TU WIEN Informatics

Readiness Assessment for the Artificial Intelligence Act

with a requirements catalogue in the field of critical infrastructure

DIPLOMA THESIS

submitted in partial fulfillment of the requirements for the degree of

Diplom-Ingenieur

in

Business Informatics

by

David Oliva, BSc.
Registration Number 01634050

to the Faculty of Informatics

at the TU Wien

Advisor: Assistant Prof. Dipl.-Wirtsch.Inf.Univ. Dr.rer.pol. Dominik Bork

Vienna, 20th July, 2023

David Oliva	Dominik Bork

Technische Universität Wien
A-1040 Wien ▪ Karlsplatz 13 ▪ Tel. +43-1-58801-0 ▪ www.tuwien.at

Erklärung zur Verfassung der Arbeit

David Oliva, BSc.

Hiermit erkläre ich, dass ich diese Arbeit selbständig verfasst habe, dass ich die verwendeten Quellen und Hilfsmittel vollständig angegeben habe und dass ich die Stellen der Arbeit – einschließlich Tabellen, Karten und Abbildungen –, die anderen Werken oder dem Internet im Wortlaut oder dem Sinn nach entnommen sind, auf jeden Fall unter Angabe der Quelle als Entlehnung kenntlich gemacht habe.

Wien, 20. Juli 2023

David Oliva

Acknowledgements

I would like to take this opportunity to express my appreciation to several individuals who have contributed to the successful completion of my thesis.

First and foremost, I would like to extend my sincere appreciation to my thesis supervisor, Assistant Prof. Dipl.-Wirtsch.Inf.Univ. Dr.rer.pol. Dominik Bork of the Research Area Business Informatics, for his unwavering commitment, expertise, and continuous guidance. His mentorship has been instrumental in shaping my research and writing this thesis.

I would also like to extend my gratitude to Dr.-Ing. Wolfgang K. Walter from the EFS Consulting GmbH, who generously shared his professional experience and knowledge, enriching my understanding of the industry and broadening my professional network. His support and collaboration have been vital in bridging the gap between theory and practical application. In addition, I want to thank all the members of the Information Security engagement at EFS who dedicated their time and provided valuable feedback.

Finally, I would like to acknowledge the immense support and encouragement from my friends and family. Their understanding and support have been a constant source of motivation and essential in helping me achieve my goals throughout this academic journey.

Kurzfassung

Die Integration künstlicher Intelligenz (KI) birgt sowohl immense Chancen als auch erhebliche Herausforderungen, insbesondere wenn sie in kritischen Infrastrukturen genutzt wird. Die Komplexität von KI-Systemen in Verbindung mit den potenziellen Risiken bei ihrem Versagen erfordert strenge Compliance-Maßnahmen. Der Artificial Intelligence Act der EU zielt darauf ab, diese Systeme zu regulieren, jedoch kann die Einhaltung so einer Regulierung herausfordernd und ressourcenintensiv sein. In dieser Diplomarbeit wurde ein Anforderungskatalogs für Anbieter von KI-Systemen in kritischer Infrastruktur entwickelt, mit dem Ziel ein Instrument zur Beurteilung der Einhaltung des AI Acts bereitzustellen. Die Forschung untersuchte auch die Anwendungsbereiche von KI und deren Integration in verschiedene Infrastrukturbereiche wie die Versorgung von Wasser, Strom, Wärme, Gas und der Verkehrssteuerung. Als methodischer Ansatz wurde das Design Science Forschungsframework verwendet, ergänzt durch eine systematische Literaturrecherche, das Technology Acceptance Model (TAM) und die System Usability Scale (SUS). Zur Bewertung der Nützlichkeit des Katalogs wurden Interviews auf der Grundlage von TAM und eine SUS-Umfrage mit Informationssicherheitsexperten durchgeführt.

Die systematische Literaturrecherche zeigte, dass KI-Technologien in allen Bereichen kritischer Infrastrukturen integriert sind. Während die Integration von KI im Bereich der Wärme- und Gasversorgung relativ gering ist, demonstriert die Verkehrssteuerung die größte Konzentration. Der entwickelte Anforderungskatalog wurde als hilfreiches Tool mit hoher Benutzerfreundlichkeit wahrgenommen, was durch einen durchschnittlichen SUS-Wert von 92.9% bestätigt wurde. Der Katalog bietet mehrere Vorteile für Organisationen, welche von dem AI Act betroffen sind. Er stellt eine vorgefilterte Liste der relevanten Anforderungen bereit, was Zeit und Ressourcen erspart. Dabei werden die Anforderungen auf klare und verständliche Weise präsentiert. Dies fördert ein gemeinsames Verständnis unter Teammitgliedern und ermutigt zur Beteiligung am Compliance-Prozess. Der Katalog unterstützt auch bei der Kontextualisierung von Anforderungen, unter Berücksichtigung verwandter Standards, potenzieller Nachweise und spezifischer Ziele. Darüber hinaus werden die Ergebnisse mit Spinnendiagrammen und Heatmaps visualisiert. Dies verbessert die Fähigkeit, die Anforderungen effektiv zu interpretieren und Entscheidungen darüber zu treffen, wo weitere Aktionen und Maßnahmen durchgeführt werden müssen. Einer der wichtigsten Aspekte ist die Berechnung des Reifegrads, der als Indikator für die Gesamt-Compliance dient. Zusammenfassend betont die Arbeit die wertvolle Rolle des Anforderungskatalogs bei der Navigation durch die komplexe Landschaft der KI-

Anforderungen. Er erläutert nicht nur die komplexen Anforderungen, sondern fungiert auch als Kompass, der Unternehmen durch den Prozess der KI-Integration in kritische Infrastrukturen führt.

Abstract

The integration of Artificial Intelligence (AI) presents both immense opportunities as well as significant challenges, particularly when integrated into critical infrastructure. The complexity of AI systems, coupled with the potential risks associated with their failure, necessitates stringent compliance measures. The EU's Artificial Intelligence Act seeks to regulate these systems. However, the path to compliance can be challenging and resource-intensive. In the course of the thesis, a requirements catalogue was developed for providers of AI systems in critical infrastructure, aiming to provide a tool to assess compliance with the AI Act. The research further investigated the application areas of AI and its integration into various infrastructure domains such as the supply of water, electricity, heating, gas, and road traffic management. As a methodological approach, the Design Science Research framework was used, supplemented by a Systematic Literature Review (SLR), the Technology Acceptance Model (TAM), and the System Usability Scale (SUS). To evaluate if the catalogue is perceived as a useful tool, interviews based on TAM and a SUS survey were conducted with Information Security experts.

The SLR revealed that AI technologies have been incorporated into all critical infrastructure domains. While there is relatively low AI integration in the heating and gas supply sector, road traffic management shows the most significant concentration. The developed requirements catalogue itself was well accepted, with high perceived usefulness and ease of use, supported by an average SUS score of 92.9%. It offers several benefits to organisations seeking compliance. The catalogue provides a pre-filtered list of relevant requirements, saving time and resources and presents these requirements in a clear, understandable manner. This fosters a shared understanding among team members, encouraging participation in the compliance process. The catalogue also aids in contextualising requirements, considering related standards, potential proofs, and specific objectives. Further, the results were visualised with spider graphs and heatmaps. This enhances the ability to interpret the requirements effectively and decide where further actions and measures need to be implemented. Most importantly, the maturity level is calculated, which serves as an indicator of the overall compliance. In conclusion, the thesis underscores the valuable role of the requirements catalogue in navigating the complex landscape of AI requirements. It not only clarifies the complex requirements but also functions as a compass, guiding organisations through the process of AI integration into critical infrastructure.

Contents

CHAPTER 1

Introduction

1.1 General introduction to the subject area

Artificial Intelligence has become increasingly relevant and pervasive in various domains [Sta22], demonstrating its potential to transform the way we work, live and engage with technology. Its applications are rapidly expanding, not only in the hands of private individuals but also within public sectors [vNM22]. As the influence of AI continues to expand, there is an urgent need to ensure that its deployment adheres to ethical and responsible principles. Recognising this urgency, the Commission of the European Union has taken a proactive approach by proposing the Artificial Intelligence Act which seeks to establish a framework for governing AI and fostering its responsible deployment. The primary objectives of the AI Act are to enhance governance, facilitate the development of a single market for AI applications and ensure compliance with existing laws on fundamental rights and safety. This proposed regulation is a crucial step in addressing the challenges posed by Artificial Intelligence technologies and maximising its potential for the benefit of society. Establishing a set of standardised rules aims to foster innovation while protecting fundamental rights and preventing market fragmentation. Through effective enforcement mechanisms, the AI Act intends to ensure that AI applications are lawful, safe, and trustworthy. Particularly significant is the application of AI within critical infrastructure sectors. The integration in this domain presents unique opportunities to enhance efficiency, resilience, and decision-making capabilities. However, it also introduces potential risks and vulnerabilities that must be adequately addressed to safeguard critical systems and protect them against potential threats.

1.2 Problem statement and research questions

Depending on the service or product a company offers on a certain market, it has to comply with specific regulations, standards, and norms. In the critical infrastructure

sector, where failures have enormous repercussions on society, compliance with established standards becomes crucial. These businesses that provide or deploy AI solutions in the EU (independent of the company headquarters) or systems with an output that is used inside the EU have to comply with the AI Act [Cou22]. This poses a time-consuming and expensive challenge for the affected parties, especially in situations where various regulations must be addressed and noncompliance has significant financial consequences. In case of infringements of the AI Act's requirements, fines up to 20 million euros or four per cent of the company's worldwide annual revenue (whichever is higher) could face the operator of the AI system [Cou22]. However, companies that integrate AI systems into essential infrastructure lack a tool or procedure for determining if or to what extent they comply with the comprehensive set of new AI technology requirements. This presents a serious challenge for affected stakeholders who are aiming to assure compliance and prevent any legal or reputational risks. Consequently, the purpose of this thesis is to investigate the application areas of AI in critical infrastructure and to develop an assessment possibility for AI system compliance. The thesis proposes a tool in the format of a requirements catalogue for verifying compliance with the AI Act, determining the degree of fulfilment with maturity levels, identifying possible improvement areas, and further preparing companies for prospective audits.

According to Annex III [Cou22] of the AI Act, domains that belong to the area of critical infrastructure are, among others, the supply of water, gas, heating, electricity, and the management and operation of road traffic. One goal of the thesis is to gain insights into the level of integration of AI technology in critical infrastructure and discover application areas of AI in the individual domains, which leads to the following initial research question: *For what purposes is Artificial Intelligence used in critical infrastructure (supply of water, electricity, heating, gas, and road traffic management)?*

The key element of this thesis deals with the AI Act and its impact on companies and governments that provide AI systems in the field of critical infrastructure. The AI Act proposes a very comprehensive set of rules, that is not expressed in an easily comprehendible manner and does not provide possible evidence or an evaluation method for compliance. This fact makes it challenging for affected stakeholders, who are already obligated to follow other standards anyway. The indicated problem statement leads to the following formulation of the main research questions: *What specific requirements do providers of AI systems in the field of critical infrastructure have to meet concerning the AI Act and what is an appropriate means to integrate these into a requirements catalogue?*

In the process of designing the requirements catalogue, the appropriate way will be devised, taking into account factors such as understandable formulations, providing possible proofs, a clear structure and the overall layout of the catalogue. Deriving from this central research question, the following sub-research question is implied: *Which standards are associated with the requirements of the AI Act?*

Accordingly, the goal is to prepare, test, and benchmark providers of AI systems in critical infrastructure for this future regulation as best as possible. Furthermore, it should be

possible to determine the degree of fulfilment with maturity levels in order to estimate compliance for assessments and audits.

1.3 Methodological Approach

1.3.1 Process model

The initial step of this thesis involves conducting a Systematic Literature Review and a conventional literature review, which serve as essential foundations for the development of the requirements catalogue. This part aims to showcase the various application areas of AI systems in critical infrastructure domains and their level of integration. Moreover, it encompasses an examination of standards in fields such as Information Security, Cyber Security, Risk Management and, Quality Management. After gaining a deeper understanding of the developments of AI in critical infrastructure domains and the associated standards, the design of the requirements catalogue commences. To enhance the comprehensibility of the AI Act's requirements, they are filtered, condensed, organised into chapters and expressed in a more accessible manner. Subsequently, a comparative analysis is performed to align the requirements with other relevant standards and facilitate the provision of possible evidence. Additionally, control questions and requirement objectives are formulated to offer supplementary guidance for a better understanding of the fundamental essence of each requirement. To enable benchmarking and compliance verification, a fulfilment schema based on maturity levels is incorporated into the catalogue. Further, the results are visualised in multiple ways to improve interpretation and clarity. Lastly, to evaluate the intention to use this artefact, interviews and surveys are conducted with Information Security experts, which are then analysed to derive conclusions. Figure 1.1 presents an overview of the thesis process, which will be embedded in the Design Science approach by Hevner [HMPR04] in the subsequent section.

1.3.2 Design Science Approach

In Europe, especially in German-speaking nations design-oriented research has a long history, where Design Science research became a dominant Information Systems (IS) research paradigm [Win08]. According to Hevner [HMPR04], Design Science is a discipline that focuses on the creation and evaluation of IT artefacts that are meant to solve practical problems. They are reviewed based on their utility in addressing such real-world issues. In general, two concepts describe the majority of research in the field of Information Systems. On the one hand, behavioural science research with the objective of discovering the truth and research results that are theories. On the other hand, Design Science research with the goal of seeking utility and outcomes in the manner of IT artefacts. Those are frequently described as constructs, models, methods, and instantiations. Constructs serve as a source of vocabulary and symbols that are used to describe and explain phenomena in the field of IS. Models are abstractions and representations of phenomena that are used to comprehend and explain real-world systems. Methods are the procedures, techniques,

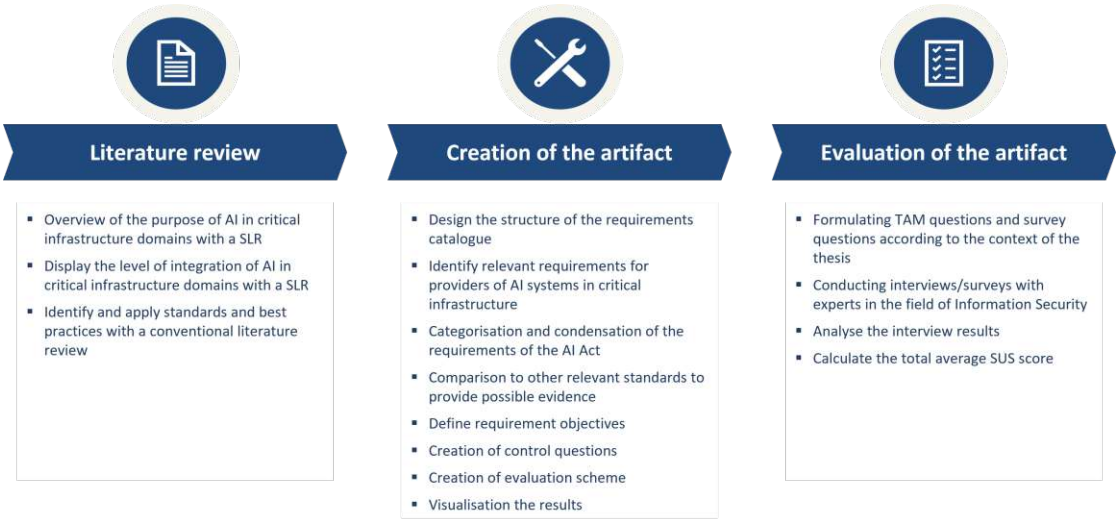


Figure 1.1: Process of the thesis

algorithms, and practices that support the Design Science research with guidance on how to address issues. Instantiations are specific implementations or examples of an artefact that are created as part of Design Science research. In the context of this thesis, the created artefact is an instantiation in the form of a requirements catalogue, which provides a tool that can be utilised in a particular problem field.

In addition, Hevner constructed the three inherent research cycles [Hev07], which are based on his IS research framework [HMPR04]. The Relevance Cycle connects the activities of Design Science with the environment of the research topic. In the context of this thesis, the main activities are the creation and evaluation of the requirements catalogue. The environment is composed of the application domain, including critical infrastructure domains, Quality management, Risk management, Information Security, and field experts, as well as further the problems and opportunities like the lack of clarity, benchmarking, and readiness assessment. On the other side, the Rigor Cycle establishes a link between the Design Science activities and the body of knowledge, including expertise, experience, and scientific foundations that underpin the research project topic. The knowledge base in this particular case consists of AI applications in critical infrastructure domains, norms, standards, best practices, and other security assessment catalogues. Finally, as already mentioned, iterating between the primary tasks of creating and evaluating the design artefact and research procedures is what the central design cycle accomplishes. In Figure 1.2, the three inherent research cycles from Hevner are shown, with the specific applications and context areas of this thesis underneath.

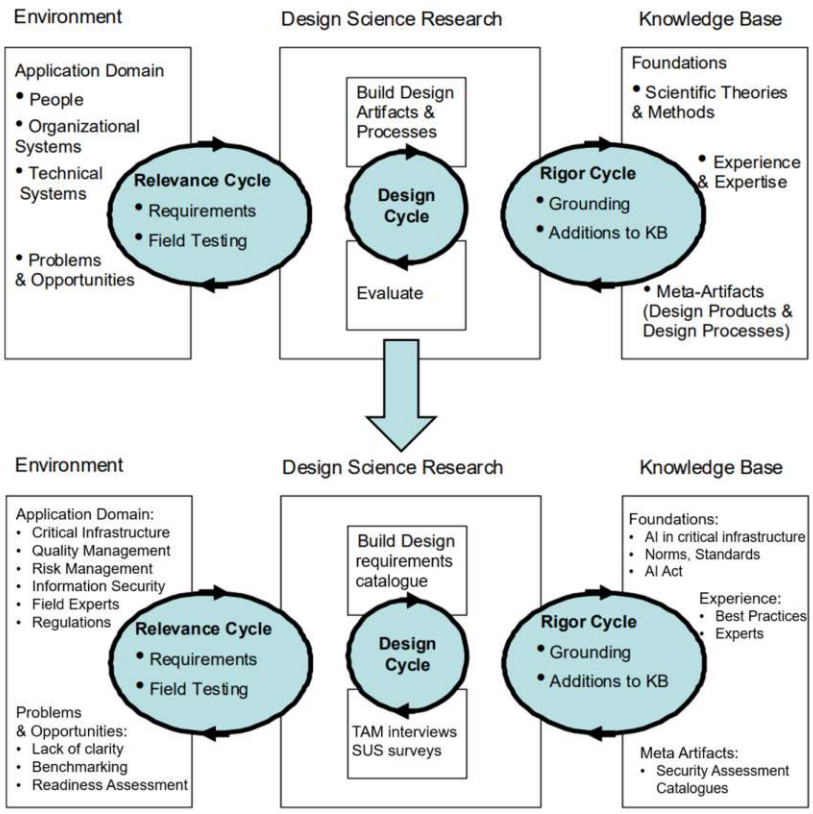


Figure 1.2: The three inherent research cycles [Hev07]

1.3.3 Literature review

The chosen methods to get a deeper understanding of the knowledge base are the Systematic Literature Review (SLR) as well as the conventional literature review, also called a narrative review. Whereas the latter has a much broader scope and does not provide the utilised databases, methodological approaches, or evaluation criteria for the inclusion or exclusion of retrieved papers [Rot07]. Although it has to be mentioned that a comprehensive SLR has the disadvantage of taking considerably more time and requiring in most steps more than one person [Rot07].

Literature reviews may be undertaken for a variety of purposes. These include establishing a theoretical foundation for future research, comprehending the scope of research on a certain subject or resolving questions by reviewing previous research on a certain topic. Frequently, research reviews are published as portions of single-subject articles or chapters of academic theses [OS10]. The first research question of this thesis is addressed through the use of a Systematic Literature Review, which provides support for the Rigour Cycle. For this purpose, the SLR follows a methodological framework that focuses on the procedure outlined by Yu Xiao and Maria Watson [XW19]. When the SLR is completed, there should be a clear picture of the different application areas of AI in

critical infrastructure domains and the level of integration. The information concerning standards in relation to the topics of AI Act is not obtained through a SLR, but rather by a conventional literature search, since it is not clearly known in advance what is being searched for. In designing the requirements catalogue, the relevant standards are investigated in an iterative process according to the requirements of the AI Act.

1.3.4 Evaluation

A critical factor and guideline of Hevner's Design Science research framework [HMPR04] is the evaluation of the created artefact. There are many factors by which an IT artefact can be evaluated, including functionality, consistency, reliability, usability, and many more [HMPR04]. In the context of this thesis, the *intention to use* and *usability* are tested, since it is important that the requirements catalogue is actually used as a tool and is further applicable in a practicable and simple way.

Fred Davis introduced 1986 the Technology Acceptance Model (TAM) [DBW89], which provides an approach to explain why people use or reject IT systems. The primary objective of TAM is to establish a foundation for examining the influence of external variables on internal intentions. TAM utilised the theoretical foundation of the Theory of Reasoned Action (TRA) [Sar83] to construct an adopted model that illustrates the associations between these variables (see Figure 1.3). The model argues that the two fundamental beliefs that significantly influence technology acceptance behaviours are the *perceived usefulness* and *perceived ease of use*. Perceived usefulness refers to the extent to which an individual believes that a particular technology will improve their job performance or make tasks simpler to complete. Users are more likely to employ a technology if they perceive that it will help them achieve their requirements or objectives. Perceived ease of use refers to the degree to which an individual believes that utilising a specific technology will be effortless. Users are more likely to employ a technology if they view it as simple to learn and easy to use. According to the TAM [DBW89], these two factors have a direct impact on the attitude towards using a technology, which in turn affects their intention to use the technology and finally leads to the actual system use.

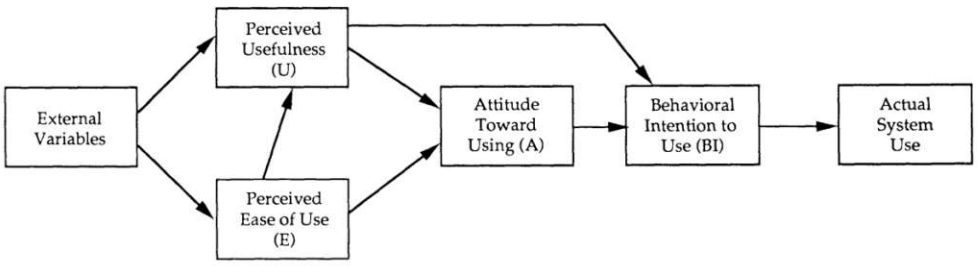


Figure 1.3: Technology Acceptance Model (TAM) [DBW89]

Further, the usability represents a substantial indicator if the artefact will actually be utilised and has the potential to achieve the intended solution. For the evaluation, a

standardised usability survey provides a tool for collecting the information of participants and offers the advantages like objectivity, replicability, and quantification [SL16]. There are several commonly applied standardised usability surveys, including:

- The Questionnaire for User Interaction Satisfaction (QUIS) [CDN88]
- The Software Usability Measurement Inventory (SUMI) [KCS93]
- The Post-Study System Usability Questionnaire (PSSUQ) [Lew92]
- The System Usability Scale (SUS) [B+96]

After taking a closer look at each of the questionnaires, the System Usability Scale of John Brooke [B+96] suits best the evaluation of the IT artefact due to its good adaptability and easy application. SUS consists of a ten-item questionnaire with a global perspective of the subjective perceived usability of a product or system and is based on a Likert scale, which represents five degrees ranging from strongly disagree to strongly agree (see Figure 1.4).

1.4 Limitations of the work

There are limitations to the thesis that must be considered. Firstly, the SLR is limited in terms of the number of searched papers. This is due to the time constraints and the scope of a thesis, as well as the absence of a team, which could have facilitated a more extensive review. In addition, employees from the EFS Consulting GmbH's Information Security engagement are involved in the evaluation of the research. The limited diversity of participants and the possibility of bias may have an impact on the validity and generalisability of the evaluation results, despite their notable expertise. Moreover, the requirements catalogue developed in the thesis is drawn up based on the AI Act version of 25th November 2022 [Cou22]. Consequently, further revisions to the regulation are not discussed in the thesis. The catalogue was tailored specifically for providers of AI systems in the field of critical infrastructure targeting the EU market and its applicability to other providers or AI systems in different domains may be limited.

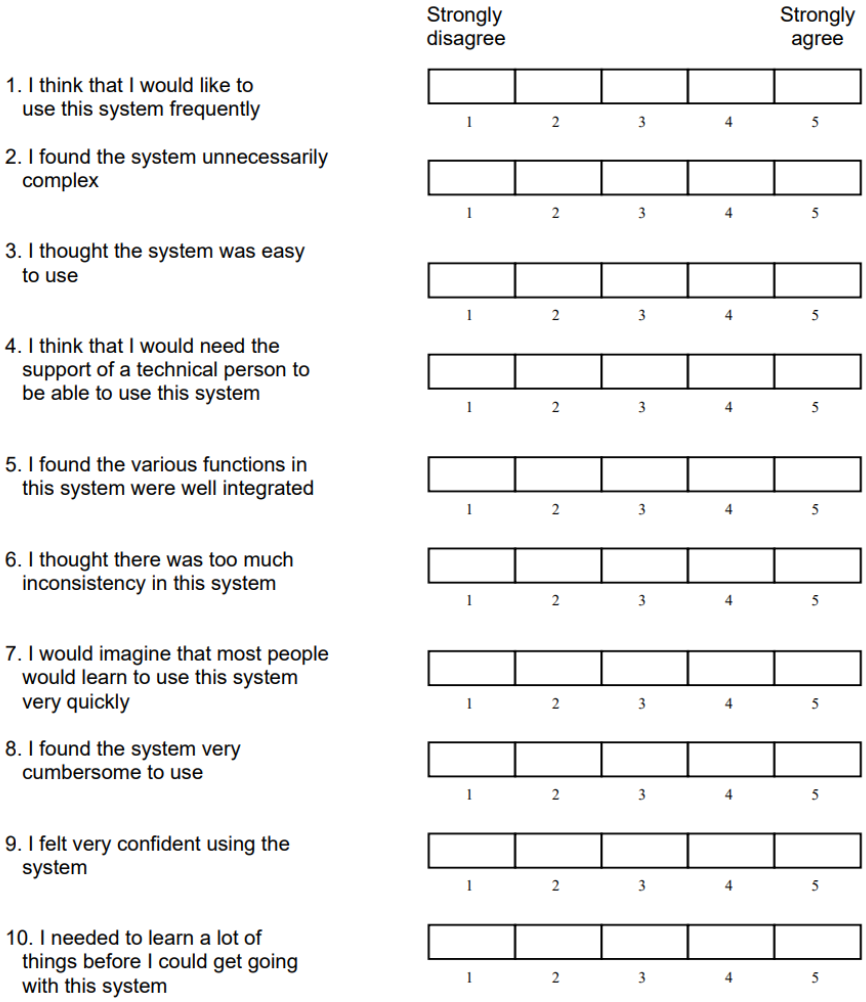


Figure 1.4: Original System Usability Scale [B+96]

CHAPTER 2

Theoretical Foundations and State of the Art

2.1 Definition of Artificial Intelligence

2.1.1 Different terms in the field of AI

The term *Artificial Intelligence* has now been mentioned in connection with many sectors [Abb21] and is one of the most promising drivers to increase innovation capabilities in organisations [YRS+20]. Therefore, the task of defining the expression is rather challenging. As a result, both inside and outside the field, the term *Artificial Intelligence* has been used in a wide range of contexts. It is impossible to anticipate a generally recognised definition of AI at the current level of research given the complexity of the term *intelligence* [Wan19]. AI is collectively referred to as this term more for historical than theoretical reasons and is now used to describe a variety of study areas, all of which have their own objectives, approaches, and use cases [Wan19]. Despite the fact that there isn't a single definition that applies to all study areas, an effort is made to define the word AI and clarify the differences between *Machine Learning (ML)* and *Deep Learning (DL)*, which are frequently mistaken for synonyms.

According to Ongsulee [Ong17] AI is used colloquially to describe situations of learning and problem-solving in which a computer imitates cognitive processes that people often identify as other human minds. In the research field of computer science, AI is often characterised as *intelligent agents*, that represent any kind of tool sensing its surroundings and taking measures to increase the chances of success. This perception of the environment could be based on collected data of human behaviour, like the strategic game of chess, as well as real-time data from sensors for self-driving cars. The goals in these examples range from winning a board game to arriving at a location safely, which demonstrates the variety of possible applications. For decades, AI has alternately been praised as a universal

solution and a vision of an overactive academic imagination. What supported both perceptions was that until around 2012, such technologies were only used by governments, research organisations and enterprises that use cutting-edge technology. Since that time, AI has moved past the speculative stage and into actual commercial solutions. The two main drivers for that were the faster and cheaper parallel processing capabilities through Graphics Processing Units (GPUs) and the simultaneous growth of storage options for data such as transactions, audio, images, videos, and plenty more.

On the other hand, the term *Machine Learning* exists in the field of computer science and builds a subset of AI [JK20]. Back in 1959, Arthur Samuel defined ML as a *field of study that gives computers the ability to learn without being explicitly programmed* [Mun14]. This is, according to Ongsule [Ong17], enabled by creating a model from data sample inputs for data-driven decision-making and predictions. Such algorithms evolved from studies of computational learning theory and pattern recognition in Artificial Intelligence. ML is used for a diverse range of automated tasks when it is difficult or impossible to implement explicit methods with high performance. Some examples of use cases like that are the detection of network intruders, email filtering, ranking data, speech recognition and computer vision. The computer-performed tasks allow to establish models that produce reliable results and reveal hidden insights by continuously learning correlations, trends, and patterns in historical data.

Jakhar and Kaur [JK20] outline three widely recognised methods in Machine Learning: supervised learning, unsupervised learning, and reinforcement learning. The first presents the most distributed approach of ML and is trained with labelled data to create and adjust a model that can predict the results of unknown data. An example would be the prediction of whether a patient has diabetes. Therefore, the model is fed with data like age, weight, and fitness level of already diagnosed people and can then make predictions for new patients with the help of their attributes. To forecast the label values on further unlabelled data, supervised learning applies patterns through techniques including classification, prediction, regression, and gradient boosting and is frequently used in situations where historical data is available. In contrast, unsupervised learning is employed for problems without past data. The objective is to examine the data and find some structure, cluster, or hidden patterns within the unlabelled datasets. One example is the clustering of customers according to their attributes, therefore it groups customers with similar properties, which can be helpful for individual marketing strategies. K-means clustering, self-organising maps, and Singular Value Decomposition (SVG) represent the most common techniques. With reinforcement learning, the algorithm discovers through trial and error which actions lead to the greatest rewards. This method is often used for better performance in gaming, autonomous navigation tasks, and the training of robots. The goal is to select actions that maximise the expected reward over a given amount of time. In order to reach the objective quicker, good guidelines and policies must be followed.

Deep Learning, as described by Jakhar and Kaur [JK20], is a significant concept that constitutes a subset of Machine Learning. This topic represents the research of Neural

Networks with multiple hidden layers that each contain several nodes. The result from one layer is used as the input for the next layer and the algorithms can be supervised or unsupervised. Deep Learning is based on learning representations of data, which can be, for instance, an observation like an image that has several possibilities to be presented, including sets of edges, regions in a specific colour, or shape, the intensity value for each pixel and many more. All these illustrations have their advantages and disadvantages in reducing the complexity of learning tasks like facial recognition. The objective of research in areas like computer vision, speech recognition, natural language processing, and bioinformatics is to develop models that can learn improved representations from massive amounts of unlabelled data. The presented concepts are summarised in Figure 2.1 for easier comprehension.

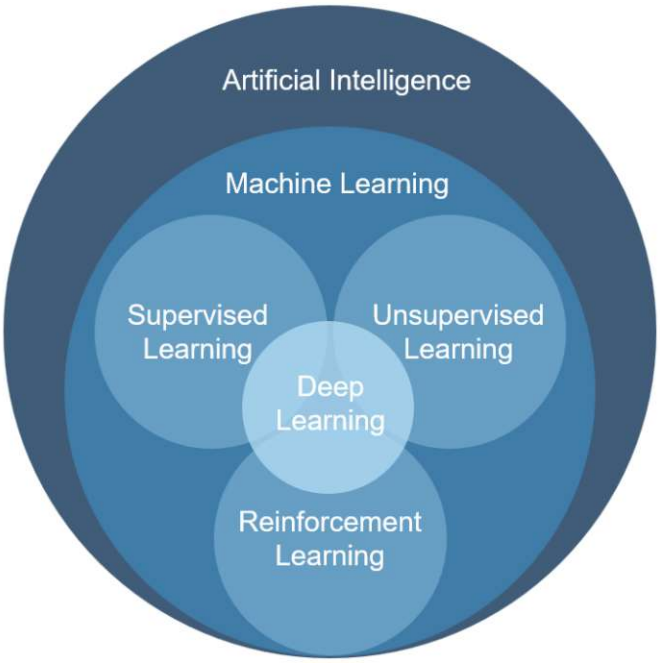


Figure 2.1: Areas of Artificial Intelligence [KP22]

2.1.2 The EU's scope of AI

The latest version of the AI Act by the Council of the EU defines an AI system as *a system that is designed to operate with elements of autonomy and that, based on machine and/or human-provided data and inputs, infers how to achieve a given set of objectives using Machine Learning and/or logic- and knowledge based approaches and produces system-generated outputs such as content (generative AI systems), predictions, recommendations or decisions, influencing the environments with which the AI system interacts* [Cou22]. The definition provided presents a more precise clarification of AI systems compared to the initial version of the Commission, which encompassed statistical

approaches, Bayesian estimation, search and optimisation methods, thus opening up a very broad scope beyond AI technologies [Eur22b].

Two important key terms in the definition are *Machine Learning* and *logic- and knowledge-based approaches*. ML was already described in the previous section but is now further outlined with the description of ML of the AI Act [Cou22]. There Machine Learning techniques refer to systems that can learn and draw conclusions from data to solve practical problems without explicit programming of each step. During the learning process, the parameters of a mathematical model that generates outputs based on input data are optimised. This process encompasses a number of approaches, including supervised learning, unsupervised learning, and reinforcement learning. These approaches can again contain methods such as Deep Learning with Neural Networks, statistical techniques including logistic regression and Bayesian estimation and search and optimisation techniques. Significantly different from the first definition, these methods may be incorporated into ML but cannot be classified as AI on their own. According to the AI Act [Cou22] logic- and knowledge-based approaches possess logical reasoning abilities and utilise knowledge to solve specific problems. In general, these systems comprise a knowledge base and an inference engine that generates outputs by applying logical reasoning to the knowledge base. This base is often constructed by human experts and represents relevant entities and logical relationships. The inference engine operates on the knowledge base and derives new information through various operations, like sorting or matching. *Logic- and knowledge based approaches include for instance knowledge representation, inductive (logic) programming, knowledge bases, inference and deductive engines, (symbolic) reasoning, expert systems and search and optimisation methods* [Cou22].

To gain a deeper understanding of the previously mentioned concepts, they will be elaborated on in greater detail in the subsequent text. The knowledge-based approach is a kind of AI that analyses data or knowledge with the goal of representing new knowledge for various scenarios [NBA+22]. With the support of AI concepts, the approach evaluates the context of the data to assist decision-making. Many effective agent-based technologies have been built on logic-based approaches, as a result, agents use logic to plan and coordinate their activities [CCDO20]. Modelling the actions of an agent is the main purpose of knowledge representation, where real-world data is displayed in a certain way to solve complex problems [SFN21]. Inductive logic programming produces logic programs from data, where this data is not represented as vectors like in most AI approaches but rather as logic programs, which promises, in comparison, better generalisation, interpretability, and only a small amount of training examples [CDM20]. Despite the fact that many contributions to Artificial Intelligence systems derive from computer science, statistics plays a significant role in the conceptual and practical understanding of AI, according to Friedrich et. al [FAB+21]. In their paper, they outline that the first examples emerged in the process of recognising the link between backpropagation and nonlinear least squares approaches. Furthermore, several AI subjects have connections to Random Forests, Support Vector Machines (SVM), nonlinear regression models, multiple multivariate, and ridge regression.

In addition, Turner and Sederberg [TS12] mention in their paper the significance of Bayesian statistics for analysing random variables, since it offers several advantages, including the ability to quantify uncertainty using the posterior distribution. Unlike conventional testing of the null hypothesis, directed examination of the posterior distribution permits statistical inference, which doesn't violate the process theory underlying the experiments that generated the data. Therefore, Bayesian estimation is recommended when it is difficult or impossible to determine the likelihood function.

The search and optimisation methods domain is extremely expansive and is already covered by the majority of the listed notions. In the field of Artificial Intelligence, finding a solution is often perceived as a search process across a spectrum of potential outcomes, wherein mathematics, including the statistical methods mentioned earlier, plays a crucial role in optimisation by striving not only to identify a solution but rather the optimal one specific to a given situation [SC14].

2.2 Overview of the AI Act

2.2.1 Emergence and current status

On the 19th October 2017, the European Council requested *a sense of urgency to address emerging trends: this includes issues such as Artificial Intelligence and blockchain technologies, while at the same time ensuring a high level of data protection, digital rights, and ethical standards. The European Council invites the Commission to put forward a European approach to Artificial Intelligence by early 2018 and calls on the Commission to put forward the necessary initiatives for strengthening the framework conditions with a view to enable the EU to explore new markets through risk-based radical innovations and to reaffirm the leading role of its industry* [Eur17]. Responding to this call, the Commission released a document titled *Coordinated Plan on Artificial Intelligence* [Eur18] on 7th December 2018. The primary goals outlined in this document included the implementation of AI initiatives, increasing investment in AI, and coordinating AI policies to prevent fragmentation. On 11th February 2019 in the Conclusions on this Coordinated Plan [Cou19], the Council emphasised the significance of upholding the rights of European individuals. It also advocated for a revision of existing legislation to align it with the new opportunities and challenges posed by AI. Subsequently, on February 19th 2020, the Commission presented the *White Paper on AI* [Eur20a], which proposed a European approach to trust in policy solutions, to expand AI usage and mitigate associated risks. Finally, on April 21st 2021, the proposal for a new regulation under the name of *Artificial Intelligence Act* [Eur22b] was published by the European Commission to set harmonised rules on AI. After several statements of committees, discussions, and opinion papers of the European Central Bank, European Data Protection Board, and European Data Protection Supervisor, the Council of the European Union adopted on 25th November 2022, the proposal [Cou22], which serves as the foundation for the requirements catalogue of this thesis.

2.2.2 Objectives

Artificial Intelligence is a rapidly expanding group of technologies that may provide a vast variety of economic, social, and environmental advantages across a large range of sectors. The EU considers, with the support of AI techniques, offering businesses and the European economy crucial competitive advantages [Eur22b]. Nevertheless, the same AI methods that enable socio-economic benefits also have the potential to introduce negative effects, new challenges and risks for individuals or whole countries. As a consequence, the EU wants to sustain technological power and enable access to future technologies, but with respect to the principles, values, and fundamental rights of the Union. For this purpose and to establish harmonised rules in the EU, the AI Act was created [Eur22b]. The Commission stated the following particular goals in their first proposal [Eur22b]:

- *ensure that AI systems placed on the Union market and used are safe and respect existing law on fundamental rights and Union values*
- *ensure legal certainty to facilitate investment and innovation in AI*
- *enhance governance and effective enforcement of existing law on fundamental rights and safety requirements applicable to AI systems*
- *facilitate the development of a single market for lawful, safe and trustworthy AI applications and prevent market fragmentation*

To reach these goals, the AI Act proposes a balanced and consistent approach to regulating AI systems. With the proposal, an attempt is made to limit requirements to a minimum for potential risks and challenges without slowing down technological progress or making it too costly to bring AI technologies to market. Therefore, on the one hand, it is complete in its basic regulatory choices, and on the other, it sets up an appropriate framework based on a well-defined risk-based approach that does not impose unwarranted trade restrictions [Eur22b].

2.2.3 Scope and structure

The AI Act's [Cou22] primary subject is to provide standardised guidelines for the placement and operation of AI systems on the European market. Additionally, the Act encompasses the prohibition of specific AI activities and imposes requirements for high-risk AI systems and obligations for their operators. Furthermore, the regulation addresses market monitoring and surveillance, as well as defines transparency criteria for AI systems that directly engage with individuals and content-producing AI programs. A fundamental principle underlying the AI Act is the differentiation of risk levels associated with the use of AI. This risk-based approach serves as the foundation for implementing a proportional set of enforceable regulations. Consequently, the requirements should match the degree and complexity of the threats that AI systems may pose. The risk-based approach categorises AI applications into four different risk levels, as shown in Figure 2.2:

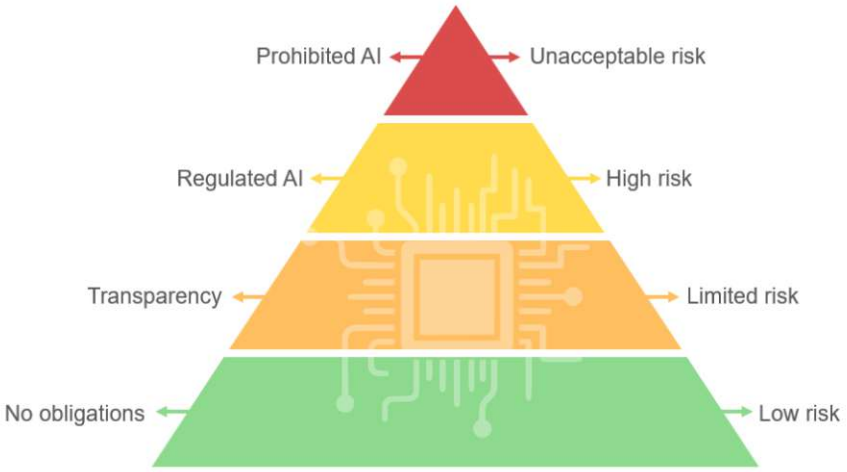


Figure 2.2: Risk categories of the AI Act [Com22b]

Starting at the top with prohibited AI Systems, these practices are not allowed, including [Cou22]:

- AI systems that employ subconscious tactics to substantially influence the behaviour of a person in a way that is most likely to result in physical or psychological damage to that individual
- AI Systems that abuse any of the weaknesses of a certain group related to their physical or mental incapacity or age to significantly change the behaviour of a member of that group
- AI systems that, on behalf of public authorities, classify or analyse the trustworthiness of real individuals over a time span based on their social behaviour or qualities, with the social score resulting in:
  - unfavourable handling of groups or individuals where the data of social situations is unrelated to the circumstances in which the information was initially created
  - unfavourable handling of groups or individuals that is not reflected by their social behaviour
- AI Systems that utilise real-time remote biometric identification methods in public areas for law enforcement purposes, with some exceptions such as the concentrated investigation of potential crime victims, proactive prevention of terrorism, avoidance of significant danger that poses a risk to critical infrastructure, human well-being or personal safety.

The next category of the risk-based approach represents AI systems that possess an elevated risk. For these high-risk AI systems to be viewed as such, one of the three subsequent requirements [Cou22] has to be met:

1. AI systems intended for use as independent products that fall under the Union harmonisation legislation mentioned in Annex II and require a third-party conformity assessment prior to market placement or use will be classified as high risk

2. AI systems intended for use as product safety components that fall under the Union harmonisation legislation mentioned in Annex II and require a third-party conformity assessment prior to market placement or use will be classified as high risk

3. AI systems mentioned in Annex III will be classified as high risk, unless the system's output is purely supplemental with regard to the intended action or decision and thus unlikely to pose a substantial risk to the well-being, safety, or fundamental rights of individuals.

This list in Annex II [Cou22] of the harmonisation legislation contains regulations and directives concerning machinery, toy safety, personal watercraft, lifts, protective systems in potentially explosive atmospheres, radio equipment, pressure equipment, cableway installations, personal protective equipment, appliances burning gaseous fuels, medical devices based on the New Legislative Framework (NLF) and civil aviation security, approval of motor vehicles, approval of two- or three-wheel vehicles, approval of agricultural and forestry vehicles, marine equipment, interoperability of the rail system based on the Old Approach legislation. Moreover, the EU has identified in Annex III [Cou22] certain application areas in which AI systems are directly classified as high-risk:

- *Biometrics*
- *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*

The next chapter focuses particularly on the management and operation of critical infrastructure, including AI systems in the distribution of water, gas, heating, electricity, and management of road traffic.

The majority of AI systems fall into the last two categories of the risk-based approach that pose limited or low risks. In the case of limited-risk technologies like chatbots, certain transparency obligations have to be fulfilled according to the AI Act [Cou22]. Providers are required to ensure that their AI system intended for human interaction is designed in a manner that clearly communicates to individuals that they are engaging with an AI technology. This only needs to be done if such awareness is not already evident to a well-informed and attentive person considering the context of use. The transparency obligation is of particular importance for AI tools that generate artificial audio, image or video content (deep fakes) that closely resembles real voices, individuals, objects or locations, which could be mistaken as authentic. Therefore, awareness is an important aspect, whereby the AI Act does not restrict AI systems that pose a low risk, for instance in spam protection filters or video games [Com22b].

2.2.4 Linked topics

As outlined in the *White Paper on Artificial Intelligence* [Eur20a], the proposed AI Act contributes to the extensive set of actions designed to address concerns faced by the development and applications of Artificial Intelligence. To make this technological progress beneficial to both people and companies and further fulfil its goal of having a climate-neutral Europe by 2050, the EU has developed a digital strategy [Com21]. More information on the strategy is provided in the communication *Shaping Europe's digital future* [Com20], where three foundational pillars concern technology that improves people's daily lives, providing a competitive but fair market and a sustainable society. It establishes a reasonable framework for trustworthy AI systems and goes along with *Europe's Digital Decade* [Com22a], which sets specific targets for 2030 in terms of infrastructure, governments, businesses, and skills of the population. AI approaches are tightly connected with data to draw conclusions or make predictions. Furthermore, the increasing amount of data from individuals encourages data-driven innovation and has the potential to bring many benefits to society across a vast array of industries. To enable save and high-quality pooling, distribution, and reuse of datasets, the EU published the *European Strategy for data* [Eur22a]. This strategy has the objective of providing reliable solutions and technologies in relation to data-driven AI and is therefore supported by the *Open Data Directive* [Off19] and the proposed *Data Governance Act* [Eur20b].

Due to the horizontal design of the AI Act, complete compliance with current existing Union laws is necessary since they apply to industries where AI systems with high risks are already deployed or will be in the coming decades. The proposal [Cou22] complements the rules for handling personal data, including processing principles, duties of controlling companies, and rights of the data owner, which are regulated in the *General Data Protection Regulation* [Eur16b] and further in the *Law Enforcement Directive* [Eur16a]. In addition, the AI Act [Cou22] will be included in the current sectoral safety regulations

for high-risk AI systems that are safety components of goods to maintain uniformity, prevent duplication, and reduce unnecessary responsibilities. Whereas in the case of the New Legislative Framework, the requirements of the AI Act will be evaluated as part of the current conformity evaluation processes under the NLF legislation but, on the other hand, do not apply directly to the Old Approach legislation. Therefore, according to the EU Commission, it is necessary to adapt these Acts under the Old Approach legislation in accordance with the rules of the AI Act to ensure that the governance and specifications of each sector are not compromised. Figure 2.3 displays and summarises the mentioned concepts and existing policies in connection with the AI Act.

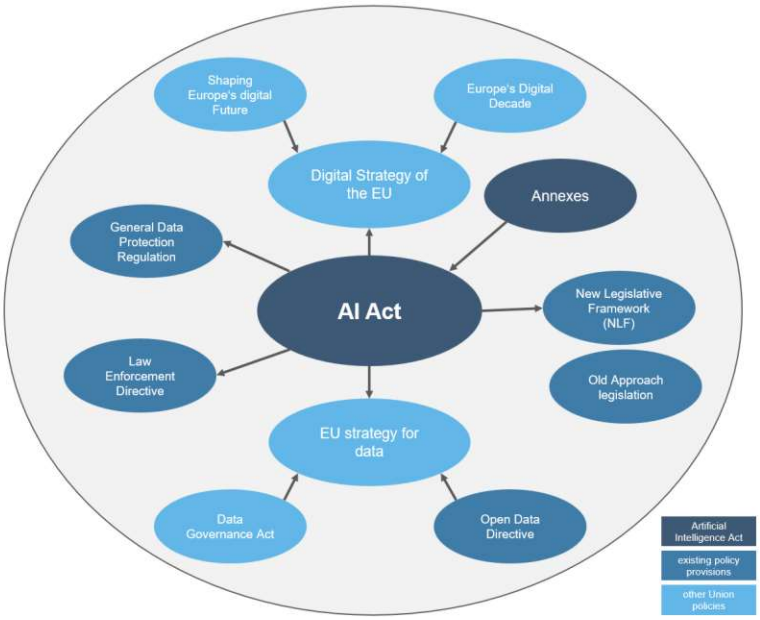


Figure 2.3: AI Act including linked policies

2.3 Purpose of AI-based applications in critical infrastructure

The Systematic Literature Review supports the Rigor Cycle by answering the first research question of this thesis. The following methodological steps of the SLR are based on the procedure of Yu Xiao and Maria Watson described in their paper *Guidance on Conducting a Systematic Literature Review* [XW19] and accordingly adapted to the scope of the thesis:

1. Step: Formulate the problem
   All literature reviews should be directed by research questions and consequently, these research questions guide the whole literature review process. As already mentioned, in this case, the direction is given by the first research question: *For*

what purposes is Artificial Intelligence used in critical infrastructure (supply of water, electricity, heating, gas, and road traffic management)?

2. Step: Develop the protocol
   The review protocol defines the procedures used to perform the review. It is essential for conducting robust systematic reviews since it decreases the researcher's bias. Further, it enhances the review's credibility since others can replicate the research using the same process of the protocol. Some aspects of the protocol are directly presented in the thesis, including the research question, used databases, inclusion and exclusion criteria, and reporting the findings. Documented in a distinct Excel file are procedures such as documenting the results of the search query, removing duplicates and screening according to inclusion and exclusion criteria.

3. Step: Search the literature
   The search procedure is an essential component of a SLR and needs also to be done in a systematic manner. In terms of the channel, online databases provide the most important source for published literature collections [PR06]. In this thesis, the chosen databases are CatalogPlus, Web of Science, and IEEE Xplore, for certain reasons. CatalogPlus is an invaluable resource for TU Vienna students as it provides full-text access to a diverse array of academic publications, such as conference papers, e-journals, and university publications [cat23]. In addition, Web of Science stands out as an exceptionally comprehensive multidisciplinary database, encompassing approximately 100 million items across various academic disciplines [Cla23]. Similarly, IEEE Xplore is a renowned database known for its specialisation in electrical engineering, computer science, and related fields [IEE23].

   For the actual research in those search engines, keywords should be defined and determined from the research question. These keywords in the thesis follow in general the same structure for all the different utility sectors (water, gas, heating, road traffic management and operation) and include search strings containing the Booleans *OR* as well as *AND*:

   Artificial Intelligence **OR** *AI*
   AND
   supply of water **OR** *supply of gas* **OR** *supply of heating* **OR** *supply of electricity* **OR** *road traffic management* **OR** *operation of road traffic*

   To refine the results of the research in advance, the following restrictions are introduced, which can be configured directly in the knowledge databases, therefore only papers are displayed, that:

   - were written in English
   - were published between the years January 2018 and March 2023
   - are available via open or university access

4. Step: Screen for Inclusion and exclusion
   After generating a list of references, the next step is to examine each paper to determine whether it should be included for further analysis. This first screening is intended to exclude publications whose content is irrelevant to the research question. To support this step, criteria for inclusion and exclusion should be formulated, like in the case of this thesis: The inclusion criteria are already covered by the above-mentioned restrictions of the search engines and the exclusion criteria reject papers that:

   - are duplicates (occur in multiple databases)
   - deal with the subject of AI (including Machine Learning, Deep Learning, logic- and knowledge based approaches), but not with the context of the utility area
   - deal with the context of the utility area but not with the subject of AI (including Machine Learning, Deep Learning, logic- and knowledge based approaches)
   - are opinion papers
   - are surveys
   - are literature reviews

5. Step: Assess Quality
   After the screening, where the estimation is mostly based on the title, abstract or conclusion, the quality evaluation refines the pool of references by going through the full text of each paper. This also provides an opportunity to understand each study in its entirety before proceeding to the next steps.

6. Step: Extracting and synthesising Data
   When the selection of papers is completed, the process of data extraction collects applicable information from the respective work. The extracted data is then used to synthesise the findings of the included work and draw conclusions about the research question being addressed by the review. This data synthesis process summarises the purposes of different academic papers to provide an overview of the current developments and application areas of AI in critical infrastructure.

7. Step: Report Findings
   The final step of the SLR is to document the findings and research process in order to be consistent and independently reproducible. Thus, the screening process is documented in Excel tables and the findings are presented directly in this thesis. In each domain, the discovered application areas will be described and accompanied by a small extract from particularly interesting articles.

Figure 2.4 provides an overview of the SLR filtering process, which resulted in 153 relevant papers that serve as the foundation for subsequent sections.

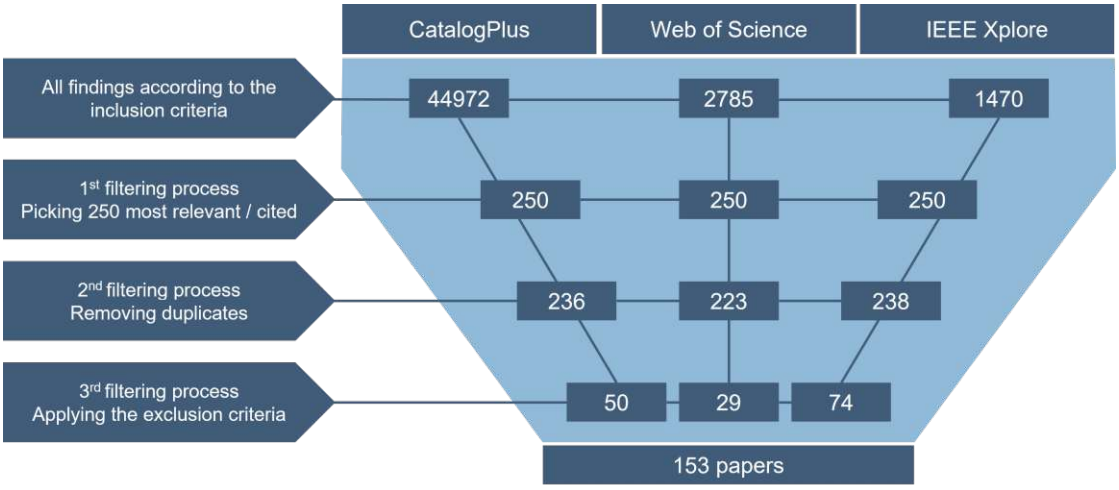


Figure 2.4: Filtering Process of the SLR

2.3.1 Supply of water

In the course of the Systematic Literature Review, 21.33% (32 out of 150) of the published articles met the selected criteria and were considered relevant to the domain *supply of water*. The application areas were identified by synthesising the data within this domain and are described in the following sections. Further, the percentage distribution of the relevant articles for each application area is shown in Figure 2.5.

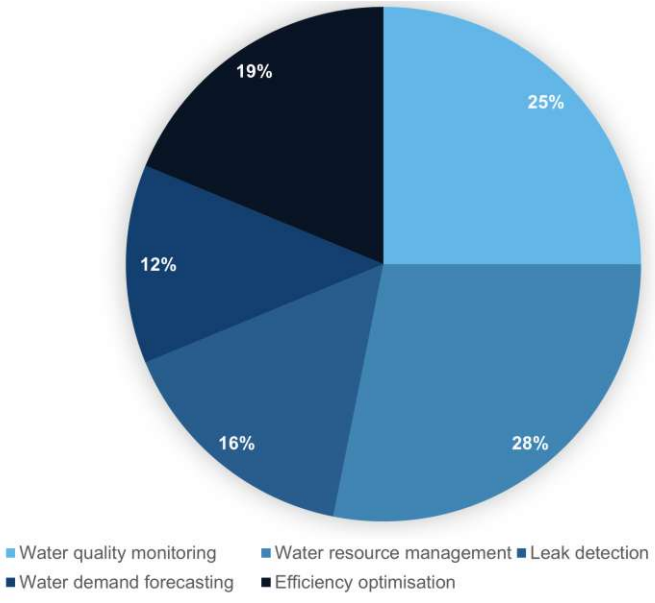


Figure 2.5: Application area distribution (supply of water)

Water quality monitoring

In this application field, Artificial Intelligence is used to analyse water samples in order to predict water turbidity [SB19], water hardness [NSN+22], pigment concentration [HSK19], and quality changes in groundwater [SSA+22]. In addition, detecting trends and anomalous events like intentional contamination of water pollutants allows for the rapid identification and resolution of any quality concerns [NSN+22] [VSR+21].

Quality predictions are the focus of the paper by Sarreshtedar et al. [SSA+22], which aims to investigate the groundwater variations the north-western Iran. The study used an ensemble of AI-based modelling techniques, including the Artificial Neural Network (ANN), Adaptive Neuro-Fuzzy Inference System (ANFIS), and Support Vector Regression (SVR) models, to forecast groundwater quality and quantity changes. The results show that the ensemble modelling technique provided more accurate predictions of groundwater quality and quantity changes compared to individual models.

Another interesting article in this field was published by Stevenson and Bravo [SB19]. They developed a decision support tool that can predict >1 Nephelometric Turbidity Unit (NTU) events up to seven days in advance so that water supply managers may take preventative action. Turbidity is a measure of the visual quality of water produced by suspended particulates, indicating poor water quality and perhaps concealing the presence of parasites like Cryptosporidium. The World Health Organisation (WHO) recommends that turbidity levels should not exceed 1 NTU prior to chlorination. The study applies Machine Learning techniques such as Generalised Linear Model (GLM) and Random Forest to predict turbidity peaking events at groundwater sources on the South Coast of England. The models obtained satisfactory results, demonstrating that ML approaches are suitable for forecasting turbidity-peaking occurrences.

The described papers contribute significantly to the development of reliable water quality monitoring and prediction models that are able to support management authorities and water treatment facilities in planning and maintaining their operations. This results in avoiding temporary shutdowns, supporting the decision-making process and ensuring the delivery of safe, high-quality water to consumers.

Water resource management

This application area, along with approaches of AI, presents possibilities to enhance water resource management by analysing vast quantities of data to give insights into groundwater levels [MMNK22a] [CBGM21], mapping of probable groundwater springs [AFPAS+20] and improving the water reservoirs security [PSA+22]. With the support of smart water metering technology [NSZ+18] and innovative water management frameworks [PSA+22], AI-driven water resource management aims to ensure sustainable access to water resources and secure ecosystems.

Climate change has a significant impact on ecosystems as well as water resource management. Focusing on this issue, Xiang et al. [XLKK21] suggest the use of Adaptive

Intelligent Dynamic Water Resource Planning (AIDWRP) for sustainable water development in urban areas. AIDWRP is an intelligent adaptive strategy that uses AI techniques to address the problem of dynamic water resource management with yearly consumption and released location-based limits. This approach enables the optimisation of various efficient policies for environmental planning and management.

Accurately predicting the levels of freshwater lakes also plays an essential role in effective water resource management. Consequently, the paper by Bonakdari et al. [BESG19] demonstrates the use of four advanced Artificial Intelligence models (Minimax Probability Machine Regression (MPMR), Relevance Vector Machine (RVM), Gaussian Process Regression (GPR), and Extreme Learning Machine (ELM)) to forecast lake level fluctuation in Lake Huron (North America) using historical datasets. The conclusion of this study is that the MPMR model can be employed as a viable computational tool for the sustainable management of Lake Michigan-Huron's resources in both present and future planning.

Water is a crucial resource for economic and social development. The presented articles contribute to the appropriate management of water resources, including the allocation of water resources, while guaranteeing the sustainability of water systems and environmental preservation.

Leak detection

AI can be applied in this field to analyse real-time data from sensors and further pinpoint leaks with algorithms, that involve signal denoising and feature extraction [ZPL+19]. Using data like pressure, flow, acoustic signals, and temperature, AI-based systems are able to discover anomalies and trends that indicate the existence of a leak [JBSP19] [ZLW19]. These leaks may vary from localised damages, such as industrial accidents, to slow leaks produced by ageing infrastructure [PML18].

Common leak detection techniques that collect data from acoustic emission or pressure sensors are difficult to use on a broad scale. The research of Massaro et al. [MPG21] suggests employing infrared thermography and drones to monitor expansive regions in real-time. The system combines several detection technologies, such as Infrared Thermography and Ground Penetrating Radar, with AI algorithms for locating water leaks and assessing hydrogeological hazards. This approach is also able to forecast the hydrogeological risk of an area by combining meteorological data with hydrogeological risk maps.

Another interesting contribution is presented by the study of Zhou et al. [ZPL+19]. It proposes a new method for detecting and locating pipeline leaks using a combination of Improved Spline-local Mean Decomposition (ISLMD) and Deep Neural Networks. ISLMD effectively removes noise interference from a signal and decomposes the signal, making it possible to obtain reasonable extraction of leak information. In addition, a DNN employs a method of Deep Learning to recognise patterns in the data and precisely pinpoint the leak.

The presented techniques can efficiently identify and localise leaks in pipelines, which assists in lowering the risk of environmental harm as well as the economic loss that is caused by pipeline failures and water wastage.

Water demand forecasting

In conjunction with AI and smart water metering technology [RG20], this application area is able to evaluate complex information, recognise trends and create reliable projections. Models can be trained using historical water consumption data or real-time data from sensors [NRH20] to predict seasonal water availability [FG19] and future water demand on different time horizons [ZMR23].

One of the articles on this subject is written by Zanfei et al. [ZMR23] and discusses the use of an ANN to construct a model for anticipating the water demand of an urban water distribution system. The model is trained using actual water demand time series, augmented with historical data, meteorological, and calendar factors. It is intended to be adaptable, allowing for the accurate forecast of water demand across a range of time horizons. In addition, this model is able to guide water utilities in their daily operations and decision-making processes.

Another compelling study by Fleming and Goodbody [FG19] addresses the integration of several Machine Learning, statistical, and optimisation approaches to update and enhance an existing principle components regression framework for water supply forecasting in the West of the United States. This territory is significantly complex owing to conflicting demands for hydroelectric power generation, agribusiness, and residential water consumption. The new method was designed to be more precise than previous systems, in order to manage diverse prediction errors and to be inexpensive to implement. The US Department of Agriculture has adopted the resulting prediction engine, which has the potential for application in other sectors.

The articles mentioned above demonstrate that the integration of AI approaches is a viable option for regulating water demand, lowering the risk of water shortages and maximising resource utilisation. In addition, by using AI-based water demand forecasting, water utilities can enhance their planning and decision-making, optimise their operations and lower the costs connected with water supply and distribution.

Efficiency optimisation

The last identified application area in the water supply domain is efficiency optimisation. Here, AI can be used to optimise the operations of water treatment plants like seawater purification [PD21] and Fluoride removal [TMK+19]. Furthermore, AI techniques are able to support utilities by maintaining constant pressure for water supply [LF20] and providing optimum water transfer solutions to reduce freshwater imbalance [AZH+22]. This helps utilities improve their overall efficiency and reduce costs.

Severe and long-lasting social as well as financial consequences can be initiated among others by Cyber-attacks against critical water system infrastructure. To mitigate these risks, intrusion detection systems serve as a supplementary security mechanism to identify assaults. The paper by Ramotsoela et al. [RHAM20] focuses on the development of a behavioural intrusion detection approach using predictive Neural Network architectures and a novel voting-based ensemble technique. The research uncovered that when multiple algorithms collaborate, they can overcome their individual limitations and create a more resilient algorithm that yields better outcomes.

An additional factor that minimises costs and improves efficiency in water supply systems is the prevention of failures. Pérez-Padillo et al. [PPPMM22] outline in their article the creation and deployment of a web application for controlling breakdowns in aged water supply systems, which was tested successfully in a Spanish water supply company. The system captures real-time data using wireless water pressure sensors and transmits it to an IoT platform. The platform incorporates a rule-based decision algorithm that categorises faults based on pressure measurements and delivers repair warnings. This instrument can assist in the effective management of problems in aged water supply systems, allowing for prompt repairs.

These two contributions demonstrate the potential to optimise the security and resilience of water distribution systems against cyber-attacks and the effective detection of failures to minimise supply interruptions, along with cost reduction.

2.3.2 Supply of electricity

In the course of the Systematic Literature Review, 27.33% (41 out of 150) of the published articles met the selected criteria and were considered relevant to the domain *supply of electricity*. The application areas were identified by synthesising the data within this domain and are described in the following sections. Further, the percentage distribution of the relevant articles for each application area is shown in Figure 2.6.

Load and price forecasting

In this application field, AI systems analyse patterns and trends in energy production and electricity consumer attributes [WBL+21]. They take into account variables such as historical data, seasonality, time of day, weather conditions, and other factors [BNP20]. The different data enables predictions regarding electricity prices [PMI+19] [LHZ18] [PMM+21] and provides decision support for the selection of price plans [LCM+20]. Another especially valuable aspect for power suppliers is the forecast of short- and long-term electricity loads enabled by AI solutions [FABE21] [Sol20] [ZQS18] [MMNK22b] [AK20] [HKA+23].

One interesting article in this field by Behm et al. [BNP20] presents a technique for generating synthetic European electricity load profiles using Artificial Neural Networks. The objective is to develop long-term predictions by training dense Neural Networks using historical data from Germany. The input parameters consist of astronomical data,

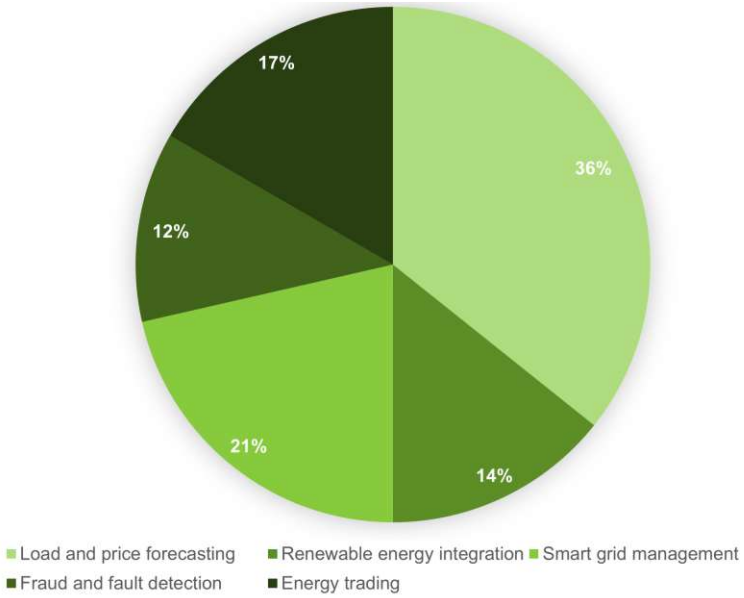


Figure 2.6: Application area distribution (supply of electricity)

yearly peak loads, and meteorological information. Additionally, the research evaluates the significant effects of external temperature and wind speed on electrical demand. The findings demonstrate a considerable accuracy increase compared to other existing state-of-the-art approaches. The synthetic load profiles are used to provide forecasts for the year 2025 for Germany, Sweden, Spain and France, illustrating their pan-European relevance.

In contrast, the article by Lu et al. [LCM+20] tries to assist smart grid end users in selecting electricity pricing plans. The system uses a Markov Decision Process (MDP) with an improved state framework to model the decision problem and a Kernel approximator-integrated batch Q-learning algorithm to solve it. The algorithm extracts hidden features from the time-varying pricing plans and can predict a precise policy for individual users and therefore reduce their costs. The suggested technique was evaluated using actual historical pricing plans and the results demonstrate that it can successfully optimise consumption cost portfolios.

Overall, the contributions to this topic presented above aid energy suppliers in optimising their supply and avoiding over- or under-production. Additionally, by selecting the most appropriate price plans, more end customers are able to reduce their power expenses and usage concerns.

Renewable energy integration

AI solutions have the potential to revolutionise the integration of renewable energy sources into energy supply systems. In the case of solar energy, AI can aid in choosing the optimal

location for photovoltaic solar systems by considering factors such as sunlight exposure, shading, and terrain [WDB+21] and further aid in the prediction of daily global solar radiation [AGB21]. In the case of wind power, AI can assist in wind speed forecasting for wind power station planning, enabling better decision-making and reducing the likelihood of power grid disruptions due to fluctuations in wind energy supply [ZWZ19]. The achievement of energy sustainability in smart cities [CLV18] and the integration of renewable energy supply for buildings [NMZ+21] are additional key factors in this application field.

The energy supply for megacities, especially in developing nations such as Iran, is a key research topic in urban management. Combining Artificial Intelligence, renewable energy and Transformational Participation (TP), the smart city concept is a unique attempt to help cities produce energy responsibly. The study by Ghadami et al. [GGK+21] analyses Mashhad's (Iran) electrical energy consumption by using AI and statistical analysis. In addition, a photovoltaic technology system simulation programme is used to evaluate the solar energy potential over the course of one year. In conclusion, three primary incentive strategies are identified for solar energy production in the short-, medium- and long-term planning horizons.

The forecast of daily worldwide solar radiation represents an additional important component for the design, operation, and integration of solar energy conversion systems, as well as for the selection of regions and investment strategies. In the article by Ağbulut et al. [AGB21], the daily global solar radiation data of four provinces in Turkey with various solar radiation distributions are predicted using four Machine Learning algorithms: Support Vector Machine (SVM), Artificial Neural Network (ANN), kernel and nearest-neighbour, and Deep Learning. The algorithms are trained using the daily minimum and maximum ambient temperature, cloud cover, daily extraterrestrial solar radiation, day length, and solar radiation of these provinces. All evaluated Machine Learning algorithms can reliably anticipate daily global sun radiation statistics, according to the findings. The ANN algorithm is determined to be the best-fitting method, followed by Deep Learning, SVM, and kernel and nearest-neighbour, in that order.

The articles point out possibilities for the integration of renewable energy, with a particular focus on solar energy. In the end, this can lead to increased efficiency, a more sustainable power supply, improved investments and energy production strategies.

Smart grid management

AI systems in this field try to balance electricity generation and consumption in order to enhance the stability of smart grids [LH19] [AKK+20] [WOH+20]. Another application in this area is solving the sizing optimisation problem for hybrid microgrid systems [CDL+20]. Algorithms can determine the optimal configuration of microgrids to ensure reliable and cost-effective management of energy supply [KDV18].

The electric grid is a system that transmits electricity from power plants to consumers, consisting of communication lines, control stations, transformers, and distributors. Mas-

sive amounts of energy are generated, mandating effective management to deliver energy to many domains like households, companies, and smart cities. To solve this problem, the study by Alazab et al. [AKK+20] proposes the deployment of intelligent systems that mix IT infrastructure and physical systems. The Machine Learning module is the IT component of this system, while the power dissipation units are the physical entities. Multidirectional Long Short-Term Memory (MLSTM) presents a new technique that outperforms conventional Machine Learning approaches and has been developed to predict the stability of the smart grid network.

Maintaining reliable smart grids also requires a balance between power production and consumption. Using reinforcement learning and Deep Neural Network approaches, the article by Lu et al. [LH19] offers a real-time incentive-based demand response algorithm for smart grid systems. The algorithm is designed to allow the service provider to acquire energy resources from registered consumers in order to balance energy variations and improve grid stability. DNNs are employed to estimate future costs and energy consumption. At the same time, reinforcement learning is used to determine appropriate incentive rates for consumers, considering the profits of service providers and customers.

The presented approaches contribute to the optimised stability and reliability of smart grids by balancing energy resources. Additionally, transmission losses are reduced and the profitability of both service providers and their clients is enhanced.

Fraud and fault detection

AI assists in this application field with the fault detection of distribution networks, which helps identify the exact location of faults for faster restoration of power supply [jie20]. AI also plays a vital role in detecting anomalies and frauds in smart meters, which ensures accurate billing and reduces revenue losses [BTACRGE19]. Furthermore, it aids in the detection of electricity theft, cyber-attacks [ISNS20], and Non-Technical Loss (NTL) detection [GAA+19], enabling utilities to prevent revenue leakage.

NTL, commonly known as energy theft and fraud, accounts for considerable revenue losses for power companies. A new end-to-end solution proposed by Buzau et al. [BTACRGE19] employs a hybrid DNN to identify abnormalities and frauds in smart meters. Training the network using raw, unprocessed data eliminates the requirement for manual feature engineering. The suggested architecture is comprised of a network with long-term memory and a network with multi-layer perceptions. The first network examines the raw daily energy usage history, while the second network incorporates non-sequential variables such as contracted power or geographical information. In NTL identification, the hybrid Neural Network outperforms contemporary classifiers and previous DL models. The model was trained and validated using actual smart meter data from Endesa, the biggest power provider in Spain.

Manual detection of NTL is expensive and labour-intensive, as highlighted in the paper by Ghori et al. [GAA+19]. It examines 15 known Machine Learning classifiers using a real-world dataset from a Pakistani power provider. The findings demonstrate that ensemble

approaches and ANN perform better than other classifiers for NTL identification. Further, a process is developed to determine the top fourteen characteristics that contribute most to predicting NTL.

The articles describe AI techniques that are able to detect possible fraud at an early stage, enabling utilities to take preventative measures and minimise revenue losses. In addition, AI can assist energy providers in gaining a deeper understanding of consumer behaviour and preferences to enhance customer service.

Energy trading

The use of AI systems is transforming the energy trading landscape by enabling more efficient and effective energy trading strategies. AI-based approaches are able to model the bidding strategies of generation companies, taking into account factors such as production costs and market trends to optimise bidding decisions [YQS+19] [LGDH20]. An energy trading scheme can be designed to choose the most suitable electric energy trading policy according to the predicted future renewable energy generation, maximising the utilisation of renewable energy resources [LXX+19]. Furthermore, a community management approach capable of implementing customer-to-customer trading allows communities to trade energy among themselves, fostering local energy markets [ZHG+19] [CS18].

In energy markets, companies try to make bid selections that provide the maximum profit. There are a variety of strategies, including bi-level optimisation and reinforcement learning, for modelling these decisions. However, these methods show limitations in their effectiveness due to their constraints. Some disregard crucial operating aspects of market participants, while other approaches must simplify the problem to the point that it is no longer accurate. To overcome these restrictions, the paper by Ye et al. [YQS+19] introduces a unique deep reinforcement learning method that combines a deep deterministic policy gradient technique with a prioritised experience replay strategy. Utilising multidimensional continuous state and action spaces that take into account non-convex operational features, this method enables accurate feedback and lucrative bidding selections within various case studies.

The authors Zhou et al. take a different approach in their paper [ZHG+19] by providing an innovative method for managing energy in residential communities that permits Peer-to-Peer (P2P) trade. It involves the establishment of a local energy pool in which residential users can trade energy and get inexpensive renewable energy without installing new energy generation technology. The energy pool gathers surplus power from consumers and renewable resources in order to sell it at a price above the Feed-in-Tariff but below the retail price. The price level is controlled by a real-time demand/supply ratio, which is impacted by retail pricing, user behaviour, and the availability of renewable energy. The energy trading process is described as a Markov Decision Process and an algorithm for reinforcement learning is utilised to determine the MDP's best decision. To address the continuous state-space issue, Q-learning is implemented using the fuzzy inference system.

In a community, a numerical analysis is undertaken, where the results indicate that the suggested system is able to significantly reduce electricity expenses while regulating energy demand efficiently.

The presented approaches outperform previous state-of-the-art strategies in terms of profitability and computational performance. They thus allow traders to make more informed choices about when to purchase and sell energy, which can result in improved earnings and lower risk. In addition, AI systems assist in automating trade procedures, hence minimising the need for human interaction and increasing productivity.

2.3.3 Supply of heating

In the course of the Systematic Literature Review, 10% (15 out of 150) of the published articles met the selected criteria and were considered relevant to the domain *supply of heating*. The application areas were identified by synthesising the data within this domain and are described in the following sections. Further, the percentage distribution of the relevant articles for each application area is shown in Figure 2.7.

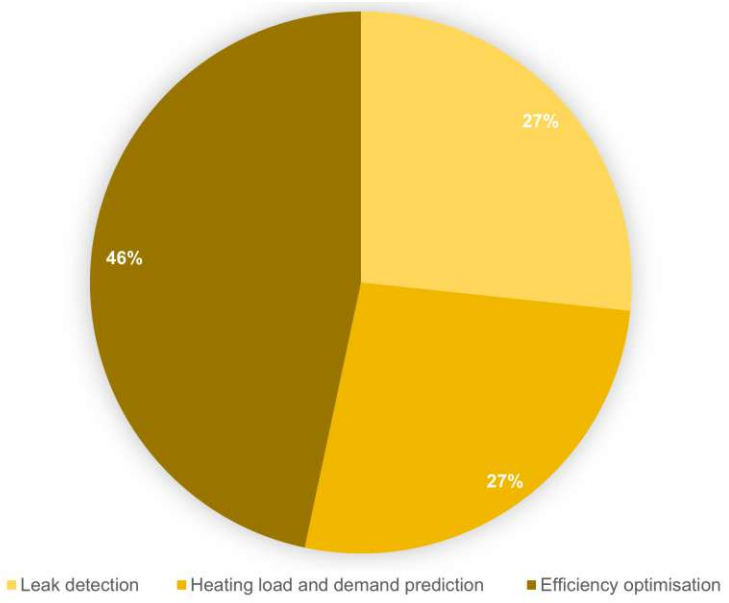


Figure 2.7: Application area distribution (supply of heating)

Leak detection

In this application field, AI can be applied to discover leakages in heating supply systems by implementing an anomaly detection scheme to identify deviations from normal operating conditions [PMH20]. This can be combined with leakage localisation techniques to pinpoint the exact location of the leak source [PVG+20].

Leakages in heating pipes are a regular issue for power supply providers and can remain undetected for an extended period of time. This demonstrates the significance of monitoring these heating networks. A very unique solution for this problem is proposed by Hossain et al. [HVF20]. Using infrared photographs collected by an Unmanned Aerial Vehicle (UAV), it automatically identifies energy leakages in the subsurface pipes of district heating systems. This process involves an area extraction algorithm and classification with a Convolutional Neural Network (CNN) and eight traditional Machine Learning classifiers. The suggested technique obtained an average weighted accuracy of 0.872%, identified about 98.6% of the real leakages and was found to be competitive when compared to other state-of-the-art practices.

A different approach is presented by Fan et al. [FGWL19]. Their research provides a two-level diagnostic model for identifying leakage defects in secondary networks for central heating using Deep Belief Networks (DBN) under constant and modest supply flow quality management. Utilising a hydraulic computation approach based on graph theory, the model calculates network pressure changes. The first-level diagnostic model employs a DBN to locate defective pipe segments, whereas the second-level model predicts the precise location of the leak. On a branch-pipe network and a loop-pipe network, the model demonstrated more accuracy than conventional approaches such as Back Propagation Neural Network and Support Vector Machine.

The presented papers illustrate that, by employing AI solutions, it is possible to identify patterns and trends in the data that may not be apparent to human operators. This enables more accurate leak detection and diagnosis, which improves the reliability of heat supply systems and reduces maintenance costs and downtime.

Heating load and demand prediction

AI systems enable in this application area the prediction of heating load and demand in buildings [LJL+21] as well as heat exchange stations [SXP+20]. Furthermore, such systems help to determine the optimal thermostat configurations, taking into account factors such as outdoor temperature information and daily consumption of electricity [Wan21].

Concerning heating load prediction, the authors Song et al. [SXP+20] present a prediction model based on a temporal CNN for Smart District Heating Systems (SDHS). SDHS plays a significant role in the future of energy conservation and pleasant heating. Nevertheless, heating load prediction is a complex nonlinear optimisation issue with low forecast accuracy due to the weak nonlinear expression capability of traditional prediction algorithms. To rapidly extract complex data characteristics, the suggested model combines the parallel feature processing of Convolutional Neural Network with the time-domain modelling capacity of a recurrent Neural Network. Using engineering data from four heat exchange stations in Anyang, China, the performance of the model was tested during the 2018 heating season and showed more accurate results than state-of-the-art algorithms.

Using a two-year hourly dataset from buildings in Espoo, Finland, Eseye and Lehtonen [EL20] offer a new ML-based approach for predicting district heating system heat demand. This model incorporates empirical mode decomposition, an imperialistic competitive algorithm and a SVM, in addition to a strategy for selecting features based on a binary genetic algorithm and GPR. Compared to other forecasting models, this model exceeds the competition in terms of accuracy measures and delivers improved prediction accuracy.

Through precise forecasting of demand and heat supply, businesses can improve their resource management and minimise waste, resulting in cost savings and enhanced energy efficiency. The presented AI-based heating load and demand prediction models have been shown to outperform traditional methods, making them a promising tool for the heat supply industry.

Efficiency Optimisation

AI systems are employed in the efficiency optimisation of heating supply, with several strategies being utilised. One such strategy involves optimising the scheduling of heat and power generating units [GES+21] and a further integrated energy system [DWY+21] to improve energy efficiency. Another approach combines smart residential hot water systems with AI to reduce the required energy supply [MBGM22]. Additionally, a data-driven optimisation strategy is used to reduce energy use while maintaining thermal comfort [WWKF20].

Improving the operation of heating plants and heat distribution systems in variable climates is essential. Woźniak et al. [WKMP18] examine the use of bio-inspired techniques to maximise the efficiency of a district heating plant while decreasing its expenses. The system was calibrated using a Polar Bear optimisation approach and the results were compared to those of the Particle Swarm optimisation method. The study's findings demonstrated that the suggested technique was effective in all simulated weather and boundary circumstances. The comparison of outcomes with non-optimal parameters substantiated the need for optimum system settings.

District heating systems are prevalent throughout Northern Europe and account for the greatest proportion of the Swedish heat supply market. However, these systems often fail to operate as intended due to a variety of errors or improper procedures. Due to economic and energy efficiency considerations, the night setback control approach has been deemed inappropriate for contemporary, well-insulated buildings since it might cause abrupt morning peak difficulties for utility providers. This paper by Zhang et al. [ZBF21] presents a Neural Network with bidirectional long-short-term memory and an attention mechanism for classifying substations that often use night setbacks. The efficacy of the suggested method is assessed using information from 10 anonymous Swedish substations. The accuracy, recall, and f1 score are utilised as performance metrics and the results of out-of-sample testing indicate that the suggested method beats the study's baseline models.

The papers presented offer AI solutions as a valuable tool to allow heat supplies to optimise their systems in order to maximise heat production, reduce operation costs and improve the stability of the heat supply.

2.3.4 Supply of gas

In the course of the Systematic Literature Review, only 6.67% (10 out of 150) of the published articles met the selected criteria and were considered relevant to the domain *supply of gas*. Due to the fact that only a small percentage of articles were relevant, no further synthesis of the data was conducted. Thus, it was not possible to identify any specific application areas in the field of gas supply. The limited number of relevant academic articles within this particular utility domain involve subjects such as: proppant detection for quality management [MC19], fault detection of electrical gas generators [ALA21], optimal scheduling of integrated energy systems [DWY+21], estimating the delivery time of oxygen gas cylinders [GAMQ22], fraudulent consumer detection [KKGG21], natural gas IIoT architecture based on blockchain and AI [MZG20] [MSW+20], decision-making optimisation for gas exploration and production [ARJZS21] [RP18].

2.3.5 Road traffic management

In the course of the Systematic Literature Review, 36.67% (55 out of 150) of the published articles met the selected criteria and were considered relevant to the domain *road traffic management*. The application areas were identified by synthesising the data within this domain and are described in the following sections. Further, the percentage distribution of the relevant articles for each application area is shown in Figure 2.8.

Traffic flow prediction

AI plays a very significant role in this application field, particularly in short-time traffic flow prediction [ZLFC20] [GLX+19] [CYY+20]. By leveraging Machine Learning techniques, AI models can analyse traffic patterns and predict traffic congestion with high accuracy [SCP20]. In addition, AI-based solutions can also model the preferences of passengers making different transportation choices, such as using public transportation or ride-sharing services, which further enhances traffic flow prediction [WLB+18] [DPW+19]. Additionally, predictive cruise control can help optimise traffic flow by predicting traffic conditions ahead and adjusting vehicle speed accordingly, leading to safer and more efficient transportation systems [MG20].

For successful urban traffic management, short-term traffic forecasting is essential. Nevertheless, non-recurring events such as road closures, accidents, and severe weather might impact the accuracy of traffic prediction. Complementing traffic data with social media data, particularly Twitter datasets, can increase the accuracy of predictions. The paper by Essien et al. [EPSS21] presents a Bi-directional Long Short-Term Memory (LSTM). This Deep Learning model combines traffic and meteorological data with information

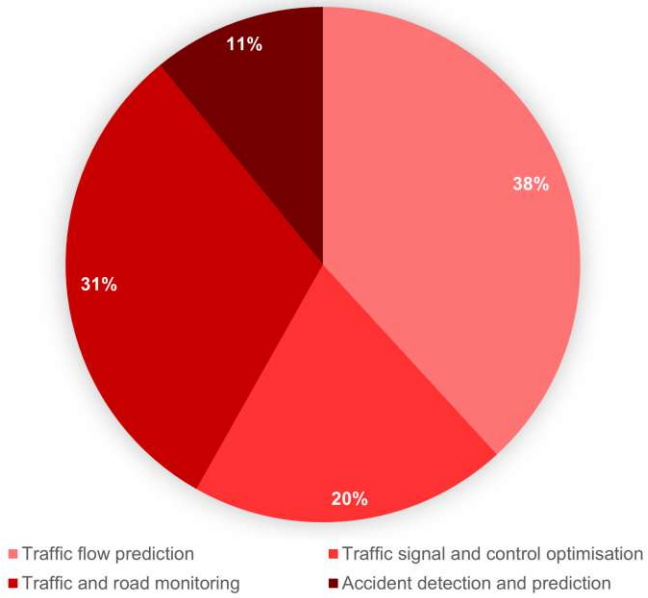


Figure 2.8: Application area distribution (Road traffic management)

collected from tweets. Using real-world data from Greater Manchester, the model was assessed and demonstrated higher accuracy than state-of-the-art models.

Predicting traffic congestion is an additional crucial aspect of optimising traffic flow and enhancing traffic management. The dynamic behaviour of vehicles inside the traffic network makes it challenging for Deep Learning algorithms to effectively anticipate traffic congestion. The authors Kothai et al. [KPD+21] present a hybrid model that efficiently predicts traffic congestion by combining the characteristics of CNN and boosted LSTM ensembles. CNN gathers features from traffic images, whereas LSTM trains and reinforces weak classifiers for congestion prediction. The suggested model is built using Tensorflow and evaluated in real-world traffic conditions. The experimental findings demonstrate a high degree of precision, recall, and accuracy, exceeding current Deep Learning models by 10% in terms of stability and performance.

The results of the presented papers show the impact of AI technologies on optimising traffic flow, reducing congestion and assisting road traffic management systems in making better decisions. As a result, these advancements led to increased safety for road users and a reduction in environmental harm.

Traffic signal and control optimisation

AI has reshaped the optimisation of traffic signals and controls by offering new approaches like a data-driven intelligent traffic management platform [NNB+19] and automatic prediction of traffic signal duration [KDR19]. Additionally, AI techniques allow the centralised coordination of autonomous cars at a junction without traffic signals [GRL+20]

and safe communication between cars with a trust management system, based on blockchain technology [ZLLH20]. Moreover, AI systems provide optimised route selection to EV charging stations with the goal of minimising overall journey time and charging costs [LWH+20] [QSWS19].

The global increase in traffic congestion requires innovative approaches to urban traffic management, including traffic signal control as a fundamental tool. Using a hierarchical multi-agent modelling framework, Jin and Ma [JM18] offer a decentralised traffic light control system. Agents represent each area of a traffic network model intersection. These intersection agents use reinforcement learning techniques to optimise timing decisions for turning actions and communicate with neighbouring agents. The traffic light control system additionally includes a phase composition procedure and turning movement priority settings. Simulations demonstrate that the suggested method enhances local and regional traffic performance.

A different approach is taken in a paper by Ashifuddin Mondal and Rehena [AMR19], where an IoT-based Intelligent Transportation System (ITS) can be used to successfully control traffic congestion. One essential aspect of intelligent traffic management is estimating and classifying the traffic congestion state of different road segments. This enables authorities to optimise traffic restrictions and passengers to pick the optimal route. Using ANN-based algorithms to analyse data acquired by in-road stationary sensors, this paper estimates and classifies the traffic congestion situation of distinct road segments inside the city of Kolkata. Based on the traffic congestion situation, ITS automatically updates traffic regulations, such as altering the waiting length at traffic lights and recommending other routes.

These articles make a valuable contribution to the enhancement of traffic management through the identification of congestion hotspots and the dynamic adjustment of traffic signal timing. In the end, this aids to optimise traffic flow and ensure a safe driving experience.

Traffic and road monitoring

AI algorithms are used in this application field for automatic extraction of roads from satellite images [BSIS18] and detecting road cracks [WFZL19], reducing the need for manual inspection. Nighttime vehicle detection is another area of interest, allowing monitoring at all times, even in poor lighting conditions [LHWL20]. UAVs equipped with AI technology can recognise traffic congestion [JLY+19] and 360-degree cameras are able to track vehicles and pedestrians [YBY+20]. Furthermore, AI can aid in vehicle classification and counting [CSK+19], licence plate recognition [PPR+20] as well as automatic detection of traffic signs to improve road traffic management. An important factor for safety reasons builds the traffic-related event monitoring based on social media data [AKM20] and the recognition of abnormal driving behaviour [JHL+20].

The use of AI-based image processing methodologies in the creation of traffic monitoring systems is increasingly gaining traction. Contributing to this subject, Tak et al. [TLSK21]

present a technique that uses cameras positioned at an intersection to gather traffic data and a deep-learning-based methodology for vehicle recognition and classification. This is followed by the estimation of lane-by-lane vehicle trajectories by matching observed vehicle locations to a high-definition map. The approach calculates the traffic volumes and queue lengths of each lane-by-lane travel direction based on the anticipated trajectories. The suggested technique was evaluated using thousands of samples. The findings indicate a vehicle detection rate of 99% with less than 20% inaccuracy in identifying vehicle types and calculating lane-by-lane travel volume.

An article by Grabowski and Czyżewski [GC20] focuses less on vehicles than directly on the condition of the road. It offers a novel approach to enhance road safety by integrating already installed cameras to identify slippery road conditions. Using CNN and transfer learning, the system can reliably recognise the surface characteristics of dry, wet, and snowy roadways by processing pictures captured by video cameras. Additionally, the system can identify slippery road conditions in low-light circumstances, making it an efficient alternative for improving road safety during inclement weather. By leveraging existing road measurement stations and roadside cameras that are publicly accessible, this method offers a cost-effective solution to cover a wide area without incurring excessive expenses.

The combination of AI technology, camera systems, and image processing in the presented papers significantly improves current traffic monitoring systems and road safety. This is achieved by providing real-time information about vehicles and road conditions.

Accident detection and prediction

AI systems play an essential role in enhancing accident detection and prediction, thereby mitigating the frequency of accidents and fostering transportation safety. Notable developments in this field are a segmentation strategy for road network-level accident probability [NB22] and a technique for identifying accident-prone highway zones [SGB+21], allowing for focused and rapid interventions. Additionally, AI can be used to predict the risk of road vehicle-train collisions by analysing factors such as weather, train speed, crossing type, number of lanes, and driver behaviour [SJA+20].

An accident detection system for vehicles that leverages both video and audio data from dashboard cameras to enhance performance is suggested by Choi et al. [CKKL21]. Unlike the majority of current vehicle collision detection systems, which depend on single-modal data, the proposed system employs a multimodal data-based ensemble Deep Learning model. The combination of both video and audio data permits many perspectives on the same source, which impacts the level of detection. The proposed system is evaluated by comparing it to single classifiers that use video or audio data alone and is further validated by YouTube clips of car accidents. The findings provide evidence that the proposed methodology outperforms individual classifiers to a significant degree. The authors suggested that the technology should be included in an emergency road call

service that automatically identifies traffic accidents and allows quick rescue following transmission to emergency recovery organisations.

The research conducted in this particular field has primarily concentrated on hourly accident predictions, which is insufficient for highly dynamic road networks with few accident data points. A paper by Zhou et al. [ZWX+20] presents a new framework called *RiskOracle* that enhances the minute-level granularity of predictions. The architecture consists of a differential time-varying Graph Neural Network to capture instantaneous changes in traffic status and dynamic inter-subregion correlations. In addition, the framework employs multi-task and region selection algorithms to find the urban subregions with the highest accident probability. Two real-world datasets from Suzhou Industrial Park and New York City are used to verify the usefulness and scalability of the proposed architecture.

The AI solutions discussed in the previous articles demonstrate the ability to enhance road safety by facilitating prompt responses from authorities in the event of accidents and providing drivers with instantaneous alerts regarding potential hazards.

2.3.6 Domain overview

Critical infrastructure is a fundamental component of modern society, covering essential domains such as the supply of water, electricity, gas, and heating, as well as the management of road traffic. The use of Artificial Intelligence has emerged as an effective tool for managing critical infrastructure. The benefits of AI systems are numerous and significant, encompassing improved quality, threat detection, efficiency optimisation, and decision-making support. The preceding sections outlined the concrete application areas of their respective domains and answered the research question: *For what purposes is Artificial Intelligence used in critical infrastructure (supply of water, electricity, heating, gas, and road traffic management)?* Figure 2.9 illustrates the proportion of relevant articles linked to each domain, providing a comprehensive perspective on the incorporation of AI across diverse industries.

The small number of relevant papers in the heating and gas supply sector is a noteworthy result, as it implies less development concerning AI-based solutions, at least under the terms of this SLR. In contrast, the majority of relevant articles have been found in the field of road traffic management. In total, the overview presented indicates that AI technologies have been incorporated and investigated across all domains, although with varying degrees of concentration. Disruptions and malfunctions within these sectors may start as minor issues but have further the potential to cause a serious and imminent threat to public safety. This is an important factor, as according to the AI Act, systems are considered high risk if they are used as safety components whose malfunction endangers the health and safety of individuals or property [Cou22]. For the mentioned reasons, the subsequent research focuses on the field of critical infrastructure and entails the development of a requirements catalogue for this area, encompassing the requirements

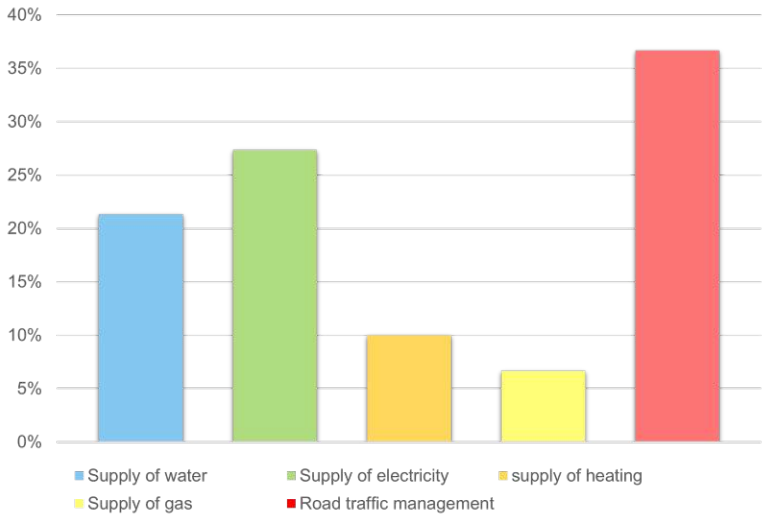


Figure 2.9: Distribution of relevant articles per domain

of the AI Act and other established norms, which often serve as a foundation for more comprehensive sector-specific regulations.

CHAPTER 3

Standards related to the AI Act

ISO norms refer to standards developed by the International Organization for Standardization (ISO). According to the Austrian standards website [Aus] ISO is an autonomous, non-governmental international institution that encompasses a total of 164 nations. This body engages in the formulation and publication of a diverse array of standards with the aim of providing a degree of uniformity and consistency across distinct sectors and industries. The ISO standards have garnered significant recognition and are implemented on a global scale, acting as benchmarks for critical parameters such as quality, safety, operational efficiency, and environmental responsibility. A principal goal of ISO standards is to foster the advancement of international trade, improve compatibility and augment the quality standard of both products and services. Guidelines and specifications are provided by these entities to both organisations and individuals with the aim of optimising processes, minimising errors, guaranteeing safety and enhancing customer satisfaction. Compliance with ISO standards offers an advantage over others by showcasing a dedication to optimal procedures and global quality benchmarks. ISO standards generally comprise documented instructions or requirements that specify optimal practices and criteria for attaining particular goals. ISO often collaborates with the International Electrotechnical Commission (IEC) [iec] and the Institute of Electrical and Electronics Engineers (IEEE) [iee] to jointly develop and publish standards that encompass a broader range of sectors and industries, ensuring compatibility between their respective standards.

3.1 Quality-, Information Security- and Risk-Management

According to a survey of 2019 [Hor], the ISO 9001 is by far the most widely certified standard. This standard [Int15a] provides a structured approach to implementing a quality management system within organisations that strive to maintain a high level of consistency in delivering products and services that align with customer expectations and adhere to regulatory requirements. The objective is to enhance customer satisfaction through

the efficient implementation of the system, which includes procedures for improvement and ensuring compliance with customer and regulatory specifications. This norm relates directly to the quality management obligations for providers of the AI Act.

ISO/IEC 27001 represents another internationally accredited standard for Information Security Management Systems (ISMS) [Int22b]. It offers comprehensive guidance to organisations aiding them in the development, application, preservation, and perpetual enhancement of an ISMS. Compliance to ISO/IEC 27001 signifies that a business has implemented a robust system for addressing risks associated with the security of their owned or managed data. Supporting this standard, ISO/IEC 27002 [Int22c] provides guidelines and best practices for implementing the controls specified in ISO/IEC 27001. These two norms serve as a complement to the areas of automatic event recording (logs), accuracy, robustness, monitoring, and incident reporting for the AI system.

The implementation of a risk management system for an AI system is one of the first requirements of the AI Act [Cou22]. In this context, the ISO 31000 standard is noteworthy. This internationally accepted guideline furnishes a schematic for risk management, facilitating organisations in realising their objectives, pinpointing potential hazards, and optimally allocating risk mitigation resources [Int22a]. Building on ISO 27001 while remaining in the context of risk management, ISO/IEC 27005 [Int22d] provides recommendations for the reduction of information security risks, including actions like information security risk assessment and treatment. Going further in the field of AI the ISO/IEC 23894 [Int23a] presents guidance for the development of AI systems, enabling them to effectively manage AI-specific risks and integrate risk management into AI-related activities.

3.2 Process capability according to ISO/IEC 33020:2019

Providers should specify to what degree procedures and measures are in place for a certain obligation or requirement. At this point, ISO/IEC 33020 [Int19c] plays a crucial role as it provides a defined measurement framework. It sets a structure to assess process capability by using process attributes, which represent measurable characteristics. The overall process capability level is determined based on the fulfilment of all the process attributes in the process profile. It is assessed on a six-point scale that ranges from *incomplete* to *innovating*. This scale reflects the increasing capability of the implemented process, progressing from not fulfilling the process purpose to consistently improving and adapting to organisational changes:

- Level 0: Incomplete process
  The implementation of the process does not fulfil its intended objectives, with limited or absent indications of accomplishing the process's purpose.
- Level 1: Performed process
  The implementation of the process successfully fulfils its intended objectives, as

evidenced by the attainment of the *process performance process attribute*, which measures the degree to which the process's purpose is accomplished.

- Level 2: Managed process
  In addition, the performed process is now executed in a controlled manner, including appropriate documented information. This achievement is demonstrated by the *performance management process attribute*, which measures the performance metrics including results, risks, responsibilities, and resources and the *documented information management process attribute*, which assesses the proper management of documented information during process execution.

- Level 3: Established process
  In addition, the managed process is now executed through a well-defined and continuously improved process. This achievement is demonstrated by the *process definition process attribute*, which evaluates the establishment and maintenance of a standardised process and the process deployment process attribute, which assesses the extent to which the standardised process is effectively implemented.

- Level 4: Predictable process
  In addition, the established process is now executed in a proactive manner. It involves the identification of quantitative management needs, collection and analysis of measurement data to identify the causes of variation and the implementation of corrective actions. This achievement is demonstrated by the two process attributes, namely the *quantitative analysis process attribute* and the *quantitative control process attribute*. The former evaluates the definition of information needs and identification of links between process elements as well as data collection, while the latter assesses the utilisation of objective data to effectively manage and control predictable process performance.

- Level 5: Innovating process
  In addition, the established process is now continuously enhanced to adapt to changes by employing approaches for process innovation. This achievement is demonstrated by the *process innovation process attribute*, which measures the identification and successful application of changes in process management based on innovative approaches. These might be drawn from internal resources or outside ideas, as long as they support the process innovation goals that have been established.

3.3 Other relevant standards

In this section, all other relevant standards that will be integrated into the requirements catalogue are listed with the corresponding requirement topics of the AI Act.

- The ISO 9241-220 [Int19b] outlines the processes and defines the desired results for the implementation of human-centred design within organisations, with the

objective of ensuring human-centred quality across the entire life cycle of interactive systems. The standard provides further guidance for the human oversight measures of the AI Act, which wants to ensure supervision of the AI system by natural persons.

- AI systems that use datasets for training or testing have to fulfil certain rules of the AI Act for their data management. ISO/IEC 25012 [Int08] acts as a useful tool for developing data quality measurements and data quality assessments in a variety of scenarios, including data generation, collection, integration, and improvement activities.

- The ISO/IEC/IEEE 29119-2 [Int21] describes test processes for governing, managing, and executing software testing that support the specific requirements of the AI Act, which concern the testing of the AI system in the field of risk management.

- Providers have to create instructions for use by their AI systems to ensure transparent instructions. The IEC/IEEE 82079-1 [Int19a] establishes comprehensive principles and specific criteria for the design and creation of user instructions for different kinds of products.

- Subsequent to this more detailed instruction for user information of software provides the ISO/IEC/IEEE 26514 [Int22e] includes the process of determining user information requirements, selecting appropriate presentation methods and delivering the information.

- The ISO/IEC 27032 [Int23b] offers guidance for enhancing cybersecurity, highlighting its distinct characteristics and interdependencies with other security domains, offering guidance to mitigate common cybersecurity threats and furnishing a collaborative framework to facilitate the resolution of cybersecurity issues. This norm is a valuable source for additional information on the cybersecurity requirements of the AI Act.

- Accuracy, robustness, and transparent design are important requirements of the proposed EU regulation. The ISO/IEC 25010 [Int11] defines a quality-in-use model evaluating the results of interaction for human-computer systems within certain use contexts, as well as a product quality model examining different properties including the ones mentioned.

- Also part of the AI Act's requirements is the generation and management of system events (logs). The fundamental concepts that form the basis for the generation, collection and management of records are provided by ISO 15489-1 [Int16].

- To comply with the AI Act, providers also need to create technical documentation. ISO/IEC/IEEE 15289 [Int19d] provides further guidance by outlining the content and purpose of various documentation throughout the software lifecycle.

- One of the obligations of AI system providers is to conduct a conformity assessment, supporting this process the ISO/IEC 33002 [Int15b] outlines a set of fundamental criteria for performing assessments to ensure objectivity, consistency, and repeatability of the evaluated procedures.

This chapter answers the research question: *Which standards are associated with the requirements of the AI Act?* Moreover, Figure 3.1 illustrates the correlations between the topics addressed in the AI Act and the previously stated standards. In addition to this, the EU declaration of conformity includes a link to the official template, while the CE marking of conformity references a specific article within an EU Regulation that outlines the fundamental principles of the CE marking.

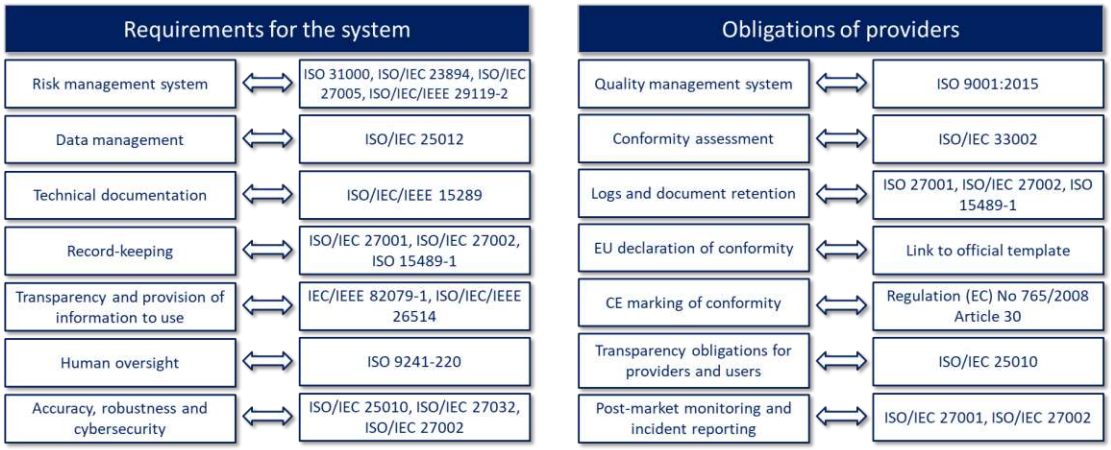


Figure 3.1: Related standards to the subjects of the AI Act

CHAPTER 4

Design and development of the requirements catalogue

4.1 Design process

As an initial step, it was necessary to determine which technology or tool would be used to realise the requirements catalogue. Multiple factors led to the selection of the Microsoft Excel application. First, it is a standard product widely used by businesses. As a result, a large number of employees are already familiar with Excel and have fundamental skills for its operation, which eliminates the need for additional training. Furthermore, Excel offers a high degree of flexibility and customisation, allowing it to be tailored to the needs and format of the anticipated catalogue. The capabilities of the tool enable the construction of charts and diagrams to visualise data and the sorting and filtering of data according to the user's preferences. Moreover, files provide the advantage of facile sharing and collaborative editing by numerous users. This is a major advantage since it is expected that several individuals or teams will collaborate on the same requirements catalogue or want to share it with other stakeholders for review. Finally, Excel is relatively inexpensive, especially compared to specialised management tools. This can make the requirements catalogue a more accessible option for smaller businesses.

There are different catalogues or checklists available online, each with its own structure. One example that particularly leverages multiple worksheets to organise and categorise essential data is the VDA ISA catalogue [VDA20]. This document consists of widely accepted specifications for the automotive sector corresponding to Information Security and serves as the fundamental framework for evaluations aimed at determining the level of Information Security. None of the included information or specifications in this catalogue will be used, it serves as a source of inspiration for the fundamental structure of the requirements catalogue of the AI Act.

The AI Act's requirements catalogue is structured into multiple worksheets that are categorised into information support, controls with corresponding fields for completion, and the presentation of results. Figure 4.1 illustrates the mentioned categories along with their corresponding worksheets, which will be explained in greater detail in the subsequent section.

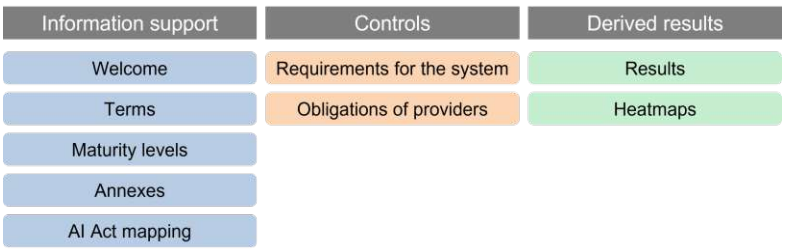


Figure 4.1: Worksheets Overview

4.2 Development process

4.2.1 Worksheet overview

Information support

All the blue-coloured worksheets in the requirements catalogue provide support to better understand the catalogue's content and to assist in the completion process. The first worksheet that users are presented with is the *Welcome* page. Here, each individual worksheet and its content are described in greater detail, as in the sections below.

The worksheet, titled *Terms*, plays a crucial role in clarifying the descriptions of requirements. It acknowledges that some terms may require further explanation and provides a comprehensive overview by incorporating relevant definitions directly from the AI Act. This worksheet ensures that all stakeholders have a common understanding of the terminology used throughout the compliance evaluation.

Next, the *Maturity levels* worksheet focuses on assessing the compliance of individual requirements. To facilitate this evaluation, each requirement is assigned a maturity level ranging from 0 to 5 by the users. These maturity levels correspond to the process capability levels of ISO/IEC 33020 [Int19c], which were also further outlined in Section 3.2. By referring to this worksheet, providers of high-risk AI systems can determine the extent to which a requirement is fulfilled. It is important to note that a score of 3 is considered sufficient to comply with the AI Act for both specific requirements and the overall compliance score.

The fourth worksheet, *Annexes*, addresses certain requirements and obligations that are associated with the content found in the AI Act's Annexes. By including them as a separate worksheet, users can conveniently access the relevant information without having

to navigate between multiple documents. This facilitates a seamless understanding of the AI Act's supplementary content.

Last, in terms of informative support, the *AI Act mapping* worksheet serves as a comprehensive reference point for aligning the individual requirements from the AI Act with their corresponding entries in the compliance evaluation catalogue. This worksheet lists the requirements in their original order and provides the new ID assigned to each requirement within the catalogue. By consulting this mapping, stakeholders can quickly determine which requirements of the catalogue correspond to those outlined in the AI Act. Furthermore, the comment column offers additional explanations as to why certain requirements were not included in the catalogue.

Controls

In this category, the orange-coloured worksheets *Requirements for the system* and *Obligations for providers* form the centrepiece of the requirements catalogue. As such, they are explained in greater detail at the level of individual columns. Although the two worksheets display disparate content, they share identical columns, which can be distinguished by the ones that provide context for the requirements and the ones that must be filled out by the users of the catalogue.

The informational columns within the two worksheets enable a comprehensive understanding of the AI Act's requirements. The *Chapter ID* column assigns a unique identification number to each requirement, enabling easy referencing. The *Control question* column formulates concise questions summarising each requirement, aiding in the evaluation process. Depending on the type of question, the maturity level can range from 0 to 5 or be selected between *no* and *yes*, which corresponds to maturity levels of 0 and 3. If a question is answered *yes*, it means something is in place or done, like signing a document, therefore, it is mapped to a 3 and not higher because it indicates sufficient compliance but does not indicate a predictable or innovative process. The *Requirement objective* column highlights the specific objectives of the requirements, providing a clear understanding of the intended outcomes. The *Requirement description* column offers detailed information in a very comprehensible manner. The *Possible proof for fulfilment* column suggests processes, systems, documents, or certificates that are, among other things, a possible ways to demonstrate compliance with the AI Act. This information is based on the column *Reference to other standards and regulations*, which contains the ISO norms and EU regulations relevant to the specific requirement, as outlined in the previous chapter. Lastly, the column *Reference to AI Act*, provides references to the original chapters as described in the worksheet *AI Act mapping*. To enhance comprehension, a specific example is provided:

- Chapter ID: 1.2
- Control question: To what extent are the procedures of the risk management system integrated into the system's life cycle?

- Requirement objective:
  - Identification and analysis of potential risks
  - Evaluation of additional risks based on post-market monitoring data
  - Adoption of suitable risk management measures
- Requirement description: The risk management system must be an ongoing process throughout the system's entire life cycle. This includes identifying and analysing potential risks to health, safety, and fundamental rights and further evaluating additional risks based on post-market monitoring data. Suitable risk management measures must then also be adopted to mitigate or eliminate the identified risks through system development, design or technical information provision.
- Possible proof of requirement:
  - ISO 31000 certification
  - A documented procedure for identifying, assessing, and addressing risks within the system's lifecycle
  - A catalogue of risk criteria including the likelihood and potential impact of a risk event
  - Documented measures for dealing with risks and their responsible parties
- Reference to AI Act: Article 9(2)
- Reference to other standards or regulations: ISO 31000, ISO/IEC 23894:2023, ISO/IEC 27001:2022

The columns in the Excel worksheets that mandate completion are crucial for users to record their progress towards compliance and are the basis for the illustration of the results. The *Maturity level* column holds significant importance, as providers must fill in the maturity level achieved for each requirement, contributing to the overall compliance assessment. The *Description of implementation* and *Date of assessment* columns enable providers to record the processes, systems, and measures they have implemented to meet each requirement, along with the date of assessing them. The *Responsible department* and *Contact* columns identify the specific department or individuals within the organisation responsible for implementing each requirement, ensuring accountability and effective communication. The *Future planned measures* and *Due Date* columns allow providers to outline upcoming measures intended to enhance compliance maturity levels, along with anticipated completion dates. Finally, the *Further information* column serves as an optional space for providers to record additional notes, remarks, or relevant information associated with each requirement, providing additional context and clarity during the compliance evaluation process.

Derived results

Based on the maturity levels provided in the two previous worksheets, the overall results are displayed in the *Results* worksheet. Each requirement's *Chapter ID*, *Control question*, *Maturity level*, and *capped Maturity level* are summarised in tables. The *Target maturity level* can be selected, indicating the desired maturity level to be attained by the user. A maturity level of 3 is recommended because, at this level, you can be confident that the measures are adequately documented and demonstrate compliance with the AI Act. Anything below this level is not indicative and anything above is just an additional improvement. The above-mentioned capped maturity level reduces all values that exceed the target maturity level. This means that if a target maturity level of 3, for instance, is selected and the maturity level of a requirement is rated at 5, the capped maturity level is 3. The purpose of this shortening is intended to ensure that in the calculation of the total average maturity level, very good results in one area don't compensate for poor results in another area because all requirements must be sufficiently fulfilled for full compliance. Moreover, the spider graphs shown in Figure 4.2 demonstrate the maturity level of the main chapters, with the target maturity level indicated by the green line. There are two variations of the spider graphs, based on the original maturity level and the capped Maturity level. As a final result, the overall maturity level is shown in the upper right corner next to the corresponding spider graph.

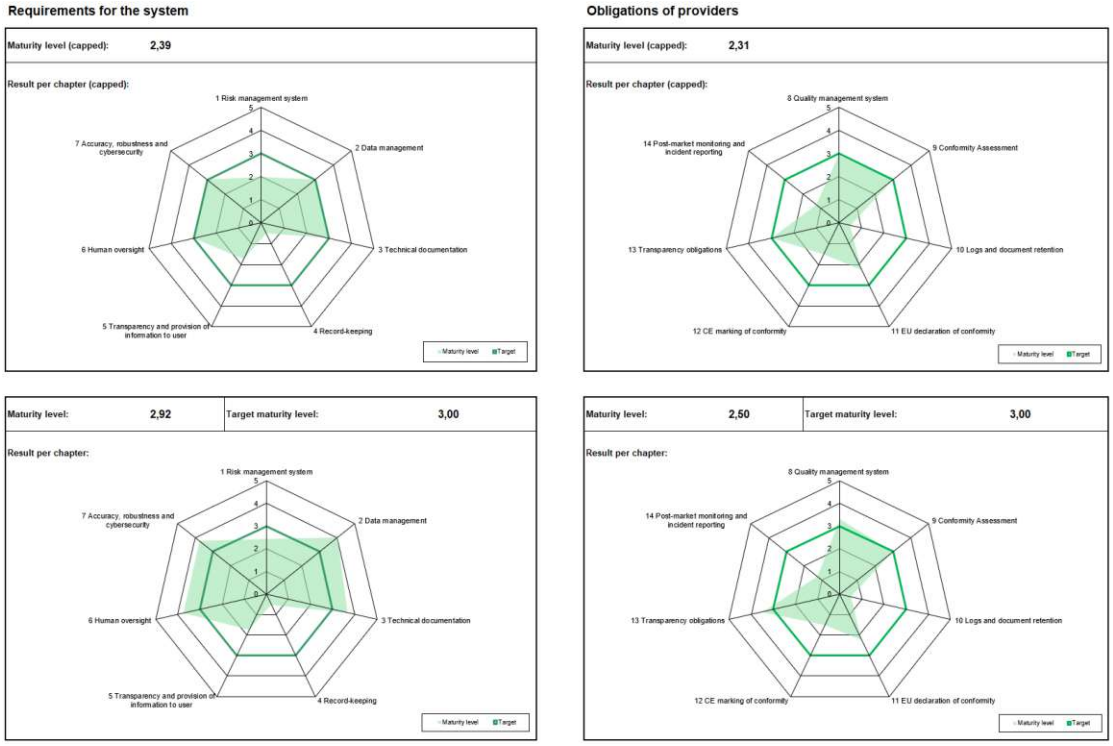


Figure 4.2: Spider graphs

In addition to the presentation of the results described previously, the worksheet *Heatmaps* displays the results in an alternative visual format. The heatmaps illustrate the dominant presence or distribution of maturity levels within a specific chapter through the intensity of colour. They provide an intuitive way of identifying areas with high and low values, allowing users to make informed decisions regarding further enhancements based on the visualised data. Looking at Figure 4.3, it can be seen, for instance, that very comprehensive measures have been implemented for the requirements of accuracy, robustness, and cybersecurity, but that there is still plenty of potential for improvement in the area of risk management.

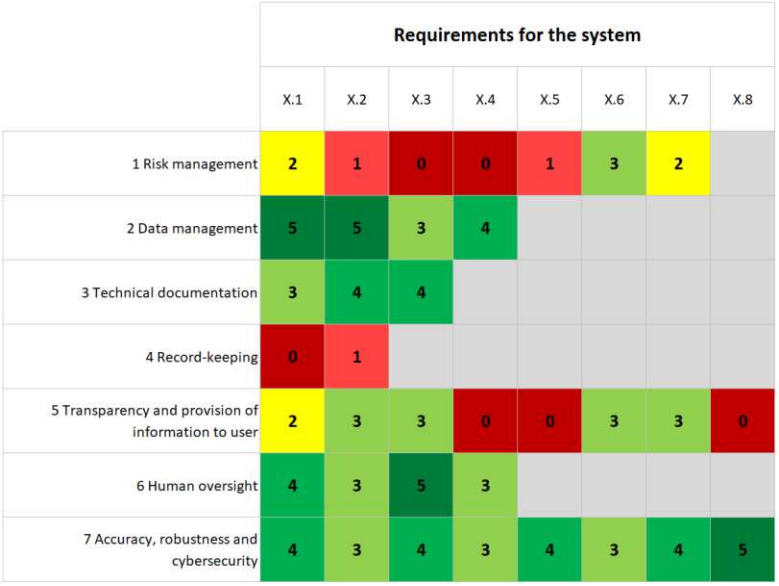


Figure 4.3: Heatmap of the requirements for the system

The described design and development of the requirements catalogue, in conjunction with the following two sections about the requirements for the system and obligations of providers, constitute the response to the research question: *What specific requirements do providers of AI systems in the field of critical infrastructure have to meet concerning the AI Act and what is an appropriate means to integrate these into a requirements catalogue?* More detailed views of the catalogue can be found in Appendix A (6.2).

4.2.2 Requirements for the system

The following subsections outline the summarised requirements that are related to the AI system itself and are relevant to the providers of AI systems in critical infrastructure. These requirements are formulated in a manner that enhances comprehensibility and are based on the AI Act [Cou22].

Risk management

1. Providers are required to establish, implement, document, and maintain a risk management system that is consistent with the intended purpose of the AI system. This means that the risks associated with the specific application of the AI system, including the potential impacts on individuals, society, and the environment must be considered.

2. The risk management system must be an ongoing process throughout the system's entire life cycle. This includes identifying and analysing potential risks to health, safety, and fundamental rights and further evaluating additional risks based on post-market monitoring data. Suitable risk management measures must then also be adopted to mitigate or eliminate the identified risks through system development, design, or technical information provision.

3. Appropriate risk management measures must be implemented in accordance with the following requirements. These measures must take into account the potential effects and interactions resulting from the combined application of all of the catalogue's requirements.

4. The risk management measures must ensure that any residual risks associated with each hazard, as well as the overall residual risk of the system, are considered acceptable. To identify the most appropriate risk management measures, the system must be adequately designed and developed to eliminate or reduce identified risks. Where risks cannot be eliminated, adequate mitigation and control measures must be implemented. Users must be provided with information regarding these measures and if necessary, appropriate training must be given. Technical knowledge, experience, education, and training expected by the user and the environment where the system will be used must be considered when eliminating or reducing risks related to the system's use.

5. AI systems must be tested to ensure that they operate as intended and meet the requirements outlined in this catalogue. This may include testing under real-life conditions.

6. AI systems must be tested at different steps of development and before they are used or put on the market. During testing, measures and probability thresholds that match the system's intended use must be used to make sure it operates as expected and meets the required standards.

7. The risk management system outlined in the points 1 to 6 must consider whether the AI system will be accessed or affect individuals under the age of 18.

Data management

When developing AI systems employing data-driven model training techniques, high-quality training, validation and testing datasets that meet the quality criteria enumerated

in points 1 to 4 are required. These requirements only apply to the testing datasets for the development AI systems that do not use techniques involving the training of models.

1. Appropriate data governance and management practices must be implemented to assure the quality of training, validation and testing datasets used in AI systems. These practices shall include a variety of aspects, such as design decisions, data collection processes, data preparation, formulation of assumptions, prior assessment of the availability and suitability of datasets, examination for potential biases, identification of data gaps and methods for addressing them.

2. The datasets used for training, validation and testing of AI systems must be appropriate, error-free and comprehensive, as well as possess the required statistical properties applicable to the intended consumers. Individual datasets or their combination may be used to accomplish the required statistical properties.

3. The data used in the training, validation and testing must consider the unique characteristics of the specific geographical, functional, or behavioural context in which the system is intended to be used. Therefore the data must be representative of the real-world context in which the system will be used and this context may differ based on the particular application or use case.

4. System providers may engage in the processing of specialised categories of personal data for the monitoring and correction of biases, provided that they maintain sufficient protections for essential rights and freedoms. The providers must use technical measures to safeguard the data, such as replacing identifying information with pseudonyms or encrypting the data to prevent unauthorised access.

Technical documentation

1. The technical documentation of an AI system must be prepared and kept up-to-date before the system is made available.

2. The technical documentation must be compliant with the requirements outlined in this catalogue and provide relevant authorities with comprehensive information to assess the system's compliance. The documentation must include the minimal elements listed in Annex IV or in the case of SMEs comparable documentation that serves the same purpose unless the relevant authority deems it unsuitable.

3. A single technical documentation containing all the required information, as specified in Annex IV and information of relevant legal acts listed in Annex II, Section A, must be created when an AI system associated with a product under those acts is placed on the market or put into service.

Record-keeping

1. AI systems must have the technical capability to automatically record events (logs) throughout the entire lifespan of the system. The logs must be recorded in a standardised format, with clear identification of the time, date, and participants involved in the activity or event.

2. These logs must be able to capture events that are relevant to identifying potential risks like threats to the health, safety, or to fundamental rights of persons, facilitating post-market monitoring and monitoring the system's operation.

Transparency and provision of information to user

1. AI systems must be developed so that their operation is adequately transparent to comply with the relevant obligations, as specified in Section 4.2.3. Users must be able to understand the AI operation of the system and based on that knowledge be able to use it.

2. The instructions for use for AI systems must be concise, complete, accurate and must be provided in a suitable digital or other format. These instructions must provide users with information that is relevant, accessible, simple to comprehend, and free from unnecessary technical language. Visual aids, diagrams, and practical examples may be used.

   The information specified by point 2. must include the following aspects:

3. The provider's identity and contact information and, if applicable, of the authorised representative must be included in the instructions for use.

4. Information regarding the characteristics, performance, and limitations of an AI system, including its intended purpose (including the particular geographical, functional, or behavioural setting), level of accuracy (including its metrics, robustness, and cybersecurity referred to in Section 4.2.2) and any known or foreseeable risks to health, safety, or fundamental rights (see point 2 of Section 4.2.2) must be included in the instructions for use. It also must include, when appropriate, the behaviour of the system towards specific persons or groups, specifications for input data, validation and testing datasets and the expected output of the system.

5. The changes to the AI system and predetermined performance during the initial conformity assessment must be communicated by the provider.

6. Measures to ensure human oversight of the AI system, as specified in Section 4.2.2 and details about the technical measures that are used to assist users in interpreting the outputs of the AI system must be included in the instructions for use.

7. Information regarding the required computational and hardware resources, the anticipated lifetime of the AI system and the frequency of any mandatory maintenance and care measures must be included in the instructions for use.

8. A description of the tool that enables users to appropriately collect, store, and interpret relevant logs within the AI system must be included in the instructions for use.

Human oversight

1. The design and development of AI systems must include suitable tools for human machine interaction to ensure that natural persons can effectively supervise AI systems during the operation.

2. Human oversight must be implemented to avoid or reduce risks to health, safety, or fundamental rights that may arise from using AI systems as intended or under reasonably predictable misuse.

3. Human oversight must be ensured through measures built into the AI system by the provider or measures identified by the provider and implemented by the user before the system is put into service.

4. Referring to the points 1 to 3, the AI system must be provided to the user in a way that enables them to understand its capabilities and limitations and monitor its operation. Users must be aware of the possibility of automation bias and be able to correctly interpret the system's output. Further, they must be able to override or invert the output, intervene in the system's operation or stop it using a *stop* mechanism or similar method.

Accuracy, robustness and cybersecurity

1. AI systems must be created to achieve a proper level of accuracy, robustness, and cybersecurity for their intended application. In addition, the AI system must maintain consistent performance in these areas throughout its lifetime, even if unexpected inputs, exceptions, errors, or variations in the operating environment occur.

2. Instructions for using the AI systems must include a declaration of their accuracy levels and relevant accuracy metrics.

3. AI systems must be designed to resist errors, faults, or contradictions that may occur within the system or its operating environment, especially as a result of their engagement with natural persons or systems.

4. To ensure the robustness of AI systems, technical redundancy solutions may be used, such as backup or fail-safe plans.

5. AI systems that proceed to learn after being put into service must be designed to minimise the risk of biased outputs impacting input for future operations (feedback loops). Appropriate mitigation measures must be taken to address this issue.

6. AI systems must be designed to resist unauthorised attempts by third parties to modify their performance or use by exploiting system vulnerabilities.

7. To ensure their cybersecurity, AI systems must have technical solutions that are suitable and proportional to the specific circumstances and risks involved.

8. The technical measures to address the specific vulnerabilities of AI systems must be included where relevant, such as measures to prevent and control attacks aimed at manipulating the training dataset (data poisoning), input designed to deceive the model (adversarial examples) or model flaws.

4.2.3 Obligations of Providers

The following subsections outline the summarised obligations that the providers of AI systems in critical infrastructure have to fulfil. These obligations are formulated in a manner that enhances comprehensibility and are based on the AI Act [Cou22].

Quality management

Providers of high-risk AI systems must have a quality management system that is implemented in a manner consistent with the provider's organisational size. At least the following elements must be included in the quality management system's structured documentation in the form of written policies, processes, and instructions:

1. a strategy for compliance that addresses conformity assessment procedures and how to manage changes made to the AI system

2. specific techniques and processes must be employed for its design, design control, and design verification, as well as its development, quality control and quality assurance

3. examination, testing, and validation processes to be performed before, during and after the development of a AI system, as well as the frequency with which they must be performed

4. technical specifications, including standards, to be used and when the appropriate uniform standards are not fully applied, the mechanisms to guarantee that the AI system conforms with requirements in Section 4.2.2

5. systems and procedures for data management, including data collection, analysis, labelling, storage, filtration, mining, aggregation, retention, and any other data operation performed before the deployment of AI systems

6. the risk management system referred to in Section 4.2.2

7. the post-market monitoring system referred to in Section 4.2.3

8. serious incident reporting processes referred to in Section 4.2.3

9. the management of communication with national competent authorities, including sectoral ones, providing data access, notified bodies, other operators, customers, or other interested parties

10. processes for keeping track of all necessary information and documentation

11. resource management, including measures related to ensuring the security of supply

12. a framework for accountability describing the duties of the management and other personnel with reference to each of the previous items in this section.

Conformity assessment

The internal control-based conformity assessment procedure includes the points 1 to 3.

1. The provider ensures that the quality management system established meets the standards set in Chapter 8.

2. The provider reviews the technical documentation to evaluate whether the AI system conforms with the necessary essential requirements specified in the requirements of the Section 4.2.2

3. The provider confirms that the AI system's post-market monitoring, design and development processes conform to the technical documentation.

Logs and document retention

1. Providers are required to maintain the logs generated by their AI systems that are under their control due to a contractual agreement with the user or by law. These logs must be kept for a minimum of six months unless specified differently by applicable Union or national laws, especially laws related to personal data protection.

2. The provider of an AI system is required to maintain the technical documentation, the EU declaration of conformity, documentation of the quality management, documentation of approved changes and other decisions of notified bodies for a period of 10 years after the system has been placed on the market or put into service. During this time, the provider must make this information available to national competent authorities upon request.

EU declaration of conformity

1. The provider is responsible for creating a signed EU declaration of conformity in written or electronic form for each AI system. It must be kept for ten years following the system's market introduction or start of service. The EU declaration of conformity must identify the specific AI system for which it is created and a copy of the declaration must be submitted to national competent authorities upon request.

2. The EU declaration of conformity must confirm that the AI system meets the necessary requirements mentioned in Section 4.2.2. The declaration must include the details as specified in Annex V and it must be translated into a language that is easily understandable by the national competent authorities in which the AI system is available.

3. In cases where systems are subject to multiple Union harmonisation legislations that demand an EU declaration of conformity, a single declaration must be created for all relevant Union regulations. The EU declaration of conformity must include all the information required to identify the Union harmonisation legislation to which it applies.

4. The provider is responsible for ensuring that the AI system complies with the requirements outlined in Section 4.2.2, by creating the EU declaration of conformity. The provider must keep the declaration updated whenever necessary.

CE marking of conformity

1. The CE marking that indicates the conformity with this regulation must comply with the fundamental principles stated in Article 30 of Regulation (EC) No 765/2008.

2. The CE marking must be attached clearly visible, permanent, and readable. If this is not practicable the marking can be displayed on the packaging or accompanying documents.

3. The CE marking of conformity must, if applicable, include the identification number of the notified body responsible for carrying out the conformity assessment procedures described in Article 43. This identification number must also be included in any documentation claiming that the AI system fulfils the CE marking requirements.

4. Providers of AI systems must provide their name, registered trade name, or trademark, as well as their contact information on the system. If direct labelling on the AI system is not practicable, this information must be included on its packaging or the accompanying documentation.

Transparency obligations

1. Providers are required to ensure that AI systems which are meant to interact with natural persons must be designed and developed in a manner that it is clear to them that they are interacting with an AI system. This requirement does not apply if it is already obvious to a reasonably well-informed person, considering the circumstances and context of use. This does not affect the other requirements in this catalogue and other transparency obligations that users of AI systems may have under existing Union or national laws.

2. Natural Persons must be informed no later than the moment of their first engagement or exposure with the system that they are interacting with AI technology, unless it is already obvious to a reasonably well-informed person, considering the circumstances and context of use.

Post-market monitoring and incident reporting

1. After the AI system is placed on the market, providers are required to develop a system for monitoring and keeping track of it. The scope of the post-market monitoring system must be appropriate to the level of risk posed by the AI system.

2. The post-market monitoring system must collect, document, and analyse data on the performance of the AI system over the course of its life cycle. The data may come from users or other sources. However, it must not cover sensitive operational data of users that are law enforcement authorities. The purpose is to evaluate compliance with the requirements set out in Section 4.2.2.

3. The post-market monitoring system must be based on a post-market monitoring plan, which has to be included in the technical documentation referred to in Annex IV.

4. The provider must report any serious incidents involving the AI system to the market surveillance authorities of the Member State in which the incident occurred. The provider is required to report the incident as soon as they confirm a connection between the AI system and the incident, or if there is a reasonable possibility of such a connection. This has to be done not later than 15 days after becoming aware of the incident.

CHAPTER 5

Evaluation

5.1 Choice of the evaluation methods

To conduct a meaningful evaluation within Hevner's Design Science Research Framework [HMPR04] of the requirements catalogue, a combination of qualitative and quantitative methods was chosen. This mixed-methods approach facilitates insight into a comprehensive understanding of the intention to use and usability of the artefact from diverse perspectives.

The first method of evaluation is the Technology Acceptance Model (TAM) of Fred Davis [DBW89]. This method is particularly useful for understanding the factors that influence the acceptance and intention to use new technologies. The evaluation focuses on two primary variables of TAM: *Perceived Usefulness* and *Perceived Ease of Use*. These two factors are elaborated on in qualitative interviews with Information Security experts from the Austrian EFS Consulting GmbH. They have extensive experience in Information Security assessments and are familiar with the aspects that are significant when it comes to a requirements catalogue. An adapted set of TAM questions to assess the two primary variables is created:

Perceived Usefulness:

- In what way do you think the requirements catalogue can support AI system providers in evaluating compliance with the requirements of the AI Act?
- What specific benefits do you expect from using the requirements catalogue in a compliance audit?
- How do you overall assess the usefulness of the requirements catalogue in the compliance audit of the AI Act?

Perceived Ease of Use:

- How do you evaluate the usability when filling out the requirements catalogue?
- How do you assess the clarity and ease of interpretation of the presentation of results?
- Do you see any challenges in using the requirements catalogue, if any?
- How do you overall assess the usability of the requirements catalogue for assessing compliance with the AI Act? What factors contribute to this?

The second method of evaluation is the System Usability Scale (SUS) of John Brooke [B+96]. This tool is renowned for its good adaptability and easy application, providing a numerical score indicative of a system's overall usability. The overall score is based on the answers of the Information Security experts and complements the qualitative insights derived from the TAM-based interviews. The original SUS questions (see Figure 1.4) are modified to obtain quantitative feedback on certain usability aspects of the requirements catalogue. The questions are answered by choosing one value on a Likert scale, which ranges from strongly disagree to strongly agree. The SUS score is calculated by summing the scores for all 10 items, where the odd items are positively formulated and subtracted by 1 and the even items are negatively formulated and subtracted from 5. Then the sum is multiplied by 2.5 to scale the score to a range of 0-100, with higher scores indicating higher perceived usability. The adopted survey consists of the following questions, with the tested usability aspects in brackets:

1. I thought all the relevant requirements of the AI Act were integrated completely. *(Completeness)*
2. I thought the requirements were enumerated several times (duplicates). *(Uniqueness)*
3. I thought the requirements were formulated in an understandable and practical way. *(Practicability)*
4. I thought that certain requirements were conflicting with each other. *(Consistency)*
5. I thought the requirements could be evaluated according to an understandable fulfilment degree. *(Verifiability)*
6. I thought the requirement catalogue was unstructured and unnecessary complex. *(Usability)*
7. I feel confident that the requirements catalogue will support the development of an AI system that is compliant with the AI Act. *(Effectiveness)*

8. I thought further instructions would be necessary to fully use the requirements catalogue. *(Accessibility)*

9. I thought the calculated results of the maturity level were presented in a clear manner. *(Usability)*

10. I thought in the requirements catalogue important aspects for a compliance assessment were missing. *(Completeness)*

After each survey has been filled out, the final scoring is calculated, which contributes to the ease of use and therefore also to the acceptance of the artefact. All the scores of the surveys are summed up and divided by the number of participants to get the total average SUS score. To interpret the results the Adjective Rating Scale of Bangor et al. [BKM09] is applied, which matches the SUS scores to the school grade system and the acceptable ranges, as shown in Figure 5.1.

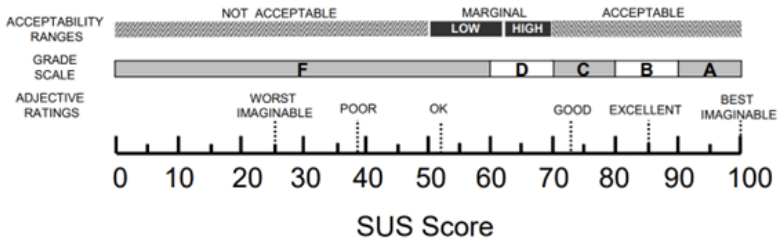


Figure 5.1: Adjective Rating Scale [BKM09]

According to this Adjective Rating Scale, a SUS score over 70 falls within the acceptable range. For this thesis, an average SUS score over 80 or respectively at least a grade B characterises one success criteria. To get a better understanding of the interview participants, Table 5.1 provides an overview that outlines their respective job positions at EFS, along with a brief description of their educational background and professional experience.

5.2 Evaluation process

The two chosen methods are integrated into the overall evaluation process, which is divided into three main stages:

Stage 1: Artefact demonstration and interviews
The first stage involves the presentation of the requirements catalogue to the participants. During a meeting with each individual, the worksheets and functions of the requirements catalogue are demonstrated. This is followed by an interview based on the TAM questions listed before. The interviews are conducted in a manner that encourages honest feedback and are recorded for subsequent analysis. The transcripts were translated into English by the author and can be found in Appendix B (6.2).

Interviewee	Job title	Background, Professional experience
Interviewee 1	Expert	Business Informatics, six years of experience in Information Security consulting, compliance reviews and risk assessments for the security release of applications, ISMS and certification support, management of supplier audits, ISO 27001 Lead Auditor
Interviewee 2	Senior Consultant	Business Engineering, three years of experience in Information Security consulting, experience in the use of requirements catalogues in the context of SUMS, readiness assessment CSMS and SUMS, IT Service Continuity Management, Shopfloor Cyber Security
Interviewee 3	Senior Consultant	Business Economics, three years of experience in Information Security consulting, audit and improvement of internal control and risk management systems, Information Security governance for ISMS and CSMS, preparation for IT security certifications, endpoint security strategy, requirements management of cyber security tools
Interviewee 4	Senior Consultant	Business Engineering, three years of experience in Information Security consulting, experience in the use of requirements catalogues in the context of CSMS, risk management, vulnerability management
Interviewee 5	Consultant	General and Digital Forensics, two years of experience in Information Security consulting, TISAX certification project, preparation for IT security certifications (e.g. ISO/IEC 27001, TISAX) including support of auditing processes
Interviewee 6	Consultant	Business law, one year of experience in Information Security consulting, risk management, design and implementation of holistic management systems, preparation of compliance documents (country-specific information)

Table 5.1: Overview of interview participants

Stage 2: SUS Questionnaire
After the interview, the participants are provided with the catalogue for a more detailed examination. They independently examine the catalogue in detail and complete the SUS questionnaire. The participant's responses to the questionnaire supplement the

qualitative data obtained from the interviews, providing a more rounded view of the catalogue's usability.

Stage 3: Analysis
The final stage of the evaluation process involves a thorough analysis of the data collected from the interviews and the SUS questionnaires. The interview transcriptions are analysed to identify main themes, commonalities, contradictions, and trends in the participants' perceptions of the catalogue's perceived usefulness and ease of use. The SUS scores are calculated to provide a quantitative measure of the catalogue's usability. Based on the analysis results, conclusions are drawn regarding the effectiveness of the catalogue in assisting AI system providers in evaluating compliance with the AI Act, identifying potential barriers to its use, and pinpointing factors that support its adoption.

5.3 Evaluation results

All interviewees concurred that the requirements catalogue is beneficial in assisting AI system providers to conform with the AI Act, but there were different factors that led to that conclusion. Interviewee 1 highlighted its ability to make the requirements more understandable and to break them down into manageable tasks by saying: *It's pointless to have requirements that are only understandable to specific departments like the legal or compliance departments or where only individuals have a mental model of what needs to be done. Companies need to break down the tasks based on their organisation and hierarchy and I think the catalogue will support that.* Interviewee 2 emphasised the catalogue's ability to *quickly get a full overview of the current state of compliance* and *identify areas that require immediate attention.* Interviewee 3 appreciated the catalogue's capability to contextualise the requirements by *not just listing the individual requirements, but rather creating connections with other norms or regulations and suggesting possible evidence.* Interviewees 4 and 5 noted the catalogue's ability to guide providers through the process of compliance, allowing them to work through the requirements step by step and continuously improve their product or system. Interviewee 6 also saw the catalogue as a valuable tool to make the requirements of the AI Act understandable to people who are not legally versed, similar to the answer of interviewee 1. The participants also viewed the use of the catalogue during compliance audits as beneficial. Interviewee 1 highlighted that *the pre-defined requirements catalogue saves me from having to read through all the requirements myself and filter out which ones apply to me as a provider or manufacturer and which ones are only relevant to other authorities.* This quick understanding of the regulation and filtering of the relevant requirements was also mentioned by interviewee 6. Interviewees 2 and 3 underlined that the requirements catalogue can serve as a shared tool for exchange between providers, auditors, and consultants. Interviewee 4 appreciated its ability to track progress and evaluate compliance with maturity levels. Interviewees 2 and 5 considered the catalogue especially useful in the preparation phase to proactively align with the legal framework, whereas interviewee 5 pointed out not to blindly rely on it.

The positive estimation of perceived ease of use can be attributed to the well-structured and user-friendly design. Interviewee 1 appreciated the use of familiar tools like Excel and the integration of a maturity model, which is often already used within a company. Interviewees 1 and 2 mentioned that the mapping of chapters between the AI Act and the requirements catalogue is crucial for them, as interviewee 1 stated: *I think it contributes to usability that the catalogue has a recognition value and that the law is included and cross-referenced to the chapters, which is an advantage I haven't seen in other requirements catalogues.* Interviewees 4 and 5 especially liked the use of colour within the catalogue as it supports usability by clustering the worksheets and indicating the degree of maturity levels. The welcome page, which introduces the worksheets of the requirements catalogue, increased the usability according to interviewees 3, 4, and 6. The clarity of the presentation of the results, particularly the use of heatmaps and spider graphs, was praised by all interviewees. Some interviewees suggested improvements to the terminology by exchanging certain words. Interviewees 2, 3, and 5 recommended highlighting the presence of the capped maturity level even more, as they consider it more relevant to the user. Further, interviewee 1 pointed out that *the spider graphs could have a more transparent area to better visualise the underlying lines of each maturity level.* These suggested changes were implemented according to the input of the participants. The identified challenges were primarily associated with maintaining the catalogue up-to-date with the rapid changes of the AI Act and that users should not rely solely on the accuracy of the catalogue, like interviewee 4 said: *So, despite using the catalogue, one should still maintain a necessary level of scepticism and not blindly rely on it.*

Strongly connected with the ease of use are the usability aspects that were additionally estimated with the SUS surveys. The answers of the interview participant were calculated and resulted in a value of 92.9%. This score is in the upper acceptable range and a grade *A* according to the Adjective Rating Scale of Bangor et al. [BKM09], as shown in Figure 5.2. This high score signifies that the interviewees perceived the requirements catalogue as highly usable. This aligns with the findings from the Technology Acceptance Model, which also indicated a high perceived ease of use of the catalogue. The convergence of these results from both the SUS and TAM suggests that the requirements catalogue is likely to be adopted by AI system providers for compliance with the AI Act, given its high perceived usefulness, perceived ease of use, and usability.

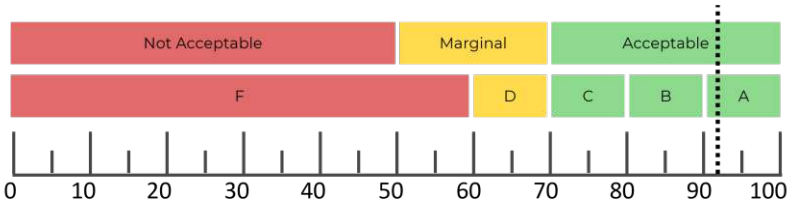


Figure 5.2: Total average SUS score on the Adjective Rating Scale [BKM09]

The results indicate that the requirements catalogue will be perceived as a valuable tool

for AI system providers to comply with the AI Act. In conclusion, the concrete benefits that can be derived from the evaluation are:

- Pre-filtered list of relevant requirements
- Summarised requirements in a clear and understandable manner for all employees
- Contextualisation of the requirements from the AI Act (related standards, possible proof, requirement objective)
- Assignment of tasks to individuals or departments
- Identification of areas that require immediate attention
- Exchange tool between providers of AI systems, consultants, and auditors
- Transparency of the current maturity level and the improvement potentials
- Tracking of the progress, responsibilities, and planned measures
- Supporting decision-making of the management with the various presentations of results
- Clear graphical presentation of results
- Enhanced structure and clustering through a colour scheme
- Enabling a continuous improvement process by outlining future planned measures
- Storing and retention of the sensitive data in the Excel file, not on a server or database

But of course, some challenges and concerns were also identified regarding the requirements catalogue for the AI Act. One challenge relates to potential incompleteness, where there is apprehension about whether all requirements from the AI Act are adequately covered in the catalogue. This raises the possibility of some requirements being missed or not properly addressed. Even if there is a mapping to the original requirements and comments that state the reason why a certain requirement is not covered by the catalogue. Another challenge highlighted is the need for regular updates to the catalogue to align with the evolving AI Act. Given the fast-paced nature of legal changes, failure to reflect the latest updates in the catalogue could result in compliance issues. Additionally, the extensive information presented in the catalogue may be seen as burdensome by some individuals who prefer a more streamlined approach. Lastly, the challenge of responsibility mapping arises, particularly in companies with diverse structures, as the catalogue may not provide explicit clarification on the individuals responsible for each requirement. Although not directly covered, this mapping challenge can still pose difficulties during the implementation process. These challenges emphasise the importance of continuous review and improvement of the requirements catalogue to ensure its effectiveness in supporting compliance with the AI Act in the future.

CHAPTER 6

Conclusion

6.1 Summary and Results

This thesis was centred on the development of a requirements catalogue for the Artificial Intelligence Act, particularly for providers of AI systems in critical infrastructure. The aim was to address the challenge faced by companies integrating AI systems into essential infrastructure, as they lack a tool or procedure for determining compliance with the new AI technology requirements, which poses legal and reputational risks. Moreover, the application areas of AI in critical infrastructure domains such as water supply, electricity, gas, heating, and road traffic management were explored to comprehend the purpose and level of integration of AI technology. To support the Rigor Cycle, the Design Science Framework of Hevner was utilised, along with conventional and systematic literature research. In addition, the evaluation involved the use of the Technology Acceptance Model to conduct interviews with Information Security experts from EFS Consulting GmbH to assess perceived usefulness and ease of use. This was supplemented by a SUS survey to estimate usability factors.

The results of the SLR revealed, that AI technologies have been incorporated and explored in varying degrees across all critical infrastructure domains. While the heating and gas supply sector had a very limited number of relevant papers, the field of road traffic management showed the highest concentration. Disruptions and malfunctions within these sectors have the potential to pose serious threats to public safety, making them high-risk areas according to the AI Act. Next, the related standards were investigated to establish links to additional information and provide possible evidence for complying with the requirements of the AI Act.

Eventually, a requirements catalogue was designed and developed, which includes multiple coloured worksheets. Beginning with blue-coloured worksheets that serve as informative support. They outline important terms, the maturity levels, the annexes of the AI

Act and establish a mapping to the chapters of the original regulation. The orange-coloured worksheets built the centrepiece, as they provide comprehensive information about the summarised requirements. They include identification numbers, control questions, requirement objectives, requirement descriptions, possible proof for fulfilment and references to other standards. The users are also able to track their progress using completion columns in the worksheets, where the estimation of the maturity level is the most important aspect. Further, the description of implementation, date of assessment, responsible department and contact columns capture implemented processes and responsible parties. The future planned measures and corresponding due date columns capture the upcoming actions to improve the maturity level. The green-coloured worksheets summarise the overall results and illustrate them with spider graphs and heatmaps.

The interviews showed high perceived usefulness and ease of use, supplemented by the total average SUS score of 92.9%. Utilising the requirements catalogue has thus numerous benefits for organisations seeking compliance. Firstly, it provides a pre-filtered list of relevant requirements, which saves time and resources. Further, these requirements are presented in a clear and understandable manner for all employees. By presenting the requirements in a digestible format, organisations foster a shared understanding among team members, encouraging even less experienced employees to participate in the compliance process. Contextualising the requirements within the catalogue is another benefit. This involves considering related standards, potential proofs, and the specific objectives of each requirement. As a result, the ability to interpret and implement the requirements effectively is improved. Assigning tasks to individuals or departments based on the requirements catalogue ensures clear accountability and proper execution. This not only facilitates efficient workflow management but also fosters a sense of ownership and responsibility within the company. Furthermore, the catalogue helps to identify areas that require immediate attention. This enables proactive measures to be taken promptly, addressing compliance gaps, or risks before they escalate. The catalogue also serves as a valuable tool during audits and as an exchange platform between AI system providers, consultants, and auditors. It supports transparency by providing a structured basis for discussions and assessments, enhancing the audit process. The clear graphical presentation of results derived from the catalogue improves their accessibility and comprehension. This visual representation facilitates the communication of compliance status and progress to stakeholders at various levels within the company. A very important benefit is the insight into the organisation's current maturity level and improvement potential regarding AI compliance. This transparency allows for informed decision-making and targeted measures to enhance compliance and ensures a safe market entry for products.

Summarising, the thesis has provided valuable insights regarding the use of AI in critical infrastructure and the development of a requirements catalogue for the AI Act. The catalogue has the potential to serve as a crucial tool for AI system providers preparing companies for the AI regulation.

6.2 Outlook

This thesis has explored the development of a requirements catalogue that is able to substantially streamline the compliance process, reducing the burden on system providers and ensuring a higher level of adherence to the AI Act. Looking forward, there are a number of promising opportunities for future research. As the AI Act continues to evolve, research should concentrate on how the requirements catalogue can be adapted to accommodate these changes. This will ensure that the tool remains pertinent and effective despite legislative changes. In addition, the process of conducting a Systematic Literature Review may be expanded to include a team-based analysis of all papers resulting from the search queries. This may result in a more robust understanding of the field and an in-depth analysis of AI applications. Future research could also consider the integration of a preliminary questionnaire into the requirements catalogue. A feature like this would aid in determining whether a system provider is subject to the requirements of the AI Act and how the system is classified according to the risk-based approach. Consequently, only applicable requirements would be displayed, further streamlining the compliance process. Assigning particular roles for the requirements could be an additional aspect of future enhancements. Despite the fact that this is a difficult task due to the varying sizes, structures, and responsibilities across companies, such a recommendation for responsibility has the potential to improve the usability of the requirements catalogue.

In conclusion, this research represents a significant step forward in the development of tools to assist AI system providers in critical infrastructure to assess compliance with the AI Act. It is anticipated that this work will serve as a foundation for future research in this area, thereby contributing to the ongoing development and refinement of AI compliance tools. However, ongoing improvement and updating of the catalogue are necessary to address the challenges identified. The rapid development of Artificial Intelligence and its impending regulation under the AI Act highlight the significance of this work. As we move forward, it is evident that the role of AI in our society will continue to evolve and it is imperative that we continue to develop comprehensive tools and frameworks to manage this evolution effectively.

List of Figures

List of Tables

Acronyms

AI Artificial Intelligence. ix, 1–7, 9–20, 22–38, 40–43, 45–60, 63–65, 67–69, 71

AIDWRP Adaptive Intelligent Dynamic Water Resource Planning. 22, 23

ANFIS Adaptive Neuro-Fuzzy Inference System. 22

ANN Artificial Neural Network. 22, 24, 25, 27, 29, 35

CNN Convolutional Neural Network. 31, 34, 36

DBN Deep Belief Networks. 31

DL Deep Learning. 9–12, 20, 23, 27, 28, 33, 34, 36

DNN Deep Neural Network. 23, 28

DTGN Graph Neural Network. 37

ELM Extreme Learning Machine. 23

GLM Generalised Linear Model. 22

GPR Gaussian Process Regression. 23, 32

GPUs Graphics Processing Units. 10

IEC International Electrotechnical Commission. xii, 39, 40, 42, 43, 46, 48, 62

IEEE Institute of Electrical and Electronics Engineers. 19, 39, 42

IS Information Systems. 3, 4

ISA Information Security Assessment. 45

ISLMD Improved Spline-local Mean Decomposition. 23

ISMS Information Security Management Systems. 40

ISO International Organization for Standardization. xii, 39–43, 46–48, 62

ITS Intelligent Transportation System. 35

LSTM Long Short-Term Memory. 33, 34

MDP Markov Decision Process. 26, 29

ML Machine Learning. 9–12, 20, 22, 24, 27, 28, 31–33

MLSTM Multidirectional Long Short-Term Memory. 28

MPMR Minimax Probability Machine Regression. 23

NLF New Legislative Framework. 16, 18

NTL Non-Technical Loss. 28, 29

NTU Nephelometric Turbidity Unit. 22

P2P Peer-to-Peer. 29

RF Random Forest. 12, 22

RVM Relevance Vector Machine. 23

SDHS Smart District Heating Systems. 31

SLR Systematic Literature Review. ix, 3, 5–7, 18–21, 25, 30, 33, 37, 67, 69

SUS System Usability Scale. ix, 7, 60–64, 67, 68

SVG Singular Value Decomposition. 10

SVM Support Vector Machines. 12

SVM Support Vector Machine. 27, 31, 32

SVR Support Vector Regression. 22

TAM Technology Acceptance Model. ix, 6, 59–61, 64, 67

TISAX Trusted Information Security Assessment Exchange. 62

TP Transformational Participation. 27

TRA Theory of Reasoned Action. 6

UAV Unmanned Aerial Vehicle. 31, 35

VDA Verband der Automobilindustrie. 45

WHO World Health Organisation. 22

Bibliography

[Abb21] Hussein Abbass. What is artificial intelligence? *IEEE Transactions on Artificial Intelligence*, 2(2):94–95, 2021.

[AFPAS+20] A'kif Al-Fugara, Hamid Reza Pourghasemi, Abdel Rahman Al-Shabeeb, Maan Habib, Rida Al-Adamat, Hani Al-Amoush, and Adrian L Collins. A comparison of machine learning models for the mapping of groundwater spring potential. *Environmental Earth Sciences*, 79:1–19, 2020.

[AGB21] Ümit Ağbulut, Ali Etem Gürel, and Yunus Biçen. Prediction of daily global solar radiation using different machine learning algorithms: Evaluation and comparison. *Renewable and Sustainable Energy Reviews*, 135:110114, 2021.

[AK20] Mostafa Askari and Farshid Keynia. Mid-term electricity load forecasting by a new composite method based on optimal learning mlp algorithm. *IET Generation, Transmission & Distribution*, 14(5):845–852, 2020.

[AKK+20] Mamoun Alazab, Suleman Khan, Somayaji Siva Rama Krishnan, Quoc-Viet Pham, M Praveen Kumar Reddy, and Thippa Reddy Gadekallu. A multidirectional lstm model for predicting the stability of a smart grid. *IEEE Access*, 8:85454–85463, 2020.

[AKM20] Ebtesam Alomari, Iyad Katib, and Rashid Mehmood. Iktishaf: A big data road-traffic event detection tool using twitter and spark machine learning. *Mobile Networks and Applications*, pages 1–16, 2020.

[ALA21] Moath Alrifaey, Wei Hong Lim, and Chun Kit Ang. A novel deep learning framework based rnn-sae for fault detection of electrical gas generator. *IEEE Access*, 9:21433–21442, 2021.

[AMR19] Md Ashifuddin Mondal and Zeenat Rehena. Intelligent traffic congestion classification system using artificial neural network. In *Companion Proceedings of The 2019 World Wide Web Conference*, pages 110–116, 2019.

[ARJZS21] Mohd Shahrizan Abd Rahman, Nor Azliana Akmal Jamaludin, Zuraini Zainol, and Tengku Mohd Tengku Sembok. Machine learning algorithm model for improving business decisions making in upstream oil & gas. In *2021 International Congress of Advanced Technology and Engineering (ICOTEN)*, pages 1–5. IEEE, 2021.

[Aus] Austrian Standards. ISO. `https://www.austrian-standards.at/en/standardization/why-standards/basic-terms/iso-standards`. [Accessed: May 26, 2023].

[AZH+22] Chao Ai, Lu Zhao, Mengyao Han, Siyuan Liu, and Zhongyang Wang. Mitigating water imbalance between coastal and inland areas through seawater desalination within china. *Journal of Cleaner Production*, 371:133418, 2022.

[B+96] John Brooke et al. Sus-a quick and dirty usability scale. *Usability evaluation in industry*, 189(194):4–7, 1996.

[BESG19] Hossein Bonakdari, Isa Ebtehaj, Pijush Samui, and Bahram Gharabaghi. Lake water-level fluctuations forecasting using minimax probability machine regression, relevance vector machine, gaussian process regression, and extreme learning machine. *Water Resources Management*, 33:3965–3984, 2019.

[BKM09] Aaron Bangor, Philip Kortum, and James Miller. Determining what individual sus scores mean: Adding an adjective rating scale. *Journal of usability studies*, 4(3):114–123, 2009.

[BNP20] Christian Behm, Lars Nolting, and Aaron Praktiknjo. How to model european electricity load profiles using artificial neural networks. *Applied Energy*, 277:115564, 2020.

[BSIS18] Alexander Buslaev, Selim Seferbekov, Vladimir Iglovikov, and Alexey Shvets. Fully convolutional network for automatic road extraction from satellite imagery. In *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, pages 207–210, 2018.

[BTACRGE19] Madalina-Mihaela Buzau, Javier Tejedor-Aguilera, Pedro Cruz-Romero, and Antonio Gomez-Exposito. Hybrid deep neural networks for detection of non-technical losses in electricity smart meters. *IEEE Transactions on Power Systems*, 35(2):1254–1263, 2019.

[cat23] CatalogPlus TU Wien. `https://catalogplus.tuwien.at/primo-explore/search?vid=UTW`, 2023. Accessed: January 21, 2023.

[CBGM21] Debaditya Chakraborty, Hakan Başağaoğlu, Lilianna Gutierrez, and Ali Mirchi. Explainable ai reveals new hydroclimatic insights for ecosystem-centric groundwater management. *Environmental Research Letters*, 16(11):114024, 2021.

[CCDO20] Roberta Calegari, Giovanni Ciatto, Enrico Denti, and Andrea Omicini. Logic-based technologies for intelligent systems: State of the art and perspectives. *Information*, 11(3):167, 2020.

[CDL+20] Bin Cao, Weinan Dong, Zhihan Lv, Yu Gu, Surjit Singh, and Pawan Kumar. Hybrid microgrid many-objective sizing optimization with fuzzy decision. *IEEE Transactions on Fuzzy Systems*, 28(11):2702–2710, 2020.

[CDM20] Andrew Cropper, Sebastijan Dumančić, and Stephen H Muggleton. Turning 30: New ideas in inductive logic programming. *arXiv preprint arXiv:2002.11002*, 2020.

[CDN88] John P Chin, Virginia A Diehl, and Kent L Norman. Development of an instrument measuring user satisfaction of the human-computer interface. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 213–218, 1988.

[CKKL21] Jae Gyeong Choi, Chan Woo Kong, Gyeongho Kim, and Sunghoon Lim. Car crash detection using ensemble deep learning and multimodal data from dashboard cameras. *Expert Systems with Applications*, 183:115400, 2021.

[Cla23] Clarivate. `https://www.webofscience.com/wos/woscc/basic-search`, 2023. Accessed: January 21, 2023.

[CLV18] Kwok Tai Chui, Miltiadis D Lytras, and Anna Visvizi. Energy sustainability in smart cities: Artificial intelligence, smart monitoring, and optimization of energy consumption. *Energies*, 11(11):2869, 2018.

[Com20] European Commission. Communication: Shaping europe's digital future, February 2020.

[Com21] European Commission. A europe fit for the digital age, March 2021.

[Com22a] European Commission. Europe's digital decade: Digital targets for 2030, July 2022.

[Com22b] European Commission. Regulatory framework proposal on artificial intelligence, September 2022.

[Cou19] Council of the European Union. Artificial intelligence - Conclusions on the coordinated plan on artificial intelligence - Adoption, 2019.

[Cou22] Council of the European Union. 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 - General approach, 2022.

[CS18] Tao Chen and Wencong Su. Indirect customer-to-customer energy trading with reinforcement learning. *IEEE Transactions on Smart Grid*, 10(4):4338–4348, 2018.

[CSK+19] Mayank Singh Chauhan, Arshdeep Singh, Mansi Khemka, Arneish Prateek, and Rijurekha Sen. Embedded cnn based vehicle classification and counting in non-laned road traffic. In *Proceedings of the tenth international conference on information and communication technologies and development*, pages 1–11, 2019.

[CYY+20] Weihong Cai, Junjie Yang, Yidan Yu, Youyi Song, Teng Zhou, and Jing Qin. Pso-elm: A hybrid learning model for short-term traffic flow forecasting. *IEEE access*, 8:6505–6514, 2020.

[DBW89] Fred D Davis, Richard P Bagozzi, and Paul R Warshaw. User acceptance of computer technology: A comparison of two theoretical models. *Management science*, 35(8):982–1003, 1989.

[DPW+19] Bowen Du, Hao Peng, Senzhang Wang, Md Zakirul Alam Bhuiyan, Lihong Wang, Qiran Gong, Lin Liu, and Jing Li. Deep irregular convolutional residual lstm for urban traffic passenger flows prediction. *IEEE Transactions on Intelligent Transportation Systems*, 21(3):972–985, 2019.

[DWY+21] Jian Dong, Haixin Wang, Junyou Yang, Xinyi Lu, Liu Gao, and Xiran Zhou. Optimal scheduling framework of electricity-gas-heat integrated energy system based on asynchronous advantage actor-critic algorithm. *IEEE Access*, 9:139685–139696, 2021.

[EL20] Abinet Tesfaye Eseye and Matti Lehtonen. Short-term forecasting of heat demand of buildings for efficient and optimal energy management based on integrated machine learning models. *IEEE Transactions on Industrial Informatics*, 16(12):7743–7755, 2020.

[EPSS21] Aniekan Essien, Ilias Petrounias, Pedro Sampaio, and Sandra Sampaio. A deep-learning model for urban traffic flow prediction with traffic events mined from twitter. *World Wide Web*, 24(4):1345–1368, 2021.

[Eur16a] European Parliament, Council of the European Union. Directive (EU) 2016/680 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data by competent authorities for the purposes of the prevention, investigation, detection or prosecution of criminal offences or

the execution of criminal penalties, and on the free movement of such data, and repealing Council Framework Decision 2008/977/JHA, 2016.

[Eur16b] European Parliament, Council of the European Union. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), 2016.

[Eur17] European Council. European Council meeting (19 October 2017) – Conclusions, 2017.

[Eur18] European Commission. communication from the commission to the european parliament, the european council, the council, the european economic and social committee and the committee of the regions coordinated plan on artificial intelligence, 2018.

[Eur20a] European Commission. On Artificial Intelligence - A European approach to excellence and trust, 2020.

[Eur20b] European Commission. Proposal for a Regulation of the European Parliament and of the Council on European data governance (Data Governance Act), 2020.

[Eur22a] European Commission. A European Strategy for data, 2022.

[Eur22b] European Commission, Directorate-General for Communications Networks, Content and Technology. proposal for a regulation of the european parliament and of the council laying down harmonised rules on artificial intelligence (artificial intelligence act), 2022.

[FAB+21] Sarah Friedrich, Gerd Antes, Sigrid Behr, Harald Binder, Werner Brannath, Florian Dumpert, Katja Ickstadt, Hans A Kestler, Johannes Lederer, Heinz Leitgöb, et al. Is there a role for statistics in artificial intelligence? *Advances in Data Analysis and Classification*, pages 1–24, 2021.

[FABE21] Behnam Farsi, Manar Amayri, Nizar Bouguila, and Ursula Eicker. On short-term load forecasting using machine learning techniques and a novel parallel deep lstm-cnn approach. *IEEE Access*, 9:31191–31212, 2021.

[FG19] Sean W Fleming and Angus G Goodbody. A machine learning metasystem for robust probabilistic nonlinear regression-based forecasting of seasonal water availability in the us west. *IEEE Access*, 7:119943–119964, 2019.

[FGWL19] Qingwu Fan, Yiliang Guo, Shaoen Wu, and Xudong Liu. Two-level diagnosis of heating pipe network leakage based on deep belief network. *IEEE Access*, 7:182983–182992, 2019.

[GAA+19] Khawaja Moyeezullah Ghori, Rabeeh Ayaz Abbasi, Muhammad Awais, Muhammad Imran, Ata Ullah, and Laszlo Szathmary. Performance analysis of different types of machine learning classifiers for non-technical loss detection. *IEEE Access*, 8:16033–16048, 2019.

[GAMQ22] Ahmed M Ghaithan, Ibrahim Alarfaj, Awsan Mohammed, and Osaid Qasim. A neural network-based model for estimating the delivery time of oxygen gas cylinders during covid-19 pandemic. *Neural Computing and Applications*, 34(13):11213–11231, 2022.

[GC20] Dariusz Grabowski and Andrzej Czyżewski. System for monitoring road slippery based on cctv cameras and convolutional neural networks. *Journal of Intelligent Information Systems*, 55(3):521–534, 2020.

[GES+21] Ahmed R Ginidi, Abdallah M Elsayed, Abdullah M Shaheen, Ehab E Elattar, and Ragab A El-Sehiemy. A novel heap-based optimizer for scheduling of large-scale combined heat and power economic dispatch. *IEEE Access*, 9:83695–83708, 2021.

[GGK+21] Nasim Ghadami, Mohammad Gheibi, Zahra Kian, Mahdieh G Faramarz, Reza Naghedi, Mohammad Eftekhari, Amir M Fathollahi-Fard, Maxim A Dulebenets, and Guangdong Tian. Implementation of solar energy in smart cities using an integration of artificial neural network, photovoltaic system and classical delphi methods. *Sustainable Cities and Society*, 74:103149, 2021.

[GLX+19] Yuanli Gu, Wenqi Lu, Xinyue Xu, Lingqiao Qin, Zhuangzhuang Shao, and Hanyu Zhang. An improved bayesian combination model for short-term traffic prediction with deep learning. *IEEE Transactions on Intelligent Transportation Systems*, 21(3):1332–1342, 2019.

[GRL+20] Yang Guan, Yangang Ren, Shengbo Eben Li, Qi Sun, Laiquan Luo, and Keqiang Li. Centralized cooperation for connected and automated vehicles at intersections by proximal policy optimization. *IEEE Transactions on Vehicular Technology*, 69(11):12597–12608, 2020.

[Hev07] Alan R Hevner. A three cycle view of design science research. *Scandinavian journal of information systems*, 19(2):4, 2007.

[HKA+23] Marcus Harris, Elizabeth Kirby, Ameeta Agrawal, Rhitabrat Pokharel, Francis Puyleart, and Martin Zwick. Machine learning predictions of electricity capacity. *Energies*, 16(1):187, 2023.

[HMPR04] Alan R Hevner, Salvatore T March, Jinsoo Park, and Sudha Ram. Design science in information systems research. *MIS quarterly*, pages 75–105, 2004.

[Hor] Liza Horielikova. Which iso standards are the most popular? analysis of iso 2019 survey. https://advisera.com/articles/which-iso-standards-are-the-most-popular-analysis-/of-iso-2019-survey/. [Accessed: May 26, 2023].

[HSK19] Salim Heddam, Hadi Sanikhani, and Ozgur Kisi. Application of artificial intelligence to estimate phycocyanin pigment concentration using water quality data: a comparative study. *Applied Water Science*, 9:1–16, 2019.

[HVF20] Kabir Hossain, Frederik Villebro, and Søren Forchhammer. Uav image analysis for leakage detection in district heating systems using machine learning. *Pattern Recognition Letters*, 140:158–164, 2020.

[iec] International Electrotechnical Commission (IEC). https://iec.ch/who-we-are. [Online; accessed 26-May-2023].

[iee] IEEE: At a Glance. https://www.ieee.org/about/at-a-glance.html. [Online; accessed 26-May-2023].

[IEE23] IEEE Xplore. https://ieeexplore.ieee.org/Xplore/home.jsp, 2023. Accessed: January 21, 2023.

[Int08] International Organization for Standardization/International Electrotechnical Commission. ISO/IEC 25012:2008 Systems and software engineering – Software product Quality Requirements and Evaluation (SQuaRE) – Data quality model, 2008.

[Int11] International Organization for Standardization/International Electrotechnical Commission. ISO/IEC 25010:2011 Systems and software engineering – Systems and software Quality Requirements and Evaluation (SQuaRE) – System and software quality models, 2011.

[Int15a] International Organization for Standardization. ISO 9001:2015: Quality management systems - requirements, 2015.

[Int15b] International Organization for Standardization/International Electrotechnical Commission. ISO/IEC 33002:2015 Systems and software engineering – Process assessment – Requirements for process assessment models and assessment processes, 2015.

[Int16] International Organization for Standardization. ISO 15489-1:2016 Information and documentation – Records management – Part 1: Concepts and principles, 2016.

[Int19a] International Electrotechnical Commission/Institute of Electrical and Electronics Engineers. IEC/IEEE 82079-1:2019 Preparation of information for use (instructions for use) of products – Part 1: Principles and general requirements, 2019.

[Int19b] International Organization for Standardization. ISO 9241-220:2019 Ergonomics of human-system interaction – Part 220: Processes for enabling, executing and assessing human-centred design throughout the life cycle, 2019.

[Int19c] International Organization for Standardization/International Electrotechnical Commission. ISO/IEC 33020:2019 Systems and software engineering – Process measurement framework for assessment and improvement, 2019.

[Int19d] International Organization for Standardization/International Electrotechnical Commission/Institute of Electrical and Electronics Engineers. ISO/IEC/IEEE 15289:2019 Systems and software engineering – Content of life-cycle information items (documentation), 2019.

[Int21] International Organization for Standardization/International Electrotechnical Commission/Institute of Electrical and Electronics Engineers. ISO/IEC/IEEE 29119-2:2021 Software and systems engineering – Software testing – Part 2: Test processes, 2021.

[Int22a] International Organization for Standardization. ISO 31000: Risk Management, 2022.

[Int22b] International Organization for Standardization. ISO/IEC 27001:2022: Information technology - Security techniques - Information security management systems - Requirements, 2022.

[Int22c] International Organization for Standardization. ISO/IEC 27002:2022: Information security, cybersecurity and privacy protection — Information security controls, 2022.

[Int22d] International Organization for Standardization. ISO/IEC 27005:2022: Information security, cybersecurity and privacy protection — Guidance on managing information security risks, 2022.

[Int22e] International Organization for Standardization/International Electrotechnical Commission/Institute of Electrical and Electronics Engineers. ISO/IEC/IEEE 26514:2022 Systems and software engineering – Requirements for managers of user documentation, 2022.

[Int23a] International Organization for Standardization. ISO/IEC 23894:2023 Information technology — Artificial intelligence — Guidance on risk management, 2023.

[Int23b] International Organization for Standardization/International Electrotechnical Commission. ISO/IEC 27032:2012 Cybersecurity — Guidelines for Internet security, 2023.

[ISNS20] Muhammad Ismail, Mostafa F Shaaban, Mahesh Naidu, and Erchin Serpedin. Deep learning detection of electricity theft cyber-attacks in renewable distributed generation. *IEEE Transactions on Smart Grid*, 11(4):3428–3437, 2020.

[JBSP19] Mohammadreza Javadiha, Joaquim Blesa, Adria Soldevila, and Vicenç Puig. Leak localization in water distribution networks using deep learning. In *2019 6th International Conference on Control, Decision and Information Technologies (CoDIT)*, pages 1426–1431. IEEE, 2019.

[JHL+20] Shuo Jia, Fei Hui, Shining Li, Xiangmo Zhao, and Asad J Khattak. Long short-term memory and convolutional neural network for abnormal driving behaviour recognition. *IET Intelligent Transport Systems*, 14(5):306–312, 2020.

[jie20] YU jie. Application of artificial bee colony algorithms in smart grid. In *Journal of Physics: Conference Series*, volume 1453, page 012083. IOP Publishing, 2020.

[JK20] Deepack Jakhar and Ishmeet Kaur. Artificial intelligence, machine learning and deep learning: definitions and differences. *Clinical and experimental dermatology*, 45(1):131–132, 2020.

[JLY+19] Lihua Jian, Zhen Li, Xiaomin Yang, Wei Wu, Awais Ahmad, and Gwanggil Jeon. Combining unmanned aerial vehicles with artificial-intelligence technology for traffic-congestion recognition: electronic eyes in the skies to spot clogged roads. *IEEE Consumer Electronics Magazine*, 8(3):81–86, 2019.

[JM18] Junchen Jin and Xiaoliang Ma. Hierarchical multi-agent control of traffic lights based on collective learning. *Engineering applications of artificial intelligence*, 68:236–248, 2018.

[KCS93] J Kirakowski, M Corbett, and M Sumi. The software usability measurement inventory. *Br J Educ Technol*, 24(3):210–2, 1993.

[KDR19] Santhosh Kelathodi Kumaran, Debi Prosad Dogra, and Partha Pratim Roy. Queuing theory guided intelligent traffic scheduling through video analysis using dirichlet process mixture model. *Expert systems with applications*, 118:169–181, 2019.

[KDV18] P Kofinas, AI Dounis, and GA Vouros. Fuzzy q-learning for multi-agent decentralized energy management in microgrids. *Applied energy*, 219:53–67, 2018.

[KKGG21] Ammar Yousaf Kharal, Hassan Abdullah Khalid, Adel Gastli, and Josep M Guerrero. A novel features-based multivariate gaussian distribution method for the fraudulent consumers detection in the power utilities of developing countries. *IEEE Access*, 9:81057–81067, 2021.

[KP22] Akshaya Karthikeyan and U Priyakumar. Artificial intelligence: machine learning for chemical sciences. *Journal of Chemical Sciences*, 134, 03 2022.

[KPD+21] G Kothai, E Poovammal, Gaurav Dhiman, Kadiyala Ramana, Ashutosh Sharma, Mohammed A AlZain, Gurjot Singh Gaba, and Mehedi Masud. A new hybrid deep learning algorithm for prediction of wide traffic congestion in smart cities. *Wireless Communications and Mobile Computing*, 2021:1–13, 2021.

[LCM+20] Tianguang Lu, Xinyu Chen, Michael B McElroy, Chris P Nielsen, Qiuwei Wu, and Qian Ai. A reinforcement learning-based decision system for electricity pricing plan selection by smart grid end users. *IEEE Transactions on Smart Grid*, 12(3):2176–2187, 2020.

[Lew92] James R Lewis. Psychometric evaluation of the post-study system usability questionnaire: The pssuq. In *Proceedings of the human factors society annual meeting*, volume 36, pages 1259–1260. Sage Publications Sage CA: Los Angeles, CA, 1992.

[LF20] Mingzhu Li and Xi Feng. Research on adaptive comprehensive learning artificial bee colony algorithm and its application in constant pressure water supply. In *Journal of Physics: Conference Series*, volume 1650, page 032150. IOP Publishing, 2020.

[LGDH20] Yanchang Liang, Chunlin Guo, Zhaohao Ding, and Huichun Hua. Agent-based modeling in electricity market using deep deterministic policy gradient algorithm. *IEEE transactions on power systems*, 35(6):4180–4192, 2020.

[LH19] Renzhi Lu and Seung Ho Hong. Incentive-based demand response for smart grid with reinforcement learning and deep neural network. *Applied energy*, 236:937–949, 2019.

[LHWL20] Che-Tsung Lin, Sheng-Wei Huang, Yen-Yi Wu, and Shang-Hong Lai. Gan-based day-to-night image style transfer for nighttime vehicle detection. *IEEE Transactions on Intelligent Transportation Systems*, 22(2):951–963, 2020.

[LHZ18] Renzhi Lu, Seung Ho Hong, and Xiongfeng Zhang. A dynamic pricing demand response algorithm for smart grid: Reinforcement learning approach. *Applied Energy*, 220:220–230, 2018.

[LJL+21] Chanuk Lee, Dong Eun Jung, Donghoon Lee, Kee Han Kim, and Sung Lok Do. Prediction performance analysis of artificial neural network model by input variable combination for residential heating loads. *Energies*, 14(3):756, 2021.

[LWH+20] Peng Liu, Chaoyu Wang, Jia Hu, Tingting Fu, Nan Cheng, Ning Zhang, and Xuemin Shen. Joint route selection and charging discharging scheduling of evs in v2g energy network. *IEEE Transactions on Vehicular Technology*, 69(10):10630–10641, 2020.

[LXX+19] Xiaozhen Lu, Xingyu Xiao, Liang Xiao, Canhuang Dai, Mugen Peng, and H Vincent Poor. Reinforcement learning-based microgrid energy trading with a reduced power plant schedule. *IEEE Internet of Things Journal*, 6(6):10728–10737, 2019.

[MBGM22] Salvador Merino, Alfredo Burrieza, Francisco Guzman, and Javier Martinez. Smart sensorization using propositional dynamic logic. *Sensors*, 22(10):3899, 2022.

[MC19] Debotyam Maity and Jordan Ciezobka. Designing a robust proppant detection and classification workflow using machine learning for subsurface fractured rock samples post hydraulic fracturing operations. *Journal of Petroleum Science and Engineering*, 172:588–606, 2019.

[MG20] Mohammed Mynuddin and Weinan Gao. Distributed predictive cruise control based on reinforcement learning and validation on microscopic traffic simulation. *IET Intelligent Transport Systems*, 14(5):270–277, 2020.

[MMNK22a] Yueling Ma, Carsten Montzka, Bibi S Naz, and Stefan Kollet. Advancing ai-based pan-european groundwater monitoring. *Environmental Research Letters*, 17(11):114037, 2022.

[MMNK22b] Wulfran Fendzi Mbasso, Reagan Jean Jacques Molu, Serge Raoul Dzonde Naoussi, and Saatong Kenfack. Demand-supply forecasting based on deep learning for electricity balance in cameroon. *International Journal of Energy Economics and Policy*, 12(4):99–103, 2022.

[MPG21] Alessandro Massaro, Antonio Panarese, and Angelo Galiano. Technological platform for hydrogeological risk computation and water leakage detection based on a convolutional neural network. In *2021 IEEE International Workshop on Metrology for Industry 4.0 & IoT (MetroInd4.0&IoT)*, pages 225–230. IEEE, 2021.

[MSW+20] Yiming Miao, Jeungeun Song, Haoquan Wang, Long Hu, Mohammad Mehedi Hassan, and Min Chen. Smart micro-gas: A cognitive

micro natural gas industrial ecosystem based on mixed blockchain and edge computing. *IEEE Internet of Things Journal*, 8(4):2289–2299, 2020.

[Mun14] Andres Munoz. Machine learning and optimization. *URL: https://www. cims. nyu. edu/˜ munoz/files/ml_optimization. pdf [accessed 2016-03-02][WebCite Cache ID 6fiLfZvnG]*, 2014.

[MZG20] Yiming Miao, Ming Zhou, and Ahmed Ghoneim. Blockchain and ai-based natural gas industrial iot system: architecture and design issues. *IEEE Network*, 34(5):84–90, 2020.

[NB22] Shilpa R Nair and BK Bhavathrathan. Hybrid segmentation approach to identify crash susceptible locations in large road networks. *Safety science*, 145:105515, 2022.

[NBA+22] Omar A Nasseef, Abdullah M Baabdullah, Ali Abdallah Alalwan, Banita Lal, and Yogesh K Dwivedi. Artificial intelligence-based public healthcare systems: G2g knowledge-based exchange to enhance the decision-making process. *Government Information Quarterly*, 39(4):101618, 2022.

[NMZ+21] Seyed Azad Nabavi, Naser Hossein Motlagh, Martha Arbayani Zaidan, Alireza Aslani, and Behnam Zakeri. Deep learning in energy modeling: application in smart buildings with distributed energy generation. *IEEE Access*, 9:125439–125461, 2021.

[NNB+19] Dinithi Nallaperuma, Rashmika Nawaratne, Tharindu Bandaragoda, Achini Adikari, Su Nguyen, Thimal Kempitiya, Daswin De Silva, Damminda Alahakoon, and Dakshan Pothuhera. Online incremental machine learning platform for big data-driven smart traffic management. *IEEE Transactions on Intelligent Transportation Systems*, 20(12):4679–4690, 2019.

[NRH20] Ahmed Abdel Nasser, Magdi Z Rashad, and Sherif E Hussein. A two-layer water demand prediction system in urban areas based on micro-services and lstm neural networks. *IEEE Access*, 8:147647–147661, 2020.

[NSN+22] Pushpa Bhakuni Negi, Sandeep Kumar Sunori, Hansi Negi, Amit Mittal, Shweta Arora, Pradeep Juneja, and Anita Rana. Ai and ml based prediction of water hardness. In *2022 2nd International Conference on Intelligent Technologies (CONIT)*, pages 1–5. IEEE, 2022.

[NSZ+18] Khoi A Nguyen, Rodney A Stewart, Hong Zhang, Oz Sahin, and Nilmini Siriwardene. Re-engineering traditional urban water management practices with smart metering and informatics. *Environmental modelling & software*, 101:256–267, 2018.

[Off19] Official Journal of the European Union. Directive (EU) 2019/1024 of the European Parliament and of the Council of 20 June 2019 on open data and the re-use of public sector information, 2019.

[Ong17] Pariwat Ongsulee. Artificial intelligence, machine learning and deep learning. In *2017 15th international conference on ICT and knowledge engineering (ICT&KE)*, pages 1–6. IEEE, 2017.

[OS10] Chitu Okoli and Kira Schabram. A guide to conducting a systematic literature review of information systems research. 2010.

[PD21] Vishwa P Parmar and Akshit J Dhruv. Efficient sea water purification using hybrid nanofiltration system and ml for optimization. In *2021 International Conference on Artificial Intelligence and Machine Vision (AIMV)*, pages 1–6. IEEE, 2021.

[PMH20] Sungwoo Park, Jihoon Moon, and Eenjun Hwang. Explainable anomaly detection for district heating based on shapley additive explanations. In *2020 International Conference on Data Mining Workshops (ICDMW)*, pages 762–765. IEEE, 2020.

[PMI+19] Alireza Pourdaryaei, Hazlie Mokhlis, Hazlee Azil Illias, S HR Aghay Kaboli, Shameem Ahmad, and Swee Peng Ang. Hybrid ann and artificial cooperative search algorithm to forecast short-term electricity price in de-regulated electricity market. *Ieee Access*, 7:125369–125386, 2019.

[PML18] Alexandru Predescu, Mariana Mocanu, and Ciprian Lupu. A modern approach for leak detection in water distribution systems. In *2018 22nd International Conference on System Theory, Control and Computing (ICSTCC)*, pages 486–491. IEEE, 2018.

[PMM+21] Alireza Pourdaryaei, Mohammad Mohammadi, MunirAzam Muhammad, Junaid Bin Fakhrul Islam, Mazaher Karimi, and Amidaddin Shahriari. An efficient framework for short-term electricity price forecasting in deregulated power market. *IEEE Access*, 2021.

[PPPMM22] José Pérez-Padillo, Francisco Puig, Jorge García Morillo, and Pilar Montesinos. Iot platform for failure management in water transmission systems. *Expert Systems with Applications*, 199:116974, 2022.

[PPR+20] Irina Valeryevna Pustokhina, Denis Alexandrovich Pustokhin, Joel JPC Rodrigues, Deepak Gupta, Ashish Khanna, K Shankar, Changho Seo, and Gyanendra Prasad Joshi. Automatic vehicle license plate recognition using optimal k-means with convolutional neural network for intelligent transportation systems. *Ieee Access*, 8:92907–92917, 2020.

[PR06] Mark Petticrew and Helen Roberts. *Systematic reviews in the social sciences: A practical guide*. John Wiley & Sons, 2006.

[PSA+22] Prachi Palsodkar, Rishabh Shrivastav, Anvesh Ayangar, Sayali Atkare, Suraj Yadav, and Prasanna Palsodkar. Sensor cloudlet interconnecting system for water reservoirs security. In *2022 IEEE 10th Region 10 Humanitarian Technology Conference (R10-HTC)*, pages 67–70. IEEE, 2022.

[PVG+20] Dennis Pierl, Kai Vahldiek, Julia Geißler, Bernd Rüger, Kai Michels, Frank Klawonn, and Andreas Nürnberger. Online model-and data-based leakage localization in district heating networks-impact of random measurement errors. In *2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, pages 2331–2338. IEEE, 2020.

[QSWS19] Tao Qian, Chengcheng Shao, Xiuli Wang, and Mohammad Shahidehpour. Deep reinforcement learning for ev charging navigation by coordinating smart grid and intelligent transportation system. *IEEE transactions on smart grid*, 11(2):1714–1723, 2019.

[RG20] Aritra Ray and Shreemoyee Goswami. Iot and cloud computing based smart water metering system. In *2020 International Conference on Power Electronics & IoT Applications in Renewable Energy and its Control (PARC)*, pages 308–313. IEEE, 2020.

[RHAM20] Tsotsope Daniel Ramotsoela, Gerhard Petrus Hancke, and Adnan M Abu-Mahfouz. Behavioural intrusion detection in water distribution systems using neural networks. *IEEE Access*, 8:190403–190416, 2020.

[Rot07] Edna Terezinha Rother. Systematic literature review x narrative review. *Acta Paulista de Enfermagem*, 20:v–vi, 2007.

[RP18] Hamid Rahmanifard and Tatyana Plaksina. Application of fast analytical approach and ai optimization techniques to hydraulic fracture stage placement in shale gas reservoirs. *Journal of Natural Gas Science and Engineering*, 52:367–378, 2018.

[Sar83] Vernon T Sarver. Ajzen and fishbein's" theory of reasoned action": A critical assessment. 1983.

[SB19] Matthew Stevenson and Cristián Bravo. Advanced turbidity prediction for operational water supply planning. *Decision Support Systems*, 119:72–84, 2019.

[SC14] Manu Sood and Ashwani Chandel. Searching and optimization techniques in artificial intelligence: A comparative study & complexity analysis. 03 2014.

[SCP20] Dong-Hoon Shin, Kyungyong Chung, and Roy C Park. Prediction of traffic congestion based on lstm through correction of missing temporal and spatial data. *IEEE Access*, 8:150784–150796, 2020.

[SFN21] Iqbal H Sarker, Md Hasan Furhad, and Raza Nowrozy. Ai-driven cybersecurity: an overview, security intelligence modeling and research directions. *SN Computer Science*, 2(3):1–18, 2021.

[SGB+21] Mamoudou Sangare, Sharut Gupta, Samia Bouzefrane, Soumya Banerjee, and Paul Muhlethaler. Exploring the forecasting approach for road accidents: Analytical measures with hybrid machine learning. *Expert Systems with Applications*, 167:113855, 2021.

[SJA+20] Vivek Singhal, SS Jain, Divya Anand, Aman Singh, Sahil Verma, Joel JPC Rodrigues, Noor Zaman Jhanjhi, Uttam Ghosh, Ohyun Jo, Celestine Iwendi, et al. Artificial intelligence enabled road vehicle-train collision risk assessment framework for unmanned railway level crossings. *IEEE Access*, 8:113790–113806, 2020.

[SL16] Jeff Sauro and James R. Lewis. Chapter 8 - standardized usability questionnaires. In Jeff Sauro and James R. Lewis, editors, *Quantifying the User Experience (Second Edition)*, pages 185–248. Morgan Kaufmann, Boston, second edition edition, 2016.

[Sol20] Davut Solyali. A comparative analysis of machine learning approaches for short-/long-term electricity load forecasting in cyprus. *Sustainability*, 12(9):3612, 2020.

[SSA+22] Ayda Sarreshtedar, Elnaz Sharghi, Amin Afkhaminia, Vahid Nourani, and Anne Ng. Investigation of quantitative and qualitative changes in groundwater of ardebil plain using ensemble artificial intelligence-based modeling. *Water Supply*, 22(9):7140–7157, 2022.

[Sta22] Statista. Market size and revenue comparison for artificial intelligence worldwide from 2018 to 2030 (in billion u.s. dollars) [graph]. In Statista, June 27 2022. Retrieved May 20, 2023.

[SXP+20] Jiancai Song, Guixiang Xue, Xuhua Pan, Yunpeng Ma, and Han Li. Hourly heat load prediction model based on temporal convolutional neural network. *IEEE Access*, 8:16726–16741, 2020.

[TLSK21] Sehyun Tak, Jong-Deok Lee, Jeongheon Song, and Sunghoon Kim. Development of ai-based vehicle detection and tracking system for c-its application. *Journal of advanced transportation*, 2021:1–15, 2021.

[TMK+19] Athanasia K Tolkou, Manassis Mitrakas, Ioannis A Katsoyiannis, Mathias Ernst, and Anastasios I Zouboulis. Fluoride removal from water by

composite al/fe/si/mg pre-polymerized coagulants: Characterization and application. *Chemosphere*, 231:528–537, 2019.

[TS12] Brandon M Turner and Per B Sederberg. Approximate bayesian computation with differential evolution. *Journal of Mathematical Psychology*, 56(5):375–385, 2012.

[VDA20] Informationssicherheit in unternehmen. `https://www.vda.de/de/themen/digitalisierung/daten/informationssicherheit`, 2020.

[vNM22] Colin van Noordt and Gianluca Misuraca. Artificial intelligence for the public sector: results of landscaping the use of ai in government across the european union. *Government Information Quarterly*, 39(3):101714, 2022.

[VSR+21] A Harsha Vardhan, Bramhadevara Subramanyam, M Lakshmi Reddy, G Gowtham Reddy, and A Ramesh. Anomaly detection in water distribution systems. 2021.

[Wan19] Pei Wang. On defining artificial intelligence. *Journal of Artificial General Intelligence*, 10(2):1–37, 2019.

[Wan21] Ziyi Wang. An ai to help reduce heating bills. In *2021 3rd International Conference on Machine Learning, Big Data and Business Intelligence (MLBDBI)*, pages 581–584. IEEE, 2021.

[WBL+21] Yi Wang, Imane Lahmam Bennani, Xiufeng Liu, Mingyang Sun, and Yao Zhou. Electricity consumer characteristics identification: A federated learning approach. *IEEE Transactions on Smart Grid*, 12(4):3637–3647, 2021.

[WDB+21] Chia-Nan Wang, Thanh-Tuan Dang, Julius Bayer, et al. A two-stage multiple criteria decision making for site selection of solar photovoltaic (pv) power plant: A case study in taiwan. *IEEE Access*, 9:75509–75525, 2021.

[WFZL19] Siyuan Wu, Jie Fang, Xiangtao Zheng, and Xijie Li. Sample and structure-guided network for road crack detection. *IEEE Access*, 7:130032–130043, 2019.

[Win08] Robert Winter. Design science research in europe. *European Journal of Information Systems*, 17(5):470–475, 2008.

[WKMP18] Marcin Woźniak, Kamil Książek, Jakub Marciniec, and Dawid Połap. Heat production optimization using bio-inspired algorithms. *Engineering Applications of Artificial Intelligence*, 76:185–201, 2018.

[WLB+18] Guojun Wu, Yanhua Li, Jie Bao, Yu Zheng, Jieping Ye, and Jun Luo. Human-centric urban transit evaluation and planning. In *2018 IEEE International Conference on Data Mining (ICDM)*, pages 547–556. IEEE, 2018.

[WOH+20] Zhishang Wang, Mark Ogbodo, Huakun Huang, Chen Qiu, Masayuki Hisada, and Abderazek Ben Abdallah. Aebis: Ai-enabled blockchain-based electric vehicle integration system for power management in smart grid platform. *IEEE Access*, 8:226409–226421, 2020.

[WWKF20] Guiqiang Wang, Haiman Wang, Zhiqiang Kang, and Guohui Feng. Data-driven optimization for capacity control of multiple ground source heat pump system in heating mode. *Energies*, 13(14):3595, 2020.

[XLKK21] Xiaojun Xiang, Qiong Li, Shahnawaz Khan, and Osamah Ibrahim Khalaf. Urban water resource management for sustainable environment planning using artificial intelligence techniques. *Environmental Impact Assessment Review*, 86:106515, 2021.

[XW19] Yu Xiao and Maria Watson. Guidance on conducting a systematic literature review. *Journal of planning education and research*, 39(1):93–112, 2019.

[YBY+20] Shiyun Yang, Emily Bailey, Zhengye Yang, Jonatan Ostrometzky, Gil Zussman, Ivan Seskar, and Zoran Kostic. Cosmos smart intersection: Edge compute and communications for bird's eye object tracking. In *2020 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops)*, pages 1–7. IEEE, 2020.

[YQS+19] Yujian Ye, Dawei Qiu, Mingyang Sun, Dimitrios Papadaskalopoulos, and Goran Strbac. Deep reinforcement learning for strategic bidding in electricity markets. *IEEE Transactions on Smart Grid*, 11(2):1343–1355, 2019.

[YRS+20] Nina Bozic Yams, Valerie Richardson, Galina Esther Shubina, Sandor Albrecht, and Daniel Gillblad. Integrated ai and innovation management: The beginning of a beautiful friendship. *Technology Innovation Management Review*, 10(11), 2020.

[ZBF21] Fan Zhang, Chris Bales, and Hasan Fleyeh. Night setback identification of district heat substations using bidirectional long short term memory with attention mechanism. *Energy*, 224:120163, 2021.

[ZHG+19] Suyang Zhou, Zijian Hu, Wei Gu, Meng Jiang, and Xiao-Ping Zhang. Artificial intelligence based smart energy community management: A reinforcement learning approach. *CSEE Journal of Power and Energy Systems*, 5(1):1–10, 2019.

[ZLFC20] Haifeng Zheng, Feng Lin, Xinxin Feng, and Youjia Chen. A hybrid deep learning model with attention-based conv-lstm networks for short-term traffic flow prediction. *IEEE Transactions on Intelligent Transportation Systems*, 22(11):6910–6920, 2020.

[ZLLH20] Chenyue Zhang, Wenjia Li, Yuansheng Luo, and Yupeng Hu. Ait: An ai-enabled trust management system for vehicular networks using blockchain technology. *IEEE Internet of Things Journal*, 8(5):3157–3169, 2020.

[ZLW19] Bingpeng Zhou, Vincent Lau, and Xun Wang. Machine-learning-based leakage-event identification for smart water supply systems. *IEEE Internet of Things Journal*, 7(3):2277–2292, 2019.

[ZMR23] Ariele Zanfei, Andrea Menapace, and Maurizio Righetti. An artificial intelligence approach for managing water demand in water supply systems. In *IOP Conference Series: Earth and Environmental Science*, volume 1136, page 012004. IOP Publishing, 2023.

[ZPL+19] Mengfei Zhou, Zheng Pan, Yunwen Liu, Qiang Zhang, Yijun Cai, and Haitian Pan. Leak detection and location based on islmd and cnn in a pipeline. *IEEE Access*, 7:30457–30464, 2019.

[ZQS18] Wenjie Zhang, Hao Quan, and Dipti Srinivasan. An improved quantile regression neural network for probabilistic load forecasting. *IEEE Transactions on Smart Grid*, 10(4):4425–4434, 2018.

[ZWX+20] Zhengyang Zhou, Yang Wang, Xike Xie, Lianliang Chen, and Hengchang Liu. Riskoracle: a minute-level citywide traffic accident forecasting framework. In *Proceedings of the AAAI conference on artificial intelligence*, volume 34, pages 1258–1265, 2020.

[ZWZ19] Qingguo Zhou, Chen Wang, and Gaofeng Zhang. Hybrid forecasting system based on an optimal model selection strategy for different wind speed forecasting problems. *Applied Energy*, 250:1559–1580, 2019.

Appendix A

Artificial Intelligence Act - Assessment

General information

The European Commission has proposed the Artificial Intelligence Act (AI Act) to regulate the development and deployment of AI systems in the European Union. The proposed regulation is intended to ensure that AI is designed, developed and operated in a manner that is safe, ethically correct and that respects fundamental rights. The AI Act proposes a risk-based approach with several additional requirements for high-risk AI systems. These systems pose substantial risks to health, safety, or basic rights, such as those used in critical infrastructure.
This requirements catalogue represents a tool for providers of high risk AI systems in the context of critical infrastructure to check their compliance with the AI Act.

All blue coloured worksheets have the purpose of information support.
All orange worksheets contain the controls that need to be filled.
All green worksheets present the results based on the provided information.

Worksheet - Terms
The descriptions of the requirements contain terms that may need further explanation. In this worksheet these terms are described in more detail with the relevant definitions taken from the AI Act.

Worksheet - Maturity levels
A key factor of the compliance evaluation is the assignment of maturity levels to the individual requirements ranging from 0 to 5. The maturity levels are based on the ISO/IEC 33020:2019 and provide a detailed description that comprises the various maturity degrees a process. Providers of high-risk AI systems should be able to understand their fulfilment of a requirement based on the guidance of this worksheet. It is essential to note that a score of 3 is sufficient for both a particular requirement and the overall score.

Worksheet - Annexes
Some requirements and obligations of this catalogue refer to content of the AI Act's Annexes. These have been included as an additional worksheet so that the information can be accessed directly without having to transition between multiple documents.

Worksheet - AI Act Mapping
This worksheet lists the individual requirements from the AI Act in their original order. For each requirement that is relevant for providers in critical infrastructure, the new ID of the catalogue is also referenced. This makes it possible to quickly determine which requirements of the catalogue belong to the ones of the AI Act. In the comment column, an additional explanation was given as to why a requirement was not dealt with in the catalogue.

Worksheet - Requirements for the system

This worksheet contains all the summarised requirements of the AI Act that are relevant to providers and focus directly on high-risk AI systems. The individual columns are separated into those that provide information and those that must be completed.

Information:

Chapter ID - assigns every requirement an ID in accordance with the ID of its main chapter
Control question - formulates and summarises a requirement in the form of a question, supporting the evaluation process. Depending on the type of question the maturity level can range form 0 to 5 or be selected between "yes" and "no", which corresponds to maturity levels of 0 and 3.
Requirement objective - highlights the specific objectives of the individual requirements in a few bullet points
Requirement description - contains specific information regarding the requirement
Possible proof for fulfillment - indicates the processes, systems, documents or certificates that are among other things a possible way to demonstrate compliance with the AI Act. The possible proof of each row refers to the respective standards in column H.
Reference to AI Act , *Reference to other standards and Regulations* - display one or multiple chapters, standards, norms and other EU regulations that are linked to the requirement

Filled in by the provider:

Maturity level - builds the most significant factor for assessing the compliance and should be filled out in any case by the provider
Description of implementation and *Date of assessment* - describe the processes, systems and measures that the provider has implemented to meet the requirement, along with the date of assessing them
Responsible department and *Contact* - identifies the group or individuals that are responsible for the implementation of the requirement
Future planned measures and *Due Date* - outline all measures that will be implemented to increase the maturity level in the near future and to what date they are anticipated to be concluded

Worksheet - Obligations of providers

This worksheet contains all the summarised obligations of providers and is structured in the same way as the previous worksheet.

Worksheet - Results

Based on the provided maturity levels of the two previous worksheets the overall results are displayed here. Each requirement's Chapter ID, Control question, Maturity level and capped Maturity level are summarised in tables. The Target maturity level can be selected, indicating the desired maturity level to be attained by the user. A maturity level of 3 is recommended because, at this level, you can be confident that the measures are adequately documented and demonstrate compliance with the AI Act. Anything below this level is not indicative and anything above is just an additional improvement. The above-mentioned capped maturity level reduces all values that exceed the target maturity level. This means that if a target maturity level of 3, for instance, is selected and a requirement is rated a 5, the shorted maturity level is 3. The purpose of this shortening is intended to ensure that in the calculation of the total average maturity level very good results in one area don't compensate poor results in other area, because all requirements must be sufficiently fulfilled for full compliance. Moreover, the spider charts demonstrate the maturity level of the main chapters with the target maturity level indicated by green line.

Artificial Intelligence Act - Assessment

Glossary of terms

Term	Explanation
Artificial intelligence system	AI system means a system that is designed to operate with elements of autonomy and that, based on machine and/or human-provided data and inputs, infers how to achieve a given set of objectives using machine learning and/or logic- and knowledge based approaches, and produces system-generated outputs such as content (generative AI systems), predictions, recommendations or decisions, influencing the environments with which the AI system interacts.
Authorised representative	Authorised representative means any natural or legal person physically present or established in the Union who has received and accepted a written mandate from a provider of an AI system to, respectively, perform and carry out on its behalf the obligations and procedures established by this Regulation.
CE marking of conformity	*CE marking of conformity* or *CE marking* means a marking by which a provider indicates that an AI system is in conformity with the requirements set out in Title III, Chapter 2 or in Article 4b of this Regulation and other applicable Union legal act harmonising the conditions for the marketing of products ('Union harmonisation legislation') providing for its affixing.
Conformity assessment	*Conformity assessment* means the process of verifying whether the requirements set out in Title III, Chapter 2 of this Regulation relating to a high-risk AI system have been fulfilled.
Harmonised standard	*Harmonised standard* means a European standard adopted on the basis of a request made by the Commission for the application of Union harmon isation legislation [Article 2(1)(c) of Regulation (EU) No 1025/2012].
Input data	*Input data* means data provided to or directly acquired by an AI system on the basis of which the system produces an output.
Instructions for use	Instructions for use means the information provided by the provider to inform the user of in particular an AI system's intended purpose and proper use.
Intended purpose	Intended purpose means the use for which an AI system is intended by the provider, including the specific context and conditions of use, as specified in the information supplied by the provider in the instructions for use, promotional or sales materials and statements, as well as in the technical documentation.
Law enforcement	*Law enforcement* means activities carried out by law enforcement authorities or on their behalf for the prevention, investigation, detection or prosecution of criminal offences or the execution of criminal penalties, including the safeguarding against and the prevention of threats to public security.
Life cycle of an AI system	Life cycle of an AI system means the duration of an AI system, from design through retirement. Without prejudice to the powers of the market surveillance authorities, such retirement may happen at any point in time during the post-market monitoring phase upon the decision of the provider and implies that the system may not be used further. An AI system lifecycle is also ended by a substantial modification to the AI system made by the provider or any other natural or legal person, in which case the substantially modified AI system shall be considered as a new AI system.
Making available on the market	Making available on the market means any supply of an AI system for distribution or use on the Union market in the course of a commercial activity, whether in return for payment or free of charge.
National competent authority	*National competent authority* means any of the following: the notifying authority and the market surveillance authority. As regards AI systems put into service or used by EU institutions, agencies, offices and bodies, the European Data Protection Supervisor shall fulfil the responsibilities that in the Member States are entrusted to the national competent authority and, as relevant, any reference to national competent authorities or market surveillance authorities in this Regulation shall be understood as referring to the European Data Protection Supervisor.
Notifying authority	*Notifying authority* means the national authority responsible for setting up and carrying out the necessary procedures for the assessment, designation and notification of conformity assessment bodies and for their monitoring;

Notifying authority	*Notifying authority* means the national authority responsible for setting up and carrying out the necessary procedures for the assessment, designation and notification of conformity assessment bodies and for their monitoring;
Performance of an AI system	Performance of an AI system means the ability of an AI system to achieve its intended purpose.
Personal data	*Personal data* means any information relating to an identified or identifiable natural person ('data subject'); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person; [point (1) of Article 4 of Regulation (EU) 2016/679].
Placing on the market	Placing on the market means the first making available of an AI system on the Union market.
Post-market monitoring system	*Post-market monitoring system* means all activities carried out by providers of AI systems to collect and review experience gained from the use of AI systems they place on the market or put into service for the purpose of identifying any need to immediately apply any necessary corrective or preventive actions.
Provider	Provider means a natural or legal person, public authority, agency or other body that develops an AI system or that has an AI system developed and places that system on the market or puts it into service under its own name or trademark, whether for payment or free of charge.
Putting into service	Putting into service means the supply of an AI system for first use directly to the user or for own use in the Union for its intended purpose.
Real world testing plan	*Real world testing plan* means a document that describes the objectives, methodology, geographical, population and temporal scope, monitoring, organisation and conduct of testing in real world conditions.
Reasonably foreseeable misuse	Reasonably foreseeable misuse means the use of an AI system in a way that is not in accordance with its intended purpose, but which may result from reasonably foreseeable human behaviour or interaction with other systems.
Serious incident	*Serious incident* means any incident or malfunctioning of an AI system that directly or indirectly leads to any of the following: (a) the death of a person or serious damage to a person's health (b) a serious and irreversible disruption of the management and operation of critical infrastructure (c) breach of obligations under Union law intended to protect fundamental rights (d) serious damage to property or the environment
Testing data	*Testing data* means data used for providing an independent evaluation of the trained and validated AI system in order to confirm the expected performance of that system before its placing on the market or putting into service.
Testing in real world conditions	*Testing in real world conditions* means the temporary testing of an AI system for its intended purpose in real world conditions outside of a laboratory or otherwise simulated environment with a view to gathering reliable and robust data and to assessing and verifying the conformity of the AI system with the requirements of this Regulation; testing in real world conditions shall not be considered as placing the AI system on the market or putting it into service within the meaning of this Regulation, provided that all conditions under Article 53 or Article 54a are fulfilled.
Training data	*Training data* means data used for training an AI system through fitting its learnable parameters.
User	User means any natural or legal person, including a public authority, agency or other body, under whose authority the system is used.
Validation data	*Validation data* means data used for providing an evaluation of the trained AI system and for tuning its non-learnable parameters and its learning process, among other things, in order to prevent overfitting; whereas the validation dataset can be a separate dataset or part of the training dataset, either as a fixed or variable split.

All definitions of terms are taken from the AI Act

Artificial Intelligence Act - Assessment

Requirements for the system

Chapter ID	Control question	Requirement objective	Requirement description	Possible proof of requirement	Reference to AI Act	Reference to other standards or regulations
1.	Risk management					
1.1	To what extent is a risk management system implemented and documented?	- Established risk management system	Providers are required to establish, implement, document and maintain a risk management system that is consistent with the intended purpose of the AI system. This means that the risks associated with the specific application of the AI system, including the potential impacts on individuals, society and the environment must be considered.	- ISO 31000 certification - Defined policies and procedures regarding the management of risks - Records, reports and audit trails that demonstrate the effectiveness of the risk management system	Article 8(2), Article 9(1)	ISO 31000, ISO/IEC 23894:2023, ISO/IEC 27001:2022
1.2	To what extent are the procedures of the risk management system integrated into the system's life cycle?	- Identification and analysis of potential risks - Evaluation of additional risks based on post-market monitoring data - Adoption of suitable risk management measures	The risk management system must be an ongoing process throughout the system's entire life cycle. This includes identifying and analysing potential risks to health, safety and fundamental rights and further evaluating additional risks based on post-market monitoring data. Suitable risk management measures must then also be adopted to mitigate or eliminate the identified risks through system development, design or technical information provision.	- ISO 31000 certification - A documented procedure for identifying, assessing and addressing risks within the system's lifecycle - A catalogue of risk criteria including the likelihood and potential impact of a risk event - Documented measures for dealing with risks and their responsible parties	Article 9(2)	ISO 31000, ISO/IEC 23894:2023, ISO/IEC 27001:2022
1.3	To what extent have the risk management measures taken into account the potential effects resulting from the combined application of the requirements?	- Implementation of risk management measures in accordance with the requirements of the catalogue	Appropriate risk management measures must be implemented in accordance with the following requirements. These measures must take into account the potential effects and interactions resulting from the combined application of all of the catalogue's requirements.	- ISO 31000 certification - Risk assessment procedures that identify and document potential combined risks that may arise from the interaction of different risks - A catalogue of potential risk sources - Documented control measures to mitigate combined risks	Article 9(3)	ISO 31000, ISO/IEC 23894:2023, ISO/IEC 27001:2022
1.4	To what extent do the risk management measures ensure that residual risks of the system are considered acceptable?	- Assurance that residual risks associated with each hazard are considered acceptable - Adequate design and development of the system to eliminate/reduce risks - Adequate mitigation and control measures where risks cannot be eliminated - Provision of information and appropriate training to users regarding risk measures	The risk management measures must ensure that any residual risks associated with each hazard, as well as the overall residual risk of the system, are considered acceptable. To identify the most appropriate risk management measures, the system must be adequately designed and developed to eliminate or reduce identified risks. Where risks cannot be eliminated, adequate mitigation and control measures must be implemented. Users must be provided with information regarding these measures and if necessary, appropriate training must be given. Technical knowledge, experience, education, training expected by the user and the environment where the system will be used must be considered when eliminating or reducing risks related to the system's use.	- Documented risk mitigation and control measures - Decision process to whether avoid, reduce, transfer or accept risks - Training materials for employees and users	Article 9(4)	ISO 31000, ISO/IEC 23894:2023, ISO/IEC 27001:2022
1.5	To what extent has the AI system been tested to ensure that it operates as intended?	- Conducted tests to ensure the compliance with the requirements	AI systems must be tested to ensure that they operate as intended and meet the requirements outlined in this catalogue. This may include testing under real-life conditions.	- Defined frequency or times for testing - A catalogue of testing criteria including completion criteria - Documented testing procedure	Article 9(5), Article 9(6), Article 54a	ISO 31000, ISO/IEC 23894:2023, ISO/IEC/IEEE 29119-2:2021
1.6	To what extent has the AI system been tested at different steps of development?	- Conducted testes at different steps of development with appropriate measures and probability limits	AI systems must be tested at different steps of development and before they are used or put on the market. During testing, measures and probability thresholds that match the system's intended use must be used to make sure it operates as expected and meets the required standards.	- Defined frequency or times for testing - A catalogue of testing criteria including completion criteria - Defined acceptable probability range for test results (probabilistic thresholds) - Documented testing procedure	Article 9(7)	ISO 31000, ISO/IEC 23894:2023, ISO/IEC/IEEE 29119-2:2022
1.7	To what extent does the risk management system consider the potential impact on individuals under the age of 18?	- Consideration whether the AI system affects individuals under the age of 18	The risk management system outlined in the points 1.1 to 1.6 must consider whether the AI system will be accessed or affect individuals under the age of 18.	- Policies that address risks to individuals under the age of 18 - Documented risk assessment process that includes the identification of potential risks to individuals under the age of 18	Article 9(8)	ISO 31000, ISO/IEC 23894:2023, ISO/IEC 27001:2022

2.	**Data management**					
-			When developing AI systems employing data-driven model training techniques, high-quality training, validation and testing datasets that meet the quality criteria enumerated in points 2.1 to 2.4 are required. These requirements only apply to the testing datasets for the development AI systems that do not use techniques involving the training of models.		Article 10(1), Article 10(6)	
2.1	To what extent are appropriate data governance and management practices implemented to ensure the quality of datasets?	- Appropriate data governance and management practices to assure the quality of used datasets	Appropriate data governance and management practises must be implemented to assure the quality of training, validation, and testing datasets used in AI systems. These practises shall include a variety of aspects, such as design decisions, data collection processes, data preparation, formulation of assumptions, prior assessment of the availability and suitability of datasets, examination for potential biases, identification of data gaps and methods for addressing them.	- Documentation of design decisions, data collection, data preparation (e.g. labelling, cleaning) - List of assumptions that describes what the data measures and represents - Reports of prior assessments of the datasets - Defined procedures for examining potential biases in the datasets - Selection of methods for addressing identified gaps - Procedures for ongoing monitoring and updating of datasets	Article 10(2)	ISO/IEC 25012:2008
2.2	To what extent meet the datasets the requirements of being appropriate, error-free and comprehensive?	- Appropriate, error-free, comprehensive datasets	The datasets used for training, validation and testing of AI systems must be appropriate, error-free and comprehensive, as well as possess the required statistical properties applicable to the intended consumers. Individual datasets or their combination may be used to accomplish the required statistical properties.	- Defined data assessment plan - Catalogue of inherent data quality criteria including (e.g. Accuracy, Completeness, Consistency, Credibility, Currentness, Accessibility etc.)	Article 10(3)	ISO/IEC 25012:2009
2.3	To what extent do the datasets take into account the characteristics of the specific geographical, behavioural or functional context of the AI system?	- Datasets that consider the unique characteristics of the specific geographical, functional or behavioural context of the system	The data used in the training, validation and testing must consider the unique characteristics of the specific geographical, functional or behavioural context in which the system is intended to be used. Therefore the data must be representative of the real-world context in which the system will be used and this context may differ based on the particular application or use case.	- Defined data assessment plan - Defined intended purpose of the AI system - Documented process for identification of geographical, functional or behavioural settings - Catalogue of system dependent data quality criteria including (e.g. Accessibility, Compliance, Confidentiality, Efficiency, Precision, Traceability, Understandability etc.)	Article 10(4)	ISO/IEC 25012:2010
2.4	To what extent are fundamental rights and freedoms safeguarded when processing special categories of personal data?	- Appropriate safeguards for the fundamental rights and freedoms of natural persons - State-of-the-art security and privacy-preserving measures	System providers may engage in the processing of specialised categories of personal data for the monitoring and correction of biases, provided that they maintain sufficient protections for essential rights and freedoms. The providers must use technical measures to safeguard the data, such as replacing identifying information with pseudonyms or encrypting the data to prevent unauthorised access.	- Established policies and procedures for the handling of special categories of personal data - Documented processes for state-of-the-art technical measures, such as pseudonymisation, encryption or anonymisation - Defined data assessment plan to test the effectiveness of safeguards	Article 10(5)	Article 9(1) of Regulation (EU) 2016/679, Article 10 of Directive (EU) 2016/680, Article 10(1) of Regulation (EU) 2018/1725)
3.	**Technical documentation**					
3.1	To what extent is the technical documentation of the AI system prepared and regularly updated?	- Up-to date technical documentation	The technical documentation of a AI system must be prepared and kept up-to-date before the system is made available.	- Technical documentation of the AI system including version control (outdated versions of the documentation are preserved and changes to the documentation are tracked) - Defined frequency or times for updating the technical documentation	Article 11(1)	ISO/IEC/IEEE 15289:2019
3.2	To what extent is the technical documentation of an AI system compliant with the requirements outlined in the catalogue?	- Technical documentation that complies with the requirements in this catalogue - Documentation that includes minimal elements listed in Annex IV	The technical documentation must be compliant with requirements outlined in this catalogue and provide relevant authorities with comprehensive information to assess the system's compliance. The documentation must include the minimal elements listed in Annex IV or in the case of SMEs comparable documentation that serves the same purpose, unless the relevant authority deems it unsuitable.	- Technical documentation of the AI system in an organised and clearly presented format	Article 11(1)	ISO/IEC/IEEE 15289:2020
3.3	To what extent does the technical documentation contain all the required information?	- A single technical documentation that includes all required information specified in Annex IV and relevant legal acts listed in Annex II, section A	A single technical documentation containing all the required information, as specified in Annex IV and information of relevant legal acts listed in Annex II, section A, must be created when a AI system associated with a product under those acts is placed on the market or put into service.	- Policies that prevent multiple technical documentations	Article 11(2)	ISO/IEC/IEEE 15289:2021

3.	**Technical documentation**					
3.1	To what extent is the technical documentation of the AI system prepared and regularly updated?	- Up-to date technical documentation	The technical documentation of a AI system must be prepared and kept up-to-date before the system is made available.	- Technical documentation of the AI system including version control (outdated versions of the documentation are preserved and changes to the documentation are tracked) - Defined frequency or times for updating the technical documentation	Article 11(1)	ISO/IEC/IEEE 15289:2019
3.2	To what extent is the technical documentation of an AI system compliant with the requirements outlined in the catalogue?	- Technical documentation that complies with the requirements in this catalogue - Documentation that includes minimal elements listed in Annex IV	The technical documentation must be compliant with requirements outlined in this catalogue and provide relevant authorities with comprehensive information to assess the system's compliance. The documentation must include the minimal elements listed in Annex IV or in the case of SMEs comparable documentation that serves the same purpose, unless the relevant authority deems it unsuitable.	- Technical documentation of the AI system in an organised and clearly presented format	Article 11(1)	ISO/IEC/IEEE 15289:2020
3.3	To what extent does the technical documentation contain all the required information?	- A single technical documentation that includes all required information specified in Annex IV and relevant legal acts listed in Annex II, section A	A single technical documentation containing all the required information, as specified in Annex IV and information of relevant legal acts listed in Annex II, section A, must be created when a AI system associated with a product under those acts is placed on the market or put into service.	- Policies that prevent multiple technical documentations	Article 11(2)	ISO/IEC/IEEE 15289:2021
4.	**Record-keeping**					
4.1	To what extent has the AI system the technical capability to automatically record events (logs) throughout the entire lifespan?	- Automatically recorded events (logs) throughout the system's entire lifespan	AI systems must have the technical capability to automatically record events (logs) throughout the entire lifespan of the system. The logs must be recorded in a standardised format, with clear identification of the time, date and participants involved in the activity or event.	- Documentation that describes how the AI system was designed and configured to capture and store logs automatically. - Sample of the log files	Article 12(1)	ISO 27001, ISO/IEC 27002:2022, ISO 15489-1:2016
4.2	To what extent does the AI system have the capability to automatically record relevant events?	- Logs that capture all relevant events and facilitate a post-market monitoring	These logs must be able to capture events that are relevant to identifying potential risks like threats to the health, safety or to fundamental rights of persons, facilitating post-market monitoring and monitoring the system's operation.	- Lists of specific events including explanations of why those events are relevant to health, safety, or fundamental rights of persons - Sample of the log files	Article 12(2), Article 65(1), (Article 61), Article 29(4)	ISO 27001, ISO/IEC 27002:2022, ISO 15489-1:2016
5.	**Transparency and provision of information to user**					
5.1	To what extent is the AI system's operation transparent to comply with relevant obligations of providers?	- Transparent design that enables users to understand the system and how to appropriately use it	AI systems must be developed so that their operation is adequately transparent to comply with the relevant obligations, as specified in the worksheet *Obligations of providers*. Users must be able to understand the AI operation of the system and based on that knowledge be able to use it.	- Diagram of system components - List of possible data inputs - List of expected outputs - Used algorithms or models - Documented insights of the decision-making process	Article 13(1)	IEC/IEEE 82079-1:2019, ISO/IEC/IEEE 26514:2022
5.2	To what extent are the instructions concise, complete, accurate and provided in a suitable format?	- Concise instructions for use in a suitable format	The instructions for use for AI systems must be concise, complete, accurate and must be provided in a suitable digital or other format. These instructions must provide users with information that is relevant, accessible, simple to comprehend and free from unnecessary technical language. Visual aids, diagrams and practical examples may be used.	- Printed instructions for use - Digital documents like PDFs - Online user guides, tutorials, training materials - Embedded help systems	Article 13(2)	IEC/IEEE 82079-1:2019, ISO/IEC/IEEE 26514:2022
	-		The information specified by point 5.2 must include the following aspects:		Article 13(3)	
5.3	Is the providers contact information included in the instructions?	- Included identity and contact information of the provider	The provider's identity and contact information and, if applicable, of the authorised representative must be included in the instructions for use.	see possible proof of requirement 5.2	Article 13(3)	IEC/IEEE 82079-1:2019, ISO/IEC/IEEE 26514:2022
5.4	Are the capabilities, characteristics and limitations of performance of the AI system included in the instructions?	- Included characteristics, performance and limitations - Included properties of chapter 7 - Included foreseeable risks of chapter 1.2	Information regarding the characteristics, performance and limitations of an AI system, including its intended purpose (including the particular geographical, functional or behavioural setting), level of accuracy (including its metrics, robustness and cybersecurity referred to in the chapter 7) and any known or foreseeable risks to health, safety or fundamental rights (point 1.2) must be included in the instructions for use. It also must include, when appropriate, the behaviour of the system towards specific persons or groups, specifications for input data, validation and testing datasets and the expected output	see possible proof of requirement 5.2	Article 13(3)	IEC/IEEE 82079-1:2019, ISO/IEC/IEEE 26514:2022
5.5	Are the determined changes of the AI system included in the instructions?	- Included possible changes and performance	The changes to the AI system and predetermined performance during the initial conformity assessment must be communicated by the provider.	see possible proof of requirement 5.2	Article 13(3)	IEC/IEEE 82079-1:2019, ISO/IEC/IEEE 26514:2022
5.6	Are the measures regarding human oversight and technical measures for assistance of users included in the instructions?	- Included human oversight measures - Included technical measures for interpreting the output of the system	Measures to ensure human oversight of the AI system, as specified in chapter 6 and details about the technical measures that are used to assist users in interpreting the outputs of the AI system must be included in the instructions for use.	see possible proof of requirement 5.2	Article 13(3)	IEC/IEEE 82079-1:2019, ISO/IEC/IEEE 26514:2022
5.7	Are the computational/hardware resources, anticipated lifetime and frequency of maintenance included in the instructions?	- Included hardware requirements - Included excepted lifetime and care measures	Information regarding the required computational and hardware resources, the anticipated lifetime of the AI system and the frequency of any mandatory maintenance and care measures must be included in the instructions for use.	see possible proof of requirement 5.2	Article 13(3)	IEC/IEEE 82079-1:2019, ISO/IEC/IEEE 26514:2022
5.8	Is a description regarding the management of logs included in the instructions?	- Included description of the log management	A description of the tool that enables users to appropriately collect, store and interpret relevant logs within the AI system must be included in the instructions for use.	see possible proof of requirement 5.2	Article 13(3)	IEC/IEEE 82079-1:2019, ISO/IEC/IEEE 26514:2022

6.	**Human oversight**					
6.1	To what extent are suitable tools for human-machine interaction integrated?	- Appropriate system design that enables human supervision	The design and development of AI systems must include suitable tools for human machine interaction to ensure that natural persons can effectively supervise AI systems during the operation.	- Documentation or other representation of the user interface design showing clear instructions, display of information and appropriate system feedback - Defined human-centered design principles - Testimonials or survey results - User feedback - Online user guides, tutorials, training materials	Article 14(1)	ISO 9241-220:2019
6.2	To what extent is human oversight included to minimise risk?	- Appropriate system design that avoids/reduces risks by human oversight	Human oversight must be implemented to avoid or reduce risks to health, safety or fundamental rights that may arise from using AI systems as intended or under reasonably predictable misuse.	- Risk assessment procedures that identify and document potential risks to health, safety, and fundamental rights - Negative testing (How does the system react to unusual inputs?)	Article 14(2)	ISO 9241-220:2019
6.3	To what extent are human oversight measures integrated into the AI system by the provider or identified by the provider in order to be implemented by the user?	- Clarity about the responsibility of supervision between user and provider	Human oversight must be ensured through measures built into the AI system by the provider or measures identified by the provider and implemented by the user before the system is put into service.	- Defined human-centered design principles - Usability testing reports - Operating guidelines for the user	Article 14(3)	ISO 9241-220:2019
6.4	To what extent are the users of the AI system able to understand its properties and to intervene in the operation?	- Provided information to understand the system's operation and interpret the output - Possibility to stop the system	Referring to the points 6.1 to 6.3, the AI system must be provided to the user in a way that enables them to understand its capabilities and limitations and monitor its operation. Users must be aware of the possibility of automation bias and be able to correctly interpret the system's output. Further they must be able to override or invert the output, intervene in the system's operation or stop it using a "stop" mechanism or similar method.	- Online user guides, tutorials, training materials - System output interpretation guidelines including the potential for automation bias - Override or reversal procedures - Intervention and system stoppage procedures	Article 14(4)	ISO 9241-220:2019
7.	**Accuracy, robustness and cybersecurity**					
7.1	To what extent does the AI system maintain consistent performance throughout its lifetime?	- Resilient system design - Consistent performance over the lifetime	AI systems must be created to achieve a proper level of accuracy, robustness and cybersecurity for their intended application. In addition, the AI system must maintain consistent performance in these areas throughout its lifetime, even if unexpected inputs, exceptions, errors or variations in the operating environment occur.	- List of possible data inputs - List of expected outputs - Documented testing procedure - Correctness testing (Do the outputs align with the expected behaviour?) - Precision testing (Do the outputs have the required level of accuracy?) - Negative testing (How does the system react to unusual inputs?) - Documented security controls of ISO/IEC 27032 (including e.g. policies, secure handling of web sessions, input validation, secure web page scripting)	Article 15(1)	ISO/IEC 25010:2011, ISO/IEC 27032
7.2	Do the instructions for use include accuracy levels and relevant metrics?	- Included accuracy levels in the instructions for use	Instructions for using the AI systems must include a declaration of their accuracy levels and relevant accuracy metrics.	- Documented instructions for use including levels of accuracy and relevant accuracy metrics	Article 15(2)	ISO/IEC 25010:2011
7.3	To what extent is the AI system designed to resist errors, faults or contradictions?	- Appropriate design to avoid inconsistencies	AI systems must be designed to resist errors, faults or contradictions that may occur within the system or its operating environment, especially as a result of their engagement with natural persons or systems.	- Documented fault-tolerant design principles (e.g. error detection, correction/failover mechanisms) - Correctness testing (Do the outputs align with the expected behaviour?) - Precision testing (Do the outputs have the required level of accuracy?) - Negative testing (How does the system react to unusual inputs?	Article 15(3)	ISO/IEC 25010:2011
7.4	To what extent are technical redundancy solutions used?	- Integrated robustness measures	To ensure the robustness of AI systems, technical redundancy solutions may be used, such as backup or fail-safe plans.	- Documentation that outlines the technical redundancy solutions (e.g. architectural diagrams, configuration details) - Redundancy plans (what system components need redundancy solutions) - Fail-Safe plans - Backup plans	Article 15(3)	ISO 27001, ISO/IEC 27002:2022

7.5	To what extent is the AI system designed to minimise the risk of biased outputs influencing input for future operations ?	- Appropriate design to avoid biased outputs for further learning - Suitable mitigation measures	AI systems that proceed to learn after being put into service must be designed to minimise the risk of biased outputs impacting input for future operations (feedback loops). Appropriate mitigation measures must be taken to address this issue.	- Bias monitoring - Data pre-processing techniques (including minimising biases in training data and using diverse datasets) - Bias mitigation measures (e.g. adjustments of algorithms, model calibration)	Article 15(3)	ISO/IEC 25010:2011
7.6	To what extent is the AI system designed to resist unauthorised attempts by third parties?	- Appropriate authorisation measures	AI systems must be designed to resist unauthorised attempts by third parties to modify their performance or use by exploiting system vulnerabilities.	- Defined access controls (including user authentication, role-based access, permissions management) - Documented encryption techniques - Intrusion detection	Article 15(4)	ISO/IEC 25010,ISO 27001 ISO/IEC 27002:2022
7.7	To what extent is the AI system designed to ensure its cybersecurity?	- Cybersecurity measures suitable to situation and risk	To ensure their cybersecurity, AI systems must have technical solutions that are suitable and proportional to the specific circumstances and risks involved.	- Documented security controls of ISO/IEC 27032 (including e.g. policies, secure handling of web sessions, input validation, secure web page scripting)	Article 15(4)	ISO/IEC 27032
7.8	Where relevant, to what extent are technical measures included to address the specific vulnerabilities of AI systems?	- Appropriate design to prevent vulnerabilities	The technical measures to address the specific vulnerabilities of AI systems must be included where relevant, such as measures to prevent and control attacks aimed at manipulating the training dataset (data poisoning), input designed to deceive the model (adversarial examples) or model flaws.	- Vulnerability assessment (including an inventory of identified vulnerabilities) - Documented encryption techniques - Intrusion detection	Article 15(4)	ISO/IEC 25010:2011, ISO/IEC 27032

Artificial Intelligence Act - Assessment

Obligations of providers

Chapter ID	Control question	Requirement objective	Requirement description	Possible proof for fulfilment	Reference to AI Act	Reference to other standards or regulations
8.	**Quality management**					
8.1	To what extend is a quality management system in place that is appropriate for the provider's organisational size?	- Established quality management system appropriate to the organisation's size	Providers of high-risk AI systems must have a quality management system that complies with the aspects of the requirements catalogue and must be implemented in a manner consistent with the provider's organisational size. At least the following elements must be included in the quality management system's structured documentation in the form of written policies, processes, instructions:	- ISO 9001 certification - Documentation that includes the policies and procedures that are in place to ensure the effective management of quality	Article 17(1), Article 17(2)	ISO 9001:2015
8.2	To what extent does the AI system's regulatory compliance strategy address conformity assessment procedures and management of changes?	- Included compliance strategy - Included conformity assessment procedures	a strategy for compliance that addresses conformity assessment procedures and how to manage changes made to the AI system	- Defined regulatory compliance strategy - Documented conformity assessment procedures - Documentation of changes to the system (including design/development changes, review of results, authorisation of changes and	Article 17(1)	ISO 9001:2015
8.3	To what extent are procedures used for the design, design control, design verification, development, quality control and quality assurance of the AI systems?	- Included approaches for design, design control, design verification of the AI system	specific techniques and processes must be employed for its design, design control and design verification, as well as its development, quality control and quality assurance	- Defined intended results - Documented design methodologies, design control procedures and design verification activities - Documented software development methodologies including coding standards and version control - Reviews whether the outcomes of design and development processes fulfil the specified requirements	Article 17(1)	ISO 9001:2015
8.4	To what extend are procedures used for examination, testing and validation?	- Included testing procedures in the AI systems lifecycle - Defined frequency of testing	examination, testing and validation processes to be performed before, during, and after the development of a AI system, as well as the frequency with which they must be performed	- Test and validation plan - Documented test and validation procedures - Defined frequency or times for testing - A catalogue of testing criteria - Documented testing procedure	Article 17(1)	ISO 9001:2015, ISO/IEC/IEEE 29119-2:2021
8.5	To what extent are technical specifications and standards used and to what extend are mechanisms applied to conform with the risk management requirements?	- Included technical specifications - Included measures to comply with risk management requirements	technical specifications, including standards, to be used and when the appropriate uniform standards are not fully applied, the mechanisms to guarantee that the AI system conforms with requirements in chapter 1	- Defined policies and procedures regarding the management of risks - Documented risk mitigation and control measures - Decision process to whether avoid, reduce, transfer or accept risks	Article 17(1)	ISO 9001:2015, ISO 31000, ISO/IEC 27005:2022
8.6	To what extend are systems and procedures used for the data management?	- Included procedures for data management before the deployment	systems and procedures for data management, including data collection, analysis, labelling, storage, filtration, mining, aggregation, retention and any other data operation performed before the deployment of AI systems	- Documented data management procedures	Article 17(1)	ISO 9001:2015
8.7	To what extend are the procedures of the risk management included in the quality management system?	- Included risk management processes	the risk management system referred to in chapter 1	- see possible proof chapter 1	Article 17(1)	ISO 9001:2015, ISO 31000, ISO/IEC 27005:2022
8.8	To what extend are the procedures of the post-market monitoring system in place?	- Included post-market monitoring procedures	the post-market monitoring system referred to in chapter 15	- see possible proof chapter 15	Article 17(1)	ISO 9001:2015, ISO 27001, ISO/IEC 27002:2022
8.9	To what extend are procedures for the reporting of serious incidents included in the quality management system?	- Included incident reporting processes	serious incident reporting processes referred to in chapter 16	- see possible proof chapter 16	Article 17(1)	ISO 9001:2015, ISO 27001, ISO/IEC 27002:2022
8.10	To what extend is the management of communication to relevant authorities or other interested parties in place?	- Included communication procedures to relevant authorities	the management of communication with national competent authorities, including sectoral ones, providing data access, notified bodies, other operators, customers or other interested parties	- Documented communication policies and procedures - Defined roles für the communication with the mentioned parties - Documented enablement of data access for the mentioned parties	Article 17(1)	ISO 9001:2015

8.11	To what extend are procedures for keeping track of all necessary information in place?	- Included record keeping processes	processes for keeping track of all necessary information and documentation	- see possible prof chapter 10	Article 17(1)	ISO 9001:2015
8.12	To what extend is a resource management in place?	- Included resource management	resource management, including measures related to ensuring the security of supply	- Resource management plan - Resource identification catalogue (e.g. needed hardware, materials, tools, human resources) - Contingency plan (including actions to be taken in the event of a resource shortage or	Article 17(1)	ISO 9001:2015
8.13	To what extend is a framework for accountability in place?	- Included accountability framework	a framework for accountability describing the duties of the management and other personnel with reference to each of the previous items in this section	- Defined organisational structure and roles - Documented responsibilities of management and other staff - Training materials for assigned responsibilities	Article 17(1)	ISO 9001:2015
			Aspects mentioned in points 9.2 - 9.13 **may** be included in the quality management systems required by AI system providers who are subject to requirements for quality management systems under applicable sectorial Union legislation.		Article 17(2)	
9.	**Conformity Assessment**					
			The internal control-based conformity assessment procedure includes the points 2 to 4.		Article 19(1), Article 43(2), ANNEX VI	
9.1	Does the quality management system meet the standards of chapter 9?	- Quality management system according to chapter 8	The provider ensures that the quality management system established meets the standards set in chapter 8.	- Records of the conformity assessment including the date of the assessment, identification of the input and documented assessment process, resulting process profiles - see - Record proof chapter 8	Article 19(1), Article 43(2), ANNEX VI	ISO/IEC 33002:2015, 'ISO 9001:2015
9.2	Does the technical documentation conform with the specific requirements for the system?	- Technical documentation that conforms with the requirements for the system	The provider reviews the technical documentation to evaluate whether the AI system conforms with the necessary essential requirements specified in the requirements of the worksheet *Requirements for the system*.	- Records of the conformity assessment including the date of the assessment, identification of the input and documented assessment process, resulting process profiles - see possible '- Re chapter 3	Article 19(1), Article 43(2), ANNEX VI	ISO/IEC 33002:2015
9.3	Do the post-market monitoring system, design and development processes conform to the technical documentation?	- Technical documentation that is consistent with the post-market monitoring, design and development processes	The provider confirms that the AI system's post-market monitoring, design and development processes conform to the technical documentation.	- Records of the conformity assessment including the date of the assessment, identification of the input and documented assessment process, resulting process profiles - see possible proof chapter 14	Article 19(1), Article 43(2), ANNEX VI	ISO/IEC 33002:2015
10.	**Logs and document retention**					
10.1	Are the logs generated by the AI system kept for at least six months?	- Log retention for at least six months	Providers are required to maintain the logs generated by their AI systems that are under their control due to a contractual agreement with the user or by law. These logs must be kept for a minimum of six months, unless specified differently by applicable Union or national laws, especially laws related to personal data protection.	- Sample of the log files - Defined storage environment - Defined records retention schedule (at least six months)	Article 20(1)	ISO 27001, ISO/IEC 27002:2022, ISO 15489-1:2016
10.2	To what extend are procedures in place to ensure that the relevant documentation is kept for at least 10 years?	- Documentation retention for at least ten years	The provider of an AI system is required to maintain the technical documentation, the EU declaration of conformity, documentation of the quality management, documentation of approved changes and other decisions of notified bodies for a period of 10 years after the system has been placed on the market or put into service. During this time, the provider must make this information available to national competent authorities upon request.	- Technical documentation, EU declaration of conformity, documentation of the quality management, documentation of approved changes - Defined storage environment - Defined documentation retention schedule (at least ten months)	Article 18(a), Article 18(1)	ISO/IEC 27001:2022

11.	**EU declaration of conformity**					
11.1	Was a EU declaration of conformity signed?	- Signed EU declaration	The provider is responsible for creating a signed EU declaration of conformity in written or electronically form for each AI system. It must be kept for ten years following the system's market introduction or start of service. The EU declaration of conformity must identify the specific AI system for which it is created and a copy of the declaration must be submitted to national competent authorities upon request.	- A copy of the signed EU declaration of conformity	Article 48(1)	EU Declaration of Conformity - Template
11.2	Does the EU declaration of conformity include the aspects of Annex V?	- Declaration of conformity including all information listed in Annex 5	The EU declaration of conformity must confirm that the AI system meets the necessary requirements mentioned in worksheet *Requirements for the system*. The declaration must include the details as specified in Annex V and it must be translated into a language that is easily understandable by the national competent authorities in which the AI system is available.	- see possible proof chapter 11.1	Article 48(2)	-
11.3	In case of multiple EU declarations of conformity, are these aggregated in one single declaration?	- A single document that sums up all declarations	In cases where systems are subject to multiple Union harmonisation legislations that demand an EU declaration of conformity, a single declaration must be created for all relevant Union regulations. The EU declaration of conformity must include all the information required to identify the Union harmonisation legislation to which it applies.	- see possible proof chapter 11.1	Article 48(3)	-
11.4	Is the provider aware of the responsibility for compliance with the relevant requirements for the system?	- Clarity about the responsibility for the compliance - Up-to date declaration of conformity	The provider is responsible for ensuring that the AI system complies with the requirements outlined in worksheet *Requirements for the system*, by creating the EU declaration of conformity. The provider must keep the declaration updated whenever necessary.	- see possible proof chapter 11.1 - Defined frequency or times for updating the declaration of conformity	Article 48(4)	-
12.	**CE marking of conformity**					
12.1	Does the CE marking comply with Article 30 of Regulation (EC) No 765/2008?	- CE marking according to Article 30 of Regulation (EC) No 765/2008	The CE marking that indicates the conformity with this regulation must comply with the fundamental principles stated in Article 30 of Regulation (EC) No 765/2008.	- Clearly visible CE marking on the product, packaging or accompanying documents	Article 49(1)	Regulation (EC) No 765/2008 Article 30
12.2	Is the CE marking clearly attached to the system, packaging or accompanying documents?	- Clearly visible CE marking on systems, packaging, or documentation	The CE marking must be attached clearly visible, permanent and readable. If this is not practicable the marking can be displayed on the packaging or accompanying documents.	- see possible proof chapter 12.1	Article 49(2)	Regulation (EC) No 765/2008 Article 30
12.3	Does the CE marking include the identification number of the notified body?	- CE marking with identification number of the notified body	The CE marking of conformity must, if applicable, include the identification number of the notified body responsible for carrying out the conformity assessment procedures described in Article 43. This identification number must also be included in any documentation claiming that the AI system fulfils the CE marking requirements.	- Clearly visible CE marking on the product, packaging or accompanying documents and if applicable the identification number of the notified body responsible for carrying out the conformity assessment procedures	Article 49(3)	Regulation (EC) No 765/2008 Article 30
12.4	Is the name, registered trade name or trademark and contact information labelled on the AI system, packaging or documentation?	- Correctly labelled AI system	Providers of AI systems must provide their name, registered trade name or trademark, as well as their contact information on the system. If direct labelling on the AI system is not practicable, this information must be included on its packaging, or the accompanying documentation.	- Name, trade name, trademark and address on the AI system, packaging or accompanying documents	Article 16(aa)	Regulation (EC) No 765/2008 Article 30
13.	**Transparency obligations**					
13.1	To what extent is the AI system designed in a way that is clear to natural persons that they are interacting with an AI system?	- Transparency for users that clarifies the interaction to AI technology	Providers are required to ensure that AI systems which are meant to interact with natural persons must be designed and developed in a manner that it is clear to them that they are interacting with an AI system. This requirement does not apply if it is already obvious to a reasonably well-informed person, considering the circumstances and context of use. This does not affect the other requirements in this catalogue and other transparency obligations that users of AI systems may have under existing Union or national laws.	- User interface with focus on "appropriateness recognisability" - User documentation that clearly highlights that the user is interacting with an AI system - Visual or auditory cues that make it obvious the user is interacting with an AI system	Article 52(1), Article 52(3a.), Article 52(4)	ISO/IEC 25010:2011
13.2	Are persons aware that they are interacting with AI since the first contact with the AI system?	- Transparency since the first use of the AI system	Natural Persons must be informed no later than the moment of their first engagement or exposure with the system that they are interacting with AI technology, unless it is already obvious to a reasonably well-informed person, considering the circumstances and context of use.	- see possible proof chapter 13.1 - User consent to interact with an AI system at the first interaction	Article 52(1), Article 52(3a.), Article 52(4)	ISO/IEC 25010:2011

14.	**Post-market monitoring and incident reporting**					
14.1	To what extent is a post-market monitoring system developed?	- Established monitoring system suitable for the level of risk	After the AI system is placed on the market, providers are required to develop a system for monitoring and keeping track of it. The scope of the post-market monitoring system must be appropriate to the level of risk posed by the AI system.	- Documented scope of the AI system (including intended use, potential impact, data sensitivity) - Defined metrics to measure the performance and security level of the AI system - Documented monitoring procedures	Article 61(1)	ISO/IEC 27001:2022, ISO/IEC 27002:2022
14.2	To what extent does the post-market monitoring system collect the relevant data throughout the life cycle of the system?	- Enabled analysis of the performance of the AI system throughout its life cycle	The post-market monitoring system must collect, document and analyse data on the performance of the AI system over the course its life cycle. The data may come from users or other sources. However, it must not cover sensitive operational data of users that are law enforcement authorities. The purpose is to evaluate compliance with the requirements set out in the worksheet *Requirements of the system*.	- Data collection mechanisms (e.g. system logs, performance metrics, user feedback, incident reports) - Documentation process (e.g. records, reports, databases) - Data analysis mechanisms (e.g. data mining, pattern/trend recognition, statistic methods)	Article 61(2)	ISO/IEC 27001:2022, ISO/IEC 27002:2022
14.3	To what extend is a post-market monitoring plan in place?	- Established post-market monitoring plan including the technical documentation of Annex IV	The post-market monitoring system must be based on a post-market monitoring plan, which has to be included in the technical documentation referred to in Annex IV.	- Post market monitoring plan including technical documentation referred to in Annex IV - Defined frequency or times for updating the post market monitoring plan	Article 61(3)	ISO/IEC 27001:2022, ISO/IEC 27002:2022
14.4	To what extend are procedures for the reporting of serious incidents in place?	- Established incident reporting procedures	The provider must report any serious incidents involving the AI system to the market surveillance authorities of the Member State in which the incident occurred. The provider is required to report the incident as soon as they confirm a connection between the AI system and the incident, or if there is a reasonable possibility of such a connection. This has to be done not later than 15 days after becoming aware of the incident.	- Incident reporting process - Incident response process - Defined reporting timeline (not later than 15 days) - Defined communication procedure and channel to market surveillance authorities - Incident reports (including incident date, findings, actions taken, communication to market surveillance authorities)	Article 62(1)	ISO/IEC 27001:2022, ISO/IEC 27002:2022

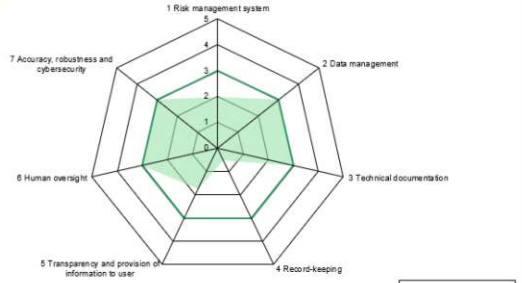
Maturity level	Description of implementation	Date of assessment	Responsible department	Contact	Future planned measures	Due Date	Further information
2							
1							
0							
4							
5							
3							

Artificial Intelligence Act - Assessment
Results

Requirements for the system

Maturity level (capped): 2,39

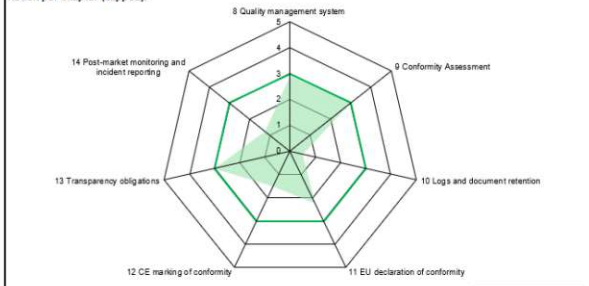
Result per chapter (capped):



Obligations of providers

Maturity level (capped): 2,31

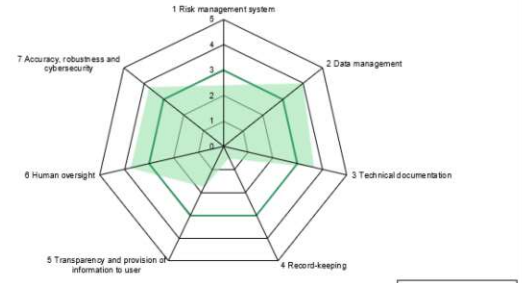
Result per chapter (capped):



Total maturity Level (capped)
2,35

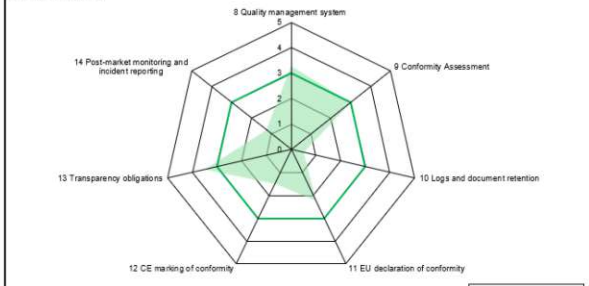
Maturity level: 2,92 | Target maturity level: 3,00

Result per chapter:



Maturity level: 2,50 | Target maturity level: 3,00

Result per chapter:



Total maturity Level
2,71

ID	Control Question	Maturity level	Maturity level (capped)
1.1	To what extent is a risk management system implemented and documented?	2	2
1.2	To what extent are the procedures of the risk management system integrated into the system's life cycle?	1	1
1.3	To what extent have the risk management measures taken into account the potential effects resulting from the combined application of the requirements?	0	0
1.4	To what extent do the risk management measures ensure that residual risks of the system are considered acceptable?	4	3
1.5	To what extent has the AI system been tested to ensure that it operates as intended?	5	3
1.6	To what extent has the AI system been tested at different steps of development?	3	3
1.7	To what extent does the risk management system consider the potential impact on individuals under the age of 18?	2	2
2.1	To what extent are appropriate data governance and management practices implemented to ensure the quality of datasets?	5	3
2.2	To what extent meet the datasets the requirements of being appropriate, error-free and comprehensive?	5	3
2.3	To what extent do the datasets take into account the characteristics of the specific geographical, behavioural or functional context of the AI system?	3	3
2.4	To what extent are fundamental rights and freedoms safeguarded when processing special categories of personal data?	4	3
3.1	To what extent is the technical documentation of the AI system prepared and regularly updated?	3	3
3.2	To what extent is the technical documentation of an AI system compliant with the requirements outlined in the catalogue?	4	3
3.3	To what extent does the technical documentation contain all the required information?	4	3
4.1	To what extent has the AI system the technical capability to automatically record events (logs) throughout the entire lifespan?	0	0
4.2	To what extent does the AI system have the capability to automatically record relevant events?	1	1
5.1	To what extent is the AI system's operation transparent to comply with relevant obligations of providers?	2	2
5.2	To what extent are the instructions concise, complete, accurate and provided in a suitable format?	3	3
5.3	Is the providers contact information included in the instructions?	3	3
5.4	Are the capabilities, characteristics and limitations of performance of the AI system included in the instructions?	0	0
5.5	Are the determined changes of the AI system included in the instructions?	0	0

ID	Control question	Maturity level	Maturity level (capped)
8.1	To what extend is a quality management system in place that is appropriate for the provider's organisational size?	3	3
8.2	To what extent does the AI system's regulatory compliance strategy address conformity assessment procedures and management of changes?	2	2
8.3	To what extent are procedures used for the design, design control, design verification, development, quality control and quality assurance of the AI systems?	3	3
8.4	To what extend are procedures used for examination, testing and validation?	4	3
8.5	To what extent are technical specifications and standards used and to what extend are mechanisms applied to conform with the risk management requirements?	3	3
8.6	To what extend are systems and procedures used for the data management?	3	3
8.7	To what extend are the procedures of the risk management included in the quality management system?	5	3
8.8	To what extend are the procedures of the post-market monitoring system in place?	3	3
8.9	To what extend are procedures for the reporting of serious incidents included in the quality management system?	3	3
8.10	To what extend is the management of communication to relevant authorities or other interested parties in place?	3	3
8.11	To what extend are procedures for keeping track of all necessary information in place?	5	3
8.12	To what extend is a resource management in place?	3	3
8.13	To what extend is a framework for accountability in place?	3	3
9.1	Does the quality management system meet the standards of chapter 3?	3	3
9.2	Does the technical documentation conform with the specific requirements for the system?	3	3
9.3	Do the post-market monitoring system, design and development processes conform to the technical documentation?	3	3
10.1	Are the logs generated by the AI system kept for at least six months?	0	0
10.2	To what extend are procedures in place to ensure that the relevant documentation is kept for at least 10 years?	1	1
11.1	Was a EU declaration of conformity signed?	3	3
11.2	Does the EU declaration of conformity include the aspects of Annex V?	3	3
11.3	In case of multiple EU declarations of conformity, are these aggregated in one single declaration?	0	0

5.5	Are the determined changes of the AI system included in the instructions?	0	0
5.6	Are the measures regarding human oversight and technical measures for assistance of users included in the instructions?	3	3
5.7	Are the computational/hardware resources, anticipated lifetime and frequency of maintenance included in the instructions?	3	3
5.8	Is a description regarding the management of logs included in the instructions?	0	0
6.1	To what extent are suitable tools for human-machine interaction integrated?	4	3
6.2	To what extent is human oversight included to minimise risk?	3	3
6.3	To what extent are human oversight measures integrated into the AI system by the provider or identified by the provider in order to be implemented by the user?	5	3
6.4	To what extent are the users of the AI system able to understand its properties and to intervene in the operation?	3	3
7.1	To what extent does the AI system maintain consistent performance throughout its lifetime?	4	3
7.2	Do the instructions for use include accuracy levels and relevant metrics?	3	3
7.3	To what extent is the AI system designed to resist errors, faults or contradictions?	4	3
7.4	To what extent are technical redundancy solutions used?	3	3
7.5	To what extent is the AI system designed to minimise the risk of biased outputs influencing input for future operations ?	4	3
7.6	To what extent is the AI system designed to resist unauthorised attempts by third parties?	3	3
7.7	To what extent is the AI system designed to ensure its cybersecurity?	4	3
7.8	Where relevant, to what extent are technical measures included to address the specific vulnerabilities of AI systems?	5	3

11.3	In case of multiple EU declarations of conformity, are these aggregated in one single declaration?	0	0
11.4	Is the provider aware of the responsibility for compliance with the relevant requirements for the system?	3	3
12.1	Does the CE marking comply with Article 30 of Regulation (EC) No 765/2008?	0	0
12.2	Is the CE marking clearly attached to the system, packaging or accompanying documents?	3	3
12.3	Does the CE marking include the identification number of the notified body?	0	0
12.4	Is the name, registered trade name or trademark and contact information labelled on the AI system, packaging or documentation?	3	3
13.1	To what extent is the AI system designed in a way that is clear to natural persons that they are interacting with an AI system?	4	3
13.2	Are persons aware that they are interacting with AI since the first contact with the AI system?	3	3
14.1	To what extent is a post-market monitoring system developed?	2	2
14.2	To what extent does the post-market monitoring system collect the relevant data throughout the life cycle of the system?	0	0
14.3	To what extend is a post-market monitoring plan in place?	1	1
14.4	To what extend are procedures for the reporting of serious incidents in place?	2	2

Artificial Intelligence Act - Assessment
Heatmap

	Requirements for the system							
	X.1	X.2	X.3	X.4	X.5	X.6	X.7	X.8
1 Risk management	2	1	0	4	5	3	2	
2 Data management	5	5	3	4				
3 Technical documentation	3	4	4					
4 Record-keeping	0	1						
5 Transparency and provision of information to user	2	3	3	0	0	3	3	0
6 Human oversight	4	3	5	3				
7 Accuracy, robustness and cybersecurity	4	3	4	3	4	3	4	5

	Obligations of providers												
	X.1	X.2	X.3	X.4	X.5	X.6	X.7	X.8	X.9	X.10	X.11	X.12	X.13
8 Quality management	3	2	3	4	3	3	5	5	3	3	5	3	3
9 Conformity Assessment	3	3	3										
10 Logs and document retention	0	1											
11 EU declaration of conformity	3	3	0	3									
12 CE marking of conformity	0	3	0	3									
13 Transparency obligations	4	3											
14 Post-market monitoring and incident reporting	2	0	1	2									

Appendix B

Appendix B contains the transcripts of the six interviews translated into English by the author as well as the individual SUS scores of the interviewees.

Interviewee 1: Transcript and SUS score

Interviewer
The first question regarding perceived usefulness would be: In what way do you think the requirements catalogue can support AI system providers in evaluating compliance with the requirements of the AI Act?

Interviewee 1
Well, I believe the value of the catalogue lies in breaking down the law into tasks that can be assigned to individuals and that's what companies need for all requirements. It's pointless to have requirements that are only understandable to specific departments like the legal or compliance departments or where only individuals have a mental model of what needs to be done. Companies need to break down the tasks based on their organisation and hierarchy and I think the catalogue will support that. So, I have separate and independent requirements that I can assign to individuals and in the end, I can demonstrate whether those individuals have done what needs to be done when I'm being audited.

Interviewer
Alright, thank you. The next question would be: What specific benefits do you expect from using the requirements catalogue in a compliance audit?

Interviewee 1
The pre-defined requirements catalogue saves me from having to read through all the requirements myself and filter out which ones apply to me as a provider or manufacturer and which ones are only relevant to other authorities. The second benefit is that it relieves me of the task of reformulating requirements into understandable language that aligns with the company's reality. Or to put it differently, the specific benefit I expect is that it saves me time and enables the responsible individuals within the company to better and more quickly understand their responsibilities.

Interviewer
Great, thank you. Then, one final question regarding perceived usefulness: How do you overall assess the usefulness of the requirements catalogue in the compliance audit of the AI Act?

Interviewee 1
I would say the usefulness is an eight out of ten for me. Overall, it is very useful.

Interviewer
Now, the next set of questions focuses on usability, so: How do you evaluate the usability when filling out the requirements catalogue?

Interviewee 1
I believe the usability is quite straightforward. I have seen many requirements catalogues and one of the values is that they all have a similar structure. If you constantly reinvent the wheel and build something completely different, it might be more fitting for a specific case, but it reduces the recognition value. Using a maturity model that is already used within the company makes it easy to use. Furthermore, people are usually familiar with Excel and can operate it. Well, I take back the statement that everyone can use Excel, but they can certainly handle parts of it.

Interviewer
I understand. Yes, it is expected and they already use it in most cases as part of their everyday work, as Excel is a common tool. The next question focuses on those green worksheets, so how do you assess the clarity and ease of interpretation of the presentation of results?

Interviewee 1
Yes, it's really great. It is also understandable and helpful for management. It can support decision-making on why certain topics require more resources. As for improvement points, I would suggest using the term capped instead of shortened maturity level as it better describes the intention. And perhaps a minor point is that the spider graphs could have a more transparent area to better visualise the underlying lines of each maturity level.

Interviewer
Do you see any challenges in using the requirements catalogue, if any?

Interviewee 1
One challenge is that when I have such a catalogue, compared to the law, I always have to fear whether every requirement is covered by the catalogue. Secondly, with such requirements catalogues, I always fear that the interpretation of the requirements from the law goes beyond what is actually demanded and I have seen that with this catalogue as well. For example, for the proposed implementation of risks, it suggests a certification according to ISO 31000. This could give the impression that I definitely need it. However, not all companies do that and there are good reasons for it. So, the biggest challenge I see is that this catalogue, in its attempt to make the legal requirements more understandable, might go beyond the goal of pure compliance.

Interviewer
That's a good point, but you have to keep in mind that this are only possible evidences and not mandatory. The final question regarding usability is: How do you overall assess the usability of the requirements catalogue for assessing compliance with the AI Act? What factors contribute to this?

Interviewee 1
I also rate the usability as an eight out of ten. I think it contributes to usability that the catalogue has a recognition value and that the law is included and cross-referenced to the chapters, which is an advantage I haven't seen in other requirements catalogues. I think there are quite a few columns that need to be filled in for each requirement. I would prefer fewer columns to reduce the documentation effort. From my perspective, it's better to start with fewer columns and introduce them only when it becomes evident that they are needed, to avoid having empty columns.

Interviewer
Those are very good points. Thank you for your insights, that concludes the questions.

Interviewee 1
You're welcome.

System Usability Scale: Interviewee 1

Statements	0 Strongly disagree	1	2	3	4 Strongly agree
I thought all the relevant requirements of the AI Act were integrated completely.				x	
I thought the requirements were enumerated several times (duplicates).	x				
I thought the requirements were formulated in an understandable and practical way.			x		
I thought that certain requirements were conflicting with each other.	x				
I thought the requirements could be evaluated according to an understandable fulfilment degree.				x	
I thought the requirement catalogue was unstructured and unnecessary complex.	x				
I feel confident that the requirements catalogue will support the development of an AI system that is compliant with the AI Act.					x
I thought further instructions would be necessary to fully use the requirements catalogue.		x			
I thought the calculated results of the maturity level were presented in a clear manner.					x
I thought in the requirements catalogue important aspects for a compliance assessment were missing.	x				

Figure 1: SUS score of interviewee 1

Interviewee 2: Transcript and SUS score

Interviewer
The first question would then be about perceived usefulness. In what way do you think the requirements catalogue can support AI system providers in evaluating compliance with the requirements of the AI Act?

Interviewee 2
Well, I believe that it provides a good way to quickly get a full overview of the current state of compliance within the company, especially for those who haven't done much experience in this area yet. With this current state and the help of features like the heatmap or other functionalities of the catalogue, one can identify areas that require immediate attention. This allows for prioritising different departments or thematic areas and starting with the ones that require the most improvement. By doing so, the maturity level can be increased and the processes can be accordingly implemented. That would be the first point. Furthermore, I can imagine that the requirements catalogue for the AI Act can serve as a basis during an audit itself, providing the auditor with an overview. It includes fields with exemplary evidence and descriptions of the current implementation, where one can write about how it is practiced in their own company. Additionally, it can reference documents that demonstrate compliance, making all necessary evidence readily available.

Interviewer
Thank you for the insights. The next question is: What specific benefits do you expect from using the requirements catalogue in a compliance audit?

Interviewee 2
I believe that having a shared step-by-step tool for auditors and reviewers is very useful. It allows both sides, the auditors or consultants and the AI system providers themselves, to benefit from the requirements catalogue.

Interviewer
The last question about the perceived usefulness would be: How do you overall assess the usefulness of the requirements catalogue in the compliance audit of the AI Act?

Interviewee 2
Yes, overall, I think it is highly useful. I think the main usefulness lies in the preparation phase, as the catalogue proves to be extremely helpful in the journey towards the audit. It can also provide guidance during the actual audit. Although using the requirements catalogue as the sole evidence for AI Act compliance might be challenging to achieve, I believe its usefulness in the preparation phase is very high.

Interviewer
The next questions will focus on usability. How do you evaluate the usability when filling out the requirements catalogue?

Interviewee 2
Overall, it's very good. The catalogue provides a good introduction, describing how to use it in general. The terminology is well explained and it makes sense to be able to switch back to it whenever needed. I think the mapping of chapters between the AI Act and the requirements catalogue is crucial because, in the end, the goal is to comply with the requirements of the AI Act. Therefore, the mapping is absolutely necessary. The user interface is also clear and understandable. While reviewing the requirements catalogue, I noticed one thing. When referring to Future planned measures, it essentially refers to

something that has not yet happened. In practice, it might be more appropriate to use the term Due Date instead of Date of completion to signify the deadline for completion.

Interviewer
Excellent points. I will move on to the next question: How do you assess the clarity and ease of interpretation of the presentation of results?

Interviewee 2
Again, overall, it is very good. It's logical and clear. I think the concept of shortened values is essential. It might be beneficial to highlight these shortened values more prominently. The heatmaps are also clear and understandable and I believe they are very helpful in identifying specific problem areas. One suggestion would be to consider adding an average value per chapter at the end, to provide a high-level view of the overall issues within different thematic areas.

Interviewer
I understand. The next question is: Do you see any challenges in using the requirements catalogue, if any?

Interviewee 2
Actually, there are very few challenges that I can think of. One challenge that might arise, especially since the AI Act is still relatively new, is determining who is responsible for each requirement. However, this is a topic that is difficult to address through the requirements catalogue, as each company has a different structure and responsibilities may lie in various areas. So, no significant challenges come to mind.

Interviewer
The final question would be: How do you overall assess the usability of the requirements catalogue for assessing compliance with the AI Act? What factors contribute to this?

Interviewee 2
Well, as I mentioned, I think it is overall really good and easy to use. Two things immediately stood out to me as positive. First is the mapping, which is crucial for quickly finding relevant sections. The new clustering you implemented in the catalogue is also beneficial because it makes it easier to implement the requirements. It allows for merging or dividing chapters as needed. I believe this approach makes sense, but it relies on having proper mapping in place. The second factor is the presentation of results, especially the heatmaps. They provide a visual representation where it's easy to identify areas that require attention, indicated by the red colour. It allows for a quick assessment of where one needs to delve deeper.

Interviewer
Thank you very much for your insights.

Interviewee 2
You're welcome.

System Usability Scale: Interviewee 2					
	0	1	2	3	4
Statements	**Strongly disagree**				**Strongly agree**
I thought all the relevant requirements of the AI Act were integrated completely.				x	
I thought the requirements were enumerated several times (duplicates).	x				
I thought the requirements were formulated in an understandable and practical way.					x
I thought that certain requirements were conflicting with each other.		x			
I thought the requirements could be evaluated according to an understandable fulfilment degree.					x
I thought the requirement catalogue was unstructured and unnecessary complex.	x				
I feel confident that the requirements catalogue will support the development of an AI system that is compliant with the AI Act.					x
I thought further instructions would be necessary to fully use the requirements catalogue.		x			
I thought the calculated results of the maturity level were presented in a clear manner.					x
I thought in the requirements catalogue important aspects for a compliance assessment were missing.	x				

Figure 2: SUS score of interviewee 2

Interviewee 3: Transcript and SUS score

Interviewer
Thank you for your time. I will now begin with the first question regarding perceived usefulness and that would be: In what way do you think the requirements catalogue can support AI system providers in evaluating compliance with the requirements of the AI Act?

Interviewee 3
In my opinion, the significant value of such a requirements catalogue is the contextualisation of the requirements from the AI Act. It's not just listing the individual requirements, but rather creating connections with other norms or regulations and suggesting possible evidence that help providers actually fulfil the requirements. And on top of that have a mechanism to evaluate the compliance.

Interviewer
Thank you. The next question is: What specific benefits do you expect from using the requirements catalogue in a compliance audit?

Interviewee 3
The requirements catalogue could be very useful in customer projects where the client, for example, a provider of a critical infrastructure AI system, could engage with us as a consulting firm to become compliant with the AI Act. Consequently, we could use this catalogue as a shared tool, providing added value to the client and support a possible audit.

Interviewer
I agree, having a shared tool for communication with the client would certainly be

beneficial. As a concluding question regarding usefulness: How do you overall assess the usefulness of the requirements catalogue in the compliance audit of the AI Act?

Interviewee 3
I see two major benefits in such requirements catalogues. First, it facilitates the tracking of compliance with the requirements and secondly, it allows for contextualisation with other implemented norms and suggested evidence. A well-designed requirements catalogue can serve as a central tool for achieving compliance with the law. It is also important, as in your requirements catalogue, to provide transparency about the current progress and the responsible parties.

Interviewer
Now, let's move on to questions about usability. The first question is: How do you evaluate the usability when filling out the requirements catalogue?

Interviewee 3
In my experience, the main challenge with Excel-based catalogues is not the initial filling out but rather ensuring that they are designed in a way that they can be sustained and reused. They should provide as much necessary information as possible without becoming too extensive, making them impractical for ongoing use. In this regard, I see that you have considered this by keeping the user-filled columns to a manageable extent.

Interviewer
The next question would be: How do you assess the clarity and ease of interpretation of the presentation of results?

Interviewee 3
I generally think the spider graphs are very clear and effective in representing different maturity levels. They are commonly used in various contexts. I also think the presentation of the shortened maturity level and the accompanying explanations on the welcome page are very useful. One improvement suggestion would be to emphasise the presentation of the shortened maturity level more. Additionally, I think the heatmaps are well-designed and easy to understand and I don't have any suggestions for improvement in that regard.

Interviewer
Thank you. The penultimate question is: Do you see any challenges in using the requirements catalogue, if any?

Interviewee 3
As I mentioned earlier, one of the key aspects for me is ensuring that the requirements catalogue is used regularly within the company and not just filled out once. For example, versioning should be considered when naming the files to address this challenge.

Interviewer
If you have nothing else to add, I have one final question. How do you overall assess the usability of the requirements catalogue for assessing compliance with the AI Act? What factors contribute to this?

Interviewee 3
Overall, I consider all the worksheets to be meaningful and well-designed, with the ones starting from Requirements for the System being particularly important. I think you have struck a good balance between providing all the necessary information and keeping it simple in terms of content and ease of use, to ensure that users are not overwhelmed.

Interviewer
If you have nothing else to add, I sincerely thank you for your time.

Interviewee 3
You're welcome.

System Usability Scale: Interviewee 3

Statements	0 Strongly disagree	1	2	3	4 Strongly agree
I thought all the relevant requirements of the AI Act were integrated completely.					x
I thought the requirements were enumerated several times (duplicates).	x				
I thought the requirements were formulated in an understandable and practical way.					x
I thought that certain requirements were conflicting with each other.		x			
I thought the requirements could be evaluated according to an understandable fulfilment degree.					x
I thought the requirement catalogue was unstructured and unnecessary complex.	x				
I feel confident that the requirements catalogue will support the development of an AI system that is compliant with the AI Act.					x
I thought further instructions would be necessary to fully use the requirements catalogue.		x			
I thought the calculated results of the maturity level were presented in a clear manner.					x
I thought in the requirements catalogue important aspects for a compliance assessment were missing.		x			

Figure 3: SUS score of interviewee 3

Interviewee 4: Transcript and SUS score

Interviewer
I will now begin with the seven questions and the first one relates to perceived usefulness. In what way do you think the requirements catalogue can support AI system providers in evaluating compliance with the requirements of the AI Act?

Interviewee 4
It is likely that you can reach your goals much faster because the catalogue guides you and provides support to the providers. They may not have to think as much or search through the legal text but rather can simply work through the requirements step by step in the requirements catalogue. So, I believe the perceived usefulness is indeed present for the providers of AI systems.

Interviewer
Thank you. The second question would be: What specific benefits do you expect from using the requirements catalogue in a compliance audit?

Interviewee 4
I think it is universally applicable. Through the catalogue, I don't have to explain the AI Act to everyone. You have this tool that supports you by providing context and offering tracking and evaluation possibilities. The requirements catalogue is definitely well-structured. Even newcomers are well introduced to the topic and its structure through the welcome page. The colours inform you where action is needed and where the results are presented. So, you are accompanied throughout the entire process and questions regarding terms, for example, are immediately answered in the glossary.

Interviewer
The last question regarding perceived usefulness is: How do you overall assess the usefulness of the requirements catalogue in the compliance audit of the AI Act?

Interviewee 4
Based on my experience with such catalogues, I see the usefulness as very high. As mentioned earlier, there are many advantages. You are faster, can rely on it and essentially have a good summary. Therefore, I consider the usefulness to be very high.

Interviewer
These first three questions focused more on usefulness and the following questions are more about usability. The first question in this regard is: How do you evaluate the usability when filling out the requirements catalogue?

Interviewee 4
Right from the beginning, when you look at just one row in the spreadsheet where you fill in the requirements, it is well-structured. It starts with the requirement itself, followed by all the additional information that may be needed. Where can I find the requirement in the AI Act? Which other standards or norms can help me? What is the actual goal? And so on. You are guided quite well throughout the entire catalogue, which should make you feel capable of completing the filling process. I think the information and the order of information are well chosen. In the Measures column, one could consider renaming it as it represents the current status. And then there are other points that can and must be filled out. They might be interesting or relevant for some requirements, but there could also be clarifications on what the provider must fill in and what is not mandatory. So, it doesn't mean that the provider has to fill in every single field. There will likely be requirements where that's the case, but also requirements where it's not necessary.

Interviewer
Very interesting points, thank you. Now, the next question is: How do you assess the clarity and ease of interpretation of the presentation of results?

Interviewee 4
From my experience, these spider graphs do make sense. They are used in other catalogues for a reason because they provide a nice overview of the individual topics. You can easily see how things are in different areas and get a good overall view at a glance. I personally think the heatmaps are very refreshing, almost innovative because they indicate even more clearly than the spider graphs where my problem areas are. You look at them, see

the red dots and also identify the clusters where there is room for improvement. But of course, with the green areas, you can clearly also see where you are right on track.

Interviewer
Now let's move on to the penultimate question. Do you see any challenges in using the requirements catalogue, if any?

Interviewee 4
There are some minor things that could be adjusted in the requirements catalogue, but they are not necessarily challenges in using it. The AI Act is currently in an early stage, which is inconvenient because there may be updates in the coming years that are not considered in the catalogue. That could be a future challenge that needs to be addressed. So, despite using the catalogue, one should still maintain a necessary level of scepticism and not blindly rely on it. It should be emphasised which version of the AI Act is being referred to.

Interviewer
And finally, the last question: How do you overall assess the usability of the requirements catalogue for assessing compliance with the AI Act? What factors contribute to this?

Interviewee 4
You are well guided from start to finish. It is immediately clear what needs to be done and what doesn't. Personally, I am a fan of colours, so the colour differentiation greatly contributes to usability. When I open the catalogue, I see blue, orange and green colours. At the first moment, when I open this catalogue my question would be what they represent, but then I immediately find the answer at the top on the welcome page. So, I think the use of colours is really meaningful. My second highlight, in line with that, would be the heatmaps. Here, at first glance, you can see how well or poorly positioned you are.

Interviewer
Great, thank you for your time and insights.

Interviewee 4
You're welcome.

System Usability Scale: Interviewee 4

Statements	0 Strongly disagree	1	2	3	4 Strongly agree
I thought all the relevant requirements of the AI Act were integrated completely.				x	
I thought the requirements were enumerated several times (duplicates).		x			
I thought the requirements were formulated in an understandable and practical way.					x
I thought that certain requirements were conflicting with each other.		x			
I thought the requirements could be evaluated according to an understandable fulfilment degree.					x
I thought the requirement catalogue was unstructured and unnecessary complex.	x				
I feel confident that the requirements catalogue will support the development of an AI system that is compliant with the AI Act.					x
I thought further instructions would be necessary to fully use the requirements catalogue.	x				
I thought the calculated results of the maturity level were presented in a clear manner.					x
I thought in the requirements catalogue important aspects for a compliance assessment were missing.	x				

Figure 4: SUS score of interviewee 4

Interviewee 5: Transcript and SUS score

Interviewer
So, I would start with the first question regarding perceived usefulness. In what way do you think the requirements catalogue can support AI system providers in evaluating compliance with the requirements of the AI Act?

Interviewee 5
Ultimately, when the product is completed or the software or system is ready, you can simply fill out this catalogue for your product and have a transparent list of where you performed well or poorly. In the next step, you can improve these areas and then re-evaluate. This means you can continuously use it to enhance your product or system.

Interviewer
So, it's a kind of tool for continuous improvement?

Interviewee 5

Yes, exactly. And once you reach Maturity level 3, you are on the safe side. And if you add any new features, you can simply go through this process again and see if it affects the end result. This way, you can ensure compliance.

Interviewer
The next question would be: What specific benefits do you expect from using the requirements catalogue in a compliance audit?

Interviewee 5
Okay, I would say there are concrete benefits for all parties involved. As a user, my advantage is that the system I use handles my data in a compliant manner and I have

nothing to worry about in that regard. If I consider the role of a consultant, it gives me the specific advantage of being able to support customers and provide advice in this regard because I can gain certain know-how from the requirements catalogue and apply it immediately. As a provider, I would say the advantage is that you need to allocate fewer resources afterwards to ensure the system's compliance because you have already proactively started aligning with the legal framework.

Interviewer
Yes, that's a good point, these preventive measures are important because they allow you to prepare in advance and not wait until the law is already in place, where it may be too late to act. The final question regarding perceived usefulness is: How do you overall assess the usefulness of the requirements catalogue in the compliance audit of the AI Act?

Interviewee 5
I would say the requirements catalogue has overall high usefulness, but one should not rely on it blindly, especially considering that laws can change quickly.

Interviewer
Now, the next questions are more about usability. Here's the first question: How do you evaluate the usability when filling out the requirements catalogue?

Interviewee 5
Well, I must say it is very well-structured and organised and the colour scheme supports that. So, it's clear to me that if I don't know something, I can simply refer to the blue data sheets. I also consider it well-structured in the worksheets where I need to fill in information, with the information presented first and then the columns to be filled. I like the clear separation. One thing I would like to mention is that I would replace the term Measure with Description of Implementation because it would clearly indicate that I should write the current status of implementation as a user. I also appreciate having space to enter planned measures and being able to set a target, as it gives a sense of milestones. Overall, I would say it is very user-friendly and I quickly found my way around.

Interviewer
Very good. The next question is: How do you assess the clarity and ease of interpretation of the presentation of results?

Interviewee 5
When I look at it now, the total Maturity level is at the top. But what is actually relevant to me is the shortened version because better results should not influence the fact that some of the areas are still insufficiently met. Therefore, I would highlight the shortened Maturity level so that it is clearly visible to me as a user that it is the more relevant one. This could be done by simply reverse the order or using colours. Otherwise, I think it is well-presented and I like that the user can choose their own Target Maturity level. It could be helpful to explicitly recommend aiming for Level 3.

Interviewer
Alright. Then the next question would be: Do you see any challenges in using the requirements catalogue, if any?

Interviewee 5
Well, the only thing I see is that to work with this catalogue, you need a basic understanding of the meaning of the standards. I believe it might be a bit overwhelming for someone who hasn't had much exposure to them because they might not be familiar with all the terminologies. But that's more of a general problem and not directly related to the requirements catalogue. Otherwise, I don't see any challenges.

Interviewer
The last question is quite interesting. How do you overall assess the usability of the requirements catalogue for assessing compliance with the AI Act? What factors contribute to this?

Interviewee 5
Well, fundamentally, when you open the file and get the first impression, I think it is very good that colours are used because it gives a sense of structure. And I also appreciate that the most important terms are explained again in case I'm not familiar with them. The view of the evaluation is also very clear, so I can directly see in which areas I have achieved which level. I especially want to highlight the last worksheet because with the heatmaps and the colours, I can quickly identify the areas and clusters where I should improve. Overall, I consider the catalogue and especially results worksheets extremely well done.

Interviewer
Great, thank you very much for your insights and your time.

Interviewee 5
You're welcome.

System Usability Scale: Interviewee 5					
	0	1	2	3	4
Statements	**Strongly disagree**				**Strongly agree**
I thought all the relevant requirements of the AI Act were integrated completely.					x
I thought the requirements were enumerated several times (duplicates).	x				
I thought the requirements were formulated in an understandable and practical way.					x
I thought that certain requirements were conflicting with each other.		x			
I thought the requirements could be evaluated according to an understandable fulfilment degree.					x
I thought the requirement catalogue was unstructured and unnecessary complex.	x				
I feel confident that the requirements catalogue will support the development of an AI system that is compliant with the AI Act.					x
I thought further instructions would be necessary to fully use the requirements catalogue.	x				
I thought the calculated results of the maturity level were presented in a clear manner.				x	
I thought in the requirements catalogue important aspects for a compliance assessment were missing.	x				

Figure 5: SUS score of interviewee 5

Interviewee 6: Transcript and SUS score

Interviewer
So, if you're ready, I would like to start with the seven questions now. And the first three questions are about perceived usefulness. The first question is: In what way do you think the requirements catalogue can support AI system providers in evaluating compliance with the requirements of the AI Act?

Interviewee 6
Well, EU law is a difficult matter and a very opaque matter. I have already learned that in projects and accordingly, I think the catalogue can support companies well. Because requirements from EU regulations are very long, spanning many pages and in the catalogue, the requirements are summarised in a clear and rephrased manner. So I think it also helps to make it understandable for people who are not legally versed, what is actually required. Therefore, it is very useful for companies to determine the scope, the extent of the requirements to be met and then implement them in a structured way. So, I definitely think it is a very valuable support.

Interviewer
The next question would then be: What specific benefits do you expect from using the requirements catalogue in a compliance audit?

Interviewee 6
Basically, when new laws are enacted, they usually come with a transition period and within this transition period, companies engage in requirement management. Am I affected, yes/no? If yes, what requirements does the law impose on my company and on myself? Then I have to implement these requirements in a timely manner before the transition period is over. In any case, the catalogue supports the understanding of this

topic. You don't have to struggle through it on your own. It makes sense to read it, but you don't have to deduce everything yourself. You can work through it with the help of this catalogue and consequently filter out more easily which requirements are relevant to you. And accordingly, nothing can be overlooked if you simply don't work with this catalogue. So, at least I would assume that it is complete to the extent that all requirements are described there. That means, if I fulfill all the requirements there, I am working in compliance with the scope of the AI Act. And yes, it make it easier, more effective and faster to comply with a law.

Interviewer
Thank you for the insight. The next question is, overall: How do you overall assess the usefulness of the requirements catalogue in the compliance audit of the AI Act?

Interviewee 6
I see a very high usefulness here because we basically have requirements for these examinations and the requirements come from the AI Act and have now been integrated into this catalogue. And if I have already implemented the AI Act using this catalogue, then I can also use it for the compliance audit afterwards and effectively see if I am compliant. And if not, I can see on the heatmap, for example, in which areas there are shortcomings. Just like in your example of risk management, where you can actively counteract. The good thing is that I can see clearly where my problems are. I don't get an endlessly long result table, but I get a clear overview on two sheets with four spider graphs on one side and two heatmaps on the other, clearly showing where the problems are. It also makes it more understandable, for example, when you have to present results or decisions to the management. You don't have to prepare lengthy results, but you have direct results to show.

Interviewer
Those were the questions about usefulness. Now let's focus on the perceived ease of use. So, the first question is: How do you evaluate the usability when filling out the requirements catalogue?

Interviewee 6
Everything I have seen so far looked very clear. The only comment I would have is regarding the sheet Requirements for the System and also Obligations of Providers. The maturity level is crucial for the later evaluation, so it could be highlighted more strongly that it must be filled out. But that's a minor point. Otherwise, what I have seen looks very reasonable. So, I think if you read through the welcome sheet and look at the structure, you can work with it relatively quickly and well. Even people who have had little experience with Excel or this type of requirements catalogue will likely consider it easy to navigate.

Interviewer
Great. The next question is about clarity and ease of interpretation. How do you assess the clarity and ease of interpretation of the presentation of results?

Interviewee 6

Yes, I think the presentation of results is very good and clear, especially the division into two sheets, so that the results can be viewed separately. For example, I also think it is very good that you have the shortened maturity level, which avoids the distortion that often occurs and is sometimes used to improve the overall result. Certain areas are given higher weights where you perform well, while other areas where you perform poorly are given lower weights, thereby boosting the overall result. So, this popular watermelon reporting can be limited to some extent, allowing you to obtain a realistic result. I consider further the heatmaps very well-designed and they are excellent for presenting complex requirements and its results in a clear way. Anyone can directly see, without any prior knowledge, that, okay, we need to do something about risk management and we are already well-positioned in chapter seven, for example. So, I think the presentation of results is well-executed.

Interviewer
Thank you for your input. The penultimate question is: Do you see any challenges in using the requirements catalogue, if any?

Interviewee 6
One challenge, I think, is that this requirements catalogue tempts people to focus only on the catalogue itself and not on the actual regulation. Therefore, there is also a risk that when the AI Act is updated, the catalogue may no longer be up to date. This may not be a direct challenge, but it should definitely be considered that the AI Act is, so to speak, a living law that will be updated at much shorter intervals due to the rapid development of AI than other EU regulations. This should not be forgotten because violations can lead to significant financial burdens, among other consequences.

Interviewer
The last question is: How do you overall assess the usability of the requirements catalogue for assessing compliance with the AI Act? What factors contribute to this?

Interviewee 6
I think the requirements catalogue is very user-friendly and well-structured. The welcome page contributes a lot to this as it provides a good overview of the individual worksheets. And in terms of result presentation, I also think the heatmaps are a very well-chosen addition to the spider graphs. So overall, I rate the usability as very good.

Interviewer
Very good. Thank you very much for your time!

Interviewee 6
No problem.

System Usability Scale: Interviewee 6

Statements	0 Strongly disagree	1	2	3	4 Strongly agree
I thought all the relevant requirements of the AI Act were integrated completely.					x
I thought the requirements were enumerated several times (duplicates).	x				
I thought the requirements were formulated in an understandable and practical way.					x
I thought that certain requirements were conflicting with each other.		x			
I thought the requirements could be evaluated according to an understandable fulfilment degree.					x
I thought the requirement catalogue was unstructured and unnecessary complex.	x				
I feel confident that the requirements catalogue will support the development of an AI system that is compliant with the AI Act.					x
I thought further instructions would be necessary to fully use the requirements catalogue.	x				
I thought the calculated results of the maturity level were presented in a clear manner.					x
I thought in the requirements catalogue important aspects for a compliance assessment were missing.	x				

Figure 6: SUS score of interviewee 6