threats. But also by encrypting sensitive data, making them incomprehensible to anyone without the correct decryption keys to protect them during transmission and storage. Finally, training measures on data security should not be underestimated, so that employees of healthcare companies should be instructed on how to protect patients' sensitive data and the risks related to data security through regular training programs, simulations of cyber-attacks, and phishing tests.

7. Concluding remarks

There are no definitive conclusions to be drawn on the highlighted themes, but only some assessments that take into account the state of the art.

It is evident that AI will enhance the capabilities of human healthcare professionals. However, in preparing to better use it in the healthcare sector, ethical access to data must be ensured, involving not only data scientists and engineers but also physicians, patient advocates, economists, and policymakers. In the nascent phase of AI applied to healthcare, cooperation is of paramount importance to identify best practices.

Unfortunately, the legal framework sometimes appears "unprepared" to address the significant changes imposed by the spread of digital technologies, an almost inevitable consequence of the speed at which technology advances and transforms.

However, given its vastness, the data market, a potential source of immense and uncontrollable power in the hands of those who possess it (whether public or private entities), must necessarily be subject to a regulatory framework that ensures robust protection while

[101] "PNRR" is the Italian acronym for "Piano Nazionale di Ripresa e Resilienza," which translates to "National Recovery and Resilience Plan" in English.

maintaining the interconnection between cybersecurity and AI in the necessary perspective of data protection.

The past year, 2023, unfortunately saw the need to strengthen defences in cybersecurity due not only to the evolution of digital technologies but also to the increase in attacks, facilitated by that same technological evolution intended to improve individuals' lives. This is evidenced by the increase and refinement of techniques based on social engineering obtained by cybercriminals through generative AI, which outpace cybersecurity measures in speed.

The goal, therefore, should be to combat cybercrime with the same tools used to perpetrate it, namely AI and machine learning. Thus, adequate prevention strategies and solutions are necessary to enable these tools to facilitate the detection and protection of sensitive data, as well as to understand the user's context, identifying critical risks.

Preventing data loss by helping to identify sensitive data (such as intellectual property and trade secrets) and automatically classifying them, scanning, labelling, and protecting them wherever they are, is necessary.

The ability of AI to process vast amounts of data in real-time allows for extensive visibility into heterogeneous and distributed environments, enabling the timely identification of anomalous behaviours. This enables protection against potential cyberattacks and threats, as well as improving compliance with privacy and data security regulations.

Generative AI, integrated with purpose-built tools, can play a crucial role in cybersecurity by promptly detecting and identifying types of data that should not be exposed or shared. Through recognition algorithms, pattern analysis, and advanced analytics, it can help identify anomalies and security breaches, anticipating potential risks before they cause irreparable damage.

In other words, although the use of generative AI tools carries intrinsic risks related to the manipulation of sensitive data, its potential to identify and prevent security breaches is fundamental in modern sensitive data management. This is alongside the implementation of an increasingly strategic synergy between vertical sector specialists in Healthcare with cybersecurity analysts and professionals, from the early stages of designing new healthcare systems.

The goal, therefore, should not be to limit its evolutionary drive but to balance the maximization of the benefits derived from the use of this technology with rigorous data governance, reducing risk factors through an organic approach capable of embracing innovation while ensuring privacy and security.