CYLCOMED

Cyber-security toolbox
for connected medical devices

D2.3 Ethical and Legal Evaluation and Recommendations Report

Revision: v.1.0

Work package	WP 2
Task	Task 2.2
Due date	30/11/2025
Submission date	30/11/2025
Deliverable lead	KUL
Version	1.0
Author	Dusko Milojevic (KUL)
Coauthor	Maja Nisevic (KUL)
Project PI	Jan De Bruyne (KUL)
Reviewer	Dietmar Frey (CHARITE)
Abstract	T2.3 aims to evaluate whether and to what extent the ethical and legal requirements identified in T2.1 and T2.2 have been adequately implemented by the consortium throughout the project. This evaluation will be used to highlight shortcomings in the current regulatory framework and to provide policymakers and lawmakers with options to consider improvements to the existing rules and policies. Special attention will be

	paid to the MDCG 2019-16 Guidance on Cybersecurity for Medical Devices, and in case of the need for revision, the evaluation will provide recommendations with special attention to the steps of revision.
Keywords	Medical devices, Data and privacy protection, Cybersecurity, Artificial intelligence, Ethics

DOCUMENT REVISION HISTORY

Version	Date	Description of change	List of contributor(s)
V0.1	10/11/2025	First draft	Dusko Milojevic (KUL)
V0.2	26/09/2025	Input and comments	Juan Carlos Perez Baun (EVIDEN)
V0.3	24/10/2025	Input and comments	Esteban Armas (ATOS)
V0.4	06/11/2025	Input and comments	Hrvoje Ratkajec (XLAB)
V.0.5	10/11/2025	Implementation of Comments	Dusko Milojevic (KUL)
V0.6	12/11/2025	Comments on the First Draft	Maja Nisevic (KUL)
V0.7	13/11/2025	Implementation of Comments	Dusko Milojevic (KUL)
V0.8	17/11/2025	Input and comments	Juan Carlos Perez Baun (EVIDEN)
V0.9	20/11/2025	Implementation of Comments	Dusko Milojevic (KUL)
V1.0	24/11/2025	Internal Review	Dietmar Frey (CHARITE)

Disclaimer

Copyright notice

Project co-funded by the European Commission in the Horizon Europe Programme	
Nature of the deliverable:	ETHICS, SECURITY

Dissemination Level		
PU	*Public, fully open, e.g. web*	**x**
SEN	*Sensitive, limited under the conditions of the Grant Agreement*	
Classified R-UE/ EU-R	*EU RESTRICTED under the Commission Decision No2015/ 444*	
Classified C-UE/ EU-C	*EU CONFIDENTIAL under the Commission Decision No2015/ 444*	
Classified S-UE/ EU-S	*EU SECRET under the Commission Decision No2015/ 444*	

* *R: Document, report (excluding the periodic and final reports)*
DEM: Demonstrator, pilot, prototype, plan designs
DEC: Websites, patents filing, press & media actions, videos, etc.
DATA: Data sets, microdata, etc
DMP: Data management plan
ETHICS: Deliverables related to ethics issues.
SECURITY: Deliverables related to security issues
OTHER: Software, technical diagram, algorithms, models, etc.

Executive summary

To ensure sustainable use of novel technologies, it is crucial to identify potential legal and ethical issues and mitigate them as early as possible in the development stage. Having this in mind, the project has adopted a privacy- and data-protection-by-design approach, as well as an ethics-by-design approach. Against this background, the goal of the deliverable D2.3, which derives from task T2.3, is twofold.

Firstly, it aims to provide an assessment of the legal and ethical considerations within the project. It demonstrates how the CYLCOMED project has integrated the identified legal and ethical requirements. In other words, the deliverable D2.3 presents the evaluation of the implementation of the identified legal and ethical requirements in the Deliverable D2.1, “Analysis of Ethical, Legal and Data Protection Frameworks”, and Deliverable D2.2, “Examination of Ethical, Legal and Data Protection Requirements”. As such, it should be read in conjunction with the deliverables mentioned above.

Secondly, this evaluation will highlight shortcomings in the current regulatory framework and provide policymakers and lawmakers with recommendations to improve the existing rules and policies. In doing so, special attention will be paid to the MDCG 2019-16 Guidance on Cybersecurity for Medical Devices.

Contents

List of figures

List of tables

Abbreviations

Abbreviation	Definition
AI	Artificial Intelligence
ALTAI	The Assessment List for Trustworthy Artificial Intelligence
CAR	Cyber Resilience Act
CER	Critical Entities Resilience Directive
CTIS	Clinical Trial Information System
CSIRTs	Computer Security Incident Response Teams
CMD	Connected Medical Device
CJEU	Court of Justice of the European Union
CoE	Council of Europe
CSA	Cybersecurity Act
CTD	Clinical Trials Directive
CRT	Clinical Trial Regulation
DPIA	Data Protection Impact Assessment
DPO	Data Protection Officer
DoH	Declaration of Helsinki
EEA	European Economic Area
EC	European Commission
ECHR	European Convention on Human Rights
ECCRI	European Code of Conduct for Research Integrity
ECtHR	European Court of Human Rights
EHDS	European Health Data Space
EDPB	European Data Protection Board
EEA	European Economic Area
EHDS	The European Health Data Space
EHR	Electronic Health Record
EMA	European Medicines Agency
ENISA	European Union Agency for Cybersecurity
EU	European Union
GCP	Good Clinical Practice
GDPR	General Data Protection Regulation
HE	Horizon Europe
HIS	Health Information System
HLEG AI	High-Level Expert Group on Artificial Intelligence
ICO	UK's Information Commissioner's Office
IMP	Investigational medicinal products
IVDR	In Vitro Diagnostic Medical Devices Regulation
IRB	Institutional Review Board
MDD	Medical Device Directive
MDCG	Medical Device Coordination Group
MDR	Medical Devices Regulation
NIS2	Network and Information Security Directive
RED	Radio Equipment Directive

REC	Research Ethical Committee
SaMD	Software as Medical Device
SW	Software
TFEU	Treaty on the Functioning of the European Union
UDHR	Universal Declaration of Human Rights
WP	Work Package
WP29	Working Party 29

Funded by Horizon Europe Framework Programme of the European Union

1 Introduction

This deliverable serves a dual purpose. First, it provides a comprehensive evaluation of the extent to which the CYLCOMED project activities have adhered to the ethical and legal requirements previously identified in Work Package 2 (WP2), specifically in Tasks T2.1 “Identification of Ethical and Legal Frameworks” [1] and T2.2. “In-depth Ethical and Legal Study” [2]. These requirements were derived from an in-depth analysis of existing regulatory and ethical frameworks relevant to emerging digital health technologies. Second, building upon the foundational research and outputs of T2.1 and T2.2, this deliverable seeks to identify potential gaps, ambiguities, and areas for improvement within the current legal and ethical frameworks governing digital medical technologies. In doing so, it aims to formulate concrete, evidence-based recommendations for EU policymakers and regulatory authorities

In other words, while the primary objective is to assess the implementation of ethical and legal principles in practice, the broader ambition is to leverage these insights to inform the future development or revision of applicable rules and policies. The evaluation thus not only provides a performance check on the consortium’s compliance efforts but also identifies structural shortcomings in existing governance instruments. Moreover, particular attention is devoted to the MDCG 2019-16 Guidance on Cybersecurity for Medical Devices [3], a key regulatory reference point within the EU framework. Where limitations, inconsistencies, or outdated provisions are identified, the deliverable will offer targeted recommendations for revising the guidance, including practical considerations and procedural steps for future updates.

In summary, the objectives of this deliverable are twofold:

- To evaluate the CYLCOMED technology and project activities against the ethical and legal requirements articulated in WP2.
- To consolidate the findings from T2.1 and T2.2 into a cohesive and actionable set of legal and ethical recommendations for law- and policymakers, with a view toward strengthening the regulatory environment.

1.1 Purpose and Scope

This deliverable builds directly on the work conducted in WP2 and aims to:

- Evaluate the implementation of ethical and legal safeguards in the CYLCOMED project, using the requirements identified in Tasks T2.1 and T2.2 as benchmarks.
- Identify structural, procedural, or conceptual shortcomings in the current legal and ethical framework that may hinder effective deployment of CYLCOMED-like technologies in real-world healthcare settings.

Funded by Horizon Europe Framework Programme of the European Union

- Provide forward-looking policy recommendations to lawmakers, regulatory authorities (e.g., the European Commission, MDCG, EMA), with the aim of supporting future regulatory updates and ensuring stronger legal-ethical alignment in the context of digital medical technologies.

The findings of this deliverable contribute directly to the project's broader objectives of promoting trustworthy, human-centric, and compliant AI in medical applications, and offer a roadmap for aligning innovation with fundamental rights and patient safety across the European Research Area.

1.2 Structure

This document's legal and ethical evaluation and policy recommendations are presented in three main chapters. **Chapter 2** performs legal and ethical analysis of the CYLCOMED toolbox. **Chapter 3** examines the existing cybersecurity framework governing connected medical devices, with particular emphasis on the MDCG 2019-16 Guidance on Cybersecurity for Medical Devices. The chapter also formulates recommendations to support the future development or revision of this guidance in light of emerging digital health technologies. **Chapter 4** consolidates the findings from the legal and ethical analysis and offers broader policy recommendations concerning the practical implementation of relevant ethical guidelines and legal instruments, including the Declaration of Helsinki, CIOMS Guidelines, ICH-GCP, and the Clinical Trials Regulation. **Chapter 5** concludes the deliverable.

Funded by Horizon Europe Framework Programme of the European Union

2 Ethical and Legal Evaluation

This chapter provides an overview of whether and to what extent the ethical and legal requirements identified in D2.1 “Analysis of Ethical, Legal and Data Protection Frameworks” and D2.2 “Examination of Ethical, Legal and Data Protection Requirements” have been adequately implemented by the consortium throughout the project. Firstly, it sets the scene with a brief description of the toolbox and its architecture. Secondly, it elaborates on the methodology used for the legal and ethical evaluation. Next, it provides a brief overview of the Ethics Guidelines and the corresponding Assessment List for Trustworthy Artificial Intelligence (ALTAI), published by the European Commission’s (EC) High-Level Expert Group on AI (HLEG). Afterward, it provides an assessment of the legal and ethical considerations within the project. It demonstrates how the CYCOMED project has integrated the identified requirements, with references to relevant technical and organizational measures. Lastly, this section identifies the limitations and weaknesses of the ALTAI, followed by a summary of the findings in a set of conclusions and recommendations.

2.1 Setting the Scene: The CYLCOMED Toolbox Overview

In the current cybersecurity landscape, healthcare is one of the most targeted sectors concerning cybersecurity incidents. Developing robust, anticipatory defence strategies is critical as cyberattacks become increasingly sophisticated and pervasive. The CYLCOMED project focuses on enhancing safety, integrity, and data protection in CMDs by developing a modular toolbox that supports end-to-end risk mitigation [4]. As illustrated in Figure 1, the toolbox adopts a layered architecture encompassing real-time monitoring, device integrity assurance, secure access management, and regulatory alignment mechanisms. The toolbox uses a layered design to provide specific protections. It supports integration between components and focuses on practical use in CMD security [5].

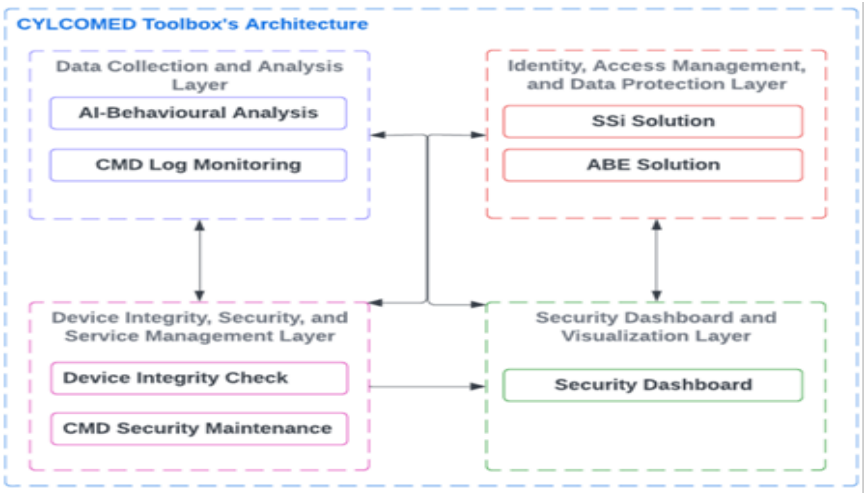

Figure 1 Cylcomed Toolbox Design

Funded by Horizon Europe Framework Programme of the European Union

The CYLCOMED toolbox encompasses various tools designed to strengthen the cybersecurity of Connected Medical Devices (CMDs). For instance, the **Data Collection and Analysis Layer** collects, monitors, and analyses data from diverse sources within the CYLCOMED ecosystem to detect potential cybersecurity threats or anomalies proactively. It integrates two core components: the CMD Log Monitoring system (**LOMOS**) and the behavioural analysis tool (**LADS** - Live Anomaly Detection System). These tools leverage artificial intelligence to examine logs and network traffic generated by connected medical devices and the platforms managing them. The aim is to identify unusual patterns or activities that could indicate a potential security incident. The primary function of this layer is to provide real-time, continuous surveillance of system activities, ensuring the prompt detection and response to emerging threats. It aggregates and processes large volumes of data from various sources, including network traffic, device logs, and other relevant streams, applying sophisticated algorithms to detect deviations from normal behaviour [6].

Next, **Device Integrity, Security and Service Management Layer's** primary objective is to uphold the operational integrity of medical devices, safeguarding them against unauthorised modifications and ensuring their lifecycle management remains secure. This layer includes mechanisms facilitating secure firmware updates and configurations aligned with healthcare standards and regulations. These protections not only prevent tampering but also ensure that updates and configurations meet strict regulatory requirements, thereby enhancing device reliability and patient safety throughout the device's lifecycle [7].

The Identity, Access Management, and Data Protection Layer establishes strong access control and data protection mechanisms tailored to healthcare's unique requirements. This layer enforces encryption and secure data-sharing protocols, acting as the system's gatekeeper by managing access through predefined policies and user roles to uphold data privacy and comply with legal and ethical standards. **The Security Dashboard and Visualisation Layer** is the topmost layer of the CYLCOMED architecture. The Security Dashboard provides a comprehensive view of the system's cybersecurity state by centralising the reporting and visualisation of security alerts, anomalies, and metrics. Serving as a centralised Security Information and Event Management (SIEM) tool, its primary function is to aggregate, analyse, and display security data from across the CYLCOMED framework, delivering real-time situational awareness for cybersecurity events and enabling timely and effective incident response. This layer is integral to sustaining the integrity and security of the CYLCOMED ecosystem, providing the analytical depth required for a resilient cybersecurity posture [8].

This deliverable will not dive deeper into toolbox functionalities and architecture, and readers are directed to the WP5 and deliverable D5.3 "CYLCOMED Final Toolbox Package", which concludes WP5 "Cybersecurity Toolbox Design and Implementation" by presenting the final architecture, packaging, and integration status of the

Funded by Horizon Europe Framework Programme of the European Union

CYLCOMED Toolbox. This document is the final iteration of the CYLCOMED Cybersecurity Toolbox, developed under WP5.

2.2 Methodology

As noted above, this chapter presents an evaluation of the legal and ethical requirements identified in Deliverable D2.2, submitted in Month 18 of the project. The analysis assesses the extent to which these requirements have been addressed in the design and development of the CYLCOMED toolbox.

Given that the CYLCOMED toolbox includes components that integrate artificial intelligence (AI) technologies, the evaluation incorporates relevant considerations derived from European Commission guidance on AI ethics. However, it is important to emphasise that this assessment is not limited to AI-based elements. The toolbox consists of multiple tools with diverse technical functions, architectures, and use cases, and the analysis reflects this diversity. Rather than evaluating each individual tool exhaustively, this chapter provides a holistic, cross-cutting assessment of the toolbox. Specific examples are used to illustrate how ethical and legal requirements, particularly those related to informed consent, data protection, fairness, transparency, and human oversight, have been addressed in practice.

The evaluation has been informed by contributions from consortium partners. Partner input was collected through regular coordination meetings, email exchanges, review of project deliverables, and technical presentations. This collaborative approach ensures that the evaluation reflects both the technical implementation and the underlying legal compliance and ethical design processes adopted throughout the project.

2.3 HLEG Ethics Guidelines & Assessment List for Trustworthy AI (ALTAI)

In 2018, the EC presented its AI for Europe plan [9]. This plan pursued three main objectives, namely: boosting the EU's technological and industrial capacity and AI uptake across the economy, preparing for socio-economic changes driven by AI and ensuring an appropriate ethical and legal framework [10]. In order to achieve these objectives, the EC created the ‘high-level expert group on AI’ (HLEG), an independent group made up of experts in the fields of computer science, law, and ethics. HLEG published its Ethics Guidelines for Trustworthy AI, which articulate **four ethical principles** for ‘Trustworthy AI, rooted in fundamental rights such as Respect for human dignity, “Respect for democracy, justice and the rule of law” and non-discrimination and solidarity. An AI system is considered ethical and robust if it adheres to four principles [11]:

Ethical Principles

Funded by Horizon Europe Framework Programme of the European Union

Respect for human autonomy	The principle of respect for human autonomy means that AI systems must not be developed in a manner that can unjustifiably subordinate, coerce, deceive, or manipulate humans. The human-centric design principle must be implemented, which will allow effective self-determination. In a manufacturing context, this means securing human oversight over work processes in AI systems. AI systems should support humans in the working environment instead of replacing them and aim for the creation of meaningful work [12].
Prevention of harm	The principle of prevention of harm requires that AI systems be secure and safe. Its design must be technically robust and secured from malicious use that might have an adverse effect on humans. In a manufacturing context, particular attention must be paid to situations where AI systems can cause or exacerbate adverse impacts due to asymmetries of power or information, such as between employers and employees [13].
Fairness	The principle of fairness requires that the development, deployment, and use of AI systems must be fair and that AI systems should not generate unfairly biased outputs.
Explicability	The principle of explicability is marked as essential in building and maintaining trust in AI systems. It is imperative that transparency is maintained in processes and that the capabilities and objectives of AI systems are clearly communicated. Whenever possible, decisions should be explained to those who are directly or indirectly impacted.

While the above principles offer some guidance, they remain rather abstract. To facilitate their practical implementation, the Ethics Guidelines translate the four ethical principles into **seven ethical requirements** that must be met to achieve 'trustworthy' AI. These requirements apply to all AI stakeholders, i.e., developers, deployers, users, as well as society in general. According to the Ethics Guidelines, developers are those who research, design, and develop AI systems. They must implement the seven ethical requirements during the development and design process. Additionally, the Ethics Guidelines define deployers as public or private organizations that use AI systems within their business processes and/or offer products and services to others. They must ensure that the systems they deploy meet the seven requirements. End-users, according to the Ethics Guidelines, are the ones who directly or indirectly engage with AI systems. They should be informed about the existence and content of the seven requirements and be able to request that they be upheld. Finally, society refers to everyone who is affected by AI systems, directly or indirectly [14]. **Seven requirements** that AI systems should implement and meet throughout their entire life cycle:

1. Human Agency and Oversight;
2. Technical Robustness and Safety;
3. Privacy and Data Governance;

Funded by Horizon Europe Framework Programme of the European Union

4. Transparency;
5. Diversity, Non-discrimination and Fairness;
6. Societal and Environmental Well-being;
7. Accountability.

Finally, the Ethics Guidelines set out a concrete and non-exhaustive Trustworthy AI assessment list to operationalize these seven requirements. This assessment list for trustworthy AI systems (ALTAI) is intended to provide AI practitioners with practical guidance for developing trustworthy AI.

Funded by Horizon Europe Framework Programme of the European Union

2.4 Ensuring Cybersecurity through CYLCOMED Toolbox: Legal and Ethical Assessment

It is important to note that the design, development, and implementation of the CYLCOMED Cybersecurity Toolbox is based on the state-of-the-art cybersecurity baseline study conducted under WP3. A preliminary set of technical, functional, and non-functional requirements had been compiled under D3.1 [15], as a first step contributing to the overall objective of developing the cybersecurity Toolbox (also in connection with legal and ethical requirements traced in WP2), so that it is user-friendly, trustworthy, reliable, and scalable. Additionally, D3.2 [16] defines functional, non-functional, and end-user requirements for the CYLCOMED toolbox that are elicited in relation to the legal and ethical requirements defined in WP2, so that the design technical solution considers different aspects regarding: needed functionalities, compliance with the current regulations, performance, user-friendly, and usability.

Building on the previous subsection, this subsection aims to evaluate the CYLCOMED cybersecurity toolbox against the seven ethical requirements provided by the HLEG. It should be noted, however, that the assessment provided in this chapter is not limited to the components involving artificial intelligence. In fact, the CYLCOMED toolbox has incorporated different tools with various technical features and procedures. As mentioned above, this deliverable aims to provide a general assessment of the toolbox, with reference to certain aspects and examples of different tools developed during the project. It does not aim to provide a detailed account of how requirements were incorporated by each tool. The evaluation of the CYLCOMED toolbox is presented in the table below, using the Ethics Guidelines for Trustworthy AI as the main reference framework. The introductory paragraph briefly outlines each requirement, while the table itself explains how the CYLCOMED toolbox meets these requirements.

The **left column** lists the requirements, grouped into the **seven thematic areas** defined by the HLEG AI, and the **right column** provides an assessment of whether and how each requirement has been addressed by the toolbox.

2.4.1 Human Agency and Oversight

This requirement entails that AI systems should support human agency and human decision-making [17]. It can be broken down into two separate dimensions: (1) ***human agency and autonomy***; and (2) ***human oversight***. The former deals with the effect AI systems can have on human behaviour in the broadest sense. It deals with the effect of AI systems that are aimed at guiding, influencing, or supporting humans in decision-making processes. It also deals with the effect on human perception and expectation when confronted with AI systems that 'act' like humans, such as chatbots. Finally, it deals with the effect of AI systems on human affection, trust, and (in)dependence [18]. To achieve human autonomy, human oversight is a prerequisite. Oversight may be achieved through governance mechanisms such as a human-in-the-loop (HITL), human-on-the-loop (HOTL), or human-in-command (HIC). Human-in-the-loop refers to the capability for human intervention in every decision cycle of the system. Human-on-the-loop refers to the capability for human intervention during the design cycle of the system and monitoring the system's operation. Human-in-command refers to the capability to oversee the overall activity of the AI system (including its broader economic, societal, legal, and ethical impact) and the ability to decide when and how to use the AI system in any situation. The latter can include the decision not to use an AI system in a particular situation to establish levels of human discretion during the use of the system, or to ensure the ability to override a decision made by an AI system [19].

LEGAL AND ETHICAL ASSESSMENT			
	REQUIREMENT	**GUIDING QUESTIONS/SUB-REQUIREMENTS**	**ASSESSMENT**
1.	**HUMAN AGENCY AND OVERSIGHT**		**CYLCOMED Toolbox**
1.1	**Human Agency and Autonomy**	• Is the AI system designed to interact, guide or take decisions by human end-users that affect humans or society? □ Could the AI system generate confusion for some or all end-users or subjects on whether a decision, content, advice or outcome is the result of an algorithmic decision?	The AI-based tools developed within the CYLCOMED Toolbox are designed to assist cybersecurity professionals by providing analytical insights and contextual information about system vulnerabilities, anomalies, and potential threats. Their purpose is to inform and guide human decision-making, not to

Funded by Horizon Europe Framework Programme of the European Union

		▫ Are end-users or other subjects adequately made aware that a decision, content, advice or outcome is the result of an algorithmic decision? • Could the AI system generate confusion for some or all end-users or subjects on whether they are interacting with a human or AI system? ▫ Are end-users or subjects informed that they are interacting with an AI system? • Could the AI system affect human autonomy by generating over-reliance by end-users? ▫ Did you put in place procedures to avoid that end-users over-rely on the AI system? • Could the AI system affect human autonomy by interfering with the end-user's decision-making process in any other unintended and undesirable way? ▫ Did you put in place any procedure to avoid that the AI system inadvertently affects human autonomy? • Does the AI system simulate social interaction with or between end-users or subjects? • Does the AI system risk creating human attachment, stimulating addictive behaviour, or manipulating user behaviour? Depending on which risks are possible or likely, please answer the questions below: ▫ Did you take measures to deal with possible negative consequences for end-users or subjects in case they develop a disproportionate attachment to the AI System? ▫ Did you take measures to minimise the risk of addiction? ▫ Did you take measures to mitigate the risk of manipulation?	replace it. The tools serve as decision-support mechanisms that enhance the end-user's situational awareness in identifying, assessing, and mitigating cybersecurity incidents affecting connected medical devices. Given their strictly analytical and non-autonomous nature, the CYLCOMED AI components are not capable of creating confusion regarding the origin of the outputs or the role of the algorithm in the decision-making process. The tools do not generate prescriptive or opaque outcomes but provide transparent, explainable analyses that can be verified by human experts. All users are clearly informed and fully aware that the outcomes are produced by AI algorithms integrated into the CYLCOMED Toolbox, and that these results must be interpreted within the broader human-led cybersecurity assessment workflow. The system design also ensures that users cannot confuse AI-generated outputs with human interaction. The tools do not simulate social behaviour, dialogue, or emotional responses, and therefore carry no risk of misleading users about whether they are engaging with a human or an AI system. All system interfaces, documentation, and user training explicitly clarify that interactions occur with AI-driven analytical software. Regarding human autonomy and potential over-reliance, the consortium has taken explicit measures to ensure that the tools reinforce rather than diminish the user's decision-making authority. Analysts retain complete control and may review, verify, and, if necessary, override all AI-generated outputs. System

Funded by Horizon Europe Framework Programme of the European Union

			logs provide detailed traceability of operations, supporting the validation of results and post-analysis auditing. In addition, users can safely stop or terminate any tool's operation at any moment through standard command-line control (docker compose stop), further guaranteeing user autonomy and operational safety. Given the non-social and non-prescriptive nature of these AI components, the risk of users developing over-reliance, attachment, or manipulation is considered negligible. The tools do not incorporate persuasive design elements, gamification, or adaptive user modelling that could trigger emotional dependence or addictive behaviour. Consequently, no dedicated mitigation procedures were deemed necessary beyond standard oversight and user training. Therefore, the CYLCOMED AI-based tools are developed and deployed within a framework that promotes transparency, accountability, and human agency. The system architecture and governance model prevent confusion, ensure full user awareness of AI involvement, and maintain the user's ability to control, interpret, and, where necessary, intervene in any AI-supported process.
1.2	**Human Oversight**	• Please determine whether the AI system (choose as many as appropriate): ☐ Is a self-learning or autonomous system; ☐ Is overseen by a Human-in-the-Loop; ☐ Is overseen by a Human-on-the-Loop; ☐ Is overseen by a Human-in-Command.	Within the CYLCOMED Toolbox, all AI-based tools operate under the direct human oversight, ensuring that ultimate decision-making authority remains with the human operator. The human supervisors responsible for this role have received specific training on how to exercise oversight, interpret system outputs,

Funded by Horizon Europe Framework Programme of the European Union

		• Have the humans (human-in-the-loop, human-on-the-loop, human-in-command) been given specific training on how to exercise oversight? • Did you establish any detection and response mechanisms for undesirable adverse effects of the AI system for the end-user or subject? • Did you ensure a 'stop button' or procedure to safely abort an operation when needed? • Did you take any specific oversight and control measures to reflect the self-learning or autonomous nature of the AI system?	and respond appropriately in case of anomalies or unexpected results. To safeguard users and maintain accountability, the AI tools have been designed with built-in monitoring capabilities. Each system generates comprehensive logs, which allow operators to detect potential undesirable or adverse effects and to implement corrective measures when needed. In line with CYLCOMED's focus on transparency and control, the architecture also incorporates a clear "stop" mechanism: any AI-based process can be safely interrupted through a command-line procedure (docker compose stop), ensuring that operations can be halted immediately if necessary. While the CYLCOMED AI tools are not self-learning or autonomous, the project recognises the importance of maintaining oversight principles consistent with such systems. Therefore, the implemented measures, human oversight, traceability through system logs, and clear shutdown procedures, collectively ensure responsible human control and operational safety throughout the system's lifecycle. Hence, the CYLCOMED AI-based components operate within a human-centred governance model. Oversight is structured, documented, and actionable, ensuring that every AI-driven process remains transparent, reversible, and under the control of qualified human experts. These measures collectively uphold the principles of trustworthy AI, namely, human agency and oversight.

Funded by Horizon Europe Framework Programme of the European Union

2.4.2 Technical Robustness and Safety

This requirement is closely linked to the principle of prevention of harm. The ***concept of robustness*** requires that AI systems should be developed with a preventative approach to risks and in a manner such that they reliably behave as intended while minimizing unintentional and unexpected harm and preventing unacceptable harm. This should also apply to potential changes in their operating environment or the presence of other agents (human and artificial) that may interact with the system in an adversarial manner [20]. In addition, AI systems must be resilient against adversarial attacks. A fallback plan must be in place in case problems arise. Moreover, AI systems must achieve a sufficient level of accuracy. To mitigate the risks of incorrect predictions, it is important to put an explicit and well-formed development and evaluation process in place [21]. Finally, the output of AI systems must be *reproducible and reliable*. An AI system is reliable if it works properly with a range of inputs and in a range of situations. Reliability is needed to prevent unintended harm [22]. Overall, the level of safety measures that must be implemented to meet the requirement of *technical robustness and safety* depends on the magnitude of the risk posed by the AI system in question, which depends on the system's capabilities and the environment in which it is deployed [23].

LEGAL AND ETHICAL ASSESSMENT			
	REQUIREMENT	**GUIDING QUESTIONS/SUB-REQUIREMENTS**	**ASSESMENT**
2.	**TECHNICAL ROBUSTNESS AND SAFETY**		**CYLCOMED Toolbox**

Funded by Horizon Europe Framework Programme of the European Union

2.1	**Resilience to Attack and Security**	• Could the AI system have adversarial, critical or damaging effects (e.g. to human or societal safety) in case of risks or threats such as design or technical faults, defects, outages, attacks, misuse, inappropriate or malicious use? • Is the AI system certified for cybersecurity (e.g. the certification scheme created by the Cybersecurity Act in Europe)19 or is it compliant with specific security standards? • How exposed is the AI system to cyber-attacks? Did you assess potential forms of attacks to which the AI system could be vulnerable? Did you consider different types of vulnerabilities and potential entry points for attacks such as: ▪ Data poisoning (i.e. manipulation of training data); ▪ Model evasion (i.e. classifying the data according to the attacker's will); ▪ Model inversion (i.e. infer the model parameters) • Did you put measures in place to ensure the integrity, robustness and overall security of the AI system against potential attacks over its lifecycle? • Did you red-team/pentest the system? • Did you inform end-users of the duration of security coverage and updates? □ What length is the expected timeframe within which you provide security updates for the AI system?	The AI-based tools of the CYLCOMED Toolbox have been developed with a strong emphasis on security-by-design, ensuring that no critical or damaging effects to human or societal safety could occur, even in the event of system faults, outages, or malicious interference. The tools operate within strictly controlled environments, providing cybersecurity analysts with diagnostic support and contextual information, without the capability to independently initiate actions that could directly impact patients, healthcare infrastructure, or public safety. For instance, the LADS, currently at Technology Readiness Level (TRL) 6, has not yet obtained a formal cybersecurity certification under the European Cybersecurity Act, its design, implementation, and validation processes follow ATOS's internal quality and security assurance standards, aligned with internationally recognised best practices (e.g., ISO/IEC 27001 and ISO/IEC 62443 principles). These standards guide both software development and deployment, ensuring that confidentiality, integrity, and availability are preserved throughout the system lifecycle. Next, the CYLCOMED AI tools are developed and operated in secure, monitored environments, minimising exposure to external cyber threats. As part of the project's development process, a static code analysis and vulnerability assessment were conducted to identify and mitigate potential weaknesses. All software dependencies are automatically verified and updated, reducing the risk of vulnerabilities propagating into the deployed environment. Furthermore, the development team conducted a vulnerability and exposure review to assess possible attack vectors,

Funded by Horizon Europe Framework Programme of the European Union

			including data poisoning, model evasion, and model inversion, ensuring that the systems remain robust against such adversarial techniques. Each tool is equipped with integrity verification mechanisms, ensuring the consistency and reliability of outputs, while continuous logging and monitoring enable rapid detection of any anomalous behaviour. These preventive measures are supported by controlled deployment environments and access management policies that restrict unauthorised use or modification. Although a full red-teaming or penetration-testing exercise has not yet been completed at the current TRL level, preliminary internal testing and dependency scanning have been performed to validate resilience against known classes of cyber threats. These activities will be expanded in the subsequent stages of the project as the tools progress toward higher readiness levels and integration within the CYLCOMED ecosystem. End-users are fully informed about the scope and duration of security coverage, including update mechanisms and maintenance responsibilities during the project period. Security updates are provided throughout the lifecycle of the tools within the project's implementation phase, ensuring that all software components remain compliant with the latest vulnerability disclosures and patches. In summary, the CYLCOMED Toolbox embodies the principles of robustness, technical resilience, and trustworthy AI, ensuring that all AI-based functionalities operate safely under human control, with clear safeguards against both accidental and malicious threats.

Funded by Horizon Europe Framework Programme of the European Union

2.2	**General Safety**	• Did you define risks, risk metrics and risk levels of the AI system in each specific use case?
		  Did you put in place a process to continuously measure and assess risks?
		  Did you inform end-users and subjects of existing or potential risks?
		• Did you identify the possible threats to the AI system (design faults, technical faults, environmental threats) and the possible consequences?
		  Did you assess the risk of possible malicious use, misuse or inappropriate use of the AI system?
		  Did you define safety criticality levels (e.g. related to human integrity) of the possible consequences of faults or misuse of the AI system?
		• Did you assess the dependency of a critical AI system's decisions on its stable and reliable behaviour?
		  Did you align the reliability/testing requirements to the appropriate levels of stability and reliability?
		• Did you plan fault tolerance via, e.g. a duplicated system or another parallel system (AI-based or 'conventional')?
		• Did you develop a mechanism to evaluate when the AI system has been changed to merit a new review of its technical robustness and safety?
			The CYLCOMED Toolbox incorporates a structured and continuous approach to risk identification, monitoring, and mitigation throughout its design and development phases. Each AI-based component has been evaluated for potential operational, technical, and ethical risks in alignment with the consortium's internal quality assurance procedures and the overarching principles of trustworthy and human-centric AI.
			During development, all software components underwent systematic debugging, static code analysis, and vulnerability scanning to identify and quantify possible risks. Detected issues, such as software defects or configuration weaknesses, were logged, evaluated for criticality, and resolved within the project's continuous integration and deployment (CI/CD) framework. This process ensured that all potential risks were not only detected early but also continuously monitored and reassessed as the system evolved.
			End-users were informed of potential risks during training and deployment phases, including operational limitations, dependencies on data quality, and system stability boundaries. Given that the CYLCOMED AI tools are used in controlled cybersecurity environments and operate under direct human supervision, the risk of misuse or malicious application is considered negligible. The tools are designed for professional use by trained analysts and are not accessible to the public, further reducing exposure to inappropriate or unsafe use.
			Due to their deterministic and explainable design, the CYLCOMED AI tools are not susceptible to unpredictable behaviour or self-modification. Each

Funded by Horizon Europe Framework Programme of the European Union

			module was tested for stability and reliability, ensuring consistent performance across operational conditions. Should external dependencies, such as data inputs or network connections, become unavailable, the systems are configured to fail safely, preserving data integrity and avoiding any cascading system effects. Testing and validation requirements were carefully aligned with each tool's functional criticality and stability level, following the internal quality standards of the technical partners. Fault tolerance is achieved through containerized deployment architectures, allowing for rapid rollback or system recovery in case of anomalies. This approach supports operational continuity and prevents the propagation of faults across the system ecosystem. Finally, the project employs a formal change management process to ensure continuous oversight of system evolution. All modifications to the AI tools, whether algorithmic, functional, or structural, are tracked through version control (Git) and CI/CD pipelines. Any significant updates trigger an internal review to assess whether the changes warrant a new robustness and safety evaluation. In summary, risk management in CYLCOMED is proactive, transparent, and iterative, ensuring that the AI-based components maintain high standards of safety, reliability, and accountability throughout their lifecycle. This approach aligns with the AI Act's requirements on technical robustness and risk management, ensuring that all AI functionalities operate within well-defined, human-controlled, and ethically sound boundaries.

Funded by Horizon Europe Framework Programme of the European Union

2.3	**Accuracy**	• Could a low level of accuracy of the AI system result in critical, adversarial or damaging consequences? • Did you put in place measures to ensure that the data (including training data) used to develop the AI system is up-to-date, of high quality, complete and representative of the environment the system will be deployed in? • Did you put in place a series of steps to monitor, and document the AI system's accuracy? • Did you consider whether the AI system's operation can invalidate the data or assumptions it was trained on, and how this might lead to adversarial effects? • Did you put processes in place to ensure that the level of accuracy of the AI system to be expected by end-users and/or subjects is properly communicated?	The CYLCOMED Toolbox follows a rigorous approach to ensure the accuracy, reliability, and data quality of its AI-based components. Given that the tools operate as decision-support systems rather than autonomous agents, their outputs are always subject to human verification, preventing any potential negative consequences that could arise from inaccurate or misleading results. Human analysts retain full authority over interpretation and action, maintaining a clear human oversight throughout the system's operation. To uphold data integrity, all AI modules consume up-to-date and contextually relevant data drawn from the monitored environment. This ensures that the tools operate on current threat intelligence, system logs, and network traffic information, maintaining their analytical validity over time. Data inputs are validated and pre-processed according to established quality assurance protocols implemented by the technical partners. A system of continuous monitoring and validation has been integrated into the CYLCOMED development lifecycle. The accuracy of model outputs is tracked and verified using system logs accessible through the command line, which facilitate real-time performance observation and retrospective evaluation. This enables the consortium to detect any deviations in analytical behaviour and to take corrective measures promptly. In addition, ongoing consistency checks are conducted to ensure that model outputs remain aligned with the data context and assumptions on which the systems were trained. This prevents the occurrence of "model drift" or adversarial effects resulting from outdated or invalid data relationships.

Funded by Horizon Europe Framework Programme of the European Union

			The expected accuracy and operational limitations of each AI-based tool are clearly communicated to end-users through technical documentation and interface guidelines. This transparency ensures that users have realistic expectations of system performance and understand the supporting, not substitutive, role of AI in their analytical workflows. Therefore, the CYLCOMED Toolbox embeds a continuous validation process grounded in data quality, model transparency, and human oversight.

2.4	**Reliability, Fall-back plans and Reproducibility**	• Could the AI system cause critical, adversarial, or damaging consequences (e.g. pertaining to human safety) in case of low reliability and/or reproducibility?
		  Did you put in place a well-defined process to monitor if the AI system is meeting the intended goals?
		  Did you test whether specific contexts or conditions need to be taken into account to ensure reproducibility?
		• Did you put in place verification and validation methods and documentation (e.g. logging) to evaluate and ensure different aspects of the AI system's reliability and reproducibility?
		  Did you clearly document and operationalise processes for the testing and verification of the reliability and reproducibility of the AI system?
		• Did you define tested failsafe fallback plans to address AI system errors of whatever origin and put governance procedures in place to trigger them?
		• Did you put in place a proper procedure for handling the cases where the AI system yields results with a low confidence score?
			The AI-based tools within the CYLCOMED Toolbox are designed as analytical and decision-support components that cannot directly affect patients or physical systems. Consequently, even in the event of low reliability or reproducibility, the tools cannot cause critical, adversarial, or damaging consequences to human safety or societal well-being. Human experts remain in full control of interpretation and operational decisions.
			A continuous monitoring and validation process is implemented throughout the tools' development and piloting phases to ensure that system outputs remain consistent with the defined functional, ethical, and performance objectives. Each component is tested against clear acceptance criteria, ensuring that performance metrics, such as reliability, reproducibility, and accuracy, align with the expected operational parameters.
			To verify reproducibility, the CYLCOMED tools were tested across multiple datasets and deployment environments, confirming consistent analytical behaviour under diverse conditions. Verification and validation procedures include comprehensive logging,

Funded by Horizon Europe Framework Programme of the European Union

		• Is your AI system using (online) continual learning? Did you consider potential negative consequences from the AI system learning novel or unusual methods to score well on its objective function?	documentation, and traceability mechanisms, enabling full reconstruction of test results and operational performance. All related procedures, test protocols, and outcomes are documented within project deliverables to maintain transparency and auditability. The consortium has also implemented failsafe and fallback mechanisms to address potential system errors or anomalies. In practice, this includes the ability to manually override or stop operations immediately, as well as automatic safeguards that ensure the system fails safely without compromising integrity or stability. Where the tools yield low-confidence results, these outputs are explicitly flagged to the user, who can review, re-process, or discard them as appropriate, ensuring that critical judgments are never automated or hidden from human oversight. The CYLCOMED Toolbox does not employ continual or online learning. All models operate on static, validated datasets, eliminating the risk of unintended behavioural drift or the autonomous adoption of novel or unvalidated analytical strategies. This design choice ensures stability, reproducibility, and compliance with the principles of traceability and technical robustness under the EU's trustworthy AI framework.

2.4.3 Privacy and Data Governance

The principle of prevention of harm also requires adequate ***privacy and data governance***. AI systems must guarantee privacy and data protection throughout a system's entire lifecycle. This includes the information initially provided by the user, as well as the information generated about the user over the course of their interaction with the system. In addition, the data sets used to train, test,

Funded by Horizon Europe Framework Programme of the European Union

and evaluate AI systems must be of sufficient quality. The societal biases, inaccuracies, errors, and mistakes contained in a data set must be addressed prior to training. Finally, data protocols governing data access should be put in place. These protocols should outline who can access data and under which circumstances [24].

LEGAL AND ETHICAL ASSESSMENT			
	REQUIREMENT	**GUIDING QUESTIONS/SUB-REQUIREMENTS**	**ASSESMENT**
3.	**PRIVACY, DATA PROTECTION & GOVERNANCE**		**CYLCOMED Toolbox**
3.1	**Privacy**	• Did you consider the impact of the AI system on the right to privacy, the right to physical, mental and/or moral integrity and the right to data protection? • Depending on the use case, did you establish mechanisms that allow flagging issues related to privacy concerning the AI system?	The idea of privacy derives from concepts such as human dignity and the rule of law, and it generally refers to the protection of an individual's "personal space" [25]. The CYLCOMED toolbox integrates various tools and functionalities to ensure privacy and data protection. Moreover, the design and development of the toolbox is based on the principles of Data Protection by Design and by Default. The principles of Data Protection by Design and by Default serve as fundamental safeguards for individual privacy and data protection. The core obligation is the implementation of appropriate measures and necessary safeguards that provide effective implementation of the data protection principles and, consequentially, data subjects' rights and freedoms by design and by default [26]. While primarily directed towards controllers, the General Data Protection Regulation (GDPR) encourages producers of products, services, and applications to take into account the right to data protection when they develop, design, select, and use applications, services,

Funded by Horizon Europe Framework Programme of the European Union

			and products (data protection by design) [27]. In doing so, these producers are encouraged to give due consideration to the state of the art [28]. In other words, they should take into account the current progress in technology that is available in the market [29]. The Cylcomed's data protection layer protects the health data collected by the CMDs by encrypting the patient's health data with the FE4MED (the Attribute-Based Encryption (ABE) data protection solution). This protected data is stored in the hospital system. The medical staff proceeds to authenticate by presenting Verifiable Credentials (VC) issued by the hospital, containing user role and attributes, stored in the user LuS4MED mobile app wallet. The role and attributes included in the VC are validated by the LuS4MED Agent and then provided to the authorisation solution (C-OPA in pilot 1 and legacy Keycloak in pilot 2) which grants the users to access the allowed services. Based on the user role and attributes, the FE4MED decrypts the requested data that the user is allowed to see. The use of the data protection layer is monitored by the AI-Log monitoring tool (the LOMOS solution) for detecting anomalies during the encryption/decryption processes. In this manner, both the authentication solution (LuS4MED) and the protection solution (FE4MED) are integrated for protecting the access to the user's data and preserving the privacy of these sensitive data [30]. The end user organisation of the CYLCOMED toolbox (e.g., a hospital) will be a 'data controller' and will be responsible for fulfilling data protection obligations. One of these obligations is the data protection by design and by default, which requires putting in place technical and organisational measures at the design

Funded by Horizon Europe Framework Programme of the European Union

			stage with a view to implementing data protection principles. Product and service providers' roles in designing the system are, therefore, crucial to make sure that final users can comply with their obligation. In order to enable data controllers to fulfil by-design obligation and other data protection requirements, the developers of the CYLCOMED toolbox have taken into account the data protection principles in the design and development of the toolbox, giving due regard to the state of the art. These principles are examined more concretely and separately below.
3.2	**Data Governance**	• Is your AI system being trained, or was it developed, by using or processing personal data (including special categories of personal data)? • Did you consider the privacy and data protection implications of the AI system's non-personal training-data or other processed non-personal data?	The AI-based tools developed within the CYLCOMED Toolbox were not trained on, nor do they process, any form of personal data, including special categories of personal data as defined under Article 9 of the GDPR. The development and operation of the CYLCOMED components are based entirely on and derived from non-personal technical sources (e.g., network logs, device telemetry, or simulated attack vectors). This ensures that no direct or indirect identification of individuals is possible at any stage of the system lifecycle. Although personal data is not involved, the consortium has explicitly assessed the privacy and data protection implications of using non-personal and technical datasets. In particular, partners evaluated whether metadata, device identifiers, or aggregated telemetry could inadvertently lead to re-identification or disclosure of sensitive operational patterns. Appropriate data minimisation and pseudonymisation principles were applied to prevent any residual privacy risk.

Funded by Horizon Europe Framework Programme of the European Union

			All partners adhere to GDPR-compliant research ethics standards and follow institutional and project-level data governance protocols, ensuring that data sources, access permissions, and processing operations are clearly documented and traceable. The Data Management Plan (DMP) established for CYLCOMED provides an overarching structure for data lifecycle management, defining safeguards for both personal and non-personal data, including storage, sharing, and deletion practices.
3.2.1	**DPO and DPIA**	• Did you put in place any of the following measures some of which are mandatory under the General Data Protection Regulation (GDPR), or a non-European equivalent? □ Data Protection Impact Assessment (DPIA)23; □ Designate a Data Protection Officer (DPO)24 and include them at an early state in the development, procurement or use phase of the AI system;	The GDPR recognises the Data Protection Officer (DPO) as a central figure in the data governance framework and sets out specific conditions for their appointment, position, and tasks. The role of the DPO is critical in ensuring compliance with data protection laws and promoting a culture of data privacy within organisations [31]. All Hospitals participating in the CYLCOMED project have appointed DPOs, who were internally engaged. For instance, Ospedale Pediatrico Bambin Gesù Hospital's DPO was continuously involved in various stages of the project implementation. These include, inter alia, monitoring the compliance of CYLCOMED's procedures with the current privacy policies, involvement in the preparation of the DPIA, and clinical protocol for the Territorial Ethical Committee. Moreover, the DPO validated the toolbox's GDPR alignment, particularly through pseudonymization and access control logs. Besides DPO, a DPIA is a critical tool for accountability, as it helps controllers not only to comply with the requirements of the GDPR but also to

Funded by Horizon Europe Framework Programme of the European Union

			demonstrate that appropriate measures have been taken to ensure compliance with the Regulation. A DPIA is a process designed to describe the processing, assess the necessity and proportionality of a processing, and help manage the risks to the rights and freedoms of data subjects resulting from the processing of personal data. A Data Protection Impact Assessment (DPIA) was carried out as part of the toolbox's integration into hospital settings and its validation within PILOT2. The DPIA process was initiated to assess the innovative features of the cybersecurity toolbox, ensuring compliance with data protection requirements while evaluating its practical deployment [32].
3.2.2	**Identity and access management**	• Did you put in place any of the following measures some of which are mandatory under the General Data Protection Regulation (GDPR), or a non-European equivalent? ☐ Oversight mechanisms for data processing (including limiting access to qualified personnel, mechanisms for logging data access and making modifications);	The Cylcomed toolbox architecture encompasses a dedicated layer addressing Identity and Access Management. This layer features LuS4MED (SSI-based authentication solution), C-OPA (policy-driven authorisation), and FE4MED (data protection via CP-ABE encryption). LuS4MED is fully implemented for SSI-based VC issuance/verification, integrated with C-OPA for policy-driven authorisation; FE4MED provides modular CP-ABE encryption/decryption, ensuring policy-driven access and GDPR-compliant data protection across pilots [33]. In the Identity, Access Management & Data Protection layer, LuS4MED handles SSI. It issues and verifies VCs using the LuS4MED Agent API, with a mobile wallet for users. C-OPA is the authorisation proxy using Envoy and OPA with Rego policies. Its Policy & Bundle API checks tokens and roles to control access to services like FE4MED. FE4MED uses CP-ABE encryption scheme

Funded by Horizon Europe Framework Programme of the European Union

			for attribute-based access, providing fine-grained access control. Its Encrypt/Decrypt REST API works at the edge (RBPI/Android) and server for decryption by allowed roles. These parts provide full data protection with decentralised identity [34].
3.2.3	**Data security (Integrity and Confidentiality)**	• Did you put in place any of the following measures some of which are mandatory under the General Data Protection Regulation (GDPR), or a non-European equivalent? □ Measures to achieve privacy-by-design and default (e.g. encryption, pseudonymisation, aggregation, anonymisation);	The principle of integrity and confidentiality includes protection against unauthorised or unlawful processing and against accidental loss, destruction or damage, using appropriate technical or organisational measures [35]. GDPR Article 32(1) states that state-of-the-art, the costs of implementation, and the nature, scope, context, and purpose of processing should be taken into account when assessing which technical or organisational measures should be put in place. The CYLCOMED toolbox integrates various tools and functionalities to ensure protection against unauthorised or unlawful processing and against accidental loss, destruction, or damage. Some of the integrated facets that the toolbox encompasses include, inter alia, the following: • Robust encryption methods, • Identity and Access Management, • Regular backups, • Update management, • Secure transfers, • Log monitoring.
	Data minimisation	• Did you put in place any of the following measures some of which are mandatory under the General Data Protection Regulation (GDPR), or a non-European equivalent?	According to this principle, personal data should be processed as long as it is adequate, relevant, and necessary for the purpose of processing [36]. The EDPB in the Guidelines on Article 25 Data Protection by Design and by Default [37] sheds more light on the

Funded by Horizon Europe Framework Programme of the European Union

		Data minimisation, in particular personal data (including special categories of data);	implementation of the data minimisation principle. Some of the key design and default data minimisation elements suggested by the EDPB may include data avoidance, access limitation, and "state-of-the-art". Taking this into account, the CYLCOMED cybersecurity toolbox integrates these design features, and avoids processing personal data altogether, and shapes the data flow in a way that a minimal number of people need access to data to perform their duties, and limit access accordingly. The toolbox integrates Identity and Access Management (IAM) and Data Protection tools (LuS4MED and C-OPA, and FE4MED, respectively) for protecting the privacy of the patient's data (from connected medical devices and smartphones) and the access control. The "State of the art" of these tools is described in the deliverable D5.3, and will not be further elaborated here.
3.2.4	**International standards**	• Did you align the AI system with relevant standards (e.g. ISO25, IEEE26) or widely adopted protocols for (daily) data management and governance?	The toolbox is aligned with the international standards on information security and Security and Privacy Controls for Information Systems and Organizations (ISO 2700115 standard on information security, NIST 800-53 Security and Privacy Controls for Information Systems and Organizations).
3.2.4	**Data breach notifications**		Critical infrastructure operators often need to meet strict deadlines to notify component authorities about data breaches. The CYLCOMED cybersecurity toolbox offers an automated solution that provides support for detecting and analysing anomalies. More specifically, the dashboard provides a comprehensive overview of network traffic data and anomaly detection results. Each section serves a specific purpose to help inspect,

Funded by Horizon Europe Framework Programme of the European Union

			visualize, and interpret data. Charts and tables support operational monitoring and investigation [38]. In that sense, it provides added value to speed up the notification processes. The toolbox does not notify component authorities automatically. This is, in fact, desirable because security and/or compliance officers in the end-user organisation will be able to analyse whether conditions of a notification obligation are met and determine which particular information or data will be shared to make sure that all applicable law (e.g., data protection) is respected.

2.4.4 Transparency

The requirement of ***transparency*** is closely linked with the principle of explicability [39]. Transparency encompasses three elements: (1) traceability, (2) explainability, and (3) communication about the limits of the AI system. Traceability requires that the data and processes that yield the AI system's output are properly documented. Explainability concerns the ability to explain both the technical processes of the AI system and the reasoning behind the decisions or predictions that the AI system makes. The degree to which explainability is needed depends on the context and the severity of the consequences of erroneous or otherwise inaccurate output to human life. Finally, an AI system's capabilities and limitations should be communicated to deployers and users in an appropriate way. In addition, humans should be informed that they are interacting with an AI system. They should have the option to decide against such an interaction if fundamental rights are at stake [40].

LEGAL AND ETHICAL ASSESSMENT			
	REQUIREMENT	**GUIDING QUESTIONS/SUB-REQUIREMENTS**	**ASSESMENT**
4.	**TRANSPARENCY**		**CYLCOMED Toolbox**

Funded by Horizon Europe Framework Programme of the European Union

• **4.1**	**Traceability**	• Did you put in place measures that address the traceability of the AI system during its entire lifecycle?
		  Did you put in place measures to continuously assess the quality of the input data to the AI system?
		  Can you trace back which data was used by the AI system to make a certain decision(s) or recommendation(s)?
		  Can you trace back which AI model or rules led to the decision(s) or recommendation(s) of the AI system?
		  Did you put in place measures to continuously assess the quality of the output(s) of the AI system?
		  Did you put adequate logging practices in place to record the decision(s) or recommendation(s) of the AI system?
			The CYLCOMED Toolbox ensures comprehensive traceability and auditability across the entire lifecycle of its AI-based components, from design and development to deployment and monitoring.
			Full traceability is achieved through rigorous version control, documentation, and configuration management. All changes to code, algorithms, and model parameters are tracked using Git-based repositories, allowing developers and auditors to reconstruct the exact state of the system at any given point in time. This versioning framework ensures that every model update or rule modification is properly documented and reviewable.
			The quality of the input data used by the AI components is continuously validated and monitored to ensure consistency, completeness, and integrity. Automated checks are implemented to detect anomalies, missing values, or irregularities in data streams originating from network logs, system events, or other monitored components. This process guarantees that the AI tools operate on accurate and up-to-date information.
			All data processed by the system is logged and traceable, enabling analysts to identify which specific datasets contributed to a given analytical output or recommendation. Similarly, the model version and rule set that produced a specific result are recorded and linked to each logged decision event, ensuring end-to-end transparency of the decision-making process.
			
			To complement input traceability, continuous output validation and performance monitoring are

Funded by Horizon Europe Framework Programme of the European Union

			conducted to identify potential anomalies or degradations in model behaviour. Any irregular outputs are flagged for human review, ensuring that system recommendations remain consistent with the intended functionality and ethical boundaries. Logging practices are implemented at both the application and infrastructure levels, supporting full auditability and accountability. These logs capture operational events, data flows, user interactions, and AI-generated outputs in a secure and immutable format. Access to logs is restricted and governed by internal data management and cybersecurity policies, ensuring confidentiality and integrity while allowing for independent verification and internal audits. In summary, the CYLCOMED AI-based components are designed to be fully traceable and auditable, with robust mechanisms in place for logging, monitoring, and validation at every stage of the system's lifecycle.
• **4.2**	**Explainability**	• Did you explain the decision(s) of the AI system to the users? • Do you continuously survey the users if they understand the decision(s) of the AI system?	The AI-based tools integrated within the CYLCOMED Toolbox have been designed to support transparency and explainability, ensuring that users can interpret and validate the results produced by the system. While the underlying models are technical in nature, the tools present their outputs through clear result visualisations and contextual data representations, helping users understand the factors influencing the AI's analysis. These visualisations allow cybersecurity analysts to trace how specific inputs, parameters, or detected anomalies contributed to the generated outcomes.* Explainability is achieved primarily through interpretable visual outputs rather than textual or rule-

Funded by Horizon Europe Framework Programme of the European Union

			based explanations. For instance, dashboards and log visualisations enable users to review contributing variables, anomaly scores, or risk indicators, thereby understanding the rationale behind the AI's suggestions. This supports informed human decision-making. At the current stage of development, continuous user comprehension surveys or feedback assessments have not been implemented. Nevertheless, user feedback from pilot testing and demonstration phases is collected.
•	**4.3**	**Communication**	• In cases of interactive AI systems (e.g., chatbots, robo-lawyers), do you communicate to users that they are interacting with an AI system instead of a human? • Did you establish mechanisms to inform users about the purpose, criteria and limitations of the decision(s) generated by the AI system? □ Did you communicate the benefits of the AI system to users? □ Did you communicate the technical limitations and potential risks of the AI system to users, such as its level of accuracy and/ or error rates? □ Did you provide appropriate training material and disclaimers to users on how to adequately use the AI system?	The CYLCOMED Toolbox does not include any interactive or conversational AI components such as chatbots, virtual assistants, or human-like interfaces. Accordingly, there is no possibility of users mistaking the system for a human interlocutor. All tools are technical and analytical in nature, operating through dashboards, logs, and command-line interfaces intended for use by trained cybersecurity professionals. Transparency toward users is ensured through clear communication of the purpose, functionality, and operational scope of each AI-based tool. The interface and accompanying documentation allow users to review the data and contextual information supporting each result, enabling them to understand the basis of the AI-generated analysis. This transparency is further supported by detailed technical deliverables describing each tool's intended use and integration within the CYLCOMED ecosystem. The benefits and added value of the CYLCOMED AI tools are explicitly communicated to users and

Funded by Horizon Europe Framework Programme of the European Union

			stakeholders through the project's public deliverables, training sessions, and dissemination materials. These highlight how the tools enhance cybersecurity monitoring, detection, and analysis capacities for connected medical devices while supporting compliance with EU regulatory and ethical standards. Equally important, the consortium ensures that users are informed of the technical limitations, operational boundaries, and potential risks of the AI components. Metrics related to performance, accuracy, and reliability are clearly stated in the project documentation, ensuring realistic expectations and responsible system use.

2.4.5 Inclusion and Diversity

Trustworthy AI enables ***inclusion and diversity***. Since AI systems are trained on data from the real world, there is a risk that they inherit the bias and discrimination that exists in the real world. As a result, the output of AI systems can be discriminatory against certain groups or people, thereby exacerbating prejudice and marginalization. Identifiable and discriminatory bias should therefore be removed in the collection phase where possible. Oversight mechanisms can be put in place to analyse an AI system's output and decisions. AI systems should also be accessible. They should be designed in a way that allows all people to use them, regardless of their age, gender, abilities or characteristics. Accessibility to this technology for persons with disabilities, which are present in all societal groups, is of particular importance. It is important to consider Universal Design principles addressing the widest possible range of users, following relevant accessibility standards. Finally, it is advisable to consult stakeholders who may directly or indirectly be affected by the system. In a manufacturing context, it is advisable to solicit regular feedback by ensuring workers' information, consultation and participation throughout the whole process of implementing AI systems at organizations [41].

LEGAL AND ETHICAL ASSESSMENT

Funded by Horizon Europe Framework Programme of the European Union

	REQUIREMENT	GUIDING QUESTIONS/SUB-REQUIREMENTS	ASSESSMENT
5.	**DIVERSITY, NON-DISCRIMINATION AND FAIRNESS**		**CYLCOMED Toolbox**

5.1	**Avoidance of Unfair Bias**	• Did you establish a strategy or a set of procedures to avoid creating or reinforcing unfair bias in the AI system, both regarding the use of input data as well as for the algorithm design?
		• Did you consider diversity and representativeness of end-users and/or subjects in the data?
		  Did you test for specific target groups or problematic use cases?
		  Did you research and use publicly available technical tools, that are state-of-the-art, to improve your understanding of the data, model and performance?
		  Did you assess and put in place processes to test and monitor for potential biases during the entire lifecycle of the AI system (e.g. biases due to possible limitations stemming from the composition of the used data sets (lack of diversity, non-representativeness)?
		  Where relevant, did you consider diversity and representativeness of end-users and or subjects in the data?
		• Did you put in place educational and awareness initiatives to help AI designers and AI developers be more aware of the possible bias they can inject in designing and developing the AI system?
		• Did you ensure a mechanism that allows for the flagging of issues related to bias, discrimination or poor performance of the AI system?
			The CYLCOMED Toolbox was developed with a deliberate focus on preventing unfair bias and ensuring the fairness, transparency, and reliability of all AI-based components. Although the system operates primarily on technical and system-level data rather than on personal or human-subject datasets, the consortium implemented a clear strategy to detect and mitigate potential algorithmic bias during all phases of data processing and model development.
			During development, project partners established procedures for bias detection and mitigation, including structured code reviews, data validation, and statistical consistency checks to ensure that no imbalance or artefact in the data could inadvertently influence the system's analytical outputs. The datasets used by CYLCOMED tools consist mainly of network logs, telemetry data, and cybersecurity event traces, which are inherently non-personal and not linked to protected or sensitive human attributes such as gender, ethnicity, or socioeconomic status. Consequently, the likelihood of bias in the social or demographic sense is negligible.
			Nevertheless, the consortium adopted a precautionary approach by performing testing and validation during the project's pilot phases to assess system performance across different operational contexts and use cases, ensuring that the AI-based components behave consistently and reliably in varying technical environments. The team also relied on open-source,

Funded by Horizon Europe Framework Programme of the European Union

		▯ Did you establish clear steps and ways of communicating on how and to whom such issues can be raised? ▯ Did you identify the subjects that could potentially be (in)directly affected by the AI system, in addition to the (end-)users and/or subjects?	research-grade tools for model evaluation and validation, ensuring adherence to state-of-the-art practices in explainability, accuracy verification, and model robustness assessment. Throughout the development lifecycle, continuous validation and review mechanisms were employed to monitor the models for potential bias or performance degradation.
5.2	**Accessibility and Universal Design**	• Did you ensure that the AI system corresponds to the variety of preferences and abilities in society? • Did you assess whether the AI system's user interface is usable by those with special needs or disabilities or those at risk of exclusion? ▯ Did you ensure that information about, and the AI system's user interface of, the AI system is accessible and usable also to users of assistive technologies (such as screen readers)? ▯ Did you involve or consult with end-users or subjects in need for assistive technology during the planning and development phase of the AI system? • Did you ensure that Universal Design principles are taken into account during every step of the planning and development process, if applicable? • Did you take the impact of the AI system on the potential end-users and/or subjects into account? ▯ Did you assess whether the team involved in building the AI system engaged with the possible target end-users and/or subjects? ▯ Did you assess whether there could be groups who might be disproportionately affected by the outcomes of the AI system?	The CYLCOMED Toolbox has been designed primarily as a technical solution for cybersecurity professionals and system administrators responsible for the protection of connected medical devices. Given this highly specialised user group and the current Technology Readiness Level (TRL 6), the tools are not intended for public use or deployment in settings requiring accessibility features for users with disabilities. Consequently, accessibility and assistive-technology compatibility were not included among the functional requirements at this development stage. Nevertheless, the design and development processes adhered to general usability and clarity principles, ensuring that all interfaces and command-line interactions are intuitive, transparent, and easily interpretable by trained operators. The impact on potential end-users and subjects has been assessed through the project's pilot activities within the connected medical device use cases, ensuring that the system's outputs and functionalities align with the operational needs and cybersecurity responsibilities of healthcare organisations. Engagement with target end-users has been achieved through collaboration with pilot partners, who provided practical feedback on

Funded by Horizon Europe Framework Programme of the European Union

		□ Did you assess the risk of the possible unfairness of the system onto the end-user's or subject's communities?	system usability, interpretability of results, and integration into existing clinical or IT workflows. The consortium also assessed potential disproportionate effects on user groups and broader stakeholder communities. Given the nature of the AI tools, focused on network and log anomaly detection rather than individual profiling or automated decision-making, the risk of unfair or adverse impacts on any social or demographic group is considered minimal.
5.3	**Stakeholder Participation**	• Did you consider a mechanism to include the participation of the widest range of possible stakeholders in the AI system's design and development?	CYLCOMED adopts an inclusive, interdisciplinary perspective that reflects the complex and multifaceted nature of cybersecurity in healthcare. The project is grounded in recognising that cybersecurity cannot be understood and addressed only through technical lenses. Instead, CYLCOMED embraces an inclusive framework that reflects cybersecurity's multi-dimensional and interdisciplinary nature, integrating legal, ethical, and human-centric perspectives alongside technical innovations. The design and development of the toolbox were based on broad stakeholder participation, which included medical device manufacturers, software developers, healthcare professionals, and legal and ethical stakeholders. The baseline cybersecurity requirements reflect this broad collaboration, which has informed the development of the toolbox. A central strength of the toolbox is its bottom-up, human-centred validation strategy. Rather than relying solely on standards-driven frameworks, the project actively engaged clinicians, engineers, data protection officers, and quality managers through structured

Funded by Horizon Europe Framework Programme of the European Union

			interviews. This participatory process revealed both enablers and barriers to adoption, including the need for streamlined emergency access, enhanced patient communication, and increased transparency in data governance. By grounding validation in real-world practice, the approach complements high-level models such as the WHO 2025 Cybersecurity and Privacy Maturity Framework, providing a field-tested perspective that bridges the gap between policy ambitions and practical implementation [42].

2.4.6 Societal and Environmental Well-Being

The sixth requirement is ***societal and environmental well-being***. AI systems must work in the most environmentally friendly way possible. An AI system's development, deployment and use process, as well as its entire supply chain, should be assessed in this regard (e.g. via a critical examination of resource usage and energy consumption during training, opting for less net negative choices). Furthermore, AI systems should support humans in the working environment and aim for the creation of meaningful work. While AI systems can be used to enhance social skills, they can equally contribute to their deterioration. The effects of these systems on people's physical and mental wellbeing must therefore be carefully monitored and considered. Finally, the impact of an AI system's development, deployment and use on individuals should also be assessed from a societal perspective, taking into account its effect on institutions, democracy and society at large [43].

LEGAL AND ETHICAL ASSESSMENT			
•	**REQUIREMENT**	**GUIDING QUESTIONS/SUB-REQUIREMENTS**	**ASSESSMENT**

Funded by Horizon Europe Framework Programme of the European Union

• 6.	**SOCIETAL AND ENVIRONMENTAL WELL-BEING**		**CYLCOMED Toolbox**

•	**Environmental Well-being**	• Are there potential negative impacts of the AI system on the environment?
		  Which potential impact(s) do you identify?
		• Where possible, did you establish mechanisms to evaluate the environmental impact of the AI system's development, deployment and/or use (for example, the amount of energy used and carbon emissions)?
		  Did you define measures to reduce the environmental impact of the AI system throughout its lifecycle?
			The CYLCOMED Toolbox has been designed with a light computational footprint and therefore presents no significant negative environmental impact arising from its development, deployment, or use. The AI-based components operate within existing cybersecurity infrastructures and do not require intensive data processing, high-performance computing, or large-scale storage resources, which are typically associated with substantial energy consumption or carbon emissions in AI systems.
			The consortium identified power consumption as the most relevant potential environmental factor. However, the system's architecture was deliberately optimised for resource efficiency and scalability, allowing it to operate effectively even on low-power hardware, such as single-board computers (e.g., Raspberry Pi). This design choice minimises the system's energy demand and reduces its environmental footprint while maintaining performance and reliability.
			Throughout the development phase, project partners monitored energy use, processing time, and resource allocation to ensure efficient utilisation of computational infrastructure. These evaluations helped confirm that the CYLCOMED tools remain lightweight and suitable for integration in energy-constrained environments, such as hospital IT networks or edge devices connected to medical systems.

6.1	**Impact on Work and Skills**	• Does the AI system impact human work and work arrangements?	The deployment of the AI-based tools within the CYLCOMED Toolbox has a limited but positive impact on human work, primarily by enhancing the capabilities

Funded by Horizon Europe Framework Programme of the European Union

		• Did you pave the way for the introduction of the AI system in your organisation by informing and consulting with impacted workers and their representatives (trade unions, (European) work councils) in advance?
		• Did you adopt measures to ensure that the impacts of the AI system on human work are well understood?
		  Did you ensure that workers understand how the AI system operates, which capabilities it has and which it does not have?
		• Could the AI system create the risk of de-skilling of the workforce?
		  Did you take measures to counteract de-skilling risks?
		• Does the system promote or require new (digital) skills?
		  Did you provide training opportunities and materials for re- and up-skilling?
			of cybersecurity professionals rather than replacing or automating their functions. The tools are designed as decision-support systems, offering analytical insights that assist experts in detecting, assessing, and responding to cybersecurity incidents involving connected medical devices. They do not perform autonomous actions or make operational decisions independently, ensuring that the human role remains central in all critical processes.
			Given the analytical and assistive nature of the CYLCOMED tools, the risk of workforce de-skilling is considered minimal. The system does not replace existing competencies but rather supports and augments the analytical capacities of human operators. Accordingly, no specific countermeasures for de-skilling were necessary within the project's scope.
			
			While the CYLCOMED Toolbox does not inherently promote or require new digital skills beyond those already held by cybersecurity professionals, the project's collaborative activities have indirectly supported skills development by exposing participants to state-of-the-art AI-based methods in threat detection and network monitoring.

6.2	**Impact on Society at large or Democracy**	• Could the AI system have a negative impact on society at large or democracy?
		  Did you assess the societal impact of the AI system's use beyond the (end-)user and subject, such as potentially indirectly affected stakeholders or society at large? o Did you take action to minimize potential societal harm of the AI system?
		  Did you take measures that ensure that the AI system does not negatively impact democracy?
			The CYLCOMED Toolbox is a highly specialised cybersecurity solution designed for use by technical experts within healthcare and research organisations. Its functionalities are confined to the detection and analysis of cybersecurity threats targeting connected medical devices and related infrastructures. Given this technical and domain-specific scope, the system does not exert any direct or indirect influence on society at large or on democratic processes.

Funded by Horizon Europe Framework Programme of the European Union

			The consortium conducted an ethical review of the project's objectives and confirmed that the AI-based tools do not involve content generation, decision-making affecting citizens, or mechanisms capable of shaping public opinion or behaviour. As such, the potential for societal or democratic harm is effectively non-existent. The system does not process personal, political, or social data, nor does it engage with users or stakeholders outside the professional cybersecurity context. Because the tools are not deployed in environments where societal or political influence could occur, no additional mitigation measures were deemed necessary.

2.4.7 Accountability

The final requirement, ***accountability***, is closely related to the principle of fairness. Mechanisms should be put in place to ensure responsibility and accountability for AI systems and their outcomes, both before and after their development, deployment and use. It must be possible for AI systems to be audited as well as to access records on said audits that can contribute to Trustworthy AI. In addition, accountability requires proper risk management [44]. The potential negative impacts of AI systems should be identified, assessed, documented and minimized. Due protection must be available for whistle-blowers, NGOs, trade unions or other entities when reporting legitimate concerns about an AI system [45]. When unjust adverse impact occurs, accessible mechanisms should be foreseen that ensure adequate redress. This is key to ensure trust.

LEGAL AND ETHICAL ASSESSMENT

Funded by Horizon Europe Framework Programme of the European Union

•	REQUIREMENT	GUIDING QUESTIONS/SUB-REQUIREMENTS	ASSESSMENT
• 7.	**ACCOUNTABILITY**		**CYLCOMED Toolbox**
• 7.1	**Auditability**	• Did you establish mechanisms that facilitate the AI system's auditability (e.g. traceability of the development process, the sourcing of training data and the logging of the AI system's processes, outcomes, positive and negative impact)? • Did you ensure that the AI system can be audited by independent third parties?	The CYLCOMED Toolbox has been developed to ensure full auditability and transparency across all stages of its lifecycle. To address it, mechanisms have been established to document, monitor, and verify the development process, the provenance of data, and the performance of each AI-based component. Auditability is achieved through rigorous traceability measures, including detailed technical documentation, systematic logging of processes and outputs, and version control via secure Git-based repositories. These records enable developers, evaluators, and auditors to reconstruct the system's development history, validate design choices, and review how each AI component was trained, configured, and deployed. Logs are maintained both at the application and infrastructure levels, providing a complete record of the system's operations, performance, and any anomalies detected during testing or use. To ensure accountability and openness, the CYLCOMED architecture and documentation are structured to allow independent third-party audits. External reviewers can access the relevant repositories, version histories, and logged outputs under appropriate confidentiality and data-management conditions.

Funded by Horizon Europe Framework Programme of the European Union

•	**7.2**	**Risk Management**	• Did you foresee any kind of external guidance or third-party auditing processes to oversee ethical concerns and accountability measures?
			  Does the involvement of these third parties go beyond the development phase?
			• Did you organise risk training and, if so, does this also inform about the potential legal framework applicable to the AI system?
			• Did you consider establishing an AI ethics review board or a similar mechanism to discuss the overall accountability and ethics practices, including potential unclear grey areas?
			• Did you establish a process to discuss and continuously monitor and assess the AI system's adherence to this Assessment List for Trustworthy AI (ALTAI)?
			  Does this process include identification and documentation of conflicts between the 6 aforementioned requirements or between different ethical principles and explanation of the 'trade-off' decisions made?
			  Did you provide appropriate training to those involved in such a process and does this also cover the legal framework applicable to the AI system?
			• Did you establish a process for third parties (e.g. suppliers, end-users, subjects, distributors/vendors or workers) to report potential vulnerabilities, risks or biases in the AI system?
			  Does this process foster revision of the risk management process?
				The CYLCOMED project is rooted in the “ethics by design” approach which entails a systematic and comprehensive inclusion of ethical considerations in the design and development process. The central idea behind it is that technology can be made ethical through its design process [46]. CYLCOMED adhered to this approach across the project lifespan. Ethics Guidelines for Trustworthy AI provided guiding ethical principles from design to deployment stage of the toolbox.
				
				CYLCOMED project has dedicated legal and ethical WP which provided input and guidance for the overall project’s ethical adherence. This adherence has been subject of the assessment of the independent Ethical Committee prior the integration of the toolbox into hospital premisses.
				
				Next, the CYLCOMED system includes a dedicated framework for comprehensive cybersecurity risk management, specifically designed to support the identification and handling of risks associated with Connected Medical Devices (CMDs). As detailed in D4.3 “Risk Benefit Scheme and Risk Management Tool”, a risk management tool for connected medical devices has been developed, as well as a benefit-risk analysis process. The objective of the risk management tool is to manage the cybersecurity, privacy and safety risks of CMDs, taking into account:
				
				• Cybersecurity standards, such as ISO 27001, NIST CSF 2, NIST RMF;
				• Privacy standards, such as the ISO 27701, which is compatible with GDPR.

Funded by Horizon Europe Framework Programme of the European Union

		• For applications that can adversely affect individuals, have redress by design mechanisms been put in place?	• Safety standard, namely the most prominent standard used by medical device manufacturers, the ISO 14971:2019 standard, which is compatible with the MDR and the IVDR regulations [47]. • The tool provides a structured, modular interface that facilitates collaboration among stakeholders, including manufacturers, hospitals, and patients, in identifying threats, assessing vulnerabilities, and evaluating risks. It further supports risk-benefit analyses by mapping cybersecurity controls to device functions and assessing their impact on safety and performance. Some of the key functionalities include: • Asset, threat, and vulnerability management; • Risk assessment, evaluation, and treatment planning; • Visualization of risk levels and control effectiveness; • Role-based dashboards tailored to stakeholder needs; • Integration with hospital and manufacturer workflows.

Funded by Horizon Europe Framework Programme of the European Union

2.5 Concluding Remarks on Legal and Ethical Compliance of the CYLCOMED Toolbox

Based on the extensive legal and ethical assessment provided, the development and implementation of the CYLCOMED Cybersecurity Toolbox demonstrates a high degree of alignment with the legal and ethical standards articulated by the EU's Ethics Guidelines for Trustworthy AI and relevant regulatory frameworks, such as the General Data Protection Regulation (GDPR) and cybersecurity regulatory frameworks. The assessment across the seven key ethical requirements outlined by the High-Level Expert Group on AI (HLEG) confirms the following:

2.5.1 Human Agency and Oversight

The CYLCOMED Toolbox operates strictly as a decision-support system under direct human control, with no autonomous decision-making capacity. Human operators remain fully responsible for the interpretation and application of AI-generated outputs. Oversight mechanisms, including human-in-command design, training provisions, and manual override procedures, effectively safeguard user autonomy and prevent over-reliance on the system.

2.5.2 Technical Robustness and Safety

The toolbox embodies a security-by-design philosophy and integrates a range of technical safeguards to ensure system robustness, resilience, and fault tolerance. Risk assessment, vulnerability management, and fallback procedures are systematically implemented. The tools are tested for stability, accuracy, and reproducibility, and are deployed in secure environments, mitigating risks of unintended or adversarial impacts.

2.5.3 Privacy and Data Governance

CYLCOMED complies with GDPR principles by incorporating privacy-by-design and data minimisation measures throughout its architecture. The tools do not process personal data directly; instead, robust encryption (FE4MED), authentication (LuS4MED), and access control mechanisms (C-OPA) ensure that sensitive data from connected medical devices is securely handled. Additionally, a Data Protection Impact Assessment (DPIA) was conducted, and Data Protection Officers (DPOs) were actively engaged, reinforcing institutional compliance.

2.5.4 Transparency

The Toolbox ensures traceability, explainability, and clear communication of AI functionalities and limitations. Logging, version control, and documentation support comprehensive auditability. Outputs are interpretable through visual dashboards and technical documentation, and end-users are clearly informed about the scope, benefits, and limitations of the AI systems.

2.5.5 Diversity, Non-Discrimination, and Fairness

While the tools do not process data that could introduce social bias, the consortium implemented procedures for bias detection and prevention. Stakeholder engagement during development and pilot testing ensured alignment with diverse operational contexts. Though accessibility for persons with disabilities is not addressed at this stage due to the Toolbox's technical user base, universal design principles have guided usability considerations.

2.5.6 Societal and Environmental Well-being

The system has minimal environmental impact, owing to its lightweight computational footprint and deployment on low-energy hardware. Societal risks are negligible, as the system neither interacts with the public nor engages in decision-making affecting individuals. It instead enhances the capabilities of cybersecurity professionals and contributes positively to the resilience of healthcare infrastructures.

2.5.7 Accountability

Mechanisms for auditability, risk management, and third-party assessment are well established. The project follows an "ethics by design" approach, with legal and ethical compliance overseen by a dedicated work package and an independent ethical review. A modular risk-benefit analysis tool (developed under WP4) provides additional assurance regarding the toolbox's safety, privacy, and cybersecurity compliance, in line with relevant ISO, NIST, and MDR/IVDR standards.

2.5.8 Final Assessment

In conclusion, the CYLCOMED Toolbox exhibits strong legal and ethical compliance across all relevant dimensions. The system is not only technically sound but also designed with a deep commitment to human-centric, trustworthy, and rights-respecting AI principles. Its architecture, governance model, and operational safeguards offer a compelling example of how complex cybersecurity tools can be ethically aligned and legally compliant within critical sectors such as healthcare. These findings position the CYLCOMED Toolbox as a robust and scalable solution fit for further integration and deployment within EU healthcare environments.

Funded by Horizon Europe Framework Programme of the European Union

2.6 Identified Limitations of the ALTAI Framework

The HLEG's *Ethics Guidelines* and the ALTAI provide an important foundation for advancing the development and deployment of trustworthy AI in European Union. By translating high-level ethical principles into concrete requirements and practical guidance, they help operationalise concepts such as fairness, safety, and accountability. As such, they offer a valuable framework for AI developers to align AI innovation with fundamental rights and societal values. However, while the ALTAI offers a practical and structured framework for supporting the alignment of AI systems with ethical principles and societal values, challenges encountered in its practical application highlight areas for improvement. These will serve as the basis for recommendations aimed at informing its future refinement.

Firstly, a central scholarly critique concerns ALTAI's dependence on self-assessment, which inherently limits accountability and comparability across organisations [48][49]. Since no external auditing, verification, or evidence-weighting mechanisms are embedded into the tool, organisations may complete the checklist without demonstrating substantive compliance. Several scholars argue that this self-evaluation approach creates risks of "ethics washing" and inconsistent interpretations of what constitutes trustworthy AI [38][39]. Without external validation, the tool may remain more aspirational than enforceable.

Secondly, ALTAI has been criticised by scholars for lacking operational precision, particularly in terms of what constitutes adequate evidence, documentation, or measurable thresholds for compliance [50], which is also CYLCOMED's experience. In other words, while the ALTA checklist translates abstract ethical principles into procedural questions, it does not provide concrete metrics, benchmarks, or methodological guidance to evaluate fairness, robustness, or human oversight in practice [51]. This results in an "interpretation gap", where stakeholders might interpret requirements differently, which consequently undermining consistent approach to self-assessment.

Thirdly, while ALTAI is designed to be sector-agnostic, empirical studies show that its effectiveness varies significantly depending on the context of use. In high-risk or highly regulated domains, such as healthcare, neuro-informatics, or public-sector applications, ALTAI often requires substantial adaptation and complementary processes to capture domain-specific risks [52][53]. This limits its immediate applicability for complex, multi-stakeholder environments and highlights the need for sector-tailored guidance.

Next, ALTAI has a broad material scope. However, artificial Intelligence comprises a diverse set of technologies, ranging from statistical models to deep learning architectures, with each class of systems presenting distinct ethical, legal, and technical risks [54]. As such, developing a single assessment framework that is equally applicable across all AI applications remains inherently challenging [55]. Although

Funded by Horizon Europe Framework Programme of the European Union

ALTAI aims to provide a comprehensive ethical foundation, its generic design does not always align with the specific risk profiles of all AI systems. In practice, the relevance of its questions and recommendations varies significantly depending on the system's purpose, design, and operational environment. For example, many of the checklist items are tailored toward large-scale AI models used in high-stakes decision-making, such as those found in healthcare, where issues of privacy, discrimination, and accountability are more pronounced [56]. In contrast, the checklist offers limited flexibility for low-risk, domain-specific applications, such as those deployed in industrial settings. This was evident in the case of the CYLCOMED project. In this context, the AI systems did not process personal data, nor did they engage in decisions that could meaningfully affect individual rights or lead to discriminatory outcomes. Consequently, many of the ALTAI items, such as those addressing bias mitigation, data governance, were either irrelevant or difficult to apply meaningfully. However, the current version of ALTAI does not provide a mechanism to indicate the irrelevance of specific questions, nor does it support conditional tailoring based on the risk level, sector, or application type. This lack of adaptability can hinder the practical usability of the checklist, especially for small-scale projects or non-critical systems, and may inadvertently contribute to assessment fatigue or superficial compliance [57].

2.7 Conclusion and Policy Recommendation

The ALTAI and the Ethics Guidelines developed by the High-Level Expert Group on AI represent an essential step toward embedding ethical considerations into the design and deployment of AI systems across Europe. By articulating a structured approach to operationalising principles such as fairness, transparency, and accountability, ALTAI has contributed significantly to the discourse on trustworthy AI. However, practical experiences, such as those from the CYLCOMED project, as well as findings from empirical and academic studies, reveal several systemic limitations that undermine its full potential. The reliance on self-assessment, the absence of methodological precision, and the lack of context-sensitive flexibility point to a need for refinement and evolution of the framework. To ensure that ALTAI can serve as an effective and enforceable instrument for guiding ethical AI innovation, especially within the context of the EU AI Act, policymakers should consider targeted updates that strengthen its robustness, adaptability, and applicability across diverse domains and risk levels.

Policy Recommendations

It could be beneficial for the ALTAI framework to be complemented by third-party auditing or certification schemes. Such mechanisms would strengthen the reliability of self-assessments by introducing independent verification, thereby enhancing transparency and accountability in the implementation of ethical AI practices.

Funded by Horizon Europe Framework Programme of the European Union

It could be valuable to support the harmonisation of assessment criteria across EU Member States. Such harmonisation would promote a consistent interpretation and application of the ALTAI framework, reducing the risk of regulatory fragmentation and ensuring coherence in how trustworthiness is assessed throughout the Union.
It could be beneficial for policymakers to collaborate with standards development organisations, such as CEN/CENELEC and ISO, to define operational metrics, thresholds, and evidence types for key ALTAI dimensions. These additions would help clarify expectations around fairness, robustness, and human oversight, and provide stakeholders with clearer benchmarks to guide ethical implementation.
It could be valuable for future iterations of the ALTAI to include practical guidance toolkits and sector-specific examples. Such additions would offer developers more concrete support in interpreting the checklist and documenting compliance, especially in applied contexts where ethical considerations may manifest differently.
It could be beneficial to create sector-specific variants of the ALTAI checklist tailored to particular domains such as healthcare, public services, industrial automation, and education. These adaptations would ensure greater contextual relevance and usability, allowing AI developers and deployers to align ethical assessments more precisely with the operational and risk characteristics of their domain.

Funded by Horizon Europe Framework Programme of the European Union

3 Regulatory Frameworks and Policy Recommendations

As stated in the executive summary, the second objective of this deliverable is to highlight shortcomings in the current regulatory framework and to provide policymakers and lawmakers with options to consider improvements to the existing rules and policies. In doing so, special attention will be paid to the MDCG 2019-16 Guidance on Cybersecurity for Medical Devices [58]. Additionally, it is essential to note that Insights from legal and ethical perspectives related to the MDCG 2019-16 Guidance on Cybersecurity for Medical Devices have been provided as a contribution to the deliverable D3.4 “Guidance improvement – final report”. Therefore, this analysis tends to offer more nuanced, in-depth recommendations from a legal standpoint.

Building on this, this section undertakes a critical analysis of the regulatory challenges posed by the cybersecurity regulatory frameworks, using the MDCG Guidance as an interpretative lens. This approach enables a closer examination of how high-level regulatory provisions translate into practical obligations for manufacturers, healthcare providers, and other stakeholders. Therefore, the following section addresses the regulatory challenges arising from applicable legal frameworks, including the Medical Device Regulation (MDR) [59], Cybersecurity Act (CSA) [60], the Network and Information Systems Directive (NIS2 Directive) [61], the General Data Protection Regulation (GDPR) [62], and the Artificial Intelligence Act (AI Act) [63], all of which are relevant to ensuring the cybersecurity of medical devices. The analysis demonstrates that significant challenges persist as a result of regulatory specialization, which has produced overlaps between instruments, risks of fragmentation, legal uncertainty, and duplication. The chapter concludes with final remarks and presents recommendations aimed at assisting regulators in addressing these challenges and in fostering a more coherent and effective framework for the cybersecurity of medical devices in the European Union. Additionally, since the D2.2 [64] extensively discussed these regulatory frameworks, the following subsections will briefly provide the context and highlight regulatory challenges and recommendations for the improvements.

3.1 Medical Device Regulation (MDR)

As discussed in the previous deliverables [65], the **MDR** is the key legislative framework governing the cybersecurity and safety of medical devices in the EU. Adopted in 2017 and fully applicable since May 2021, the MDR introduces significantly more stringent requirements than the former Medical Devices Directive, harmonising rules across the Union through its direct applicability. Despite the long transitional period, full practical implementation continues to face challenges, prompting the Commission to launch a targeted evaluation in 2024 [66], well ahead of the formal review deadline of 2027, to assess effectiveness, coherence (including with digital and environmental legislation), and overall EU added value.

Funded by Horizon Europe Framework Programme of the European Union

The MDR's scope is centred on the concept of **"intended purpose"**, which remains the decisive criterion for determining whether a product qualifies as a medical device. As confirmed by CJEU case law, the manufacturer's stated intention, expressed through labelling, instructions for use, and promotional materials, continues to be the primary determinant for classification [67].

3.1.1 Cybersecurity and the MDR: Regulatory Challenges Interpreted Through the MDCG Guidance on Medical Devices

In order to ensure a high level of safety and performance, particularly for devices incorporating electronic programmable systems and for software that qualifies as a medical device in its own right, the MDR requires demonstrable compliance with the cybersecurity obligations set out in the General Safety and Performance Requirements of Annex I (Article 5(2)). To assist device manufacturers in complying with the essential requirements in Annex I of the MDR concerning cybersecurity, the European Commission's Medical Device Coordination Group endorsed the Guidance on Cybersecurity for Medical Devices (MDCG 2019-16 Rev. 1). Being the first guidance on medical device cybersecurity, it marked a significant advancement in facilitating the implementation of MDR cybersecurity provisions. However, the landscape of medical device cybersecurity is highly dynamic and has evolved considerably since the MDCG Guidance was endorsed in 2019, creating new challenges for both regulators and stakeholders in ensuring coherent and effective implementation.

Firstly, it is interesting to point out that neither MDR nor the MDCG Guidance explicitly refers to cybersecurity [68]. Namely, the MDCG Guidance does not include explicit definitions nor reference to terms such as "cybersecurity," "security-by-design," and "security-by-default." Instead, the Guidance outlines provisions related to cybersecurity for medical devices and emphasises conceptual connections between safety and security. However, leaving these terms theoretical and undefined may present challenges for stakeholders seeking to implement practical cybersecurity measures effectively. Clear and concrete definitions would enhance understanding and support the practical implementation of cybersecurity principles in the context of medical devices [69].

Policy Recommendation
To enhance regulatory clarity and facilitate consistent implementation, the European Commission and the Medical Device Coordination Group should consider complementing the **MDR** and **MDCG** Guidance with clear definitions of key concepts, such as *cybersecurity*, *security-by-design*, and *security-by-default*. These definitions should be **aligned with existing EU regulatory instruments** (e.g., the NIS2 Directive, the Cyber Resilience Act [70], and the GDPR) to ensure coherence across frameworks. Harmonised terminology would not only reduce ambiguity for manufacturers, hospitals, and notified bodies but also improve compliance.

Funded by Horizon Europe Framework Programme of the European Union

It should also be noted that, as explicitly stated at the outset, the MDCG Guidance reflects the views of the Coordination Group and is not legally binding [71]. In practice, this means that medical device manufacturers are not strictly obliged to follow the Guidance and may interpret or implement its recommendations differently. Such divergence in interpretation risks undermining the harmonisation objectives pursued by the MDR and IVDR.

Policy Recommendation
As a soft-law instrument, the MDCG Guidance could be strengthened by explicitly clarifying how its provisions on cybersecurity are to be applied in conformity assessments. Explicitly linking the assessment criteria to the guidance would reduce interpretative uncertainty and encourage more consistent adherence by medical device manufacturers, thereby supporting the harmonisation objectives of the MDR and IVDR.

The CSA, adopted in 2019, was the first EU legislative instrument to introduce an explicit definition of cybersecurity and to establish a common understanding of related concepts within the Union's regulatory framework [72]. This marked an important milestone in shaping the EU's cybersecurity policy landscape and laid the foundations for greater coherence across sectors. Notably, however, the subsequent MDCG Guidance on Cybersecurity for Medical Devices (MDCG 2019-16 Rev.1) does not make reference to the definition set out in the CSA. Establishing an explicit link between the Guidance, as soft-law instrument, and the CSA, a hard-law framework, would have several benefits. First, it would reduce terminological ambiguity and promote greater legal certainty. Second, it would strengthen coherence across the EU's broader cybersecurity regulatory architecture, which is increasingly complex with the introduction of instruments such as NIS2 Directive and the Cyber Resilience Act. Third, it would provide manufacturers with clearer interpretative tools, thereby facilitating more consistent implementation of the MDR's cybersecurity requirements in practice.

Policy Recommendation
It could be valuable for future iterations of the MDCG Guidance to consider referencing the definitions established in the CSA. Aligning the soft-law Guidance with this hard-law framework would foster greater conceptual clarity, promote coherence across EU regulations, and provide stakeholders with more consistent interpretative tools when implementing MDR cybersecurity requirements.

Funded by Horizon Europe Framework Programme of the European Union

As noted above, the medical device cybersecurity landscape is highly dynamic and has evolved significantly since the endorsement of the MDCG Guidance in 2019. While now somewhat outdated in light of more recent EU legislative developments, Figure 2 illustrates the key cybersecurity regulatory frameworks that manufacturers should be particularly aware of, as well as their interrelation with other legal frameworks. In particular, Section 6 of the MDCG Guidance provides an overview of how the MDR intersects with parallel frameworks, most notably the CSA, the GDPR, and the NIS2 Directive.

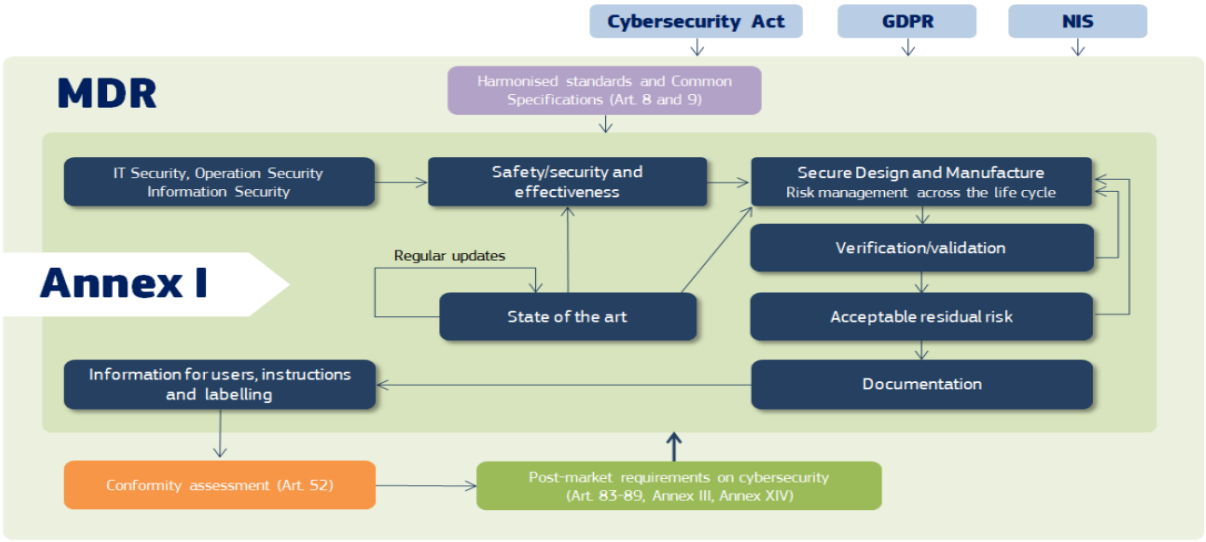

Figure 2 Cybersecurity Requirements in the MDR. Source: MDCG Guidance, p6. [73].

However, since the publication of MDCG 2019-16 in December 2019 and its subsequent revision in July 2020, the EU legislative landscape governing medical device cybersecurity has undergone substantial transformation. Key developments include the adoption of the NIS2 Directive and the Cyber Resilience Act, the adoption of the Data Act [74], new initiatives in data governance such as the European Health Data Space [75], and the emergence of regulatory instruments addressing artificial intelligence, most notably the AI Act. Despite these significant changes, MDCG 2019-16 has not been updated to reflect the evolving legislative environment.

Funded by Horizon Europe Framework Programme of the European Union

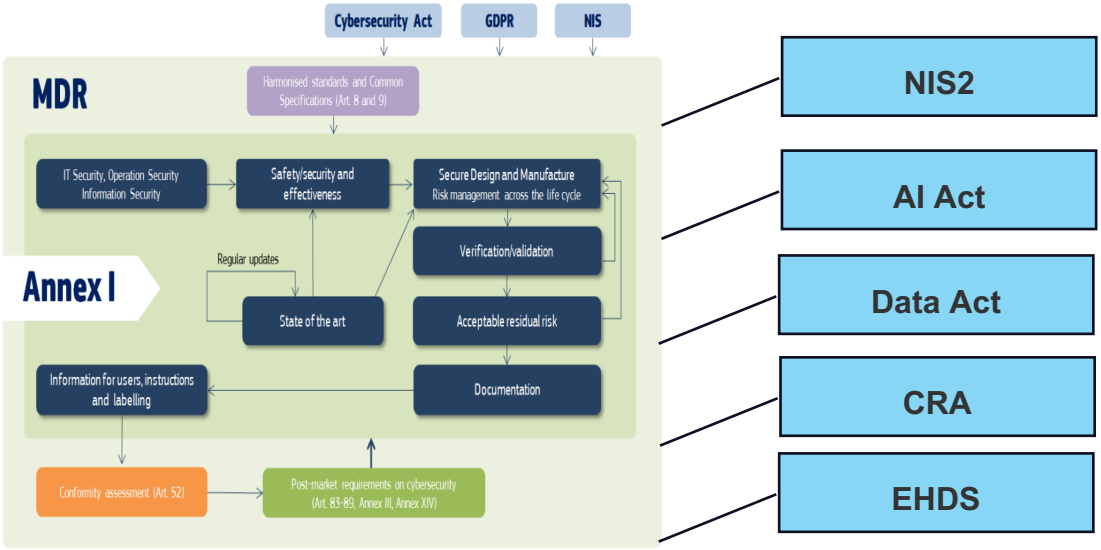

Figure 3 Evolving Regulatory Landscape and MDCG

Policy Recommendation
It could be valuable for the Guidance to be periodically updated to reflect the evolving legal and regulatory landscape, thereby ensuring its continued relevance and supporting manufacturers in addressing emerging cybersecurity requirements effectively.

Although MDCG Guidance has not been updated to reflect the evolving legal landscape, in current version it only briefly outlines the primary purposes of the legal frameworks presented in the previous figure 3. the Guidance does not provide a detailed analysis of the potential overlaps, gaps, or conflicts that may emerge in practice, nor does it establish closer links with these instruments. This is particularly relevant taking into account interplay between MDR, CSA, NIS2 and GDPR, which will be briefly outlined below.

Area where practical implementation becomes particularly challenging is the **incident reporting obligation**, especially in situations where multiple legal frameworks apply concurrently, creating regulatory overlap and legal uncertainty [76]. It has already been recognised that certain categories of cyberattacks may simultaneously trigger reporting duties under both the MDR and NIS2, thereby placing manufacturers and healthcare providers in a complex position when determining which obligations apply, to whom, and by which deadlines [77]. This regulatory uncertainty is further exacerbated by the recently adopted AI Act, which introduces its own set of transparency, monitoring, and incident-reporting obligations for high-risk AI systems. The complexity becomes even more pronounced when a cyberattack involves the exfiltration of personal data, thereby triggering additional breach-notification requirements under the GDPR, alongside the sector-specific obligations established by the MDR, NIS2, and the AI Act. To illustrate

Funded by Horizon Europe Framework Programme of the European Union

these interlocking regulatory challenges, we consider a hypothetical yet increasingly plausible scenario involving a Class IIb AI-based medical imaging system used for the early detection of lung cancer through CT-scan analysis, which is used in the manuscript academic article entitled "Cybersecurity of Connected Medical Devices: Interdisciplinary Perspectives on Ethical, Regulatory, and Practical Challenges in the EU [78]. The device is network-enabled, trained using advanced deep-learning models, and integrated into the hospital's electronic health record (EHR) system, automatically generating diagnostic reports for radiologists and clinicians. In this scenario, a sophisticated threat actor exploits a vulnerability in the device's communication protocol to gain remote access. The attacker conducts a dual-pronged operation, injecting adversarial inputs that subtly alter the internal logic of the AI model. As a result, the system misclassifies malignant lesions as benign, causing delayed interventions and potentially serious clinical consequences for multiple patients. Simultaneously, the attacker exfiltrates imaging files and associated personal identifiers, compromising sensitive health data [79].

Such an attack would almost certainly trigger a multilayered and overlapping regulatory response under EU law, engaging several legal instruments simultaneously. **First**, the manipulation of diagnostic outputs and unauthorised access to patient data would constitute a personal data breach under Article 4(12) GDPR, thereby triggering the controller's obligation to notify the supervisory authority within 72 hours (GDPR Art. 33). **Second**, as a Class IIb medical device, the system falls under the MDR's post-market surveillance regime. The erroneous diagnosis resulting from adversarial manipulation would qualify as a serious incident under Article 2(65) MDR, given the associated patient-safety risks. Consequently, the manufacturer must report the incident to the competent authority without undue delay (MDR Art. 87) [80].

Third, under the AI Act, the imaging system is categorised as a high-risk AI system used in medical diagnosis (see Section 5.5). The AI Act imposes an obligation to report serious incidents to the relevant market surveillance authorities (AI Act Art. 73). In our hypothetical scenario, if the malfunction or manipulation of the AI system directly or indirectly leads to serious harm to a person's health, this would activate reporting obligations under the AI Act as well. **Fourth**, under NIS2, hospitals are classified as essential entities (Annex I NIS2). A cyberattack of this nature, affecting device functionality, diagnostic service availability, and compromising patient data, would likely constitute a significant incident impacting the provision of essential healthcare services. Accordingly, the hospital would be required to issue an early warning, followed by an incident notification and a final report to the CSIRT or competent authority (NIS2 Art. 23). Compounding this complexity is the fact that each of these legal frameworks relies on distinct conceptualisations of key terms such as serious incident, reporting threshold, materiality, harm, and timeliness. They also differ regarding competent reporting authorities and procedural requirements, further complicating compliance for operators [81].

This hypothetical attack on an AI-enabled medical imaging device highlights fundamental tensions within the EU's regulatory architecture for cybersecurity and confirms the urgent need for enhanced legal interoperability. While a full comparative

Funded by Horizon Europe Framework Programme of the European Union

analysis of all reporting obligations lies beyond the scope of this discussion, the scenario illustrates the intricate compliance landscape faced by healthcare providers, device manufacturers, and cybersecurity professionals operating at the intersection of medical regulation, data protection, and critical-infrastructure security [82]. As a result, stakeholders such as manufacturers, healthcare providers, and notified bodies are left without clear guidance on reconciling parallel obligations, which may lead to divergent interpretations and inconsistent implementation across Member States.

Next, moving to CSA, which establishes the first EU-wide cybersecurity certification framework. Certification plays a dual role: it enhances user and consumer confidence in digital technologies and supports manufacturers and service providers by providing a harmonised, recognisable mechanism for demonstrating compliance. As emphasised in CSA Recital 69, certification is expected to contribute to higher levels of cybersecurity assurance while simultaneously enabling a more integrated market for ICT solutions, particularly in sectors where trust and reliability are critical, such as healthcare, communications, and critical infrastructure. However, despite its harmonisation ambitions, the CSA introduces an important nuance. Namely, under Article 56(2) of the CSA, cybersecurity certification remains voluntary, unless made mandatory by other EU legislation or by national laws. This voluntary nature, while it creates flexibility, also introduces regulatory uncertainty [83]. On the one hand, the CSA aims to harmonise cybersecurity certification across the EU; on the other, it implicitly permits Member States to mandate certification in specific contexts. Such national initiatives, motivated by differing risk perceptions, sectoral priorities, or political considerations, may inadvertently reintroduce the fragmentation that the CSA sought to eliminate. In other words, if certain Member States require mandatory cybersecurity certification for specific devices or services, manufacturers may need to obtain certification merely to access those markets, while other Member States impose no such obligation [84]. This scenario risks creating a patchwork of national requirements, undermining the very objective of ensuring a harmonized level of cybersecurity of network and information systems, communication networks, services, and devices within the Union [85]. For manufacturers of CE-marked devices, such as medical devices regulated under the MDR, this could result in duplicative conformity processes, forcing them to comply simultaneously with safety and performance-oriented assessments under the MDR and cybersecurity-oriented certification under national or sector-specific rules. This duplication not only increases compliance costs but also complicates market access, potentially slowing down innovation and the deployment of critical healthcare technologies [86].

The regulatory landscape becomes even more intricate with the introduction of the AI Act, which establishes its own conformity assessment, quality-management, and post-market monitoring requirements for high-risk AI systems, including those embedded in medical devices [87]. The coexistence of CSA-based cybersecurity certification and AI Act conformity assessment raises questions about regulatory interoperability, assurance level consistency, and the cumulative burden on manufacturers. Without clear alignment, manufacturers may face overlapping audits, parallel documentation

Funded by Horizon Europe Framework Programme of the European Union

requirements, and ambiguous compliance pathways, especially for AI-enabled devices that fall simultaneously under the MDR, NIS2, CSA, and the AI Act.

Policy Recommendation
Future iterations of the MDCG Guidance could consider offering more detailed clarification on how MDR cybersecurity requirements intersect with other EU legal frameworks. This could include mapping areas of potential overlap, highlighting complementary provisions, and identifying points of divergence. Such practical guidance would support stakeholders in compliance with complex regulatory landscapes while reinforcing coherence across the EU regulatory framework. This is also reinforced by many policy documents, such as those of the World Economic Forum (WEF), which state that “Regulatory provisions need well-structured guidelines to have real-world effects” [88].

Next, while the MDCG primarily aims to provide manufacturers with practical guidance on meeting the relevant General Safety and Performance Requirements (GSPRs) of the MDR regarding cybersecurity, the document also acknowledges that cybersecurity responsibilities extend beyond manufacturers [89]. It recognises that effective cybersecurity requires coordinated action by multiple stakeholders across the medical device lifecycle, including integrators, operators, healthcare professionals, and end-users. The Guidance highlights that clinicians play an essential role in configuring device parameters, integrating devices into clinical workflows, and ensuring appropriate use. Next, patients are encouraged to practise “cyber-smart” behaviour, such as being attentive to privacy and recognising potential digital threats.

This broad framing reflects an important regulatory recognition that cybersecurity is a shared responsibility. However, from a regulatory-governance perspective, the operational content of the Guidance remains more extensively developed for manufacturers, whereas the sections concerning healthcare professionals, operators, and patients are addressed at a more high-level. As a result, some stakeholders may find the Guidance less detailed in terms of practical expectations or specific measures relevant to their roles [90].

The limited attention given to human-centric aspects of cybersecurity is noteworthy given the extensive evidence documented across the human-factors literature and security-economics research that user behaviour and organisational context significantly influence cybersecurity outcomes in healthcare environments [91][92]. Human factors such as misconfigurations, insecure workarounds, time pressures, and limited cybersecurity literacy frequently contribute to vulnerabilities [93][94]. Although the MDCG acknowledges these actors as part of the cybersecurity ecosystem, it stops short of offering more operational guidance on how healthcare providers and patients

Funded by Horizon Europe Framework Programme of the European Union

can effectively fulfil these expectations, or on how organisations can support secure practices through training, workflow design, or governance structures.

In addition, the Guidance does not fully explore how the responsibilities of different stakeholders interact with obligations arising under other horizontal legislation applicable to medical devices [95]. Medical devices increasingly operate at the intersection of several regulatory frameworks, including the MDR, GDPR, NIS2, the AI Act, and the CSA. Each framework introduces distinct obligations related to cybersecurity, incident reporting, data governance, and oversight. Providing more explicit clarification on how these frameworks intersect could help stakeholders more confidently interpret their responsibilities and manage potential overlaps or uncertainties.

Considering these intersections more explicitly would enhance the Guidance's value as a practical tool and could support a more coherent cybersecurity governance model across the EU's digital regulatory landscape.

Policy Recommendation
It could be beneficial for the Guidance to include more detailed and role-specific recommendations for healthcare professionals, operators, and integrators. Such additions would better reflect their central role in the secure deployment and use of medical devices and would support these stakeholders in fulfilling their cybersecurity responsibilities in practice.
It could be valuable for the Guidance to integrate a more human-centric approach by offering targeted recommendations on user awareness, training, and operational security practices. Such additions would strengthen cyber threat prevention and mitigation efforts and enhance the overall preparedness of stakeholders who interact with medical devices in clinical and operational settings.

Funded by Horizon Europe Framework Programme of the European Union

4 Ethical Instruments in Paediatric Clinical Research: Frameworks, Implementation Challenges, and Policy Directions in the Age of Digital Health

Paediatric clinical research is an ethically sensitive domain due to the vulnerability and evolving autonomy of child participants. Despite a robust ecosystem of international instruments, such as the Declaration of Helsinki, the CIOMS Guidelines, ICH-GCP, and the Clinical Trials Regulation, significant challenges persist in their practical application, particularly as digital health tools, such as remote monitoring and AI, enter clinical settings. This section briefly reiterates the findings elaborated in D2.2 [96] regarding the ethical imperatives enshrined in international instruments, examines their practical implementation challenges in paediatric trials, and outlines key policy recommendations to enhance ethical conduct and harmonisation across jurisdictions.

4.1 Ethical Instruments Guiding Paediatric Research

Children, as a population with limited legal and decision-making capacity, require specific safeguards to protect their welfare while enabling participation in studies that may lead to crucial therapeutic advances. Over the past decades, a comprehensive body of international ethical instruments has emerged, intended to provide normative guidance and legal boundaries for researchers, institutions, and regulators. Given the large number of international documents and guidance with varying legal force, context, and focus, the true challenge lies in translating ethical principles into clinical research and establishing procedures that align with international human rights law while maintaining ethical integrity. In other words, despite these formal frameworks, the actual implementation of ethical requirements remains uneven, particularly given emerging digital tools that introduce new layers of complexity regarding data protection, consent, and autonomy [97].

4.1.1 Declaration of Helsinki

The Declaration of Helsinki (DoH) [98], first adopted by the World Medical Association in 1964 and last updated in 2013, remains one of the most influential ethical documents for biomedical research [99]. It establishes foundational principles such as voluntary informed consent, the prioritisation of participant welfare, and independent ethics review. While the DoH uses the umbrella term “vulnerable populations,” it does not provide detailed guidance on assent and dissent from minors. This generality can hinder consistent application in paediatric contexts.

Funded by Horizon Europe Framework Programme of the European Union

4.1.2 International Ethical Guidelines for Health-Related Research Involving Humans (CIOMS Guidelines)

The CIOMS International Ethical Guidelines for Health-related Research Involving Humans [100], co-developed with the World Health Organization, provide more granular guidance on research involving minors. Guideline 17 explicitly requires both the permission of a legally authorised representative and the assent of the child, commensurate with the child's capacity to understand the research. The CIOMS commentary also outlines practical expectations, such as the use of age-appropriate materials and the respect for children's dissent. These detailed provisions make CIOMS one of the more operationally useful instruments for paediatric research, but enforcement remains dependent on local implementation.

4.1.3 ICH Guideline for Good Clinical Practice (E6)

The International Council for Harmonisation's Good Clinical Practice (ICH-GCP) Guidelines [101] aim to provide a unified ethical and scientific standard for clinical trials. While ICH-GCP should be followed when generating clinical trial data intended for submission to regulatory authorities, the principles established in this guideline may also be applied to other clinical research settings, as its primary purpose is to safeguard human rights and ensure the safety and well-being of trial subjects [102]. Although ICH-GCP contains a reference to the "vulnerable subjects" by giving examples of vulnerable categories, including minors and subjects incapable of giving informed consent, it is interesting to note that ICH-GCP does not contain any specific reference to children/minors in terms of participation in a clinical research study and the possibility of giving assent. It is, therefore, silent on the extent to which children may be considered capable of giving informed consent for themselves. Moreover, it does not address the concept of assent and dissent at all. Therefore, ICH-GCP provides limited guidance on operationalising assent processes or on handling conflicts between parental consent and a child's dissent.

4.1.4 EU Clinical Trials Regulation (CTR)

The EU Clinical Trials Regulation [103] introduced several advancements in ethical oversight, including mandatory assent from minors capable of forming an opinion. The CTR mandates that information be provided in age-appropriate language and from staff trained in communicating with children. Still, CTR leaves much to the discretion of Member States, resulting in heterogeneous practices across the EU, particularly around the age at which assent becomes necessary and the legal competence of adolescents.

Funded by Horizon Europe Framework Programme of the European Union

4.2 Practical Challenges in Implementation

The concept of assent has been comprehensively addressed in Deliverable D2.2, and readers are therefore referred to that document for an in-depth examination of its ethical and legal underpinnings. In line with this, the present section does not reiterate the conceptual foundations previously established. Rather, its purpose is twofold: to outline and analyse the practical challenges associated with operationalising assent in paediatric clinical research, and to develop corresponding policy recommendations to support more consistent and effective implementation across research settings.

Firstly, it is essential to note that the concept of **assent** occupies a central, yet underdeveloped space in the ethical governance of paediatric clinical research. At its core, assent serves as an ethical mechanism intended to respect the developing autonomy of children who, while not legally competent to give full informed consent, possess sufficient maturity to form an opinion about their participation in a study [104]. However, the **operationalization of assent remains ambiguous**, both in theory and practice. The foundational difficulty in implementing assent lies in the **absence of a universally accepted definition** across major ethical frameworks delineated above. Although instruments such as the Declaration of Helsinki, ICH-GCP, CIOMS Guidelines, and the EU Clinical Trials Regulation all address the involvement of minors in the research decision-making process, none **provides an explicit definition of assent** [105]. This definitional silence is striking, given the concept's centrality in paediatric ethics. A critical examination of these documents reveals a striking omission: none of them provides a clear definition of what constitutes assent, except CTR, as illustrated in Table 1. This definitional gap has practical implications. Without a shared understanding of assent's purpose and features, there is little guidance for researchers on how to design or evaluate the assent process, and even less consistency in how ethics committees assess its adequacy.

Next, part of the complexity lies in the varying ethical rationales underpinning assent. One perspective treats assent as a developmentally adapted form of informed consent, intended to inform and protect the child through simplified explanations [106][107]. In this view, the emphasis is on ensuring that children receive information suitable to their age and can agree voluntarily to participate. A second perspective sees assent as serving a broader developmental and relational function: fostering children's autonomy and ensuring that they are respected as moral agents, even if they are not legally competent to consent [108]. Here, assent is less about procedural compliance and more about engaging children in dialogue, understanding their preferences, and creating space for meaningful participation. These divergent approaches create inconsistencies and raise concerns about equity, accountability, and procedural fairness. Children participating in similar studies may experience very different levels of involvement and respect for their views, depending on the sponsor, institution, or country where the research is conducted. Such disparities undoubtedly undermine the ethical coherence of paediatric research governance. Namely, the result is a fragmented landscape in which children's experiences vary widely across studies and jurisdictions. While this inconsistency undermines ethical coherence, it also places an

Funded by Horizon Europe Framework Programme of the European Union

undue burden on investigators to interpret ambiguous ethical guidance and may disadvantage institutions with fewer resources or less experience in paediatric research ethics.

	Declaration of Helsinki	ICH-GCP	COIMS	CTR
Requirements for Informed Consent	✓	✓	✓	**Depends on Member State**
Contains specific provisions on Minors/Children?	**No specific reference** (differentiate capable and incapable of giving consent)	**No specific reference**	✓ **(Guidelines 17)**	✓ **(Chapter V)**
Defines Minors/Children?	**Not defined**	**Not defined**	**Not defined**	✓
Do minors take part in the Informed consent process?	✓	**No specific reference**	✓	✓
Requirements for Assent?	**Must be sought**	**Not defined**	**Must be sought**	**Depends on Member State**
Age threshold for Assent?	**Not defined**	**Not defined**	**Not defined**	**Depends on Member State**
Dissent of minors able to form an opinion?	**Should be respected**	**Not defined**	**Must be respected** (unless, in exceptional circumstances, research participation is considered the best medical option for a child or adolescent)	**Should be respected**

Table 1 Assent Across Ethical Instruments

Next, although international guidelines encourage the involvement of minors in the decision-making process, there is no global consensus on age thresholds for assent. National regulations differ significantly, some countries require assent from children as young as seven, while others do not recognise assent at all. This lack of harmonisation complicates the design and approval of multinational studies and undermines the ethical principle of respecting developing autonomy.

A further complication arises in determining the **content and delivery of assent** [109]. Whereas the requirements for adult informed consent are extensively codified,

Funded by Horizon Europe Framework Programme of the European Union

detailing what information must be disclosed and how it should be communicated, ethical guidelines are largely silent on the specifics of assent. Among the main frameworks, only the **CIOMS Guidelines** explicitly call for the use of **"age-appropriate"** information, as illustrated in the table above. Yet the meaning of this term is left undefined. What is appropriate for a 6-year-old may be insufficient for a 12-year-old, and within any given age group, cognitive capacities can vary widely depending on developmental, educational, and contextual factors. Moreover, many studies involve complex biomedical or technological interventions that are challenging to explain even to adults. Communicating such concepts in a manner that children can understand, and doing so ethically and accurately, requires more than simplified language. It demands tailored approaches, educational creativity, and time, all of which are often constrained in real-world research settings.

Assent is further complicated by its intersection with **parental informed consent** and **data protection requirements** [110]. While assent and informed consent are ethical obligations tied to research participation, **they are distinct from legal consent required under the General Data Protection Regulation (GDPR)** [111].

This distinction often leads to **multiple overlapping consent processes**:

- Ethical informed consent from the parent or legal guardian
- Age-appropriate assent from the child
- Legal GDPR-compliant consent for data processing.

The administrative implications are substantial. In multi-country trials or those involving various age groups, documentation may run into **hundreds of pages**, spanning different languages, legal systems, and ethics review bodies. The effort required to manage, review, and reconcile these materials is considerable, particularly for academic or non-commercial sponsors with limited administrative support.

Furthermore, introduction of novel technologies in clinical settings poses novel risks [112]. Remote monitoring technologies (RMTs), and AI-assisted diagnostics, for instance, are reshaping paediatric research. While these technologies offer real-time data collection and personalised insights, they also amplify concerns around data privacy, algorithmic bias, and parental overreach [113]. The ethical instruments cited above were not designed with these digital realities in mind, and their current provisions often fail to address challenges such as re-consent in dynamic data environments or the opacity of AI decision-making [114].

4.3 Conclusion and Policy Recommendations

Ethical instruments such as the Declaration of Helsinki, CIOMS Guidelines, ICH-GCP, and the EU Clinical Trials Regulation provide a strong normative foundation for protecting children in research. Assent is a cornerstone of ethical paediatric research, yet its current application is undermined by definitional vagueness, practical uncertainty, and regulatory fragmentation. A more coherent, operationally grounded, and child-centred approach is urgently needed. By investing in clearer definitions, harmonised tools, and structured guidance, the research community can move beyond

Funded by Horizon Europe Framework Programme of the European Union

symbolic compliance and toward a model of paediatric research that is not only ethically sound but also genuinely participatory. Respecting the voices of children in research is not merely an obligation, it is a reflection of our collective commitment to dignity, justice, and trust in science. Additionally, rapid technological advances have outpaced the evolution of ethical guidance. Therefore, a coordinated, pragmatic approach is needed to translate these ethical imperatives into action, one that respects children's autonomy, accommodates national diversity, and addresses the ethical risks inherent in emerging digital methodologies. Hence, in order to enhance clarity and strengthen minors participation in clinical studies, policymakers should consider targeted updates that should include the following:

Policy Recommendations
It could be beneficial to introduce a harmonised definition of assent across EU and international ethical frameworks. Such a definition would clarify both the conceptual purpose and operational requirements of assent, reducing ambiguity in ethical review and implementation. Clear and shared terminology would also support more consistent practices across multinational trials and improve understanding among investigators, and ethics committees.
It could be valuable to develop guidance on age-appropriate communication standards for obtaining assent. These guidelines should include practical recommendations for tailoring content based on cognitive development stages and cultural contexts. This would help ensure that children are genuinely informed, not merely spoken to in simplified language, and that their ability to engage meaningfully with the research process is respected.
It could be beneficial to support the harmonisation of age thresholds for assent across Member States. Greater alignment in this area would reduce complexity for cross-border research and promote a fairer and more coherent ethical landscape, where children in different jurisdictions enjoy comparable standards of engagement and protection.
It could be valuable for policymakers to promote the development of integrated consent templates that align ethical assent and informed consent processes with GDPR compliance. Such integrated tools would reduce administrative burden and enhance legal clarity for researchers by consolidating multiple forms into a coherent, participant-centred consent pathway, particularly important in complex, multi-country studies.
It could be beneficial to provide practical toolkits, training modules, and age-adapted assent materials for researchers and ethics committees. These resources would support more effective and respectful engagement with paediatric participants and

Funded by Horizon Europe Framework Programme of the European Union

increase capacity for ethical oversight, especially in under-resourced settings or emerging research centres.
It could be valuable to update existing ethical instruments to better address the risks introduced by digital health technologies. This includes ensuring that provisions for assent, re-consent, and data transparency are adapted for remote monitoring, AI-supported diagnostics, and dynamic consent environments. Such updates would help ethical frameworks remain relevant and actionable in the face of rapidly evolving technological methodologies.

Funded by Horizon Europe Framework Programme of the European Union

5 CONCLUSION

This deliverable has assessed the ethical and legal dimensions of the CYLCOMED project, with particular attention to the implementation of requirements previously identified in Deliverables D2.1 and D2.2. The evaluation demonstrates that CYLCOMED has successfully integrated core ethical and legal principles into the design and development of its cybersecurity toolbox for connected medical devices, aligning with both binding regulatory instruments, such as the GDPR, MDR, and NIS2 Directive, and soft-law frameworks like the Ethics Guidelines for Trustworthy AI and MDCG 2019-16 Guidance. Across all seven ethical dimensions articulated by the High-Level Expert Group on AI, the CYLCOMED toolbox performs strongly, with special emphasis placed on human oversight, transparency, technical robustness, and data protection. The toolbox embodies a security- and ethics-by-design approach that prioritises accountability and resilience without compromising the autonomy or safety of end-users. Furthermore, partner contributions, documentation reviews, and pilot feedback confirm that relevant technical and organisational safeguards have been adopted throughout the development lifecycle.

Despite these achievements, this deliverable also identifies broader limitations in the regulatory and ethical frameworks themselves, particularly with respect to ALTAI's operational limitations, the MDCG's outdated guidance in light of emerging legislation (e.g. AI Act, Cyber Resilience Act), and challenges surrounding the implementation of ethical requirements in paediatric research, especially assent. A lack of harmonised definitions, fragmented reporting obligations, and divergent interpretations across Member States continue to pose risks to coherence, efficiency, and stakeholder confidence.

To address these challenges, the report proposes evidence-based, forward-looking policy recommendations. These include calls for clearer definitions, sector-specific ethical guidance, stronger cross-regulatory alignment, harmonised assent processes, and updated instruments that reflect technological realities such as AI, remote monitoring, and dynamic data environments. Collectively, these recommendations aim to guide European policymakers, and regulatory bodies, toward a more integrated, transparent, and ethically sound governance framework for digital medical technologies.

In sum, CYLCOMED not only meets its compliance obligations but also offers a blueprint for the trustworthy and rights-respecting development of cybersecurity tools in healthcare. Its experience highlights the urgent need for evolving policy instruments that keep pace with innovation while safeguarding public trust, patient safety, and fundamental rights in the digital age.

Funded by Horizon Europe Framework Programme of the European Union

Reference

[1] D2.1 Analysis of Ethical, Legal and Data Protection Frameworks

[2] D2.2 Examination of Ethical, Legal and Data Protection Requirements

[3] Medical Device Coordination Group. December, 2019. "Guidance on Cybersecurity for medical devices". (MDCG 2019-16 Rev.1). Available at: md_cybersecurity_en.pdf (europa.eu)

[4] Milojevic, D., & Nisevic, M. (2024, December). Navigating Cybersecurity Challenges in Healthcare: Challenges, Innovations, and EU Legal Framework for Connected Medical Devices. In International Conference on Wearables in Healthcare (pp. 159-181). Cham: Springer Nature Switzerland.

[5] D5.1 CYLCOMED toolbox prototype

[6] D5.1 CYLCOMED toolbox prototype

[7] D5.1 CYLCOMED toolbox prototype

[8] D5.1 CYLCOMED toolbox prototype

[9] COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT, THE EUROPEAN COUNCIL, THE COUNCIL, THE EUROPEAN ECONOMIC AND SOCIAL COMMITTEE AND THE COMMITTEE OF THE REGIONS Artificial Intelligence for Europe, COM (2018) 237 final.

[10] COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT, THE EUROPEAN COUNCIL, THE COUNCIL, THE EUROPEAN ECONOMIC AND SOCIAL COMMITTEE AND THE COMMITTEE OF THE REGIONS Artificial Intelligence for Europe, COM (2018) 237 final.

[11]Independent High-Level Expert Group on Artificial Intelligence, "Ethics Guidelines for Trustworthy AI", 2019, 2.

[12] Independent High-Level Expert Group on Artificial Intelligence, "Ethics Guidelines for Trustworthy AI", 2019, 12

[13] Independent High-Level Expert Group on Artificial Intelligence, "Ethics Guidelines for Trustworthy AI", 2019, 12

[14] Independent High-Level Expert Group on Artificial Intelligence, "Ethics Guidelines for Trustworthy AI", 2019, 14.

[15] D3.1 Baseline analysis, requirements and specifications

[16] D3.2 Requirements and specifications consolidation

[17] Independent High-Level Expert Group on Artificial Intelligence, "Ethics Guidelines for Trustworthy AI", 2019, 15.

[18] Independent High-Level Expert Group on Artificial Intelligence, "The Assessment List For Trustworthy Artificial Intelligence (ALTAI): for self-assessment.", 2020, 7.

[19] Independent High-Level Expert Group on Artificial Intelligence, "The Assessment List For Trustworthy Artificial Intelligence (ALTAI): for

self-assessment.", 2020, 8.

[20] Independent High-Level Expert Group on Artificial Intelligence, "The Assessment List For Trustworthy Artificial Intelligence (ALTAI): for

self-assessment.", 2020, 9.

[21] Independent High-Level Expert Group on Artificial Intelligence, "Ethics Guidelines for Trustworthy AI", 2019, 17.

[22] Independent High-Level Expert Group on Artificial Intelligence, "Ethics Guidelines for Trustworthy AI", 2019, 17.

[23] Independent High-Level Expert Group on Artificial Intelligence, "Ethics Guidelines for Trustworthy AI", 2019, 17.

[24] Independent High-Level Expert Group on Artificial Intelligence, "Ethics Guidelines for Trustworthy AI", 2019, 17.

[25] González Fuster, G. (2014). Privacy and the Protection of Personal Data Avant la Lettre . In: The Emergence of Personal Data Protection as a Fundamental Right of the EU. Law, Governance and Technology Series(), vol 16. Springer, Cham.

[26] European Data Protection Board, "Guidelines 4/2019 on Article 25 Data Protection by Design and by Default", 20 October 2020.

[27] Recital 78, GDPR

[28] Article 25(1), GDPR.

Funded by Horizon Europe Framework Programme of the European Union

[29] European Data Protection Board, "Guidelines 4/2019 on Article 25 Data Protection by Design and by Default", 20 October 2020.

[30] D5.3 Final CYLCOMED final toolbox package

[31] WP29, 'Guidelines on Data Protection Officers ('DPOs')', 16/EN WP 243 rev.01, 5 April 2017, p. 5.

[32] D6.2 Pilots initial phase report

[33] D5.3 Final CYLCOMED final toolbox package

[34] D5.3 Final CYLCOMED final toolbox package

[35] European Data Protection Board, "Guidelines 4/2019 on Article 25 Data Protection by Design and by Default", 20 October 2020.

[36] Article 5(1)(c) GDPR.

[37] European Data Protection Board, "Guidelines 4/2019 on Article 25 Data Protection by Design and by Default", 20 October 2020.

[38] D5.3 Final CYLCOMED final toolbox package

[39] WP29 Guidelines on transparency under Regulation 2016/679. Available at file:///C:/Users/u0161701/Downloads/20180413_article_29_wp_transparency_guidelines_7B894B16-B8B9-B044-ED400A6DBAA4FA60_51025%20(4).pdf

[40] Independent High-Level Expert Group on Artificial Intelligence, "Ethics Guidelines for Trustworthy AI", 2019, 17-18; Independent High-Level Expert Group on Artificial Intelligence, "The Assessment List For Trustworthy Artificial Intelligence (ALTAI): for self-assessment.",2020, 14-15.

[41] Independent High-Level Expert Group on Artificial Intelligence, "Ethics Guidelines for Trustworthy AI", 2019, 18-19; Independent High-Level Expert Group on Artificial Intelligence, "The Assessment List For Trustworthy Artificial Intelligence (ALTAI): for self-assessment.", 2020,16-18.

[42] D6.3 Pilots final phase report

[43] Independent High-Level Expert Group on Artificial Intelligence, "The Assessment List For Trustworthy Artificial Intelligence (ALTAI): for self-assessment.", 2020, 19-20.

[44] Papagiannidis, E., Mikalef, P., & Conboy, K. (2025). Responsible artificial intelligence governance: A review and research framework. *The Journal of Strategic Information Systems*, *34*(2), 101885.

[45] Independent High-Level Expert Group on Artificial Intelligence, "The Assessment List For Trustworthy Artificial Intelligence (ALTAI): for self-assessment.", 2020, 21.

[46] Brey, P., & Dainow, B. (2024). Ethics by design for artificial intelligence. AI and Ethics, 4(4), 1265-1277.

[47] D4.3 "Risk Benefit Scheme and Risk Management Tool"

[48] Radclyffe, C., Ribeiro, M., & Wortham, R. H. (2023). The assessment list for trustworthy artificial intelligence: A review and recommendations. Frontiers in artificial intelligence, 6, 1020592.

[49] Zicari, R. V., Amann, J., Bruneault, F., Coffee, M., Düdder, B., Hickman, E., ... & Wurth, R. (2022). How to assess trustworthy AI in practice. arXiv preprint arXiv:2206.09887.

[50] Vetter, D., Amann, J., Bruneault, F., Coffee, M., Düdder, B., Gallucci, A., ... & Z-Inspection® initiative (2022). (2023). Lessons learned from assessing trustworthy AI in practice. Digital Society, 2(3), 35.

[51] Vetter, D., Amann, J., Bruneault, F., Coffee, M., Düdder, B., Gallucci, A., ... & Z-Inspection® initiative (2022). (2023). Lessons learned from assessing trustworthy AI in practice. Digital Society, 2(3), 35.

[52] Zicari, R. V., Amann, J., Bruneault, F., Coffee, M., Düdder, B., Hickman, E., ... & Wurth, R. (2022). How to assess trustworthy AI in practice. arXiv preprint arXiv:2206.09887.

[53] Vetter, D., Amann, J., Bruneault, F., Coffee, M., Düdder, B., Gallucci, A., ... & Z-Inspection® initiative (2022). (2023). Lessons learned from assessing trustworthy AI in practice. Digital Society, 2(3), 35.

[54] Mantelero, A. (2022). *Beyond data: Human rights, ethical and social impact assessment in AI* (p. 200). Springer Nature.

[55] Floridi, L., Holweg, M., Taddeo, M., Amaya, J., Mökander, J., & Wen, Y. (2022). CapAI-A procedure for conducting conformity assessment of AI systems in line with the EU artificial intelligence act. Available at SSRN 4064091.

[56] Fedele, A., Punzi, C., & Tramacere, S. (2024). The ALTAI checklist as a tool to assess ethical and legal implications for a trustworthy AI development in education. Computer Law & Security Review, 53, 105986.

[57] Radclyffe, C., Ribeiro, M., & Wortham, R. H. (2023). The assessment list for trustworthy artificial intelligence: A review and recommendations. Frontiers in artificial intelligence, 6, 1020592.

[58] Medical Device Coordination Group. (2019). *MDCG 2019-16: Guidance on cybersecurity for medical devices.* European Commission.

[59] European Parliament, & Council of the European Union. (2017). *Regulation (EU) 2017/745 of the European Parliament and of the Council of 5 April 2017 on medical devices, amending Directive 2001/83/EC, Regulation (EC) No 178/2002 and Regulation (EC) No 1223/2009, and repealing Council Directives 90/385/EEC and 93/42/EEC. Official Journal of the European Union*, L 117, 1–175.

[60] European Parliament, & Council of the European Union. (2019). *Regulation (EU) 2019/881 of the European Parliament and of the Council of 17 April 2019 on ENISA (the European Union Agency for Cybersecurity) and on information and communications technology cybersecurity certification and repealing Regulation (EU) No 526/2013 (Cybersecurity Act). Official Journal of the European Union*, L 151, 15–69.

[61] European Parliament, & Council of the European Union. (2022). *Directive (EU) 2022/2555 of 14 December 2022 on measures for a high common level of cybersecurity across the Union, amending Regulation (EU) No 910/2014 and Directive (EU) 2018/1972, and repealing Directive (EU) 2016/1148 (NIS2 Directive). Official Journal of the European Union*, L 333, 80–152.

[62] European Parliament, & Council of the European Union. (2016). *Regulation (EU) 2016/679 of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation). Official Journal of the European Union*, L 119, 1–88.

[63] European Parliament, & Council of the European Union. (2024). *Regulation (EU) 2024/1689 of 12 July 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and Directives 2014/90/EU and (EU) 2016/797 (Artificial Intelligence Act). Official Journal of the European Union*, L 231, 1–139.

[64] D2.1 Analysis of Ethical, Legal and Data Protection Frameworks; D2.2 Examination of Ethical, Legal and Data Protection Requirements

[65] D2.1 Analysis of Ethical, Legal and Data Protection Frameworks; D2.2 Examination of Ethical, Legal and Data Protection Requirements

[66] European Commission: "Call for evidence for an evaluation - Ares(2024)8905137", available at EU rules on medical devices and in vitro diagnostics – targeted evaluation

[67] Minssen, T., Mimler, M., & Mak, V. (2020). When does stand-alone software qualify as a medical device in the European Union?—The CJEU's decision in Snitem and what it implies for the next generation of medical devices. Medical Law Review, 28(3), 615-624.

[68] Biasin, E., & Kamenjasevic, E. (2022). Cybersecurity of Medical Devices: Regulatory Challenges in the European Union. In I. Cohen, T. Minssen, W. Price II, C. Robertson, & C. Shachar (Eds.), The Future of Medical Device Regulation: Innovation and Protection (pp. 51-62). Cambridge: Cambridge University Press. doi:10.1017/9781108975452.005

[69] Biasin, Elisabetta and Kamenjasevic, Erik, Cybersecurity of Medical Devices: Regulatory Challenges in the EU (September 30, 2020). The Future of Medical Device Regulation: Innovation and Protection, Cambridge University Press, 2020, Available at SSRN: https://ssrn.com/abstract=3855491

[70] European Parliament, & Council of the European Union. (2024). *Regulation (EU) 2024/... of 2024 on horizontal cybersecurity requirements for products with digital elements (Cyber Resilience Act). Official Journal of the European Union.*

[71] Biasin, Elisabetta and Kamenjasevic, Erik, Cybersecurity of Medical Devices: Regulatory Challenges in the EU (September 30, 2020). The Future of Medical Device Regulation: Innovation and Protection, Cambridge University Press, 2020, Available at SSRN: https://ssrn.com/abstract=3855491

[72] European Parliament, & Council of the European Union. (2019). *Regulation (EU) 2019/881 of the European Parliament and of the Council of 17 April 2019 on ENISA (the European Union Agency for Cybersecurity) and on information and communications technology cybersecurity certification and repealing Regulation (EU) No 526/2013 (Cybersecurity Act). Official Journal of the European Union*, L 151, 15–69.

[73] Biasin, Elisabetta and Kamenjasevic, Erik, Cybersecurity of Medical Devices: Regulatory Challenges in the EU (September 30, 2020). The Future of Medical Device Regulation: Innovation and Protection, Cambridge University Press, 2020, Available at SSRN: https://ssrn.com/abstract=3855491.

[74] European Parliament, & Council of the European Union. (2023). Regulation (EU) 2023/2854 of the European Parliament and of the Council of 13 December 2023 on harmonised rules on fair access to and use of data (Data Act). Official Journal of the European Union, L 2023/2854, 1–91

[75] European Parliament & Council of the European Union. (2025, March 5). Regulation (EU) 2025/327 of 11 February 2025 on the European Health Data Space and amending Directive 2011/24/EU and Regulation (EU) 2024/2847. Official Journal of the European Union, L 63, 1–38.

Funded by Horizon Europe Framework Programme of the European Union

[76] Maja Nisevic, Dusko Milojevic, João Rodrigues and Goncalo Cadete. 2025. “Connected Medical Devices: The Cybersecurity Nexus of the AI Act and MDR” in Holistic IoT Security, Privacy and Safety: Integrated Approaches Protecting A Highly Connected World. Edited by Konstantinos Loupos. pp. 333–362. Now Publishers. DOI: 10.1561/9781638285076.ch15.

[77] Biasin, E., & Kamenjasevic, E. (2020). Cybersecurity of medical devices: regulatory chal-lenges in the EU. https://dx.doi.org/10.2139/ssrn.3855491

[78] Milojevic, D., Nisevic, M., & De Bruyne, J. (in press). Cybersecurity of connected medical devices: Interdisciplinary perspectives on ethical, regulatory, and practical challenges in the EU. Journal of Cybersecurity. Manuscript ID CyberSec-2025-190.

[79] Milojevic, D., Nisevic, M., & De Bruyne, J. (in press). Cybersecurity of connected medical devices: Interdisciplinary perspectives on ethical, regulatory, and practical challenges in the EU. Journal of Cybersecurity. Manuscript ID CyberSec-2025-190.

[80] Milojevic, D., Nisevic, M., & De Bruyne, J. (in press). Cybersecurity of connected medical devices: Interdisciplinary perspectives on ethical, regulatory, and practical challenges in the EU. Journal of Cybersecurity. Manuscript ID CyberSec-2025-190.

[81] Milojevic, D., Nisevic, M., & De Bruyne, J. (in press). Cybersecurity of connected medical devices: Interdisciplinary perspectives on ethical, regulatory, and practical challenges in the EU. Journal of Cybersecurity. Manuscript ID CyberSec-2025-190.

[82] Milojevic, D., Nisevic, M., & De Bruyne, J. (in press). Cybersecurity of connected medical devices: Interdisciplinary perspectives on ethical, regulatory, and practical challenges in the EU. Journal of Cybersecurity. Manuscript ID CyberSec-2025-190.

[83] Milojevic, D., & Nisevic, M. (2024, December). Navigating Cybersecurity Challenges in Healthcare: Challenges, Innovations, and EU Legal Framework for Connected Medical Devices. In *International Conference on Wearables in Healthcare* (pp. 159-181). Cham: Springer Nature Switzerland.

[84] Milojevic, D., & Nisevic, M. (2024, December). Navigating Cybersecurity Challenges in Healthcare: Challenges, Innovations, and EU Legal Framework for Connected Medical Devices. In *International Conference on Wearables in Healthcare* (pp. 159-181). Cham: Springer Nature Switzerland.

[85] Milojevic, D., & Nisevic, M. (2024, December). Navigating Cybersecurity Challenges in Healthcare: Challenges, Innovations, and EU Legal Framework for Connected Medical Devices. In International Conference on Wearables in Healthcare (pp. 159-181). Cham: Springer Nature Switzerland.

[86] Milojevic, D., & Nisevic, M. (2024, December). Navigating Cybersecurity Challenges in Healthcare: Challenges, Innovations, and EU Legal Framework for Connected Medical Devices. In International Conference on Wearables in Healthcare (pp. 159-181). Cham: Springer Nature Switzerland.

[87] Maja Nisevic, Dusko Milojevic, João Rodrigues and Goncalo Cadete. 2025. “Connected Medical Devices: The Cybersecurity Nexus of the AI Act and MDR” in Holistic IoT Security, Privacy and Safety: Integrated Approaches Protecting A Highly Connected World. Edited by Konstantinos Loupos. pp. 333–362. Now Publishers. DOI: 10.1561/9781638285076.ch15.

[88] World Economic Forum: “Earning Trust for AI in Health: A Collaborative Path Forward”, June 2025.

[89] Androutsos, C., Taylor, S., Bernsmed, K., Skytterholm, A. N., Epiphaniou, G., Moukafih, N., ... & Fotiadis, D. I. (2025, June). MDCG 2019-16 Guidelines: Case Study-Based Assessment and Path Forward. In *European Interdisciplinary Cybersecurity Conference* (pp. 338-355). Cham: Springer Nature Switzerland.

[90] Androutsos, C., Taylor, S., Bernsmed, K., Skytterholm, A. N., Epiphaniou, G., Moukafih, N., ... & Fotiadis, D. I. (2025, June). MDCG 2019-16 Guidelines: Case Study-Based Assessment and Path Forward. In *European Interdisciplinary Cybersecurity Conference* (pp. 338-355). Cham: Springer Nature Switzerland.

[91] Evans, M., He, Y., Maglaras, L., & Janicke, H. (2018). HEART-IS: A novel technique for evaluating human error-related information security incidents. *Computers & Security*, *80*, 74-89. https://doi.org/10.1016/j.cose.2018.09.002.

[92] Klimburg-Witjes, N., & Wentland, A. (2021). Hacking humans? Social Engineering and the construction of the “deficient user” in cybersecurity discourses. Science, Technology, & Human Values, 46(6), 1316-1339.

[93] Alohali, M., Clarke, N., Furnell, S., & Albakri, S. (2017). Information security behavior: Recognizing the influencers. In 2017 Computing Conference (pp. 844-853). IEEE.

[94] Jeong, J., Mihelcic, J., Oliver, G., & Rudolph, C. (2019). Towards an improved understanding of human factors in cybersecurity. In 2019 IEEE 5th International Conference on Collaboration and Internet Computing (CIC) (pp. 338-345). IEEE.

[95] Biasin, E., & Kamenjasevic, E. (2020). Cybersecurity of medical devices: regulatory chal-lenges in the EU. https://dx.doi.org/10.2139/ssrn.3855491

[96] D2.2 Examination of Ethical, Legal and Data Protection Requirements

[97] Martinez-Martin, Nicole, Zelun Luo, Amit Kaushal, Ehsan Adeli, Albert Haque, Sara S. Kelly, Sarah Wieten et al. "Ethical issues in using ambient intelligence in health-care settings." The lancet digital health 3, no. 2 (2021): e115-e123.

[98] World Medical Association Declaration of Helsinki – Ethical Principles for Medical Research Involving Human Subjects. Available at: WMA Declaration of Helsinki – Ethical Principles for Medical Research Involving Human Subjects – WMA – The World Medical Association

[99] Crawley, F. P. "The limits of the declaration of Helsinki." Address to Scientific Session. World Medical Association General Assembly: Helsinki (2012).

[100] Council for International Organizations of Medical Sciences. "International ethical guidelines for health-related research involving humans." (2016). Available at: 2016 International ethical guidelines for health-related research involving humans - CIOMS

[101] Available at https://www.ema.europa.eu/en/documents/scientific-guideline/ich-guideline-good-clinical-practice-e6r2-step-5_en.pdf

[102] Switula, Dorota. "Principles of good clinical practice (GCP) in clinical research." *Science and engineering ethics* 6, no. 1 (2000): 71-77.

[103] European Parliament and Council. (2014). Regulation (EU) No 536/2014 of the European Parliament and of the Council of 16 April 2014 on clinical trials on medicinal products for human use, and repealing Directive 2001/20/EC.

[104] Bioethics, Nuffield Council. "Children and clinical research: ethical issues." *London: Nuffield Council on Bioethics* (2015).

[105] Spriggs, M. (2023). Children and bioethics: clarifying consent and assent in medical and research settings. *British medical bulletin*, *145*(1), 110-119.

[106] Poston, R. D. (2016). Assent described: Exploring perspectives from the inside. *Journal of pediatric nursing*, *31*(6), e353-e365.

[107] Koonrungsesomboon, N., Charoenkwan, P., Natesirinilkul, R., Fanhchaksai, K., Sakuludomkan, W., & Morakote, N. (2022). What information and the extent of information to be provided in an informed assent/consent form of pediatric drug trials. *BMC Medical Ethics*, *23*(1), 113.

[108] Lepola, P., Needham, A., Mendum, J., Sallabank, P., Neubauer, D., & de Wildt, S. (2016). Informed consent for paediatric clinical trials in Europe. *Archives of disease in childhood*, *101*(11), 1017-1025.

[109] Bioethics, Nuffield Council. "Children and clinical research: ethical issues." *London: Nuffield Council on Bioethics* (2015).

[110] The European Data Protection Supervisor (EDPS). January, 2020. "Preliminary Opinion on Data Protection and Scientific Research".

[111] EDPB. February, 2021. "Document on response to the request from the European Commission for clarifications on the consistent application of the GDPR, focusing on health research".

[112] Maeckelberghe, E., Zdunek, K., Marceglia, S., Farsides, B., & Rigby, M. (2023). The ethical challenges of personalized digital health. *Frontiers in medicine*, *10*, 1123863.

[113] Vayena, E., Haeusermann, T., Adjekum, A., & Blasimme, A. (2018). Digital health: meeting the ethical and policy challenges. *Swiss medical weekly*, *148*, w14571.

[114] Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., ... & Vayena, E. (2018). AI4People—An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. *Minds and machines*, *28*(4), 689-707.

Funded by Horizon Europe Framework Programme of the European Union