The EU Artificial Intelligence Act Proposal - Technical Requirements and Viable Solutions for High-Risk AI Systems

Philipp Bohlen, Artur Dox, David Drexlin, Dimitrije Kovacic, Nicolai Pietrzyk, Lennart Schulze
Baden-Wuerttemberg Cooperative State University (DHBW)
Stuttgart, Germany

***Abstract*— Artificial Intelligence (AI), despite its powerful capabilities, poses severe risks to its users when employed in productive context. In response, industry, science, and politics have issued non-binding recommendations for trustworthy AI. In April 2021, the European Commission published the first-ever proposal for a binding regulation of AI systems and their stakeholders with the so-called AI Act. To ease understanding of and compliance with the technical obligations set out therein for providers of AI systems, the following contributions are made: First, formal software requirements are extracted from the proposal in a legal requirements engineering process. Second, available software solutions that assist in fulfilling the requirements are systematically identified. Third, the extent of their support it evaluated through a technical review. In total, 95 requirements were established in eight categories, for which 36 software solutions were identified. The overall requirement fulfillment support score returned low, indicating the need for adapted solutions and manual adoption efforts for developers to achieve compliance with the current version of the act. Issues to address in a revision of the proposal are presented.**

***Index Terms*—Artificial Intelligence, Trustworthiness, European Union, Regulation, Requirements Engineering, Software Solutions, Technical Review.**

I. Introduction

With affordability of data storage and the level of computational power dramatically increasing, Artificial Intelligence has been on the rise for over a decade. Providing the ability to relieve human beings from arduous work, to create new-of-a-kind insights and values, and to solve challenges previously intractable, this technology has grown to great importance in science, economy, and society.

However, Artificial Intelligence has been attributed to pose severe risks from a security, privacy, legal and ethics standpoint. These appear when personal data is processed erroneously, seminal decisions are performed autonomously, a system's behavior cannot be understood by a human being, or outputs were flawed by bias. Especially in sensitive areas such as critical infrastructure, health, public services and administration, law and justice, employment, education, and product safety these threats are severe.

Therefore, technical bodies, think tanks, intergovernmental organizations, and corporate entities have independently issued their recommendation on the development and use of AI technology to achieve trustworthiness. Since 2016, several countries from Europe, America and Asia have been publishing their own Ethical Guidelines on Artificial Intelligence to regulate the technology. As of September 2019, a total of 84 guidelines was published, which could be utilized as guidance [1].

Now, within the scope of the European Commission's theme of A Europe fit for the digital age, the European Union leaps one step ahead by establishing the first-ever binding regulation for risky AI worldwide. The long-awaited proposal from April 2021 aims at defining the rights, obligations, and constraints of the various stakeholders of AI systems in its member states [2].

A. The Artificial Intelligence Act Proposal

The scope of the act is restricted to productive AI systems, explicitly excluding military applications and research. For AI systems on the market or in use by organizations, a risk-based category system is introduced, according to which a system can either pose a prohibitive amount of risk, high risk, limited risk, or no risk. Depending on the assigned category, different measures are foreseen, with the first category being entirely banned and the last not falling subject to any restriction.

In line with this scheme, the proposal, after setting the context and legal basis in the introduction, is divided into 12 titles. Prohibitive AI systems are concerned in title II. For high-risk AI systems, an extensive list of obligations for the technical implementation of these systems, their providers, users, and other parties involved is set out in title III. The remaining titles contain general provisions and definitions of key terms used throughout the act (I), transparency obligations (IV), measures to support AI innovation (V), governance and enforcement mechanisms (VI-X) as well as final provisions (XI) and remarks (XII). Additionally, the annexes to the proposal provide further details referred to throughout the legislation.

This research specifically concerns the requirements set out for high-risk AI systems described in Title III, Chapter 2. This class of AI systems may be considered the dominant objective of the regulation, as they pose severe risks to fundamental rights, the safety of natural persons and the protection of their personal data, while equally offering large benefits [3].

In this regard, the proposal comprehensively addresses the important aspects and relevant stakeholders in the development and use of safe AI. To effectively enforce their compliance with these obligations, the current version of the act imposes

a cap of 30 million in penalties, or 6% of yearly turnover for commercial AI providers. Fines of this magnitude have already proven to give strong emphasis to new legal requirements at the time the GDPR was introduced [2].

To comply, a high-risk AI system claimed to adhere by the regulation's obligation must be assessed by an independent authority and registered in an EU-wide database before release into production as well as after every major revision.

Due to the legislative focus of a regulation, however, the proposal is limited in technical details, not clearly laying out the technical requirements nor proposing corresponding solutions for providers of AI systems. In addition, by the mere extend of the regulation, companies and the technical community may lack the resources and capacity to review the entire legal text to identify the implications it has for their AI systems.

To account for this, the proposal itself remarks that specific technical specifications or standards will be required to verify conformity in the future [2]. Their complex agreement process, nonetheless, leaves high-risk AI system providers with uncertainty about the current proposal at hand.

B. Research Objective

This paper aims to analyze the legal requirements for high-risk AI systems set out by the EU proposal for an Artificial Intelligence Act, to identify and to evaluate suitable software solutions to achieve compliance with those requirements.

The overarching objective is to contribute to the joint effort of elaborating a more detailed set of technical requirements along with corresponding software solutions that support providers of AI systems in adapting their systems to comply with the future regulation.

Consequently, to achieve the outlined objectives, the following research questions (RQ) will be answered in this paper:

1) What is the impact of legal requirements established in the EU Artificial Intelligence Act Proposal for the technical implementation of high-risk AI systems?
2) Which currently available software solutions are apt to support the satisfaction of the respective technical requirements in a high-risk AI system?
3) To what extent do these solutions support compliance with the technical requirements, which gaps remain, and what recommendations can be drawn for providers of high-risk AI systems with regard to their employment?

Finally, it is not intended to systematically evaluate the proposal and its quality nor to compare the results with other publications. Neither it is aimed to develop a sample AI system complying with the act. Instead, the findings shall serve as reference asset to the technical community.

II. Background

To position the AI Act Proposal's contribution in the vast field of AI governance, a specification of terms and summary of related publications will be provided first.

A. Terminology

Machine Learning (ML): In a broader sense, ML refers to a computer program that can learn to behave in a way that is not explicitly programmed by the author of the program [4]. In a narrower sense, ML can be defined as computational methods that detect patterns in data and use this information to make accurate predictions [5].

Explainability: In the context of AI, the purpose of eXplainable AI (XAI) is to explain the outputs from AI systems, rendering them more comprehensible to human beings and thus, rendering complex algorithms more transparent [6].

Transparency: An AI model is transparent if it is inherently understandable to human beings on its own [7]. It additionally refers to the need to describe and reproduce the procedures through which an AI system produces a decision [6], which is similar to the aim of explainability.

Interpretability: Interpretability is closely related to Explainability and is defined as the ability to provide explanations that are understandable to humans. In the ML community, interpretability is used more often than explainability [6]. These three terms are customarily used interchangeably [8] [9]. Since explainability, in an academic sense, is defined more concisely, henceforth the term explainability will be used to group the three.

Fairness: Fairness is one of the goals of XAI. An explainable ML model shows how the input leads to a certain results and thus, allows for an analysis of fairness of the given model [10][11]. Explainability can help to avoid an unfair usage of a ML model's output [6].

Trustworthiness: Trustworthiness is regarded as the main purpose of XAI [12][13]. It is considered a confidence measure of whether a ML model will act as expected on a given task. A model that behaves as expected is trustworthy. However, a trustworthy model does not necessarily imply that it can be explained on its own [7].

The EU sets trustworthiness as overarching objective for productively used high-risk AI systems [14]. Therefore, in line with related taxonomies [1], trustworthy AI, henceforth, is used to subsume the terms explainability, safe, robust, and fair AI.

AI System: Software that is developed with one or more of the following techniques: 1) ML and Deep Learning approaches including supervised, unsupervised and reinforcement learning; 2) logic- and knowledge-based approaches and (symbolic) reasoning including expert systems; 3) statistical approaches including Bayesian estimation, search and optimization methods [2].

High-Risk AI System: 1) An AI system that belongs to one of the following areas: biometric identification and categorisation of natural persons, management and operation of critical infrastructure, education and vocational training, employment, workers management and access to self-employment, access to and enjoyment of essential private services and public services and benefits, law enforcement, migration, asylum and border control management, administration of justice and democratic processes. Further details regarding these areas can be found in Annex 3 of the AI Act Proposal. 2) An AI system that is used as a safety component of a product or is itself a product and is

required to go through a conformity assessment with the intent to be put on the market, as covered by the Union harmonisation legislation listed in Annex 3 of the AI Act Proposal [15].

Supplementary definitions are provided in title I of the AI Act Proposal and in literature [1].

B. Overview over technical recommendations on trustworthy AI

In the following, a brief overview is provided on existing standards or technical recommendations for trustworthy AI. In 2017, IEEE and the IEEE Standards Association (IEEE SA) published the second version of their seminal document "Ethically Aligned Design: A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems" [16] which provides insights and recommendations, both technical and legal, for the design, development and implementation of ethical autonomous and intelligent systems (A/IS). It was created with input from multiple committees from the IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems: Ethically Aligned Design. The document provides the following recommendations for implementation: 1) Well-being metrics: Contrary to standard economical metrics, well-being metrics include psychological, social, economic fairness and environmental factors and A/IS should be tested according to these metrics to measure their impact on human well-being. 2) Embedding Values into Autonomous and Intelligent Systems: Norms of the community in which a system is intended to be used in should be embedded in the system itself. 3) Methods to Guide Ethical Research and Design: Developers should use value-based design methods to create sustainable systems. 4) Affective Computing: A/IS that are used in the context of human society should not cause harm by misusing human emotional experience. In total, Ethically Aligned Design provides very high-level recommendations that are rather visionary than practical hands-on advice for developers [16].

The International Organization for Standardization (ISO) currently develops ISO/IEC JTC 1 /SC 42 [17], a standardization for AI. It is part of the standards development environment ISO/IEC JTC 1 on Information Technology [18] and its purpose is to provide guidance to committees from the International Electrotechnical Commission (IEC) and ISO that develop AI applications. The standard covers several aspects, ranging from functional safety and AI systems, bias in AI systems and AI-aided decision-making, and assessment of the robustness of neural networks to quality evaluation guidelines for AI systems. While these standards appear promising and are more detailed than Ethically Aligned Design, they are currently still under development.

C. Overview over regulatory recommendations on trustworthy AI

Overall, there do not exist many standards with technical recommendations for trustworthy AI and more work has been done on regulatory recommendations, which will be discussed in the following. In September 2020, the United Nations Educational, Scientific and Cultural Organization (UNESCO) published a first draft of the Recommendation on the Ethics of AI [19], aimed at providing values and principles on how AI systems should work for the good of humanity, individuals and the environment, and to prevent harm. It also provides policy recommendations, emphasising on gender equality and environment protection.

In February 2020, the Pontifical Academy for Life from the Roman Catholic Church, Microsoft, IBM, the Food and Agriculture Organization of the United Nations (FAO) and the Italian Ministry of Innovation jointly signed a document titled "Call for an AI Ethics" [20], in which they outline six principles to promote ethical AI. These principles include transparency, inclusion, responsibility, impartiality, reliability and security and privacy.

The International Telecommunication Union (ITU) build a digital platform called "AI for Good" with the aim to promote the United Nations Sustainable Development Goals (SDGs) and serve as the United Nations (UN) platform on AI [21]. UN Global Pulse is the initiative of the UN Secretary-General on Big Data and AI for sustainable development, humanity and peace, with the objective to support the development and implementation of Big Data and AI ideas for the public good [22].

The 2020-2021 World Economic Forum (WEF) Global Future Council on Artificial Intelligence for Humanity is currently working on identifying technical solutions to address issues of AI fairness to be able to consult policy makers and organizations [23].

The Organization for Economic Co-operation and Development (OECD) Principles promote that AI shall be innovative, trustworthy, and respecting human rights and democratic values. They were adopted by the OECD member countries in May 2019 [24].

Finally, the purpose of the Ad-hoc Committee on AI (CAHAI) from the Council of Europe (CoE) is to examine feasibility and potential aspects of a legal framework for the development and application of AI, with regard to the standards from the Council of Europe on human rights, democracy and rule of law [25].

The AI Act Proposal examined in this paper has its origins in the establishment of a High-Level Expert Group on AI (HLEG), which consisted of 52 experts in the field, with the aim to advise the European Commission on the implementation of their strategy on AI [2].

In conclusion, while there exist many initiatives on providing recommendations and regulation for AI in similar fields of concern, a unified, binding instrument has been missing. Few provide practical recommendations that can be directly applied by organizations to comply with proposed regulation. Therefore, reaffirming the seminality of the AI Act Proposal, there exists a clear need to accompany it with recommendations for technical solutions.

III. Methodology

Subsequently, the research methodology is outlined in a global perspective followed by the detailed specification of each step contained therein.

A. Overview

While an overall objective is pursued of delivering recommendations to high-risk AI system providers regarding the choice of technical solutions apt to satisfy the AI Act Proposal's regulatory obligations, a tripartite methodology is designed corresponding to the three research objectives. In this approach, first, technical requirements are derived in structured manner from the regulatory obligations contained in the AI Act Proposal. Second, software solutions are identified in the areas covered by the requirements. Third, the solutions' effectiveness with respect to satisfaction of the obligations is assessed. The outcome of each stage is designed to constitute an independent artefact, which shall prove useful to different types of stakeholders in possession of varying levels of capacity to act upon the implications of the Act independently: While the budget-scarce small-sized company may directly start from the delivered final recommendations on software solutions to adjust their AI system for regulatory compliance, the independent AI software architect may observe the technical requirements in a first approach and design their own solution in correspondence. Herein, the selected methods follow accepted academic literature and technical standards, amended for the specifics of the AI Act Proposal.

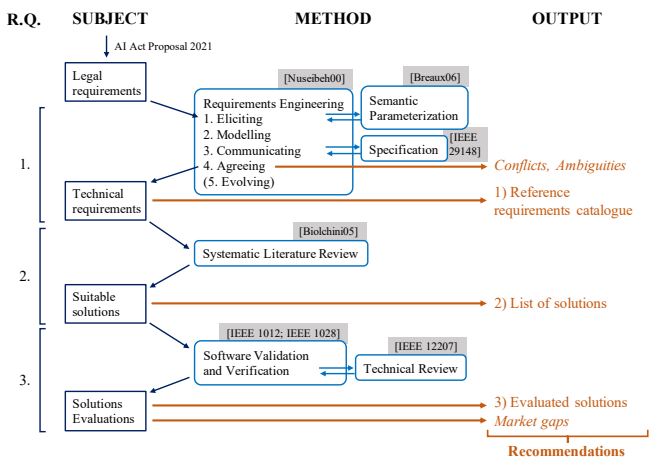

Fig. 1. Overview of the research design

To arrive at recommendations for the usage of specific technical software solutions that support the fulfillment of the requirements set out in the AI Act Proposal, a tripartite approach will be employed. First, requirements engineering [26] from the domain of software engineering allows to produce a set of technical requirements in a generalized five-step process. To render this method applicable to the first research objective its first step is substituted with a proposal by [27]. It sets out a formalised way to translate legal text, accounting for its special properties, into unambiguous demands. To formalize the requirements modeled based on them, a Software Requirements Specification according to the IEEE Standard 29148-2011 on Systems and Software Engineering [28] is produced, acknowledging its conciseness, formality, and acceptance in industry. In analysis for acceptance of the defined requirements, constituting the fourth step, overlaps, ambiguities and conflicts can be identified, accountable to shortcomes in the AI Act Proposal.

Second, to identify the most customary software solutions that principally may serve the fulfillment of the AI Act Proposal, a systematic literature review is conducted based on the technical requirements from before. The parameterized method by [29], , equally leveraging five steps, is targeted for the domain of software engineering. It is chosen due to its linear approach, allowing to control the relevance of the information extracted.

Finally, to evaluate the extent of the requirement support introduced by the capabilities of the software solution, the most scholarly established ones among the identified solutions are examined. The process of software verification, forming part of the conventional software life cycle [30], allows to verify the conformance of a software artefact with its technical requirements. It foresees the assessment of the remaining processes in the software life cycle, as governed in [31]. From different applicable approaches to realize that assessment, a limited technical review, as defined in [32], proved to possess the most favorable cost-effect ratio. There, the subject of evaluation is set to an AI system that employs the software solution of concern; the solution as an artefact itself is not verified.

From an aggregated requirement fulfillment support score computed per solution and a qualitative evaluation based thereon, final recommendations will be drawn regarding the usage of the solutions to comply with the AI Act Proposal. In this process, unaddressed requirements will be reported as potential gaps in the AI system software market.

B. Research Question 1

The approach by [26] synthesizes five steps of requirement engineering from previous literature in this field: eliciting demands, their modelling as requirements, their formal definition, and their discussion and approval, followed by their continual maintenance.

While requirements customarily are elicited from stakeholder demands subject to conflicts and negotiations, regulatory compliance is required by law. Thus, to extract the intrinsic requirements from legal text, intermediate approaches are proposed [27], [33], [34], [35], [36], [37], [38], [39]. From the two major approaches [27] and [38], both translating the legal text into intermediate representations before formally analyzing it, the former employs restricted language modelling and the latter a visual mapping in this step. Since the AI Act Proposal enforces plentiful obligations on a multitude of actors, it exceeds the capacity of visual mapping before becoming too convoluted, which is why [27] provides an appropriate approach, denoted as Semantic Parameterization.

As precondition, the articles of the EU AI Act Proposal with immediate relevance for the technical design of high-risk AI systems are identified. From these, each paragraph, amended with potential cross-references to other parts of the Act included therein, is transformed to restricted language statements in case of language ambiguity in the original text. Depending on the linguistic character of these statements,

such as the choice of verbs or use of subordinate clauses, formal obligations, rights, and constraints are extracted, the last governing the applicability of the two previous. The resulting system is interconnected through references.

In the modelling step, obligations are transformed into technical requirements according to an $m : n$ scheme, that is one obligation may yield multiple requirements and one requirement may be derived from multiple obligations. Constraints, depending on their nature, result either as part of the content of the requirement or of its applicability description. Hierarchical relations between requirements, similarly, are captured in the applicability.

As third step, the requirements are formalized as field-value pairs according to [28]. In addition to the proposed information content, type, rationale, and difficulty, the fields origin, fit criterion, applicability, and category are introduced in response to the use of the Semantic Parameterization and subsequent methods. The *fit criterion* is defined as a negatable condition entailed by the state of an AI system that fulfills the requirement. The *requirement type*, as is customarily, classifies requirements into functional, and non-functional requirements for the software artefact, and process requirements regarding its interaction with different stakeholders in its life cycle. The *rationale* explains the underlying reasoning of a requirement. The three-class *difficulty* provides an indication of the realization effort to fulfill the requirement, derived from the complexity of its content.

Finally, as part of the agreement step, each requirement is cross-checked against the set of all other requirements. Conflicts, if not solvable, as well as insufficiently precise requirements are reported as conflicts and ambiguities, respectively, that are inherent to the AI Act Proposal. The maintenance of the requirements, as iterative fifth step, is subject to updates to the AI Act Proposal or from the involved stakeholders, and is thus inapplicable.

C. Research Question 2

The final requirements specification serves as input to the systematic literature review. For this purpose, the requirements are grouped based on their content and for each group, suitable frameworks in the domain of artificial intelligence that are apt to contribute to the fulfillment of one or more of the corresponding requirements are identified.

The proposal by [29] erects five steps for a systematic review in software engineering: 1) question formularization, 2) sources selection, 3) studies selection, 4) information extraction, and 5) results summarization. First, within question formularization, the review objective, review research questions, the review approach and the measurement of the outcomes are defined. Second, applicable sources, such as publishers or conferences, in which studies shall be searched are fixed. Third, the studies, that is the publications, are selected based on inclusion and exclusion criteria. Fourth, from these, the relevant information is extracted. Based on these findings, fifth, the results are summarized in section IV.

These five steps ought to be performed within three stages: The first and the second as well as part of the third constitute the review planning. Part of the third and the fourth form the review execution. The fifth step corresponds to the results analysis stage. Review planning and review execution in turn shall be followed by an evaluation of the results from that stage, respectively.

Important parameters of review planning and evaluation and review execution are captured in a review protocol, illustrated in table I.

The result of the review consists of a set of software solutions per category, joint with an academic search results metric used as heuristic for the relevance of the solution in the scientific AI community. These outcomes will be used further in the last part of the methodology.

D. Research Question 3

To assess the conformance of a software product with some specifications, standards, or requirements software verification and validation is eligible, defined as integrated constituent of the software life cycle in [30]. Therein, tests during the implementation stage of the artefact, such as qualification and acceptance tests, may be distinguished from holistic a-posteriori approaches, such as reviews and audits [40], which assess all other life cycle stages [31].

Since qualification tests shall be performed by developers [30], acceptance tests are targeted at the acquirer of a software product [30], and audits shall be performed by independent authorities [30] - such as *notified bodies* envisioned in the AI Act Proposal [2] - technical reviews are deemed appropriate to evaluate the capacity of the software solution with respect to requirements [30].

The employed method is based on the technical review process specified in [32], from which a five-step approach towards examining software artefacts is derived: 1) provisioning of input material for the review, 2) validation that the entry criterion for the review is satisfied, 3) the software examination itself, 4) validation that the exit criterion is satisfied, and 5) output production. The objective of the review is to quantify the aptness of selected software solutions to satisfy the requirements engineered from the AI Act Proposal when used in a high-risk AI system through a manual analysis of its functioning. Thus, the solutions themselves will not be assessed for compliance with the requirements, but for their ability to support their achievement in an integrated system. There, the software product that is subject of the review is a generalized high-risk AI system that employs the respective software solution of concern, which will be evaluated against the set of applicable requirements.

For each requirement group, the three identified software solutions with the highest score for the academic relevance metric will be selected for assessment. For each solution, the following process is performed:

1) Input provisioning: The review objective as defined above, the requirements specification from subsection III-C, this procedure guidance, and the software product are provided. The last is restricted to technical documentation and complementary literature and artefacts, not however access to a running instance of the solution, owing to resource constraints.

TABLE I
SYSTEMATIC LITERATURE REVIEW PROTOCOL (EXCERPT)

Step	Value
PLANNING	
1. Question formularization	
1.1. Question focus	Identify relevant software solutions and related artefacts that could support the satisfaction of at least one technical requirement in a high-risk AI system
1.2. Question quality and amplitude	
1.2.1. Problem	Within the research landscape of ethical AI and a plethora of recommendations for AI system requirements, it is difficult to understand which software solutions are effective in satisfying the first binding requirements defined in the AI Act Proposal
1.2.2. Research question	Which published software solutions, programs, frameworks, tools, packages, or libraries could support fulfillment of at least one technical requirement from either category of the AI Act Proposal?
1.2.4. Intervention	Evaluation of software introductions, publications, reviews, comparisons, and overviews
1.2.5. Control	Reviewers' knowledge of related software solutions and their acceptance in the community
1.2.6. Effect	Set of software solutions and their relevance
1.2.7. Outcome measure	# of identified software solutions and # of search results for each framework from scholar.google.com
1.2.9. Application	Software developers of AI systems (AI Act Proposal)
2. Sources selection	
2.1. Criteria definition	Relevance for AI system developers or researchers AND ability to search through publications AND (conference OR journal OR publisher OR institution publications series)
2.3. Identification	
2.3.1. Search methods	Web search engine, sources web page search engine
2.3.2. Search string	('AI' OR 'Artificial Intelligence' OR 'Machine Learning') AND ('solution' OR 'software' OR 'framework' OR 'approach' OR 'program' OR 'algorithm' OR 'procedure' OR 'library' OR 'package') AND [REQ. CAT. NAME INCL. VARIATIONS]
2.3.3. Sources list	IEEE, ACM, NIPS, ACM SIGMOD, JMLR, Arxiv.org, Springer, Researchgate, Github.com, Stackoverflow.com, Gartner.com, SAS Publishers, Rheinwerk Verlag, Proceedings of International Conference on Machine Intelligence and Data Science Applications, IBM J. Res. Dev.
3. Studies selection	
3.1. Studies definition	
3.1.1 Inclusion and exclusion criteria definition	Includes reference to relevant software solution published by author or company AND NOT includes references to beta versions or unpublished software
3.1.2 Studies types definition	Paper, proceedings, technical reports, webpages, GitHub repositories, forum hyperlink references
3.1.3 Procedures for studies selection	1) Use 2.3.2 to search sources 2) Include studies that meet 3.1.1 criteria 3) Analyze selected study and extract information on software solutions in format of 4.2 4) Retrieve no. of academic search results for the identified solution on scholar.google.com by searching for [Solution name] + ['AI', if name does not contain explicit AI reference]
PLANNING EVALUATION	The protocol was iteratively executed with subset of sources and refined in response to recognized issues.
EXECUTION	
4. Information extraction	
4.2. Data extraction form	Solution name, category, description, publisher, academic publication, # scholarly search results

To allow for a thorough assessment nonetheless, complementary literature and artefacts can comprise of software development and architecture descriptions, maintenance manuals, release notes, source code repositories, marketing material, and user question and answer protocols, each retrieved from the original solution publisher or trusted sources.

2) Entry criterion validation: Technical documentation and complementary literature and artefacts, if necessary, are available in sufficient number, extent, and depth. Sufficiency is defined as the reviewer being able, in a preliminary assessment, to maintain that all applicable requirements can be assessed according to this procedure only from the provided information, or that a lack of information is objective evidence of failure to support the requirement.

3) Examination Procedure: Per requirement to assess against, the available technical documentation and complementary literature and artefacts are searched for relevant information. From evidence regarding the functionality and non-functional properties such as architecture, interoperability, operational or maintenance conditions, the reviewer establishes the extent to which a solution supports a high-risk AI system's compliance with the requirement along four levels.

- 0 - No support. Integration of the solution does not contribute to satisfying the requirement.
- 1 - Limited support. Integration of the solution partially contributes to satisfying the requirement but considerable effort remains to fulfill it.
- 2 - Moderate support. Integration of the solution contributes to satisfying the requirement but some effort remains to fulfill it.
- 3 - Extensive support. Integration of the solution substantially contributes to satisfying the requirement, leaving no or minimal effort to fulfill it.

Here, effort refers to the delta between the contribution of the software solution and the target state of the fulfilled requirement, which is provided by a high-risk AI system that can fulfill the fit criterion of the requirement. This delta can be closed with manual development or administration activities or with further software solutions.

4) Exit criterion validation: All requirements pertaining to the category were assigned level 1-3 or conclusively assigned level 0.

5) Output production: The evaluations per requirement are stored. In addition, for each solution a level-weighted requirement fulfillment support score is computed as

$$\sum_{r=1}^{r_{max}} \frac{l_r}{3} \times \frac{1}{r_{max}} \tag{1}$$

where r is the integer requirement ID, r_{max} is the number of requirements to consider, usually the number of applicable requirements of the category, and l_r is the fulfillment support level assigned for the requirement with ID r. Thus, the score returns the portion of requirements whose satisfaction in a high-risk AI system is fully supported when employing the software solution, where 100% equates to all requirements being evaluated as level 3.

Concluding the methodology, a qualitative analysis of the quantitative results from the third step allows to achieve the overall research objective. Thereby, recommendations for an effective use of software solutions to comply with the AI Act Proposal's technical obligations are pronounced and gaps and ambiguity-induced uncertainties that should be considered are pointed out.

IV. Results

In line with the research objective, the results from execution of the research design will be portrayed in order of the research questions.

A. Requirement Engineering

After an initial analysis of the proposal, Articles 9 to 15, in *Title III*, *Chapter 2 - Requirements for High-Risk AI systems*, were classified as relevant as they contain immediate technical obligations for high-risk AI systems and specify the conditions they must satisfy. Hence, the clauses pertaining to this chapter, which were found to be linguistically unambiguous, will be used as basis for the Semantic Parameterization. Each article, thus, produces a set of obligations, requirements, and constraints, constituting the eliciting step of the requirement engineering process.

To demonstrate how the articles in the legal text were transformed into finished requirements, the engineering of one requirement is examined in table II and table III.

Table II depicts the erection of obligations, rights, and constraints from analysis of the original legal text. There, the verb indicating whether the sentence yields an obligation or a right is highlighted in **bold and underlined**, the details about an obligation are formatted **bold**, and details about a constraint, governing the applicability of the obligation, are formatted *italic*. In this case, the requirement arises from two paragraphs in article 9. 'Shall', in the legal sense, implies an obligation (O) for the high-risk AI system, for which reason both are transformed into such, respectively. While the content of art. 9 (5), only specifies the content of the obligation, art. 9 (2d) additionally conditions the scope of its corresponding obligation, normally translated into a constraint. However, as the content of the constraint is superfluous in light of the additional requirements arising from the remaining paragraphs, it was not employed as such to restrict O9.5.

Table III subsequently shows the result of analysis of the two obligations to arrive at a requirement. The *description* as content of the requirement introduces the obligation to test the system. The directly deducible, subjective motivation for a testing procedure, next to the articles requiring it regardless of consent to it, is presented in the *rationale*. Because the existence of technical test routines is a technical requirement compared to an organizational one, but one with no functionality for the user of the system, the *type* is set to non-functional. Since testing is obligatory in any software project, the additional workload is minimal, rendering the *difficulty* low. The *applicability*, besides applying the requirement to all types of high-risk AI systems defined in the AI Act Proposal, conditions the requirement on the existence of a risk management system in the system, which is defined in another requirement. Finally, the *fit criterion* specifies a scenario resulting from a system that implements the test routine with the specified purpose, which can be probed to assess conformance in a later stage.

TABLE II
Sample Requirement: Semantic Parameterization

Art. 9 (2) d:	Art. 9 (5):
The risk management system [...] **shall comprise the following steps: [...] adoption of suitable risk management measures** *in accordance with the provisions of the following paragraphs*	High-risk AI systems **shall be tested for the purposes of identifying the most appropriate risk management measures**. [...]
↓	↓
O9.5: The risk management system comprises of suitable risk management measures	O9.15: To identify the risk management measures, the high-risk AI system is tested

TABLE III
Sample Requirement: Requirement Specification

ID	9.14
Origin	O9.5, O9.15
Description	The high-risk AI system shall be tested with the purpose of identifying appropriate risk management measures.
Rationale	Art. 9 (2)(d), Art. 9 (5); Testing a high-risk AI system reveals the risks associated with its use that are hard to expect or predict.
Difficulty	low
Fit Criterion	The risk management measures adopted in the finalised risk management system were informed by the results of a technical testing procedure performed on the high-risk AI system.
Type	Non-Functional Requirement
Applicability	All (high-risk AI systems), given req. 9.1 is fulfilled
Category	Testing

Following this scheme, a total of 95 requirements was erected from the seven articles. The total number of obligations, rights, and constraints extracted from each article is shown in table IV.

TABLE IV
Number of obligations, rights, and constraints erected from relevant articles of the AI Act Proposal

Article	# Obligations	# Rights	# Constraints
9 - Risk management system	22	-	7
10 - Data and data governance	14	2	12
11 - Technical documentation	22	1	1
12 - Record-keeping	8	-	7
13 - Transparency and provision of information to users	3	-	8
14 - Human oversight	10	-	6
15 - Accuracy, robustness and cybersecurity	10	-	2

Based on the requirements' content, regarding the aspect of the AI system they address, and on their origin among

the articles, each requirement was assigned to one of eight categories, which are shown in table V.

TABLE V
REQUIREMENT CATEGORIES

Category	Description	# Req.
Risk Management	predominantly process or functional requirements regarding the implementation of risk mitigation procedures	15
Testing	mainly non-functional requirements concerning testing routines and procedures in the high-risk AI system's life cycle	5
Dataset Properties	mostly non-functional requirements addressing the quality and content of training, validation and test sets put into the system	10
Technical Documentation	predominantly non-functional and process requirements regarding the scope of the information about the system included	23
Record Keeping	predominantly functional requirements regarding the logging of system behavior and access to these	11
Explainability	mostly process requirements on the transparency of operations of the system and the content of instructions of use	10
Human Oversight	requirements concerning interfaces and procedures for human beings to control the operation of the system	9
Accuracy, Robustness, Cybersecurity	process and non-functional requirements mitigating the proneness of the system to errors	12

The exhaustive requirement specification can be found in Appendix I.

B. Software Solution Identification

Based on the requirement categories, software solutions with the potential to support the fulfillment of the requirements were systematically searched. From a technical standpoint, the two categories *Risk Management* and *Technical Documentation* did not yield any requirements specific to AI systems compared to general IT systems. Since numerous reviews and market analyses are available in these domains, they were excluded from further research.

To demonstrate the review findings on the solution-level, in table VI, the resultant software solutions for the category *Accuracy, Robustness, Cybersecurity* are portrayed. While no AI-specific cybersecurity solution was identified, five were assessed to be relevant for the robustness- and accuracy-related requirements. Out of these, the three with the highest academic relevance score, Foolbox Native, IBM Adversarial Robustness Toolbox, and IBM CNN-Cert were selected for evaluation.

In total, 36 unique software solutions were identified. Among these, individual solutions were returned for several categories, such as Local Interpretable Model-Agnostic Explanation (LIME) (Explainability; Human Oversight), Neptune.ai (Record Keeping; Human Oversight), RuleX AI (Explainability; Human Oversight), and SHapley Additive exPlanations (SHAP) (Explainability; Human Oversight). In addition software suites contained individual tools that were relevant for different categories, such as Amazon Sage Maker (Testing; Dataset Properties; Record Keeping; Human Oversight), IBM Research Trustworthy AI 360 Toolkit (Explainability; Human Oversight; Accuracy, Robustness, Cybersecurity), IBM Watson (Testing; Dataset Properties), and Tensorflow (Dataset Properties; Human Oversight).

TABLE VI
SOFTWARE SOLUTIONS FOR SAMPLE CATEGORY *Accuracy, Robustness, Cybersecurity*

Name	Publisher	Original Publication	# Academic Search Results
CORTEX CERTIFAI	CognitiveScale	[41]	59
Foolbox Native	Rauber, J.	[42]	498
IBM Adversarial Robustness Toolbox	IBM	[43]	305
IBM CNN-Cert	IBM	[44]	85
IBM Research AI Fairness 360 Toolkit	IBM	[45]	48

The distribution over the categories, including multi-category solutions, is depicted in table VII.

TABLE VII
NUMBER OF SOFTWARE SOLUTIONS PER CATEGORY

Category	# Identified Software Solutions
Testing	3
Dataset Properties	11
Record Keeping	7
Explainability	5
Human Oversight	9
Accuracy, Robustness, Cybersecurity	5

C. Software Solution Evaluation

Finally, the three software solutions per category with the highest number of academic search results were evaluated for their aptness to satisfy the category requirements when employed in high-risk AI system. As process requirements necessitate organizational effort, they were excluded from the technical review.

Continuing the sample from subsection IV-B, table VIII shows the evaluations per requirement for the highest-relevance software solution in category *Accuracy, Robustness, Cybersecurity*: Foolbox Native. The given explanations show why different levels of fulfillment support were assigned, referencing the evidence that provided the underlying information. Out of the nine applicable requirements, using Foolbox Native would at least partially facilitate the fulfillment of seven. The level-weighted requirement fulfillment support score computes to 56%.

In table IX, the portion of applicable category requirements by evaluation level is provided for the three selected software solutions of each category along with their weighted aggregated requirement fulfillment support scores. From the 95 original requirements across eight categories, 37 across six categories were applicable. The overall rounded mean requirement fulfillment support score over all categories is 34%. On the category level, the decreasing rounded mean scores are 78% for Explainability, 46% for Accuracy, Robustness, Cybersecurity 42% for Testing, 37% for Dataset Properties, 26% for Human Oversight, and 23% for Record Keeping.

TABLE VIII
EVALUATION OF SAMPLE SOFTWARE SOLUTION *Foolbox Native* IN REQUIREMENT CATEGORY *Accuracy, Robustness, Cybersecurity*

Req. Id	Level	Explanation	Evidence
15.1	*N/A*	*Process requirement*	*N/A*
15.2	1	Foolbox provides attack models for adversarial training. There is a trade-off between robustness ('robust accuracy') and accuracy ('standard accuracy'). A consistent level of robustness through should lead to a consistent level of accuracy.	[46]
15.3	3	Foolbox provides a variety of adversarial attacks to benchmark the robustness of machine learning models.	[47]
15.4	2	Foolbox provides adversarial training, which helps mitigating adversarial attacks, but is not sufficient to achieve cybersecurity as a whole.	[48]
15.5	*N/A*	*Process requirement*	*N/A*
15.6	0	Foolbox provides adversarial training, but does not address technical redundancy or fault prevention.	[46]
15.7	0	Foolbox provides adversarial training, but does not address biased outputs through 'feedback loops'.	[46]
15.8	2	Adversarial training mitigates adversarial attacks, being a popular way of AI-System manipulation, but does not generally prevent unauthorized access by third parties.	[46]
15.9	*N/A*	*Process requirement*	*N/A*
15.10	2	Data poisoning is considered a specific strategy of adversarial attacks, which are addressed by the framework.	[47]
15.11	3	Adversarial examples are considered a specific strategy of adversarial attacks that are explicitly addressed by the framework.	[47]
15.12	2	Model flaw exploitation is considered a specific strategy of adversarial attacks, which are addressed by the framework.	[47]

TABLE IX
OVERVIEW OF EVALUATION OF SOFTWARE SOLUTIONS PER REQUIREMENT CATEGORY

Category	Software Solution	Level 0	Level 1/2/3	Score
Testing (4 req.)	Amazon Sage Maker	25%	50%/25%/0%	33%
	Watson OpenScale	0%	25%/75%/0%	58%
	Azure ML	25%	50%/25%/0%	33%
Dataset Properties (7 req.)	IBM SPSS Modeler	29%	29%/14%/29%	48%
	SAP Data Services	29%	43%/14%/14%	38%
	Informatica Data Quality	57%	14%/29%/0%	24%
Record Keeping (10 req.)	TensorBoard	50%	30%/20%/0%	23%
	Amazon CloudWatch	50%	0%/40%/10%	37%
	DataDog	80%	10%/10%/0%	10%
Explainability (1 req.)	SHapley Additive exPlanations (SHAP)	0%	0%/100%/0%	67%
	Local Interpretable Model-Agnostic Explanation (LIME)	0%	0%/100%/0%	67%
	IBM AIX360 Toolkit	0%	0%/0%/100%	100%
Human Oversight (6 req.)	SHapley Additive exPlanations (SHAP)	67%	17%/0%/17%	22%
	Local Interpretable Model-Agnostic Explanation (LIME)	67%	17%/17%/0%	17%
	MLflow	50%	0%/33%/17%	39%
Accuracy, Robustness, Cybersecurity (9 req.)	Foolbox Native	22%	11%/44%/22%	56%
	IBM Adversarial Robustness Toolbox	22%	11%/44%/22%	56%
	IBM CNN-Cert	33%	56%/11%/0%	26%

In the case of *Accuracy, Robustness, Cybersecurity*, it is recommended to employ either Foolbox Native or IBM Adversarial Robustness Toolbox in the high-risk AI system as their functionality is similar, each achieving a fulfillment support score of 55%. However, their requirement coverage is not complementary, rendering the use of both simultaneously superfluous. Part of the uncovered requirements are those that go beyond the AI-specifics robustness and explainability and instead include traditional security aspects. To fulfill these, it should be attempted to use conventional IT security practices and solutions, jointly with the novel AI-specific solutions.

Similarly, examining the assessed software solutions' individual explanations and level assignments per requirement demonstrates which solutions harmonize satisfactorily and which requirements remain entirely uncovered in each category. Thereby, recommendations on how to most effectively comply with the AI Act Proposal using established software solutions in high-risk AI systems are provided.

V. DISCUSSION

Overall, the results of the last step within the research process show that there are various technical solutions and frameworks which can be considered useful to comply with the proposed regulation on AI. Nevertheless, the individual evaluation scores indicate that few requirements and categories can be fully covered by the identified software solutions. In fact, 11 of the total 95 defined requirements were consistently evaluated as level zero, meaning their fulfilment cannot be supported by implementing the considered frameworks at all. This limitation of results can be attributed to the following factors which have become apparent in the course of the research:

During the analysis of the AI Act Proposal and the process of deriving technical requirements, various shortcomings in the level of detail have been identified. This impeded the derivation of clear technical implications for RQ1. An overview of vague or ambiguous terms and phrases has been composed and is provided in appendix IV. Without a clear definition of, for instance, what measures are considered in accordance with "recognised standards" (art. 12 (1)) or what level of transparency towards the user is "sufficient" (art. 13 (1)), the fulfilment of requirements containing such ambiguities can only be evaluated on a high level. In some cases, this has led to the respective requirement being evaluated as level zero.

The analyzed chapter is divided into 15 articles. Among these articles several overlaps and dependencies have been

identified. While explainability as defined in subsection II-A can be considered a key aspect of trustworthy AI systems [6][7], it is not explicitly mentioned in the legislative text. Instead, the concept of explainable AI appears to be covered by multiple articles, such as "Human Oversight" (art. 14), "Transparency" (art. 13) and "Record Keeping" (art. 12). These interdependencies rendered it difficult to define useful and distinguishable categories in preparation for RQ2 as described in subsection III-C. As a result, the categories and the respective names do not represent every individual requirement in the same way, returning software solutions with limited coverage in the systematic review process.

Other requirements or sets of requirements could not be covered by specific solutions due to their content being process-oriented or not specific to the AI systems special characteristic. Such process requirements need to be addressed by adequate management and governance methodologies (e.g., "Risk Management System", "Technical Documentation"). The gap in corresponding software solutions also extends to sets of requirements not considered AI-oriented in the first place: Especially in the fields of "Testing", "Record Keeping", and some traditional IT-Security aspects, only few AI-specific technical solutions were found as result from the systematic review. This may indicate a demand for stronger synergies between AI-specific and general software engineering in non-functional software areas. End-to-end ML platforms address several aspects of the ML development cycle, including important non-functional aspects, and therefore, are able to cover more requirements than task-specific solutions.

Not only with regards to the level of detail of the legislative text, but also of the information and documentation of some technical solutions, limitations have become apparent. As the systematic review described in subsection III-C included both open-source as well as proprietary software solutions, the quality of sources available to comprehend their functionality varied widely. For a practical, detailed analysis of the requirements' fulfilment, each solution would be required to be employed in the specific AI system for individual reviews in addition to technical tests. In some categories, software solutions are only applicable for specific types of ML models and data. For instance, IBM CNN-Cert is designed exclusively for certifying the robustness of Convolutional Neural Networks (CNNs), not any other neural network and ML models. While useful in the targeted cases, often reflecting technical development trends in the AI landscape, this limits the applicability of such solutions.

In addition to technology restrictions, most of the solutions solely address a certain functional or non-functional aspect, even within the assigned category, which would require combination with other solutions or manual implementation efforts to fulfill all given requirements.

Despite the limitations outlined, the results at hand are a useful foundation and guidance to understand the technical implications of the AI Act Proposal in the applicable areas and categories. For other categories, research demonstrated that further elaboration on the proposal itself, as well as case-specific evaluation for different applications and fields of AI will be necessary.

Finally, it should be noted that the scope of the act is substantially larger than the definition of obligations for the high-risk AI system itself. Only taking into consideration the rights and obligations of users, authorities, and other stakeholders will allow to estimate the total effort for AI system providers to comply with this law.

VI. Conclusion and Future Work

The main objective of this work was to analyze the legal obligations set out by the European Commission's proposal for an Artificial Intelligence Act for their technological impact on high-risk AI systems in order to identify and evaluate technical solutions that assist in achieving compliance with these requirements.

As a result, an extensive set of 95 requirements has been derived from the legislative text along with an overview of ambiguous and vague terms or phrases which require specification in a revision of the draft. A list of 36 potentially suitable software solutions has been composed through a systematic review based on six technically relevant requirement categories. For each category, the three most scholarly mentioned solutions have been selected to evaluate their suitability to support compliance with the regulation when implemented in a specific AI system. For the majority of requirement categories, the mean requirement fulfillment scores is below 50%, indicating a considerable gap between current established solutions in the market and the scope of the AI Act Proposal. If unmet, the AI Act Proposal, irrespective of the appropriateness of its measures, may require a large technical effort for high-risk AI system providers to comply.

The results of this work can be considered a contribution to the joint effort of elaborating a technical specification, derived from the AI Act Proposal, which is explicitly envisioned and encouraged by the EU Commission [2]. As is the nature of a legislative proposal, the AI Act Proposal has drawn various criticism regarding some of its crucial aspects from several parties and stakeholders [49][50][51]. The research for this work has revealed some of those shortcomings, regarding lack of technical detail, interdependencies and ambiguities, and therefore confirmed part of the criticism. When revising the proposal to arrive at a final regulation, these aspects needs to be addressed thoroughly.

Until then, this work could potentially prove useful to the technical AI community in preparing for the binding impact of the regulation. The full results are available at [52] where it is sought to maintain and extend the requirements, software solutions, and evaluations as the legislative process progresses. For this purpose, contributions are highly welcomed. On the way to trustworthy AI, the technological feasibility of international regulations will be crucial to leverage the high potential of AI in a safe, ethical, and human-centered manner.

References

[1] A. Jobin, M. Ienca, and E. Vayena, "The global landscape of ai ethics guidelines," *Nature Machine Intelligence*, vol. 1, pp. 389–399, 2019.

[2] "Proposal for a regulation of the european parliament and of the council laying down harmonised rules on artificial intelligence and amending certain union legislative acts," European Commission, 2021.

[3] J. Wolff and N. Atallah, "Early gdpr penalties: Analysis of implementation and fines through may 2020," *Journal of Information Policy*, vol. 11, pp. 63–103, 2021.

[4] A. V. Joshi, *Machine learning and artificial intelligence*. Springer, 2020.

[5] K. P. Murphy, *Machine learning: a probabilistic perspective*. MIT press, 2012.

[6] A. Adadi and M. Berrada, "Peeking inside the black-box: A survey on explainable artificial intelligence (xai)," *IEEE Access*, vol. 6, 2018.

[7] A. B. Arrieta, N. Daz-Rodrguez, J. D. Ser, A. Bennetot, S. Tabik, A. Barbado, S. Garcia, S. Gil-Lopez, D. Molina, R. Benjamins, R. Chatila, and F. Herrera, "Explainable artificial intelligence (xai): Concepts, taxonomies, opportunities and challenges toward responsible ai," *Information Fusion*, vol. 58, 6 2020.

[8] P. W. Koh and P. Liang, "Understanding black-box predictions via influence functions," in *Proceedings of the 34th International Conference on Machine Learning*, D. Precup and Y. W. Teh, Eds., vol. 70. PMLR, 7 2017, pp. 1885–1894.

[9] M. Bojarski, P. Yeres, A. Choromanska, K. Choromanski, B. Firner, L. Jackel, and U. Muller, "Explaining how a deep neural network trained with end-to-end learning steers a car," *CoRR*, 4 2017.

[10] B. Goodman and S. Flaxman, "European union regulations on algorithmic decision-making and a right to explanation," *AI Magazine*, vol. 38, 10 2017.

[11] A. Chouldechova, "Fair prediction with disparate impact: A study of bias in recidivism prediction instruments," *Big Data*, vol. 5, 6 2017.

[12] B. Kim, E. Glassman, B. Johnson, and J. Shah, "ibcm: Interactive bayesian case model empowering humans via intuitive interaction," *CSAIL Technical Reports*, 2015.

[13] M. T. Ribeiro, S. Singh, and C. Guestrin, "Why should i trust you? explaining the predictions of any classifier," *Tatra Mountains Mathematical Publications 74*, 2016.

[14] European Commission, "Excellence and trust in artificial intelligence," 2021. [Online]. Available: https://ec.europa.eu/info/strategy/priorities-2019-2024/europe-fit-digital-age/excellence-trust-artificial-intelligence_en#building-trust-through-the-first-ever-legal-framework-on-ai

[15] "Annexes to the proposal for a regulation of the european parliament and of the council laying down harmonised rules on artificial intelligence (artificial intelligence act) and amending certain union legislative acts," European Commission, 2021.

[16] K. Shahriari and M. Shahriari, "Ieee standard review ethically aligned design: A vision for prioritizing human wellbeing with artificial intelligence and autonomous systems," in *2017 IEEE Canada International Humanitarian Technology Conference (IHTC)*. IEEE, 7 2017.

[17] ISO. Iso/iec jtc 1/sc 42 artificial intelligence. [Online]. Available: https://www.iso.org/committee/6794475.html

[18] ——. Iso/iec jtc 1 information technology. [Online]. Available: https://www.iso.org/isoiec-jtc-1.html

[19] UNESCO. (2020) Outcome document: first draft of the recommendation on the ethics of artificial intelligence. Ad Hoc Expert Group for the Preparation of a Draft text of a Recommendation the Ethics of Artificial Intelligence.

[20] Romecall. (2021) Rome call for ai ethics a human-centric artificial intelligence. [Online]. Available: www.romecall.org/

[21] ITU. (2021) Ai for good. [Online]. Available: https://aiforgood.itu.int

[22] UN Global Pulse. (2021) Expert group on governance of data and ai. [Online]. Available: https://www.unglobalpulse.org/policy/expert-group-on-governance-of-data-and-ai/

[23] World Economic Forum. (2021) Global future council on artificial intelligence for humanity. [Online]. Available: https://es.weforum.org/communities/gfc-on-artificial-intelligence-for-humanity

[24] OECD. Oecd principles on ai. [Online]. Available: https://www.oecd.org/going-digital/ai/principles

[25] Council of Europe. (2021) Cahai - ad hoc committee on artificial intelligence. [Online]. Available: https://www.coe.int/en/web/artificial-intelligence/cahai

[26] B. Nuseibeh and S. Easterbrook, "Requirements engineering: a roadmap," in *Proceedings of the Conference on the Future of Software Engineering*, 2000, pp. 35–46.

[27] T. D. Breaux, M. W. Vail, and A. I. Anton, "Towards regulatory compliance: Extracting rights and obligations to align requirements with regulations," in *14th IEEE International Requirements Engineering Conference (RE'06)*, 2006, pp. 49–58.

[28] *ISO/IEC/IEEE International Standard - Systems and software engineering – Life cycle processes – Requirements engineering*, ISO/IEC/IEEE Std. 29148-2018, 2018.

[29] J. Biolchini, P. G. Mian, A. C. C. Natali, and G. H. Travassos, "Systematic review in software engineering," *System Engineering and Computer Science Department COPPE/UFRJ, Technical techreport ES*, vol. 679, p. 45, 2005.

[30] *ISO/IEC/IEEE International Standard - Systems and software engineering – Software life cycle processes*, ISO/IEC/IEEE Std., 2017.

[31] *IEEE Standard for System and Software Verification and Validation*, IEEE Std. 1012-2012, 2012.

[32] *IEEE Standard for Software Reviews and Audits*, IEEE Std. 1028-2008, 2008.

[33] P. N. Otto and A. I. Anton, "Addressing legal requirements in requirements engineering," in *15th IEEE International Requirements Engineering Conference (RE 2007)*, 2007, pp. 5–14.

[34] N. Kiyavitskaya, A. Krausov, and N. Zannone, "Why eliciting and managing legal requirements is hard," in *2008 Requirements Engineering and Law*, 2008, pp. 26–30.

[35] N. Kiyavitskaya, N. Zeni, T. D. Breaux, A. I. Antn, J. R. Cordy, L. Mich, and J. Mylopoulos, "Automating the extraction of rights and obligations for regulatory compliance," in *International Conference on Conceptual Modeling*. Springer, 2008, pp. 154–168.

[36] T. Breaux and A. Antn, "Analyzing regulatory rules for privacy and security requirements," *IEEE Transactions on Software Engineering*, vol. 34, pp. 5–20, 2008.

[37] A. Siena, J. Mylopoulos, A. Perini, and A. Susi, "From laws to requirements," in *2008 Requirements Engineering and Law*, 2008, pp. 6–10.

[38] ——, "Designing law-compliant software requirements," in *Conceptual Modeling - ER 2009*, A. H. F. Laender, S. Castano, U. Dayal, F. Casati, and J. P. M. de Oliveira, Eds. Springer Berlin Heidelberg, 2009, pp. 472–486.

[39] J. C. Maxwell, A. I. Antn, and P. Swire, "A legal cross-references taxonomy for identifying conflicting software requirements," in *2011 IEEE 19th international requirements engineering conference*. IEEE, 2011, pp. 197–206.

[40] *ANSI / IEEE Standard 1002.1987: Standard Taxonomy for Software Engineering Standards*, ANSI / IEEE Std. 1002-1987, 1987.

[41] S. Sharma, J. Henderson, and J. Ghosh, "Certifai counterfactual explanations for robustness, transparency, interpretability, and fairness of artificial intelligence models," in *arXiv preprint arXiv:1905.07857*, 2019.

[42] J. Rauber, R. Zimmermann, M. Bethge, and W. Brendel, "Foolbox native: Fast adversarial attacks to benchmark the robustness of machine learning models in pytorch, tensorflow, and jax," *Journal of Open Source Software*, vol. 5, p. 2607, 2020.

[43] M.-I. Nicolae, M. Sinn, M. N. Tran, B. Buesser, A. Rawat, M. Wistuba, V. Zantedeschi, N. Baracaldo, B. Chen, and H. Ludwig, "Adversarial robustness toolbox v1. 0.0," in *arXiv preprint arXiv:1807.01069*, 2018.

[44] A. Boopathy, T.-W. Weng, P.-Y. Chen, S. Liu, and L. Daniel, "Cnn-cert: An efficient framework for certifying robustness of convolutional neural networks," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 33, 2019, pp. 3240–3247.

[45] R. K. E. Bellamy, K. Dey, M. Hind, S. C. Hoffman, S. Houde, K. Kannan, P. Lohia, J. Martino, S. Mehta, and A. Mojsilovi, "Ai fairness 360: An extensible toolkit for detecting and mitigating algorithmic bias," *IBM Journal of Research and Development*, vol. 63, pp. 1–4, 2019.

[46] J. Rauber. (2021) Foolbox. [Online]. Available: https://foolbox.jonasrauber.de/

[47] ——. (2020) foolbox.attacks foolbox 3.3.1 documentation. [Online]. Available: https://foolbox.readthedocs.io/en/stable/modules/attacks.html

[48] R. Hamon, H. Junklewitz, and I. Sanchez, "Robustness and explainability of artificial intelligence," Publications Office of the European Union, Tech. Rep., 2020.

[49] P. Glauner, "An assessment of the ai regulation proposed by the european commission," *arXiv preprint arXiv:2105.15133*, 2021.

[50] L. Floridi, "The european legislation on AI: a brief analysis of its philosophical approach," *Philosophy & Technology*, vol. 34, no. 2, pp. 215–222, Jun. 2021.

[51] European Trade Union Institute RPS Submitter and Aida Ponce del Castillo, "The AI regulation: entering an AI regulatory winter? why an ad hoc directive on AI in employment is required," *SSRN Electronic Journal*, 2021.

[52] GitHub. Ai act propsal results wwi2018e / technical requirements and viable solutions for high risk ai-systems: The european union artificial intelligence act proposal - technical requirements

and viable solutions for high-risk ai systems. [Online]. Available: https://github.com/AI-Act-Propsal-Results-WWI2018E/Technical-Implications-and-Viable-Solutions-for-High-Risk-AI-Systems-

APPENDIX
TABLE OF APPENDICES

APPENDIX I:
REQUIREMENTS SPECIFICATION

ID	<X.X> (<Obligation ID>)
Description	<Description of the requirement>
Rationale	<Art. X (X); short rationale in own words>
Difficulty	<low, medium, high>
Fit Criterion	<As precise as possible: how will/can the requirement be evaluated?>
Type	<functional, non-functional, process>
Applicability	<All (high-risk AI systems) / Restricted (Details, Ref.)>
Category	<Category or Sub-Category, if applicable>

APPENDIX I.1: RISK MANAGEMENT SYSTEM (ART. 9)

ID	9.1 (O1)
Description	A risk management system shall exist that is maintained and documented.
Rationale	Art. 9 (1); The risks from AI systems need to be understood and controlled
Difficulty	low
Fit Criterion	A risk management system is continually operating and accessible by a user that has access to its documentation.
Type	Functional Requirement
Applicability	All (high-risk AI systems)
Category	Risk Management

ID	9.2 (O2)
Description	The risk management system shall operate through the entire lifetime of the high-risk AI system as a continuous iterative process.
Rationale	Art. 9 (2); The risks from AI systems need to be evaluated continuously
Difficulty	medium
Fit Criterion	After multiple fixed periods of operation of an AI system, respectively, the risk management system accessible by a user is still operating and updated to potentially changed circumstances with respect to the high-risk AI system.
Type	Functional Requirement
Applicability	All (high-risk AI systems), given req. 9.1 is fulfilled.
Category	Risk Management

ID	9.3 (O3)
Description	The risk management system shall have the ability to identify all known and foreseeable risks with respect to the high-risk AI system.
Rationale	Art. 9 (2)(a); The risks from AI systems need to be identified in order to be treated
Difficulty	high
Fit Criterion	The risks with respect to multiple known high-risk AI systems returned to an expert user from the risk management system match at-large the risks of these systems known beforehand.
Type	Functional Requirement
Applicability	All (high-risk AI systems), given req. 9.1 is fulfilled
Category	Risk Management

ID	9.4 (O4)
Description	The risk management system shall have the ability to evaluate and estimate the risks with respect to the high-risk AI system that arise from its purpose-conform use, reasonably foreseeable misuse, or the output from a post-market monitoring system according to requirements 12.4-12.6.
Rationale	Art. 9 (2)(b), (c); The risks from AI systems need to be evaluated and characterised in order to be treated
Difficulty	high
Fit Criterion	The evaluation of risks from ordinary use, foreseeable misuse, and post-market monitoring mechanisms with respect to multiple known high-risk AI systems returned to an expert user match at-large his evaluation of these risks.
Type	Functional Requirement
Applicability	All (high-risk AI systems), given req. 9.1 is fulfilled
Category	Risk Management

ID	9.5 (O5, O6)
Description	The risk management system shall adopt risk management measures that duly consider the effects and possible interactions from the entirety of the requirements defining the high-risk AI system in this Requirements Specification.
Rationale	Art. 9 (2)(d) + Art. 9 (3); Unexpected risks may arise from any AI system established according to a variety of independent requirements
Difficulty	high
Fit Criterion	An expert user is unable to identify any risks from the interactions of the requirements established in this Requirements Specification that define the AI system that were already identified by the risk management system.
Type	Functional Requirement
Applicability	All (high-risk AI systems), given req. 9.1 is fulfilled
Category	Risk Management

ID	9.6 (O5, O7)
Description	The risk management system shall adopt risk management measures that operate according to the industrial standard, for example through harmonised standards or common specification.
Rationale	Art. 9 (2)(d) + Art. 9 (3); Pre-existing standards and common practices in risk management system are applicable and useful to high-risk AI systems
Difficulty	medium
Fit Criterion	A proficient risk management engineer verifies that the risk management system measures conform to the most appropriate standard or common practices, if any.
Type	Functional Requirement
Applicability	Restricted (high-risk AI systems whose risks are applicable to common practices or standards), given req. 9.1 is fulfilled
Category	Risk Management

ID	9.7 (O5, O8)
Description	The risk management system shall adopt risk management measures that ensure that residual risks from a high-risk AI system used according to its purpose or under condition of reasonably foreseeable misuse associated with each hazard and overall is judged acceptable.
Rationale	Art. 9 (2)(d) + Art. 9 (4); A risk management system is only sufficiently effective when the residual, non-treatable risks are acceptable
Difficulty	high
Fit Criterion	None of the evaluations of residual risks returned from the risk management with respect to a high-risk AI system used according to its purpose or under condition of reasonably foreseeable misuse is classified worse than acceptable or some equivalent threshold.
Type	Functional Requirement
Applicability	All (high-risk AI systems), given req. 9.1 is fulfilled
Category	Risk Management

ID	9.8 (O5, O9)
Description	The risk management system shall communicate all residual risks to the user.
Rationale	Art. 9 (2)(d) + Art. 9 (4); Residual risks may only be act upon when communicated to the user of the high-risk AI system
Difficulty	low
Fit Criterion	The system returns all of the identified residual risks according to 9.7 to the user via an appropriate interface.
Type	Functional Requirement
Applicability	All (high-risk AI systems), given req. 9.1 is fulfilled
Category	Risk Management

ID	9.9 (O5, O10)
Description	The risk management system shall adopt risk management measures such that the high-risk AI system's architecture and implementation minimises the risks associated with its purpose-conform use or reasonably foreseeable misuse.
Rationale	Art. 9 (2)(d) + Art. 9 (4)(a); The use of an effective risk management system is intended to lead to the elimination of risks
Difficulty	medium
Fit Criterion	No other risk management measures can be identified by a risk management engineer the use of which would yield a further reduction of risks in the operation of the high-risk AI system.
Type	Functional Requirement
Applicability	All (high-risk AI systems), given req. 9.1 is fulfilled
Category	Risk Management

ID	9.10 (O5, O11)
Description	The risk management system shall adopt risk management measures such that the high-risk AI system's implementation includes adequate mitigation and control measures for residual risks associated with its purpose-conform use or reasonably foreseeable misuse that cannot be eliminated.
Rationale	Art. 9 (2)(d) + Art. 9 (4)(b); The use of an effective risk management system is intended to lead to the control and mitigation of risks
Difficulty	medium
Fit Criterion	No other risk management measures can be identified by a risk management engineer the use of which would yield more effective risk control and mitigation measures within the implementation of the high-risk AI system.
Type	Functional Requirement
Applicability	Restricted (High-risk AI systems with residual risks after application of risk management measures according to 9.9), given req. 9.1 is fulfilled
Category	Risk Management

ID	9.11 (O5, O12)
Description	The risk management system shall adopt risk management measures such that adequate information is provided to users about the risks associated with its purpose-conform use or reasonably foreseeable misuse of the high-risk AI system (see also requirement 13.4).
Rationale	Art. 9 (2)(d) + Art. 9 (4)(c); Risks may only be act upon when communicated to the user of the high-risk AI system
Difficulty	low
Fit Criterion	An expert user is provided with information according to requirements 13.1 through 13.10 about the risks associated with its purpose-conform use or reasonably foreseeable misuse before or shortly after beginning of their use.
Type	Functional Requirement
Applicability	All (High-risk AI systems), given req. 9.1 is fulfilled
Category	Risk Management

ID	9.12 (O5, O13)
Description	The risk management system shall adopt risk management measures such that adequate training, considering requirements 13.1 through 13.10, is provided to users.
Rationale	Art. 9 (2)(d) + Art. 9 (4)(c); Risks may only be act upon when the user of the high-risk AI system is proficient in dealing with them
Difficulty	medium
Fit Criterion	A user is provided with training before or shortly after beginning of their use.
Type	Functional Requirement
Applicability	Restricted (High-risk AI systems with risks for which training is appropriate), given req. 9.1 is fulfilled
Category	Risk Management

ID	9.13 (O5, O14)
Description	The risk management system shall adopt risk management measures such that in eliminating or reducing risks due consideration is given to the technical knowledge, experience, education, training to be expected by the user, and the environment in which the system is intended to be used.
Rationale	Art. 9 (2)(d) + Art. 9 (4); When acting upon risks in a high-risk AI system, the accumulated circumstances of use must be duly considered to allow the most accurate evaluation and the derive the most appropriate counter measures
Difficulty	high
Fit Criterion	The ways to eliminate and reduce risks of a high-risk AI system proposed by the risk management system are different between a target user with more and less technical proficiency.
Type	Process Requirement
Applicability	All (high-risk AI systems), given req. 9.1 is fulfilled
Category	Risk Management

ID	9.14 (O5, O15)
Description	The high-risk AI system shall be tested with the purpose of identifying appropriate risk management measures.
Rationale	Art. 9 (2)(d) + Art. 9 (5); Testing a high-risk AI system reveals the risks associated with its use that are hard to expect or predict
Difficulty	low
Fit Criterion	The risk management measures adopted in the finalised risk management system were informed by the results of a technical testing procedure performed on the high-risk AI system.
Type	Non-Functional Requirement
Applicability	All (high-risk AI systems), given req. 9.1 is fulfilled
Category	Testing

ID	9.15 (O15, O16)
Description	Testing procedures shall assess whether the high-risk AI system performs consistently for their intended purpose.
Rationale	Art. 9 (5); Only consistent performance of the intended objective renders an AI system reliable
Difficulty	medium
Fit Criterion	The test of multiple AI systems known to operate inconsistently showcases to the user that that is the case.
Type	Non-Functional Requirement
Applicability	All (high-risk AI systems), given req. 9.14, 15.2, and 15.3 are fulfilled
Category	Testing

ID	9.16 (O15, O18)
Description	The testing procedures shall be appropriate to the intended purpose of the high-risk AI system.
Rationale	Art. 9 (6); Testing sufficiently fulfils its intent when it relates to the intended purpose of the high-risk AI system
Difficulty	low
Fit Criterion	Each testing procedures corresponds to some aspect of the intended purpose of the high-risk AI system and all aspects of the intended purpose are covered by a test
Type	Non-Functional Requirement
Applicability	All (high-risk AI systems), given req. 9.14 is fulfilled
Category	Testing

ID	9.17 (O15, O19)
Description	The testing procedures and response to their results shall be performed before the high-risk AI system's entry into market or putting into service.
Rationale	Art. 9 (7); Testing only fulfils its intent when it allows to fix shortcomings before the high-risk AI system is used in production and affecting real users
Difficulty	low
Fit Criterion	The high-risk AI system on the market was tested beforehand.
Type	Process Requirement
Applicability	All (high-risk AI systems), given req. 9.14 is fulfilled
Category	Testing

ID	9.18 (O15, O20)
Description	The testing procedures shall be based on preliminarily defined metrics and probabilistic thresholds appropriate to the intended purpose of the high-risk AI system.
Rationale	Art. 9 (7); To ensure comparability and expressibility, testing of a high-risk AI system must be based in recognised metrics and threshold values of these metrics that determine the system's suitability
Difficulty	medium
Fit Criterion	The output of a test is presented in an industry-recognised metrics and a qualitative result associated with it is based on one or multiple threshold values of that metric
Type	Non-Functional Requirement
Applicability	All (high-risk AI systems), given req. 9.14 is fulfilled
Category	Testing

ID	9.19 (O15, O21)
Description	The risk management system shall assess and respond when the high-risk AI system is likely to be accessed by or have an impact on children.
Rationale	Art. 9 (8); A high-risk AI system affecting children imposes special risks on them that are required to be addressed and mitigated accordingly
Difficulty	medium
Fit Criterion	The output and/or behaviour of a risk management system assessing a high-risk AI system impacting children differs from that of assessing the same system without impact on children
Type	Functional Requirement
Applicability	All (high-risk AI systems), given req. 9.14 is fulfilled
Category	Risk Management

ID	9.20 (O15, O22)
Description	The risk management system shall form part of risk management procedures set out in article 74 of Directive 2013/36/EU.
Rationale	Art. 9 (9); The risks from high-risk AI systems add to intrinsic risks in the financial services industry and need to be jointly mitigated
Difficulty	low
Fit Criterion	The high-risk AI system's risk management system is included in the documentation of risk management measures and their output communicated to the authorities
Type	Process Requirement
Applicability	Restricted (high-risk AI systems deployed by credit institutions regulated by Directive 2013/36/EU), given req. 9.14 is fulfilled
Category	Risk Management

APPENDIX I.2: DATA AND DATA GOVERNANCE (ART. 10)

ID	10.1 (O3)
Description	Data governance and management practices shall concern relevant design choices (e.g., data features, AI system/data platform architecture).
Rationale	Art. 10 (2a); Design choices impact the quality and safety of the data sets which is needed to prevent attacks (e.g., adversarial examples, social engineering).
Difficulty	high
Fit Criterion	A group of experts identifies that data governance and management practices deal with relevant design choices or an appropriate standard (e.g., ISO/IEC JTC 1/SC 42) is used.
Type	Process Requirement
Applicability	All (high-risk AI systems)
Category	Dataset Properties

ID	10.2 (O4)
Description	Data governance and management practices shall concern the collection of data sets.
Rationale	Art. 10 (2b); The collection of data needs to comply with relevant data governance rules (e.g., GDPR).
Difficulty	low
Fit Criterion	The processes for collecting data comply with defined data governance rules and this is validated by a group of people responsible for data governance and management.
Type	Process Requirement
Applicability	All (high-risk AI systems)
Category	Dataset Properties

ID	10.3 (O5)
Description	Data governance and management practices shall concern relevant data preparation steps.
Rationale	Art. 10 (2c); Data preparation is a crucial step before the data is used in the AI system and all relevant operations on the data need to conform with data governance and management guidelines.
Difficulty	medium
Fit Criterion	The operations performed on the data sets during data preparation are developed and overseen by a group of experts or with use of an appropriate standard (e.g., ISO/IEC JTC 1/SC 42).
Type	Non-Functional Requirement
Applicability	All (high-risk AI systems)
Category	Dataset Properties

ID	10.4 (O6)
Description	Data governance and management practices shall concern assumptions made about the given data sets.
Rationale	Art. 10 (2d); Data governance and management practices ensure that any assumptions made regarding data are consistent over different data sets and use cases.
Difficulty	low
Fit Criterion	Any assumptions made regarding data are performed and overseen by a group of experts. Assumptions are made within the boundaries of the information the given data is supposed to measure and represent.
Type	Process Requirement
Applicability	All (high-risk AI systems)
Category	Dataset Properties

ID	10.5 (O7)
Description	Data governance and management practices shall concern the assessment of quality, availability, and suitability of the required data sets.
Rationale	Art. 10 (2e); Data governance and management practices ensure that any assumptions made regarding data are consistent over different data sets and use cases.
Difficulty	medium
Fit Criterion	All assessments regarding data are performed and overseen by a group of experts or with the use of an appropriate standard (e.g., ISO/IEC JTC 1/SC 42).
Type	Non-Functional Requirement
Applicability	All (high-risk AI systems)
Category	Dataset Properties

ID	10.6 (O8)
Description	Data governance and management practices shall concern the examination of biases in the data sets.
Rationale	Art. 10 (2f); Biases in the used data sets results in biased output of the AI system which can lead to flawed output and potentially discrimination of its users.
Difficulty	high
Fit Criterion	A group of experts verifies that current data governance and management practices can identify biases, or it is identified with the use of an appropriate standard (e.g., ISO/IEC JTC 1/SC 42).
Type	Non-Functional Requirement
Applicability	All (high-risk AI systems)
Category	Dataset Properties

ID	10.7 (O9)
Description	Data governance and management practices shall identify and address gaps and shortcomings in the data.
Rationale	Art. 10 (2g); Errors in the data sets can reduce the quality of the data, lead to biases and result in a flawed output of the system.
Difficulty	high
Fit Criterion	A group of experts verifies if current data governance and management practices can identify and address gaps and shortcomings, or it is identified with the use of an appropriate standard (e.g., ISO/IEC JTC 1/SC 42).
Type	Non-Functional Requirement
Applicability	All (high-risk AI systems)
Category	Dataset Properties

ID	10.8 (O11)
Description	Training, validation, and testing data sets shall be relevant, representative, free of errors and complete.
Rationale	Art. 10 (3); These flaws in the data sets can lead to biases, sampling errors and finally, a flawed output of the system.
Difficulty	high
Fit Criterion	A group of experts determines the correctness of the data sets based on demographic data of the persons the AI systems is used on and based on statistical analysis of the data, or it is identified with the use of an appropriate standard (e.g., ISO/IEC JTC 1/SC 42).
Type	Non-Functional Requirement
Applicability	Restricted (high-risk AI systems that perform model training with data)
Category	Dataset Properties

ID	10.9 (O12)
Description	Training, validation, and testing data sets shall have the appropriate statistical properties as regards users/groups of users.
Rationale	Art. 10 (3); Flaws in the data sets may lead to a flawed output of the system.
Difficulty	high
Fit Criterion	A group of experts performs statistical analysis to confirm that the datasets fulfil the required statistical properties, or it is identified with the use of an appropriate standard (e.g., ISO/IEC JTC 1/SC 42). Properties need to be applicable to the given use case and are only regarding the people it is intended to be used on.
Type	Non-Functional Requirement
Applicability	Restricted (high-risk AI systems that perform model training with data)
Category	Dataset Properties

ID	10.10 (O13)
Description	Training, validation, and testing data sets shall contain characteristics specific to the geographical, behavioural, or functional setting.
Rationale	Art. 10 (4); Data sets that are not representative of the AI systems' training data and specifically the environment in which the system is used in, may lead to a flawed output of the system.
Difficulty	high
Fit Criterion	A group of experts that have knowledge about the given setting the AI system is intended to be used in verify that the data sets fulfil these characteristics.
Type	Non-Functional Requirement
Applicability	Restricted (high-risk AI systems that perform model training with data)
Category	Dataset Properties

APPENDIX I.3: TECHNICAL DOCUMENTATION (ART. 11)

ID	11.1 (O1)
Description	A technical documentation shall exist for/within the high-risk AI system.
Rationale	Art. 11 (1); Authorities must be able to assess the compliance of the system with the help of the technical documentation.
Difficulty	medium
Fit Criterion	A technical documentation was drafted for the system.
Type	Non-functional Requirement
Applicability	All (high-risk AI systems)
Category	Technical Documentation

ID	11.2 (O2)
Description	The technical documentation shall be kept up to date with respect to any change that is introduced to the system.
Rationale	Art. 11 (1); The documentation needs to include every change that was made to the system.
Difficulty	medium
Fit Criterion	The accessible technical documentation contains every recent change.
Type	Process Requirement
Applicability	All (high-risk AI systems)
Category	Technical Documentation

ID	11.3 (O3)
Description	The technical documentation shall contain a general description of the AI system including its intended purpose, the person/s developing the system, the date, and the version of the system.
Rationale	Annex IV (1a); The documentation needs to be able to provide any authority with the needed basic information and complies with the standard of technical documentations regarding high-risk AI systems.
Difficulty	medium
Fit Criterion	The required specification is complete and included in the technical documentation.
Type	Non-functional Requirement
Applicability	All (high-risk AI systems)
Category	Technical Documentation

ID	11.4 (O4)
Description	The technical documentation shall contain how the AI system interacts or can be used to interact with hardware or software that is not part of the AI system itself.
Rationale	Annex IV (1b); The documentation needs to be able to provide any authority with the needed basic information and complies with the standard of technical documentations regarding high-risk AI systems.
Difficulty	medium
Fit Criterion	The required specification is complete and included in the technical documentation.
Type	Non-functional Requirement
Applicability	All (high-risk AI systems)
Category	Technical Documentation

ID	11.5 (O5)
Description	The technical documentation shall contain the versions of relevant software or firmware and any requirement related to version update.
Rationale	Annex IV (1c); The documentation needs to be able to provide any authority with the needed basic information and complies with the standard of technical documentations regarding high-risk AI systems.
Difficulty	medium
Fit Criterion	The required specification is complete and included in the technical documentation.
Type	Non-functional Requirement
Applicability	All (high-risk AI systems)
Category	Technical Documentation

ID	11.6 (O6)
Description	The technical documentation shall contain the description of all forms in which the AI system is placed on the market or put into service.
Rationale	Annex IV (1d); The documentation needs to be able to provide any authority with the needed basic information and complies with the standard of technical documentations regarding high-risk AI systems.
Difficulty	medium
Fit Criterion	The required specification is complete and included in the technical documentation.
Type	Non-Functional Requirement
Applicability	All (high-risk AI systems)
Category	Technical Documentation

ID	11.7 (O7)
Description	The technical documentation shall contain the description of hardware on which the AI system is intended to run.
Rationale	Annex IV (1e); The documentation needs to be able to provide any authority with the needed basic information and complies with the standard of technical documentations regarding high-risk AI systems.
Difficulty	medium
Fit Criterion	The required specification is complete and included in the technical documentation.
Type	Non-Functional Requirement
Applicability	All (high-risk AI systems)
Category	Technical Documentation

ID	11.8 (O8)
Description	The technical documentation shall contain where the AI system is a component of products, photographs or illustrations showing external features, marking and internal layout of those products.
Rationale	Annex IV (1f); The documentation needs to be able to provide any authority with the needed basic information and complies with the standard of technical documentations regarding high-risk AI systems.
Difficulty	medium
Fit Criterion	The required specification is complete and included in the technical documentation.
Type	Non-Functional Requirement
Applicability	All (high-risk AI systems)
Category	Technical Documentation

ID	11.9 (O9)
Description	The technical documentation shall contain instructions of use for the user as defined in the requirements 13.2- 13.10 and installation instructions.
Rationale	Annex IV (1g); The documentation needs to be able to provide any authority with the needed basic information and complies with the standard of technical documentations regarding high-risk AI systems.
Difficulty	medium
Fit Criterion	A user that accesses the technical documentation accessibly finds the instructions.
Type	Non-Functional Requirement
Applicability	All (high-risk AI systems)
Category	Technical Documentation

ID	11.10 (O10)
Description	The technical documentation shall contain a detailed description of the system development process, which needs to include all methods and steps that were performed, and all used pre-trained systems or third-party tools and how they have been used, integrated, or modified.
Rationale	Annex IV (2a); The documentation needs to be able to provide any authority with detailed information and comply with the standard of technical documentations regarding high-risk AI systems.
Difficulty	medium
Fit Criterion	The required specification is complete and included in the technical documentation.
Type	Non-Functional Requirement
Applicability	All (high-risk AI systems)
Category	Technical Documentation

ID	11.11 (O11)
Description	The technical documentation shall contain a detailed description about the design specifications of the system, namely the general logic of the AI system and of the algorithms; the key design choices including the rationale and assumptions made, also in terms of the people the system will be used on; the main classification choices; what the system is designed to optimize for and the relevance of the different parameters; decisions about any possible trade-off made to comply with other the other requirements in this Requirements Specification.
Rationale	Annex IV (2b); The documentation needs to be able to provide any authority with detailed information and comply with the standard of technical documentations regarding high-risk AI systems.
Difficulty	medium
Fit Criterion	The required specification is complete and included in the technical documentation.
Type	Non-Functional Requirement
Applicability	All (high-risk AI systems)
Category	Technical Documentation

ID	11.12 (O12)
Description	The technical documentation shall contain a detailed description of the systems architecture, explaining how software components build on or feed into each other and integrate into the overall processing and the computational resources used to develop, train, test and validate the AI system.
Rationale	Annex IV (2c); The documentation needs to be able to provide any authority with detailed information and comply with the standard of technical documentations regarding high-risk AI systems.
Difficulty	medium
Fit Criterion	The required specification is complete and included in the technical documentation.
Type	Non-Functional Requirement
Applicability	All (high-risk AI systems)
Category	Technical Documentation

ID	11.13 (O13)
Description	The technical documentation shall contain a detailed description about the data requirements in terms of datasheets describing the training methodologies and techniques and the training data sets used, including information about the provenance of those data sets, their scope, and main characteristics; how the data was obtained and selected; labelling procedures (e.g., for supervised learning), data cleaning methodologies (e.g., outlier detection).
Rationale	Annex IV (2d); The documentation needs to be able to provide any authority with detailed information and comply with the standard of technical documentations regarding high-risk AI systems.
Difficulty	medium
Fit Criterion	The required specification is complete and included in the technical documentation.
Type	Non-Functional Requirement
Applicability	All (high-risk AI systems)
Category	Technical Documentation

ID	11.14 (O14)
Description	The technical documentation shall contain a detailed description about the assessment of the demanded human oversight measures and the necessary technical measures to facilitate the interpretation of the outputs of AI systems by the users.
Rationale	Annex IV (2e); The documentation needs to be able to provide any authority with detailed information and comply with the standard of technical documentations regarding high-risk AI systems.
Difficulty	medium
Fit Criterion	The required specification is complete and included in the technical documentation.
Type	Non-Functional Requirement
Applicability	All (high-risk AI systems)
Category	Technical Documentation

ID	11.15 (O15)
Description	The technical documentation shall contain a detailed description of pre-determined changes to the AI system and its performance, together with all the relevant information related to the technical solutions adopted to ensure continuous compliance of the AI system with the relevant requirements in this Requirements Specification.
Rationale	Annex IV (2f); The documentation needs to be able to provide any authority with detailed information and comply with the standard of technical documentations regarding high-risk AI systems.
Difficulty	medium
Fit Criterion	The required specification is complete and included in the technical documentation.
Type	Non-Functional Requirement
Applicability	All (high-risk AI systems)
Category	Technical Documentation

ID	11.16 (O16)
Description	The technical documentation shall contain the validation and testing procedures used in the development of the system, including information about the validation and testing data used and their main characteristics. This also includes metrics used to measure accuracy, robustness, cybersecurity, and compliance as well as potentially discriminatory impacts. In addition to this, Test logs and all test reports dated and signed by the persons responsible.
Rationale	Annex IV (2g); The documentation needs to be able to provide any authority with detailed information and comply with the standard of technical documentations regarding high-risk AI systems.
Difficulty	medium
Fit Criterion	The required specification is complete and included in the technical documentation.
Type	Non-functional Requirement
Applicability	All (high-risk AI systems)
Category	Technical Documentation

ID	11.17 (O17)
Description	The technical documentation shall contain detailed information about the monitoring, functioning and control of the High-risk AI system, in particular with regard to: its capabilities and limitations in performance, including the degrees of accuracy for specific persons or groups of persons on which the system is intended to be used and the overall expected level of accuracy in relation to its intended purpose, as well as the foreseeable unintended outcomes and sources of risks to health and safety, fundamental rights and discrimination in view of the intended purpose of the AI system; specifications on input data.
Rationale	Annex IV (3); The documentation needs to be able to provide any authority with detailed information and comply with the standard of technical documentations regarding high-risk AI systems.
Difficulty	medium
Fit Criterion	The required specification is complete and included in the technical documentation.
Type	Non-functional Requirement
Applicability	All (high-risk AI systems)
Category	Technical Documentation

ID	11.18 (O18)
Description	The technical documentation shall contain a detailed description of the risk management system in accordance with the requirements 9.1 - 9.14.
Rationale	Annex IV (4); The documentation needs to be able to provide any authority with detailed information and comply with the standard of technical documentations regarding high-risk AI systems.
Difficulty	medium
Fit Criterion	The required specification is complete and included in the technical documentation.
Type	Non-Functional Requirement
Applicability	All (high-risk AI systems)
Category	Technical Documentation

ID	11.19 (O19)
Description	The technical documentation shall contain a description of any change made to the system through its lifecycle.
Rationale	Annex IV (5); The documentation needs to be able to provide any authority with needed basic information and comply with the standard of technical documentations regarding high-risk AI systems.
Difficulty	medium
Fit Criterion	The required specification is complete and included in the technical documentation.
Type	Non-Functional Requirement
Applicability	All (high-risk AI systems)
Category	Technical Documentation

ID	11.20 (O20)
Description	The technical documentation shall contain a list of the harmonized standards applied in full or in part the references of which have been published in the Official Journal of the European Union; where no such harmonized standards have been applied, a detailed description of the solutions adopted to meet the requirements, including a list of other relevant standards and technical specifications applied.
Rationale	Annex IV (6); The documentation needs to be able to provide any authority with detailed information and comply with the standard of technical documentations regarding high-risk AI systems.
Difficulty	medium
Fit Criterion	The required specification is complete and included in the technical documentation.
Type	Non-functional Requirement
Applicability	All (high-risk AI systems)
Category	Technical Documentation

ID	11.21 (O21)
Description	The technical documentation shall contain a copy of the EU declaration of conformity.
Rationale	Annex IV (7); The documentation needs to be able to provide any authority with needed basic information and comply with the standard of technical documentations regarding high-risk AI systems.
Difficulty	low
Fit Criterion	The required specification is complete and included in the technical documentation.
Type	Non-Functional Requirement
Applicability	All (high-risk AI systems)
Category	Technical Documentation

ID	11.22 (O22)
Description	The technical documentation shall contain a detailed description of the system in place to evaluate the AI system performance in the post-market monitoring system.
Rationale	Annex IV (8); The documentation needs to be able to provide any authority with detailed information and comply with the standard of technical documentations regarding high-risk AI systems.
Difficulty	medium
Fit Criterion	The required specification is complete and included in the technical documentation.
Type	Non-Functional Requirement
Applicability	All (high-risk AI systems)
Category	Technical Documentation

ID	11.23 (O3)
Description	The technical documentation of the high-risk AI system is combined with all the other information that is legally required to form one single technical documentation. A high-risk AI system that enters the market and is related to a product, to which the legal acts listed in Annex II, section A apply, only one single technical documentation is needed.
Rationale	Art. 11 (2); By abiding by this requirement, redundancies are avoided.
Difficulty	medium
Fit Criterion	The technical documentation of the high-risk AI system is combined with the one of the related products and can be accessed in one document.
Type	Process Requirement
Applicability	Restricted (high-risk AI systems related to a product for which the EU has already set harmonized standards.)
Category	Technical Documentation

APPENDIX I.4: RECORD KEEPING (ART. 12)

ID	12.1(O1)
Description	The high-risk AI system shall possess automatic event-recording capabilities.
Rationale	Art. 12 (1); The performance of a high-risk AI system must be reviewable in order to be trusted
Difficulty	medium
Fit Criterion	For any point in time during the operation of the high-risk AI system, its records may be accessed by a user
Type	Functional Requirement
Applicability	All (high-risk AI systems)
Category	Record Keeping

ID	12.2 (O2)
Description	All events during the AI system's entire lifecycle operation shall be recorded in a way that ensures traceability with respect to the intended purpose of the system.
Rationale	Art. 12 (2); Depending on the type of AI system, the events governing its decisions and outputs must be reviewable and understandable by an independent party.
Difficulty	low
Fit Criterion	Reviewing all steps and events of a system's operation period in the past allows a third party not present during operation to understand the system's behaviour and decisions during that period.
Type	Functional Requirement
Applicability	All (high-risk AI systems), given req. 12.1 is fulfilled
Category	Record Keeping

ID	12.3 (O1)
Description	The event-recording capability shall create and maintain its records according to an industry-acknowledged standard or common practice.
Rationale	Art. 12 (1); The records need to be interchangeable.
Difficulty	low
Fit Criterion	The records returned by the system fulfil the standard as determined by an expert user.
Type	Functional Requirement
Applicability	All (high-risk AI systems), given req. 12.1 is fulfilled
Category	Record Keeping

ID	12.4 (O4a)
Description	Depending on the type of AI system, the data provided by users or through other sources during operation shall be automatically, and systematically collected and documented such that they can be assessed against the present Requirements Specification.
Rationale	Art. 12 (3) + Art 61 (2); Depending on the type of AI system, the events governing its decisions and outputs must be reviewable and understandable by an independent party that may verify its compliance with applicable regulations.
Difficulty	medium
Fit Criterion	After a period of operation of the system, the automatically recorded, structured respective data provided in that period may be accessed by a competent user through an interface.
Type	Functional Requirement
Applicability	All (high-risk AI systems), given req. 12.1 is fulfilled
Category	Record Keeping

ID	12.5 (O4a)
Description	The data provided by users or through other sources during operation shall be automatically, and systematically analysed.
Rationale	Art. 12 (3) + Art. 61 (2); Depending on the type of AI system, the events governing its decisions and outputs must be automatically reviewed to highlight potential risks and weaknesses.
Difficulty	Medium to high
Fit Criterion	After a period of operation of the system, the automatically created, structured analysis of the respective data provided in that period may be accessed by a competent user through an interface.
Type	Functional Requirement
Applicability	All (high-risk AI systems), given req. 12.1 is fulfilled
Category	Record Keeping

ID	12.6 (O4b)
Description	A post-market monitoring plan shall be established that governs the specifics of 12.4 and 12.5.
Rationale	Art. 12 (3) + Art. 61(3); The monitoring procedure needs to be documented and reviewable to be deemed appropriate and compliant
Difficulty	low
Fit Criterion	A monitoring plan according to the template by the European Commission is included in the technical documentation of the system and adhered to.
Type	Process Requirement
Applicability	All (high-risk AI systems), given req. 12.1, 12.4, 12.5 are fulfilled
Category	Record Keeping

ID	12.7 (O3)
Description	The records of the logging capability are appropriate to monitor situations where the system a) may impose a risk to health or safety or the fundamental rights of persons or b) may lead to a substantial modification of itself.
Rationale	Art. 12 (3); Detailed review of operation periods of an AI system that are critical in the sense of previous legislation must be possible.
Difficulty	low
Fit Criterion	After occurrence of a relevant situation, the detailed records may be examined by a user through an interface
Type	Functional Requirement
Applicability	All (high-risk AI systems)
Category	Record Keeping

ID	12.8 (O5)
Description	The logging records shall include the period of each use.
Rationale	Art. 12 (4)(a); Detailed review of operation of a high-risk AI system dealing with biomedical data of human beings is critical
Difficulty	low
Fit Criterion	The periods of all past usages of the system may be examined by a user through an interface.
Type	Functional Requirement
Applicability	Restricted (high-risk AI systems intended to be used for the ‘real-time’ and ‘post’ remote biometric identification of natural persons)
Category	Record Keeping

ID	12.9 (O6)
Description	The logging records shall include the database against which the input to the model is assessed.
Rationale	Art. 12 (4)(b); Detailed review of operation of a high-risk AI system dealing with biomedical data of human beings is critical
Difficulty	low
Fit Criterion	The of the reference databases in all past usages of the system may be examined by a user through an interface.
Type	Functional Requirement
Applicability	Restricted (high-risk AI systems intended to be used for the ‘real-time’ and ‘post’ remote biometric identification of natural persons)
Category	Record Keeping

ID	12.10 (O7)
Description	The logging records shall include the input data for which the model found determined a search match.
Rationale	Art. 12 (4)(c); Detailed review of operation of a high-risk AI system dealing with biomedical data of human beings is critical
Difficulty	low
Fit Criterion	The input data points in all past biometrical identification processes carried out in the system may be examined by a user through an interface.
Type	Functional Requirement
Applicability	Restricted (high-risk AI systems intended to be used for the ‘real-time’ and ‘post’ remote biometric identification of natural persons)
Category	Record Keeping

ID	12.11 (O8)
Description	The logging records shall include the identification of the human overseer accountable according to requirements 14.5 to 14.9 during the operation of the system shall be recorded.
Rationale	Art. 12 (4)(d); Detailed review of operation of a high-risk AI system dealing with biomedical data of human beings is critical
Difficulty	low
Fit Criterion	The identification of the human overseer in all past usages of the system may be examined by a user through an interface.
Type	Functional Requirement
Applicability	Restricted (high-risk AI systems intended to be used for the 'real-time' and 'post' remote biometric identification of natural persons)
Category	Record Keeping

APPENDIX I.5: TRANSPARENCY AND PROVISION OF INFORMATION TO USERS (ART. 13)

ID	13.1 (O1)
Description	The operations executed by the AI system shall be sufficiently transparent for users to be able to interpret and appropriately use the system output.
Rationale	Art. 13 (1); The users must be able to work productively with the system outputs, and for this it is essential that they are able to trace the creation of these outputs.
Difficulty	medium
Fit Criterion	The system operations are transparent to a degree that the user can comprehend the system output.
Type	Functional Requirement
Applicability	All (high-risk AI systems)
Category	Explainability

ID	13.2 (O2)
Description	The High-Risk AI System shall be accompanied by instructions for use, in an appropriate digital format or otherwise that include concise, complete, correct, and clear information.
Rationale	Art. 13 (2); Users need relevant, accessible, and comprehensible instructions when interacting with the system.
Difficulty	medium
Fit Criterion	Every kind of instruction that is required follows this quality standard. Once the system is available to the market, every user can access a guide of instructions in which he does not miss any information he deems relevant and in which nothing is contained he deems superfluous.
Type	Process Requirement
Applicability	All (high-risk AI systems)
Category	Explainability

ID	13.3 (O3a)
Description	There shall be instructions about the identity and the contact details of the provider and its authorised representative, given there is one.
Rationale	Art. 13 (3) (a); The user should be given the opportunity to reach out to a contact person, whether for technical or legal questions.
Difficulty	low
Fit Criterion	The required instructions can be obtained and meet the quality standard for instructions from requirement 13.2.
Type	Process Requirement
Applicability	All (high-risk AI systems)
Category	Explainability

ID	13.4 (O3b, O3d)
Description	There shall be instructions about the intended purpose of the high-risk AI System and any known or foreseeable circumstances which may lead to risks to health and safety or fundamental rights when the system is used as intended or misused.
Rationale	Art. 13 (3) (b) (i) & (iii); The user should know about the scope and non-scope of the system and be informed about possible hazardous situations.
Difficulty	medium
Fit Criterion	The required instructions can be obtained and meet the quality standard for instructions from requirement 13.2.
Type	Process Requirement
Applicability	All (high-risk AI systems)
Category	Explainability

ID	13.5 (O3c, O3d)
Description	There shall be instructions about the tested and validated level of accuracy, robustness, and cybersecurity and any known or foreseeable circumstances which could impact these levels.
Rationale	Art. 13 (3) (b) (ii); The levels are intended to show the user how susceptible the system could be to errors.
Difficulty	medium
Fit Criterion	The required instructions can be obtained and meet the quality standard for instructions from requirement 13.2.
Type	Process Requirement
Applicability	All (high-risk AI systems)
Category	Explainability

ID	13.6 (O3e)
Description	There shall be instructions about the performance of the High-Risk AI system as regards its intended use cases.
Rationale	Art. 13 (3) (b) (iv); Users are informed of the default system performance for intended use.
Difficulty	low
Fit Criterion	The required instructions can be obtained and meet the quality standard for instructions from requirement 13.2.
Type	Process Requirement
Applicability	All (high-risk AI systems)
Category	Explainability

ID	13.7 (O3f)
Description	There shall be instructions about specifications for the input data, or any other relevant information in terms of the training, validation and testing data sets used.
Rationale	Art. 13 (3) (b) (v); Transparency about which data sets are processed for which purpose.
Difficulty	low
Fit Criterion	The required instructions can be obtained and meet the quality standard for instructions from requirement 13.2.
Type	Process Requirement
Applicability	Restricted (high-risk AI systems using input data or data sets when operating according to their intended use)
Category	Explainability

ID	13.8 (O3g)
Description	There shall be instructions about changes to the High-Risk AI system and its performance that were made after the initial conformity assessment.
Rationale	Art. 13 (3) (c); Timeliness and completeness of the other instruction requirements.
Difficulty	low
Fit Criterion	The required instructions can be obtained and meet the quality standard for instructions from requirement 13.2.
Type	Process Requirement
Applicability	All (high-risk AI systems)
Category	Explainability

ID	13.9 (O3h)
Description	There shall be instructions about the human oversight measures, including the applied technical measures.
Rationale	Art. 13 (3) (d); In article 14, human oversight measures are introduced, as necessary. By communicating the taken measures to the user, he may be able to better understand the system output.
Difficulty	medium
Fit Criterion	The required instructions can be obtained and meet the quality standard for instructions from requirement 13.2.
Type	Process Requirement
Applicability	All (high-risk AI systems)
Category	Explainability

ID	13.10 (O3i)
Description	There shall be instructions about the expected lifetime and any measures to ensure proper functioning.
Rationale	Art. 13 (3) (e); The user should be shown that appropriate steps are being taken to maintain the system until the end of its life cycle.
Difficulty	low
Fit Criterion	The required instructions can be obtained and meet the quality standard for instructions from requirement 13.2.
Type	Process Requirement
Applicability	All (high-risk AI systems)
Category	Explainability

APPENDIX I.6: HUMAN OVERSIGHT (ART. 14)

ID	14.1 (O1, O5)
Description	High-risk AI systems shall operate such that they can be effectively overseen by a natural person.
Rationale	Art. 14 (1); Accountability and sensitivity of the context requires the possibility of human intervention.
Difficulty	medium
Fit Criterion	Human oversight is ensured and included by design in the high-risk AI system and a human-machine interface tool can be used, the level of implementation is confirmed by an expert group.
Type	Non-Functional Requirement
Applicability	All (high-risk AI systems)
Category	Human Oversight

ID	14.2 (O2)
Description	The high-risk AI system shall integrate human oversight with the aim of preventing or minimising the risks to health, safety or fundamental rights caused by the active high-risk AI system, within its boundaries of intended purpose and under conditions of foreseeable misuse.
Rationale	Art. 14(2); Protection of human health from potential harm caused by AI systems.
Difficulty	high
Fit Criterion	The individual responsible for human oversight is able to prevent or mitigate foreseeable misuse and risk of high-risk AI systems within the scope of its intended purpose, with respect to its consequence on preservation of health, safety, or fundamental rights.
Type	Process Requirement
Applicability	All (high-risk AI systems)
Category	Human Oversight

ID	14.3 (O3)
Description	High-risk AI systems shall integrate human oversight interfaces before they are placed on or used in the market.
Rationale	Art. 14 (3a, 3b); Interfaces provide easy access to non-technical experts and allow more direct control over the AI systems with respect to interpretation and stopping mechanisms.
Difficulty	medium
Fit Criterion	Human oversight is implemented in high-risk AI systems at the point when it is ready to enter the market or put into productive service or are accompanied and outlined by the provider via instructions, so that users must implement and perform the oversight themselves.
Type	Process Requirement
Applicability	All (high-risk AI systems)
Category	Human Oversight

ID	14.4 (O4)
Description	High-risk AI systems shall operate such that the limitations and capacities of the system are clearly outlined and understood by the individuals responsible for human oversight, deviations must be detected, investigated, and properly addressed.
Rationale	Art. 14 (4a); The user must know in which scenarios, with what data and how to use the high AI system, so that misuse can be prevented.
Difficulty	high
Fit Criterion	The individual responsible for human oversight confirms their understanding of the limitations and capacities of the high-risk AI system and their ability to respond to anomalies, dysfunctions, and unexpected performance.
Type	Process Requirement
Applicability	All (high-risk AI systems)
Category	Human Oversight

ID	14.5 (O6)
Description	The high-risk AI systems shall operate such that the individuals responsible for human oversight are not over-relying on the system (automation bias) with respect to predictions and other decisions made by the system.
Rationale	Art. 14 (4b); Overreliance and consequent inattention regarding the produced output of the AI system may lead to wrong decisions as the output of the system can be flawed.
Difficulty	high
Fit Criterion	A group of experts determines that the functions and the mode of operation of the high-risk AI system sufficiently prevents its users from over-relying on its output, for instance through provision of information and warning.
Type	Non-Functional Requirement
Applicability	All (high-risk AI systems)
Category	Human Oversight

ID	14.6 (O7)
Description	The high-risk AI system shall operate such that the produced outputs and the systems' logic are transparent and can be interpreted via tools and methods by the individuals responsible for human oversight.
Rationale	Art. 14 (4c); The supervising user needs to understand how the inputs map to the outputs to prevent using a "black-box" system.
Difficulty	high
Fit Criterion	A group of experts of experts determines that the characteristics of the system are transparent, and the corresponding interpretation tools and methods are understood by the individuals responsible for human oversight.
Type	Non-Functional Requirement
Applicability	All (high-risk AI systems)
Category	Human Oversight

ID	14.7 (O8)
Description	The decisions of the high-risk AI system shall be such that they can be disregarded, overwritten, and reversed in any situation by the individuals responsible for human oversight.
Rationale	Art. 14(4d); The possibility of intervention must be guaranteed due to potential erroneous decisions arising from the results produced by the AI system.
Difficulty	low
Fit Criterion	A group of experts determines that the implementation is satisfactory regarding the ability to disregard, overwrite and reverse the decision of the high-risk AI system.
Type	Functional Requirement
Applicability	All (high-risk AI systems)
Category	Human Oversight

ID	14.8 (O9)
Description	The high-risk AI system shall operate such that the individuals responsible for human oversight can at any point interrupt or halt the program with a single procedure.
Rationale	Art. 14 (4e); The possibility of intervention must be guaranteed due to potential erroneous decisions arising from the results produced by the AI system and concomitant harm that could be caused.
Difficulty	low
Fit Criterion	A group of experts determines that the implementation is satisfactory regarding the ability to immediately stop the high-risk AI system.
Type	Functional Requirement
Applicability	All (high-risk AI systems)
Category	Human Oversight

ID	14.9 (O10)
Description	The high-risk AI system shall operate such that its decisions with respect to identification, assignment, and assessment of natural persons in educational and vocational training institutions, are confirmed by at least two natural persons.
Rationale	Art. 14 (5); Over-reliance on decisions made by AI systems in critical environments must be confirmed by natural persons, to mitigate bias and ensure an equal and fair treatment of natural persons.
Difficulty	low
Fit Criterion	At least two natural persons confirm the decisions of high-risk AI systems in the context of educational and vocational training institutions, this includes determining access of natural persons to educational and vocational training institutions or assigning natural persons thereto, assessing students in test and assessing participants in test commonly required for admission to educational institutions.
Type	Functional Requirement
Applicability	Restricted (AI systems intended to be used for the 'real-time' and 'post' remote biometric identification of natural persons.)
Category	Human Oversight

APPENDIX I.7: ACCURACY, ROBUSTNESS AND CYBERSECURITY (ART. 15)

ID	15.1
Description	Appropriate levels and metrics for the high-risk AI system's accuracy, robustness and cybersecurity shall be defined, tested, and validated.
Rationale	Art. 13 (3); Art. 15 (1); Users and maintainers need individually defined levels to verify appropriate accuracy, robustness, and cybersecurity
Difficulty	high
Fit Criterion	Clear and appropriate levels and metrics regarding the system's accuracy, robustness and cybersecurity are defined, tested, and validated based on the individual context or a commonly recognized standard.
Type	Process Requirement
Applicability	All (high-risk AI systems)
Category	Accuracy, Robustness, Cybersecurity

ID	15.2 (O1, O2)
Description	The high-risk AI system shall operate with the defined (see 15.2), consistent level of accuracy throughout its lifecycle, appropriate to its intended purpose.
Rationale	Art. 15 (1); A low or inconsistent level of accuracy poses a risk to the quality of the systems output.
Difficulty	high
Fit Criterion	The level complies with the defined specifications deemed appropriate by a subject matter expert.
Type	Non-Functional Requirement
Applicability	All (high-risk AI systems)
Category	Accuracy, Robustness, Cybersecurity

ID	15.3 (O1, O2)
Description	The high-risk AI system shall operate with the defined (see 15.2), consistent level of robustness throughout its lifecycle, appropriate to its intended purpose.
Rationale	Art. 15 (1); A low or inconsistent level of robustness poses a risk to the performance and reliability of the system.
Difficulty	high
Fit Criterion	The level complies with the defined specifications deemed appropriate by a subject matter expert.
Type	Non-Functional Requirement
Applicability	All (high-risk AI systems)
Category	Accuracy, Robustness, Cybersecurity

ID	15.4 (O1, O2)
Description	The high-risk AI system shall operate with the defined (see 15.2), consistent level of cybersecurity throughout its lifecycle, appropriate to its intended purpose.
Rationale	Art. 15 (1); A low or inconsistent level of cybersecurity poses a risk to the integrity and safety of the system.
Difficulty	high
Fit Criterion	The level complies with the defined specifications deemed appropriate by a subject matter expert.
Type	Non-Functional Requirement
Applicability	All (high-risk AI systems)
Category	Accuracy, Robustness, Cybersecurity

ID	15.5 (O3)
Description	The levels of accuracy and the respective metrics shall be declared in the accompanying instructions of use.
Rationale	Art. 15 (2); As regards quality assurance and control, users need to understand what levels of accuracy are considered acceptable.
Difficulty	low
Fit Criterion	Test-users confirm to understand all relevant metrics and levels of accuracy by consulting the instructions of use.
Type	Process Requirement
Applicability	All (high-risk AI systems)
Category	Accuracy, Robustness, Cybersecurity

ID	15.6 (O4)
Description	The high-risk AI system shall identify and mitigate or prevent errors, faults or inconsistencies within the system or its operational environment. This may be achieved through technical redundancy solutions.
Rationale	Art. 15 (3); Errors, faults or inconsistencies can pose a threat in particular towards interacting natural persons or other systems.
Difficulty	high
Fit Criterion	Testing metrics prove a high resiliency towards errors, faults, or inconsistencies.
Type	Non-Functional Requirement
Applicability	All (high-risk AI systems)
Category	Accuracy, Robustness, Cybersecurity

ID	15.7 (O5)
Description	The high-risk AI system shall duly address possibly biased outputs through ‘feedback loops’ with appropriate mitigation measures.
Rationale	Art. 15 (3); The quality and functioning of a system can be compromised by feedback loops creating biased outputs.
Difficulty	medium
Fit Criterion	Testing metrics prove a low susceptibility to biased outputs and feedback loops.
Type	Non-functional Requirement
Applicability	Restricted (high-risk AI systems that continue to learn in production)
Category	Accuracy, Robustness, Cybersecurity

ID	15.8 (O6)
Description	The high-risk AI system shall identify and mitigate or prevent attempts by unauthorised third parties to alter their use or performance.
Rationale	Art. 15 (4); Malevolent third parties can cause great damage by manipulating AI systems.
Difficulty	medium
Fit Criterion	A sufficient level of attacks performed for testing purposes is successfully identified and mitigated or prevented by the system.
Type	Non-Functional Requirement
Applicability	All (high-risk AI systems)
Category	Accuracy, Robustness, Cybersecurity

ID	15.9 (O7)
Description	The technical solutions aimed at ensuring the cybersecurity of high-risk AI systems shall be appropriate to the relevant circumstances and risks.
Rationale	Art. 15 (4); With respect to the cost of risk, the measures must be chosen appropriately.
Difficulty	medium
Fit Criterion	The measures comply with the risk metrics defined by the risk management system.
Type	Process Requirement
Applicability	All (high-risk AI systems)
Category	Accuracy, Robustness, Cybersecurity

ID	15.10 (O)
Description	The technical solutions addressing AI specific vulnerabilities shall include measures to prevent and control for attacks trying to manipulate the training dataset (‘data poisoning’).
Rationale	Art. 15 (4); Manipulation of training datasets is a common way to interfere with an AI system.
Difficulty	medium
Fit Criterion	A sufficient level of ‘data poisoning’ attacks performed for testing purposes is successfully identified and mitigated or prevented by the system.
Type	Non-Functional Requirement
Applicability	Restricted (high-risk AI systems exposed to AI specific vulnerabilities)
Category	Accuracy, Robustness, Cybersecurity

ID	15.11 (O9)
Description	The technical solutions addressing AI specific vulnerabilities shall include measures to prevent and control for inputs designed to cause the model to make a mistake (‘adversarial examples’).
Rationale	Art. 15 (4); Malicious inputs are a common way to interfere with an AI system.
Difficulty	medium
Fit Criterion	A sufficient level of ‘adversarial example’ attacks performed for testing purposes is successfully identified and mitigated or prevented by the system.
Type	Non-Functional Requirement
Applicability	Restricted (high-risk AI systems exposed to AI specific vulnerabilities)
Category	Accuracy, Robustness, Cybersecurity

ID	15.12 (O10)
Description	The technical solutions addressing AI specific vulnerabilities shall include measures to prevent and control for model flaws.
Rationale	Art. 15 (4); Unidentified and uncontrolled model flaws can significantly compromise the systems quality.
Difficulty	medium
Fit Criterion	The high-risk AI system is continuously controlled for model flaws and shows a low rate of such.
Type	Non-Functional Requirement
Applicability	Restricted (high-risk AI systems exposed to AI specific vulnerabilities)
Category	Accuracy, Robustness, Cybersecurity

Appendix II.1: Identified Software Solutions - Testing

Framework	Publisher	Category	Original Paper	Scholar Results	Alternative Source
Azure Machine Learning	Microsoft	Testing	N/A	3.110	https://docs.microsoft.com/en-us/azure/architecture/data-science-process/deploy-models-in-production#ab-testing df
Amazon SageMaker	Amazon	Testing	N/A	1.210	https://docs.aws.amazon.com/sagemaker/latest/dg/model-ab-testing.html https://aws.amazon.com/de/blogs/machine-learning/a-b-testing-ml-models-in-production-using-amazon-sagemaker/
IBM Watson OpenScale	IBM	Testing	N/A	36	https://dataplatform.cloud.ibm.com/docs/content/wsj/model/getting-started.html

Appendix II.2: Identified Software Solutions - Dataset Properties

Framework	Publisher	Category	Original Paper	Scholar Results	Alternative Source
IBM SPSS Modeler	IBM	Dataset Properties	N/A	4.840	https://www.ibm.com/products/spss-modeler
SAP Data Services	SAP	Dataset Properties	N/A	142	https://www.sap.com/products/data-services.html https://www.altexsoft.com/blog/data-quality-management-and-tools/
Informatica Data Quality	Informatica	Dataset Properties	N/A	93	https://www.informatica.com/products/data-quality/informatica-data-quality.html https://www.altexsoft.com/blog/data-quality-management-and-tools/
SAS Data Quality	SAS	Dataset Properties	N/A	82	https://www.sas.com/en_us/software/data-quality.html https://www.researchgate.net/publication/220102796_A_Survey_of_Data_Quality_Tools
TensorFlow Data Validation (TFDV)	Google	Dataset Properties	E. Caveness, P. S. GC, Z. Peng, N. Polyzotis, S. Roy, and M. Zinkevich,"Tensorflow data validation: Data analysis and validation in continuousml pipelines," inProceedings of the 2020 ACM SIGMOD InternationalConference on Management of Data, 2020, pp. 2793–2796.	39	https://www.tensorflow.org/tfx/guide/tfdv
IBM Infosphere Information Server for Data Quality	IBM	Dataset Properties	N/A	26	https://www.ibm.com/products/infosphere-info-server-for-datamgmt https://www.altexsoft.com/blog/data-quality-management-and-tools/
IBM Watson Studio AutoAI	IBM	Dataset Properties	D. Wang, P. Ram, D. K. I. Weidele, S. Liu, M. Muller, J. D. Weisz,A. Valente, A. Chaudhary, D. Torres, H. Samulowitzet al., "Autoai:Automating the end-to-end ai lifecycle with humans-in-the-loop," inProceedings of the 25th International Conference on Intelligent UserInterfaces Companion, 2020, pp. 77–78.	13	https://www.ibm.com/cloud/watson-studio/autoai
Trillium Quality	Precisely	Dataset Properties	N/A	6	https://www.precisely.com/product/precisely-trillium/trillium-quality https://www.researchgate.net/publication/220102796_A_Survey_of_Data_Quality_Tools
Amazon SageMaker Data Wrangler	Amazon	Dataset Properties	N/A	3	https://aws.amazon.com/sagemaker/data-wrangler/
Amazon Web Services Glue DataBrew	Amazon	Dataset Properties	N/A	1	https://aws.amazon.com/glue/features/databrew/
IBM InfoSphere Advanced Data Preparation	IBM	Dataset Properties	N/A	0	https://www.ibm.com/de-de/products/infosphere-advanced-data-preparation

Appendix II.3: Identified Software Solutions - Record Keeping

Framework	Publisher	Category	Original Paper	Scholar Results	Alternative Source
TensorBoard	Google	Record Keeping	N/A	3.800	https://www.tensorflow.org/tensorboard https://colab.research.google.com/github/tensorflow/tensorboard/blob/master/docs/scalars_and_keras.ipynb
Amazon CloudWatch	Amazon	Record Keeping	N/A	1.670	https://geekflare.com/ai-frameworks/
Datadog	Datadog	Record Keeping	N/A	658	https://www.datadoghq.com/
SolarWinds Loggly	Loggly	Record Keeping	N/A	14	https://sematext.com/blog/log-analysis-tools/
Amazon SageMaker Model Monitor	Amazon	Record Keeping	N/A	13	https://docs.aws.amazon.com/sagemaker/latest/dg/model-monitor.html
Sematext Logs	Sematext	Record Keeping	N/A	1	https://sematext.com/blog/log-analysis-tools/
Neptune.ai	Neptune	Record Keeping	N/A	0	https://neptune.ai/product

Appendix II.4: Identified Software Solutions - Explainability

Framework	Publisher	Category	Original Paper	Scholar Results	Alternative Source
SHapley Additive exPlanations (SHAP)	Lundberg, S. and Lee, S.	Explainability	S. M. Lundberg and S.-I. Lee, "A unified approach to interpreting modelpredictions," inProceedings of the 31st international conference onneural information processing systems, 2017, pp. 4768–4777.	9.080	N/A
Local Interpretable Model-Agnostic Explanation (LIME)	Ribero, M. and Singh, S. and Guestrin C.	Explainability	M. T. Ribeiro, S. Singh, and C. Guestrin, "" why should i trust you?"explaining the predictions of any classifier," inProceedings of the 22ndACM SIGKDD international conference on knowledge discovery anddata mining, 2016, pp. 1135–1144.	1.780	N/A
AIX360 Toolkit	IBM	Explainability	V. Arya, R. K. Bellamy, P.-Y. Chen, A. Dhurandhar, M. Hind, S. C.Hoffman, S. Houde, Q. V. Liao, R. Luss, A. Mojsilović et al., "Oneexplanation does not fit all: A toolkit and taxonomy of ai explainabilitytechniques,"arXiv preprint arXiv:1909.03012, 2019.	85	N/A
RuleX AI	Rulex	Explainability	N/A	18	https://www.rulex.ai/
IBM Research Uncertainty Quantification 360	IBM	Explainability	S. Ghosh, Q. V. Liao, K. N. Ramamurthy, J. Navratil, P. Sattigeri, K. R.Varshney, and Y. Zhang, "Uncertainty quantification 360: A holistictoolkit for quantifying and communicating the uncertainty of ai,"arXivpreprint arXiv:2106.01410, 2021.	3	http://uq360.mybluemix.net/

Appendix II.5: Identified Software Solutions - Human Oversight

Framework	Publisher	Category	Original Paper	Scholar Results	Alternative Source
SHapley Additive exPlanations (SHAP)	Lundberg, S. and Lee, S.	Human Oversight	S. M. Lundberg and S.-I. Lee, "A unified approach to interpreting modelpredictions," inProceedings of the 31st international conference onneural information processing systems, 2017, pp. 4768–4777.	9.080	N/A
Local Interpretable Model-Agnostic Explanation (LIME)	Ribero, M. and Singh, S. and Guestrin C.	Human Oversight	M. T. Ribeiro, S. Singh, and C. Guestrin, "" why should i trust you?"explaining the predictions of any classifier," inProceedings of the 22ndACM SIGKDD international conference on knowledge discovery anddata mining, 2016, pp. 1135–1144.	1.780	N/A
MLflow	MLflow	Human Oversight	N/A	409	https://github.com/mlflow/mlflow
Kubeflow	Kubeflow	Human Oversight	N/A	250	https://www.kubeflow.org/docs/about/kubeflow/
IBM Research AI Explainability 360	IBM	Human Oversight	Arya, R. K. Bellamy, P.-Y. Chen, A. Dhurandhar, M. Hind, S. C.Hoffman, S. Houde, Q. V. Liao, R. Luss, A. Mojsilovicet al., "Aiexplainability 360: An extensible toolkit for understanding data andmachine learning models."J. Mach. Learn. Res., vol. 21, no. 130, pp.1–6, 2020.	85	N/A
RuleX AI	Rulex	Human Oversight	N/A	18	https://www.rulex.ai/
Amazon SageMaker Edge Manager	Amazon	Human Oversight	N/A	3	https://aws.amazon.com/de/sagemaker/
Neptune.ai	Neptune	Human Oversight	N/A	0	https://neptune.ai/product
Tensorflow Responsible AI	Google	Human Oversight	J. Wexler, M. Pushkarna, T. Bolukbasi, M. Wattenberg, F. Vi´egas, andJ. Wilson, "The what-if tool: Interactive probing of machine learningmodels,"IEEE Transactions on Visualization and Computer Graphics,vol. 26, no. 1, pp. 56–65, 2020.	0	N/A

Appendix II.6: Identified Software Solutions - Accuracy, Robustness, Cybersecurity

Framework	Publisher	Category	Original Paper	Scholar Results	Alternative Source
Foolbox Native	Rauber, J.	Accuracy, Robustness, Cybersecurity	J. Rauber, R. Zimmermann, M. Bethge, and W. Brendel, "Foolboxnative: Fast adversarial attacks to benchmark the robustness of machinelearning models in pytorch, tensorflow, and jax,"Journal of Open SourceSoftware, vol. 5, no. 53, p. 2607, 2020.	498	https://github.com/bethgelab/foolbox
IBM Adversarial Robustness Toolbox	IBM	Accuracy, Robustness, Cybersecurity	M.-I. Nicolae, M. Sinn, M. N. Tran, B. Buesser, A. Rawat, M. Wistuba,V. Zantedeschi, N. Baracaldo, B. Chen, H. Ludwiget al., "Adversarialrobustness toolbox v1. 0.0,"arXiv preprint arXiv:1807.01069, 2018.	305	https://github.com/Trusted-AI/adversarial-robustness-toolbox
IBM CNN-Cert	IBM	Accuracy, Robustness, Cybersecurity	A. Boopathy, T.-W. Weng, P.-Y. Chen, S. Liu, and L. Daniel, "Cnn-cert: An efficient framework for certifying robustness of convolutionalneural networks," inProceedings of the AAAI Conference on ArtificialIntelligence, vol. 33, no. 01, 2019, pp. 3240–3247.	85	https://github.com/IBM/CNN-Cert
CORTEX CERTIFAI	CognitiveScale	Accuracy, Robustness, Cybersecurity	S. Sharma, J. Henderson, and J. Ghosh, "Certifai: Counterfactual explanations for robustness, transparency, interpretability, and fairness ofartificial intelligence models,"arXiv preprint arXiv:1905.07857, 2019.	59	https://www.cognitivescale.com/certifai/
IBM Research AI Fairness 360 Toolkit	IBM	Accuracy, Robustness, Cybersecurity	R. K. E. Bellamy, K. Dey, M. Hind, S. C. Hoffman, S. Houde,K. Kannan, P. Lohia, J. Martino, S. Mehta, A. Mojsilović, S. Nagar,K. N. Ramamurthy, J. Richards, D. Saha, P. Sattigeri, M. Singh, K. R.Varshney, and Y. Zhang, "Ai fairness 360: An extensible toolkit fordetecting and mitigating algorithmic bias,"IBM Journal of Researchand Development, vol. 63, no. 4/5, pp. 4:1–4:15, 2019.	48	http://aif360.mybluemix.net/?_ga=2.219990425.1482784964.1625572737-971823234.1625572737

Appendix III.1: Evaluation of Software Solutions - Testing

Req. ID	Framework	Evaluation	Comment	Source
9.14	Amazon Sage Maker	1	Amazon Sage Maker can be used to test the AI system with A/B testing and record its performance. Insights about performance differences might be used to design risk systems (with considerable effort).	https://docs.aws.amazon.com/sagemaker/latest/dg/model-ab-testing.html
9.15	Amazon Sage Maker	1	Amazon Sage Maker can be used to record performance but not to explicitly/automatically test for consistent metric values. This could be automized manually.	https://docs.aws.amazon.com/sagemaker/latest/dg/model-ab-testing.html
9.16	Amazon Sage Maker	0	Amazon Sage Maker does not provide purpose-specific testing abilities.	https://docs.aws.amazon.com/sagemaker/latest/dg/model-ab-testing.html
9.17	Amazon Sage Maker	N/A	Process requirement	N/A
9.18	Amazon Sage Maker	2	Amazon Sage Maker provides industry standardized ML model metrics. Thresholds could be applied to these metrics manually. However, no automatic check for exceeding these thresholds is provided.	https://docs.aws.amazon.com/sagemaker/latest/dg/model-ab-testing.html
9.14	Watson OpenScale	2	Monitoring of the model and reporting of customary metrics including with respect to fairness can be used to manually design an adequate risk management system. Test and assessment functions especially targeted at risk management.	https://dataplatform.cloud.ibm.com/docs/content/wsj/model/wos-insight-timechart.html https://dataplatform.cloud.ibm.com/docs/content/wsj/model/cloud-risk-wos-only.html
9.15	Watson OpenScale	1	Ability to record performance and related metrics but not explicitly/automatically test for consistent metric values; this could be automized manually	https://dataplatform.cloud.ibm.com/docs/content/wsj/model/getting-started.html
9.16	Watson OpenScale	2	Watson OpenScale provides model fairness (and quality) metrics depending on different use cases where bias is prevalent.	https://dataplatform.cloud.ibm.com/docs/content/wsj/model/getting-started.html https://dataplatform.cloud.ibm.com/docs/content/wsj/model/wos-fairness-group.html
9.17	Watson OpenScale	N/A	Process requirement	N/A
9.18	Watson OpenScale	2	Watson OpenScale provides industry standardized ML model metrics. Thresholds could be applied to these metrics manually. However, no automatic check for exceeding these thresholds is provided. Provides ability to create custom metrics.	https://dataplatform.cloud.ibm.com/docs/content/wsj/model/getting-started.html
9.14	Azure ML	1	Azure ML can be used to test the AI system with A/B testing and record its performance. Insights about performance differences might be used to design risk systems (with considerable effort).	https://docs.microsoft.com/en-us/azure/machine-learning/how-to-deploy-azure-kubernetes-service?tabs=python#deploy-models-to-aks-using-controlled-rollout-preview
9.15	Azure ML	1	Azure ML can be used to record performance but not to explicitly/automatically test for consistent metric values. This could be automized manually.	https://docs.microsoft.com/en-us/azure/machine-learning/how-to-deploy-azure-kubernetes-service?tabs=python#deploy-models-to-aks-using-controlled-rollout-preview; https://docs.microsoft.com/en-us/azure/machine-learning/how-to-enable-data-collection
9.16	Azure ML	0	Azure ML does not provide purpose-specific testing abilities.	https://docs.microsoft.com/en-us/azure/machine-learning/how-to-deploy-azure-kubernetes-service?tabs=python#deploy-models-to-aks-using-controlled-rollout-preview
9.17	Azure ML	N/A	Process requirement	N/A
9.18	Azure ML	2	Azure ML provides industry standardized ML model metrics. Thresholds could be applied to these metrics manually. However, no automatic check for exceeding these thresholds is provided.	https://docs.microsoft.com/en-us/azure/machine-learning/how-to-deploy-azure-kubernetes-service?tabs=python#deploy-models-to-aks-using-controlled-rollout-preview

Appendix III.2: Evaluation of Software Solutions - Dataset Properties 1/2

Req. ID	Framework	Evaluation	Comment	Source
10.1	IBM SPSS Modeler	N/A	Process requirement	N/A
10.2	IBM SPSS Modeler	N/A	Process requirement	N/A
10.3	IBM SPSS Modeler	3	SPSS Modeler is explicitly build around CRISP-DM process. Data Preparation is one step in the CRISP-DM process.	https://www.ibm.com/docs/en/spss-modeler/18.1.1?topic=preparation-data-overview
10.4	IBM SPSS Modeler	N/A	Process requirement	N/A
10.5	IBM SPSS Modeler	2	SPSS Modeler only provides tools for verifying data quality. Assessing availability and suitability of data sets may require additional expert knowledge.	https://www.ibm.com/docs/en/spss-modeler/18.1.1?topic=understanding-verifying-data-quality
10.6	IBM SPSS Modeler	0	No explicit functionality for detecting biases in the data sets.	N/A
10.7	IBM SPSS Modeler	3	SPSS Modeler provides tools for verifying data quality which includes detecting and addressing missing data with various methods.	https://www.ibm.com/docs/en/spss-modeler/18.1.1?topic=understanding-verifying-data-quality, https://www.ibm.com/docs/en/wsd?topic=values-handling-missing
10. 8	IBM SPSS Modeler	1	SPSS Modeler provides tools for verifying data quality which includes detecting errors in the data. Assessing whether it is relevant and representative will require additional expert knowledge.	https://www.ibm.com/docs/en/spss-modeler/18.0.0?topic=nodes-data-audit-node
10.9	IBM SPSS Modeler	1	General data exploration to analyze statistical properties of the data possible. Does not assess whether data sets have statistical properties with regard to users/groups of users.	https://www.ibm.com/docs/en/spss-modeler/18.1.1?topic=understanding-exploring-data
10.10	IBM SPSS Modeler	0	Does not provide functionality to assess whether datasets have characteristics specific to the geographical, behavioral or functional setting.	N/A
10.1	SAP Data Services	N/A	Process requirement	N/A
10.2	SAP Data Services	N/A	Process requirement	N/A
10.3	SAP Data Services	3	Provides data quality functionality that includes cleansing, enhancing, matching and consolidating data elements which is part of data preparation.	https://help.sap.com/viewer/ec06fadc50b64b6184f835e4f0e1f52f/4.2.14/en-US/572548f96d6d1014b3fc9283b0e91070.html
10.4	SAP Data Services	N/A	Process requirement	N/A
10.5	SAP Data Services	2	SAP Data Services provide data assessment tools to determine the quality of the data. Assessing availability and suitability of data sets may require additional expert knowledge.	https://help.sap.com/viewer/ec06fadc50b64b6184f835e4f0e1f52f/4.2.14/en-US/5724c17c6d6d1014b3fc9283b0e91070.html
10.6	SAP Data Services	0	No functionality for detecting biases in the data sets.	N/A
10.7	SAP Data Services	1	Detecting and filtering missing or bad values is possible, but no specific methods for handling these values are available.	https://help.sap.com/viewer/e54136ab6a4a43e6a370265bf0a2d744/4.2.14/en-US/5739b9f56d6d1014b3fc9283b0e91070.html
10. 8	SAP Data Services	1	Provides functionality as part of Data Assessment tool to identify and fix errors in the data. Assessing whether it is relevant and representative will require additional expert knowledge.	https://help.sap.com/viewer/e54136ab6a4a43e6a370265bf0a2d744/4.2.7/en-US/5724c17c6d6d1014b3fc9283b0e91070.html
10.9	SAP Data Services	1	Provides functionality to obtain statistics about the cleansing and assignment processes as part of data quality. Does not assess whether data sets have statistical properties with regard to users/groups of users.	https://help.sap.com/viewer/e54136ab6a4a43e6a370265bf0a2d744/4.2.7/en-US/24cad7e5f0fb44a8865e3451b1eff56c.html?q=data%20statistics
10.10	SAP Data Services	0	Does not provide functionality to assess whether datasets have characteristics specific to the geographical, behavioral or functional setting.	N/A

Appendix III.2: Evaluation of Software Solutions - Dataset Properties 2/2

Req. ID	Framework	Evaluation	Comment	Source
10.1	Informatica Data Quality	N/A	Process requirement	N/A
10.2	Informatica Data Quality	N/A	Process requirement	N/A
10.3	Informatica Data Quality	2	Provides data quality and profiling functionality that includes some data preparation steps.	https://docs.informatica.com/data-quality-and-governance/data-quality/10-5/data-quality-getting-started-guide/getting-started-overview/informatica-developer-overview/data-quality-and-profiling.html
10.4	Informatica Data Quality	N/A	Process requirement	N/A
10.5	Informatica Data Quality	2	Provides comprehensive "data quality capabilities". Assessing availability and suitability of data sets may require additional expert knowledge.	https://docs.informatica.com/data-quality-and-governance/data-quality/10-5/data-quality-getting-started-guide/getting-started-overview/informatica-developer-overview/data-quality-and-profiling.html
10.6	Informatica Data Quality	0	No explicit functionality for detecting biases in the data sets.	N/A
10.7	Informatica Data Quality	0	No explicit functionality mentioned in the documentation that detects and addresses missing data (gaps).	N/A
10.8	Informatica Data Quality	1	Provides functionality to remove errors as part of data standardization. Assessing whether it is relevant and representative will require additional expert knowledge.	https://docs.informatica.com/data-quality-and-governance/data-quality/10-5/data-quality-getting-started-guide/getting-started-overview/informatica-developer-overview/data-quality-and-profiling.html
10.9	Informatica Data Quality	0	Does not assess whether data sets have statistical properties with regard to users/groups of users.	N/A
10.10	Informatica Data Quality	0	Does not provide functionality to assess whether datasets have characteristics specific to the geographical, behavioral or functional setting.	N/A

Appendix III.3: Evaluation of Software Solutions - Record Keeping 1/2

Req. ID	Framework	Evaluation	Comment	Source
12.1	TensorBoard	2	TensorBoard provides callback functionalities, which ensures that logs are created and stored at a central customizable location, hence automatic event recording capabilities can be recorded and accessed.	https://www.tensorflow.org/tensorboard/get_started
12.2	TensorBoard	1	TensorBoard only provides logging and visualization of training data and therefore does not facilitate record keeping during the system's entire lifecycle.	https://pubs.rsna.org/doi/pdf/10.1148/ryai.2020200012; https://github.com/tensorflow/tensorboard/blob/master/docs/get_started.ipynb
12.3	TensorBoard	0	Insufficient information on whether TensorBoard's record keeping complies with an industry-acknowledged standard or common practice.	https://pubs.rsna.org/doi/pdf/10.1148/ryai.2020200012; https://www.tensorflow.org/tensorboard/scalars_and_keras
12.4	TensorBoard	1	TensorBoard is inherently not able to save input data. However, the closely affiliated TensorFlow framework can be used to store entire data-stream inputs and subsequently access specific groups of them directly to TensorBoard.	https://stackoverflow.com/questions/42355122/can-i-export-a-tensorflow-summary-to-csv
12.5	TensorBoard	1	TensorBoard's visualization capabilities main purpose is facilitating the manual analysis of logged training data, rather than a comprehensive, automatic data analysis.	https://pubs.rsna.org/doi/pdf/10.1148/ryai.2020200012; https://www.tensorflow.org/tensorboard/scalars_and_keras
12.6	TensorBoard	N/A	Process requirement	N/A
12.7	TensorBoard	0	Insufficient information available.	N/A
12.8	TensorBoard	2	TensorBoard's visualized datasets include timestamps, which allows to retrace all periods of use.	https://pubs.rsna.org/doi/pdf/10.1148/ryai.2020200012; https://www.tensorflow.org/tensorboard/scalars_and_keras
12.9	TensorBoard	0	Insufficient information available.	N/A
12.10	TensorBoard	0	Insufficient information available.	N/A
12.11	TensorBoard	0	Insufficient information available.	N/A
12.1	Amazon CloudWatch	3	CloudWatch provides functionality for visibility of metrics and logs data, data retention, and the ability to perform calculations on metrics.	https://s3.cn-north-1.amazonaws.com.cn/aws-dam-prod/china/pdf/acw-dg.pdf; https://aws.amazon.com/cloudwatch/features/?nc1=h_ls
12.2	Amazon CloudWatch	2	CloudWatch's comprehensive metric and logging capabilities should ensure a high level of trackability, but there is insufficient information on whether it covers an AI system's entire lifecycle.	https://s3.cn-north-1.amazonaws.com.cn/aws-dam-prod/china/pdf/acw-dg.pdf; https://aws.amazon.com/cloudwatch/features/?nc1=h_ls
12.3	Amazon CloudWatch	0	Insufficient information on whether CloudWatch's record keeping complies with an industry-acknowledged standard or common practice.	https://aws.amazon.com/cloudwatch/features/?nc1=h_ls
12.4	Amazon CloudWatch	2	The Amazon CloudWatch Logs service provides functionality to collect and store logs from resources, applications, and services in near real-time.	https://s3.cn-north-1.amazonaws.com.cn/aws-dam-prod/china/pdf/acw-dg.pdf; https://aws.amazon.com/cloudwatch/features/?nc1=h_ls
12.5	Amazon CloudWatch	2	CloudWatch custom metrics are automatically extracted from the ingested logs. Further analysis is provided by CloudWatch Logs Insights' advanced query language.	https://s3.cn-north-1.amazonaws.com.cn/aws-dam-prod/china/pdf/acw-dg.pdf; https://aws.amazon.com/cloudwatch/features/?nc1=h_ls

Appendix III.3: Evaluation of Software Solutions - Record Keeping 2/2

Req. ID	Framework	Evaluation	Comment	Source
12.6	Amazon CloudWatch	N/A	Process requirement	N/A
12.7	Amazon CloudWatch	0	Insufficient information available.	N/A
12.8	Amazon CloudWatch	2	Amazon CloudWatch provides functionality to collect custom metrics which may include timestamps or "user activity", enabling the monitoring of periods of use.	https://s3.cn-north-1.amazonaws.com.cn/aws-dam-prod/china/pdf/acw-dg.pdf; https://aws.amazon.com/cloudwatch/features/?nc1=h_ls
12.9	Amazon CloudWatch	0	Insufficient information available.	N/A
12.10	Amazon CloudWatch	0	Insufficient information available.	N/A
12.11	Amazon CloudWatch	0	Insufficient information available.	N/A
12.1	DataDog	1	DataDog provides various application monitoring functionality, building upon logs generated by the AI system, but does not facilitate the creation of logs itself.	https://skemman.is/bitstream/1946/28745/1/Project%20Report.pdf; https://www.datadoghq.com/product/log-management/
12.2	DataDog	2	DataDog's application monitoring capabilities support auditing or investigations on logs and therefore the traceability of the system.	https://skemman.is/bitstream/1946/28745/1/Project%20Report.pdf; https://www.datadoghq.com/product/log-management/
12.3	DataDog	0	Insufficient information on whether DataDog's record keeping complies with an industry-acknowledged standard or common practice.	https://skemman.is/bitstream/1946/28745/1/Project%20Report.pdf; https://www.datadoghq.com/product/log-management/
12.4	DataDog	0	Insufficient information available.	N/A
12.5	DataDog	0	Insufficient information available.	N/A
12.6	DataDog	N/A	Process requirement	N/A
12.7	DataDog	0	Insufficient information available.	N/A
12.8	DataDog	0	Insufficient information available.	N/A
12.9	DataDog	0	Insufficient information available.	N/A
12.10	DataDog	0	Insufficient information available.	N/A
12.11	DataDog	0	Insufficient information available.	N/A

Appendix III.4: Evaluation of Software Solutions - Explainability

Req. ID	Framework	Evaluation	Comment	Source
13.1	SHapley Additive exPlanations (SHAP)	2	SHAP computes Shapley values from game theory and provides feature explanations. It explains the Machine Learning model prediction of a data instance by computing the contribution (= importance) of each feature to that prediction. Makes the model more transparent as it tries to explain its output.	https://arxiv.org/pdf/1705.07874.pdf
13.1	Local Interpretable Model-Agnostic Explanation (LIME)	2	LIME is a model-agnostic explainability method that explains a complex Machine Learning model by approximating it locally with a simpler model that is in itself explainable. LIME has been mostly applied for image and text data. It makes the model more transparent as it tries to explain its output.	https://arxiv.org/pdf/1602.04938v3.pdf
13.1	AIX360 Toolkit	3	AIX360 Toolkit provides 8 state-of-the-art explanation algorithms that can be selected depending on the type of explanation needed. All algorithms may provide a high level of explainability.	https://arxiv.org/pdf/1909.03012.pdf

Appendix III.5: Evaluation of Software Solutions - Human Oversight 1/2

Req. Id	Framework	Evaluation	Comment	Source
14.1	SHapley Additive exPlanations (SHAP)	0	Out of Scope. A human-machine interface tool is not provided to monitor the continuous deployment of the model during operation. SHAP focuses on explaining the prediction (output) of a Machine Learning model.	N/A
14.2	SHapley Additive exPlanations (SHAP)	N/A	Process requirement	N/A
14.3	SHapley Additive exPlanations (SHAP)	N/A	Process requirement	N/A
14.4	SHapley Additive exPlanations (SHAP)	N/A	Process requirement	N/A
14.5	SHapley Additive exPlanations (SHAP)	1	SHAP is inherently not able to prevent automation-bias. However, it makes ML models more transparent by explaining the contribution of each feature to the model output. It could help preventing automation bias if subject matter experts are required to evaluate the produced explanations. If explanations do not make sense or are not as expected, this could be a warning that the output is flawed and hence, could help preventing over-reliance on the system.	https://github.com/slundberg/shap
14.6	SHapley Additive exPlanations (SHAP)	3	SHAP computes Shapley values from game theory and provides feature explanations. It explains the Machine Learning model prediction of a data instance by computing the contribution (= importance) of each feature to that prediction. Makes the model more transparent as it tries to explain its output.	https://stats.stackexchange.com/questions/379744/comparison-between-shap-shapley-additive-explanation-and-lime-local-interpret
14.7	SHapley Additive exPlanations (SHAP)	0	Out of Scope. A human-machine interface tool is not provided to monitor the continuous deployment of the model during operation. SHAP focuses on explaining the prediction (output) of a Machine Learning model.	N/A
14.8	SHapley Additive exPlanations (SHAP)	0	Out of Scope. A human-machine interface tool is not provided to monitor the continuous deployment of the model during operation. SHAP focuses on explaining the prediction (output) of a Machine Learning model.	N/A
14.9	SHapley Additive exPlanations (SHAP)	0	Out of Scope. SHAP cannot be used to check if the high-risk AI system is operating such that its decisions with respect to identification, assignment and assessment of natural persons in educational and vocational training institutions, are confirmed by at least two natural persons. SHAP focuses on explaining the prediction (output) of a Machine Learning model.	https://github.com/slundberg/shap
14.1	Local Interpretable Model-Agnostic Explanation (LIME)	0	Out of Scope. A human-machine interface tool is not provided to monitor the continuous deployment of the model during operation. SHAP focuses on explaining the prediction (output) of a Machine Learning model.	N/A
14.2	Local Interpretable Model-Agnostic Explanation (LIME)	N/A	Process requirement	N/A
14.3	Local Interpretable Model-Agnostic Explanation (LIME)	N/A	Process requirement	N/A
14.4	Local Interpretable Model-Agnostic Explanation (LIME)	N/A	Process requirement	N/A

Appendix III.5: Evaluation of Software Solutions - Human Oversight 2/2

Req. ID	Framework	Evaluation	Comment	Source
14.5	Local Interpretable Model-Agnostic Explanation (LIME)	1	LIME does not provide this functionality inherently. However, it could potentially be added manually via code not native to the framework.	https://github.com/marcotcr/lime
14.6	Local Interpretable Model-Agnostic Explanation (LIME)	2	LIME is a model-agnostic explainability method that explains a complex Machine Learning model by approximating it locally with a simpler model that is in itself explainable. LIME has been mostly applied for image and text data. It makes the model more transparent as it tries to explain its output.	https://github.com/marcotcr/lime
14.7	Local Interpretable Model-Agnostic Explanation (LIME)	0	Out of Scope. A human-machine interface tool is not provided to monitor the continuous deployment of the model during operation. LIME focuses on explaining the prediction (output) of a Machine Learning model.	N/A
14.8	Local Interpretable Model-Agnostic Explanation (LIME)	0	Out of Scope. A human-machine interface tool is not provided to monitor the continuous deployment of the model during operation. LIME focuses on explaining the prediction (output) of a Machine Learning model.	N/A
14.9	Local Interpretable Model-Agnostic Explanation (LIME)	0	Out of Scope. LIME cannot be used to check if the high-risk AI system is operating such that its decisions with respect to identification, assignment and assessment of natural persons in educational and vocational training institutions, are confirmed by at least two natural persons. LIME focuses on explaining the prediction (output) of a Machine Learning model.	https://github.com/marcotcr/lime/blob/master/doc/notebooks/Tutorial%20-%20Faces%20and%20GradBoost.ipynb https://github.com/marcotcr/lime/blob/master/lime/lime_text.py
14.1	MLflow	3	The MLflow Tracking component is an API and UI for logging parameters, code versions, metrics, and output files when running a Machine Learning model. It also allows to later visualize the results. As a result, it provides broad oversight capabilities.	https://mlflow.org/docs/latest/tracking.html#concepts
14.2	MLflow	N/A	Process requirement	N/A
14.3	MLflow	N/A	Process requirement	N/A
14.4	MLflow	N/A	Process requirement	N/A
14.5	MLflow	0	Out of Scope. MLflow does not provide functionality that controls or stops users from over relying on the AI-system.	N/A
14.6	MLflow	2	MLflow provides tools for general interpretation via various metrics and visualizes them in real-time in a dashboard. However, in terms of transparency, the tools available in the MLflow framework are lacking. Nonetheless, integration with other explainability frameworks such as SHAP or LIME could be possible. Using additional frameworks could help providing transparency for the high-risk AI system.	https://mlflow.org/docs/latest/tracking.html#visualizing-metrics
14.7	MLflow	0	Out of Scope. As an AI framework for managing rather than designing AI systems, there is no direct support for such functionality. However, MLflow could potentially be used in synergy with other frameworks to achieve the required objective.	N/A
14.8	MLflow	2	Partly realized via command line interruption directly within the server with Ctrl C or by using 'pkill -f gunicorn' on the server.	https://stackoverflow.com/questions/60531166/how-to-safely-shutdown-mlflow-ui https://github.com/mlflow/mlflow/issues/1856
14.9	MLflow	0	Out of Scope. As an AI framework for managing rather than designing AI systems, there is no direct support for such functionality. However, MLflow could potentially be used in synergy with other frameworks to achieve the required objective.	N/A

Appendix III.6: Evaluation of Software Solutions - Accuracy, Robustness, Cybersecurity 1/2

Req. ID	Framework	Evaluation	Comment	Source
15.1	Foolbox Native	N/A	Process requirement	N/A
15.2	Foolbox Native	1	Foolbox provides attack models for adversarial training. There is a trade-off between robustness ("robust accuracy") and accuracy ("standard accuracy"). A consistent level of robustness through should lead to a consistent level of accuracy.	https://foolbox.jonasrauber.de/guide/getting-started.html#robust-accuracy https://arxiv.org/abs/1805.12152
15.3	Foolbox Native	3	Foolbox provides a variety of "adversarial attacks to benchmark the robustness of machine learning models".	https://foolbox.readthedocs.io/en/stable/modules/attacks.html
15.4	Foolbox Native	2	Foolbox provides adversarial training, which helps mitigating adversarial attacks, but is not sufficient to achieve "cybersecurity" as a whole.	https://publications.jrc.ec.europa.eu/repository/bitstream/JRC119336/dpad_report.pdf
15.5	Foolbox Native	N/A	Process requirement	N/A
15.6	Foolbox Native	0	Foolbox provides adversarial training, but does not address technical redundancy or fault prevention.	N/A
15.7	Foolbox Native	0	Foolbox provides adversarial training, but does not address biased outputs through 'feedback loops'.	N/A
15.8	Foolbox Native	2	Adversarial training mitigates adversarial attacks, being a popular way of AI-System manipulation, but does not generally prevent unauthorized access by third parties.	https://foolbox.jonasrauber.de
15.9	Foolbox Native	N/A	Process requirement	N/A
15.10	Foolbox Native	2	Data poisoning is considered a specific strategy of adversarial attacks, which are implicitly addressed by the framework.	https://foolbox.readthedocs.io/en/stable/modules/attacks.html
15.11	Foolbox Native	3	Adversarial examples are considered a specific strategy of adversarial attacks, which are explicitly addressed by the framework.	https://foolbox.readthedocs.io/en/stable/modules/attacks.html
15.12	Foolbox Native	2	Model flaw exploitation is considered a specific strategy of adversarial attacks, which are implicitly addressed by the framework.	https://foolbox.readthedocs.io/en/stable/modules/attacks.html
15.1	IBM Adversarial Robustness Toolbox	N/A	Process requirement	N/A
15.2	IBM Adversarial Robustness Toolbox	1	Adversarial Robustness Toolbox provides attack models for adversarial training. There is a trade-off between robustness ("robust accuracy") and accuracy ("standard accuracy"). A consistent level of robustness through should lead to a consistent level of accuracy.	https://adversarial-robustness-toolbox.readthedocs.io/en/latest/ https://arxiv.org/abs/1805.1
15.3	IBM Adversarial Robustness Toolbox	3	Adversarial Robustness Toolbox includes certifying and verifying model robustness and model hardening, which ensures a consistent level of robustness.	https://adversarial-robustness-toolbox.readthedocs.io/en/latest/modules/metrics.html?highlight=robustness#empirical-robustness https://arxiv.org/pdf/1807.01069.pdf
15.4	IBM Adversarial Robustness Toolbox	2	Adversarial Robustness Toolbox provides adversarial training, which helps mitigating adversarial attacks, but is not sufficient to achieve "cybersecurity" as a whole.	https://publications.jrc.ec.europa.eu/repository/bitstream/JRC119336/dpad_report.pdf
15.5	IBM Adversarial Robustness Toolbox	N/A	Process requirement	N/A
15.6	IBM Adversarial Robustness Toolbox	0	Adversarial Robustness Toolbox provides adversarial training, but does not address technical redundancy or fault prevention.	N/A
15.7	IBM Adversarial Robustness Toolbox	0	Adversarial Robustness Toolbox provides adversarial training, but does not address biased outputs through 'feedback loops'.	N/A
15.8	IBM Adversarial Robustness Toolbox	2	Adversarial training mitigates adversarial attacks, being a popular way of AI-System manipulation, but does not generally prevent unauthorized access by third parties.	https://adversarial-robustness-toolbox.readthedocs.io/en/latest/

Appendix III.6: Evaluation of Software Solutions - Accuracy, Robustness, Cybersecurity 2/2

Req. ID	Framework	Evaluation	Comment	Source
15.9	IBM Adversarial Robustness Toolbox	N/A	Process requirement	N/A
15.10	IBM Adversarial Robustness Toolbox	2	Data poisoning is considered a specific strategy of adversarial attacks, which are addressed by the framework through a "module providing poisoning attacks under a common interface."	https://adversarial-robustness-toolbox.readthedocs.io/en/latest/modules/attacks/poisoning.html?highlight=poisoning
15.11	IBM Adversarial Robustness Toolbox	3	Adversarial examples are considered a specific strategy of adversarial attacks, which are addressed by the framework. "The attacks implemented in ART allow creating adversarial attacks against Machine Learning models which is required to test defenses with state-of-the-art threat models."	https://adversarial-robustness-toolbox.readthedocs.io/en/latest/modules/attacks.html; https://arxiv.org/pdf/1807.01069.pdf
15.12	IBM Adversarial Robustness Toolbox	2	Model flaw exploitation is considered a specific strategy of adversarial attacks, which are addressed by the framework.	https://adversarial-robustness-toolbox.readthedocs.io/en/latest/modules/attacks.html
15.1	CNN-Cert	N/A	Process requirement	N/A
15.2	CNN-Cert	1	CNN-Cert provides a mechanism for robustness certification. There is a trade-off between robustness ("robust accuracy") and accuracy ("standard accuracy"). Verification of robustness may indirectly support verification of accuracy.	https://arxiv.org/abs/1805.1
15.3	CNN-Cert	2	CNN-Cert provides Robustness verification, which passively supports ensuring a consistent level of robustness.	https://medium.com/@MITIBMLab/cnn-cert-a-certified-measure-of-robustness-for-convolutional-neural-networks-fd2ff44c6807
15.4	CNN-Cert	1	CNN-Cert provides Robustness verification, which does only assist in identifying an AI-System's potential cybersecurity risk.	https://medium.com/@MITIBMLab/cnn-cert-a-certified-measure-of-robustness-for-convolutional-neural-networks-fd2ff44c6807
15.5	CNN-Cert	N/A	Process requirement	N/A
15.6	CNN-Cert	0	CNN-Cert provides a mechanism for robustness certification, but does not address technical redundancy or fault prevention.	N/A
15.7	CNN-Cert	0	CNN-Cert provides a mechanism for robustness certification, but does not address biased outputs through 'feedback loops'.	N/A
15.8	CNN-Cert	0	Verifying the robustness of an AI-System does only passively assist in identifying a low level of robustness as a potential point of vantage for unauthorized third parties.	N/A
15.9	CNN-Cert	N/A	Process requirement	N/A
15.10	CNN-Cert	1	Robustness verification does not sufficiently prove and actively test protection against data poisoning.	https://arxiv.org/pdf/1811.12395.pdf
15.11	CNN-Cert	1	Robustness verification does not sufficiently prove and actively test protection against adversarial examples.	https://arxiv.org/pdf/1811.12395.pdf
15.12	CNN-Cert	1	Robustness verification does not sufficiently prove and actively test protection against model flaw exploitation.	https://arxiv.org/pdf/1811.12395.pdf

APPENDIX IV:
AMBIGUOUS AND VAGUE WORDINGS AND PHRASINGS

AI Act Article	Derived Req. ID	Content	Comment
9 (4c)	9.12	"adequate […] training to users"	It is unclear whether training is related to the high-risk AI system or to the risk response from users? If the former, it is not defined how any such training shall be designed or carried out.
9 (5)	9.15	"perform consistently for their intended purpose"	It is unclear what consistent performance of a purpose means. For instance, shall the same inputs passed to a high-risk AI system lead to the same outputs? In addition, intended purpose is only vaguely defined.
9 (7)	9.18	"preliminarily defined metrics and probabilistic thresholds"	It is not defined which metrics are considered suitable, in which way thresholds should be defined, and what are appropriate levels for thresholds regarding specific high-risk AI system purposes for customary metrics.
10 (2a)	10.1	"[...] relevant design choices."	Design choices regarding data features, AI system/data platform architecture? No sufficient information given, what design choices are about.
10 (2d)	10.4	"[...] relevant assumptions" (about the given data sets)	No examples are given for assumptions. What are possible assumptions?
10 (2e)	10.5	"suitability of the data sets that are needed."	What data sets classify as suitable?
10 (3)	10.9	"appropriate statistical properties […] as regards the persons or groups of persons […]"	What are statistical properties regarding users/groups of users are deemed appropriate?
10 (4)	10.10	"[...] characteristics or elements that are particular to the specific geographical, behavioural or functional setting."	What are characteristics that are specific to the mentioned settings? No examples are given.
11 (1)	11.1	"[...] all the necessary information to assess the compliance of the AI system [...]"	What information is considered necessary?
Annex IV (2d)	11.13	"[...] where relevant, the data requirements in terms of datasheets [...]"	What is considered relevant?
61 (2)	12.5	"systematically […] analyse relevant data"	It is not defined what a systematic analysis of data provided by users comprises of (e.g., aspects of input to consider, output) or how it should be carried out (e.g., tools, frequency, output storage) within the post-market monitoring system.
12 (4b)	12.9	"shall provide […] the reference database against which input data has been checked"	It is unclear whether the inclusion of a database in logging records refers to storing a reference to the database (e.g., ID, hyperlink), metadata about the database, or the database itself with the last option carrying the highest cost and being the least technically feasible
13 (1)	13.1	"their operation is sufficiently transparent to enable users to interpret the system's output"	It is unclear which level of transparency is required and how it should be achieved since the ability of users to interpret and use results is a vague objective.
15 (1)	15.1	"appropriate level of accuracy, robustness and cybersecurity"	What level of accuracy, robustness and cybersecurity is considered appropriate? In what metric are these levels measured?
15 (2)	15.5	"relevant accuracy metrics"	What accuracy metrics are considered relevant?
15 (4)	15.9	"appropriate to the relevant circumstances and the risks"	How can this appropriateness be specifically measured?