Ref. Ares(2025)3460929 - 29/04/2025

Artificial Intelligence in Manufacturing leading to Sustainability and Industry5.0

Project Acronym:

AIMS5.0

Grant agreement no: 101112089

Deliverable no. and title	**D7.2 – Guidelines incl. checklist and self-assessment of how to apply AI**	
Work package	WP 7	Human Acceptance, trust and ethics for digital workspaces
Task	T 7.1, 7.2, 7.3	T7.1 - Personas, roles and acceptance
Subtasks involved	N/A	
Lead contractor	Infineon Technologies AG Thomas Gutt, mailto:thomas.gutt@infineon.com	
Deliverable responsible	RSA, Markus Tauber, markus.tauber@researchstudio.at Philips, Olga Kattan, olga.kattan@philips.com RSA, Yasin Ghafourian, yasin.ghafourian@researchstudio.at	
Version number	V1.0	
Date	29/04/2025	
Status	Final	
Dissemination level	Public	

Copyright: AIMS5.0 Project Consortium, 2025

Authors

Participant no.	Part. name	Author name
1	Research Studios Austria	Yasin Ghafourian
2	Research Studios Austria	Markus Tauber
3	Philips	Olga Kattan
4	Zittau/Görlitz University of Applied Sciences	Fabian Lindner
5	Harokopio University of Athens	Konstantina Karathanasopoulou
6	Harokopio University of Athens	George Dimitrakopoulos
7	Budapest University of Technology and Economics	Ádám Kővári
8	Budapest University of Technology and Economics	Csaba Hegedűs
9	Research Studios Austria	Belal Abu Naim
10	Budapest University of Technology and Economics	Kitti Mezei
11	Budapest University of Technology and Economics	Pál Varga
12	AIT Austrian Institute of Technology	Erwin Schoitsch

Document History

Version	Date	Author name	Reason
V0.1	10.02.2025	Yasin Ghafourian	Initial Template
	13.04.2025	Germar Schneider	Review
V1.0	29.04.2025	Yasin Ghafourian	Integration of review feedback, finalization and submission

Publishable Executive Summary

This deliverable, D7.2 of the AIMS5.0 project, focuses on developing practical tools and insights to support human acceptance, ethical integration, and regulatory compliance of Artificial Intelligence (AI) in industrial environments, in alignment with Industry 5.0 principles. It shows the work carried out under Tasks T7.1, T7.2, and T7.3, and also marks the conclusion of Tasks T7.1 and T7.2. We introduce a persona-driven approach to understanding stakeholders in industrial AI and presents a user-centric self-assessment tool—an AI-powered chatbot—to help navigate the European Union's AI Act and related compliance frameworks.

A central part of the work involved a survey of 61 participants across European AI research and development projects. This survey gathered insights into AI professionals' familiarity with AI, their application domains, and their use and understanding of existing AI guidelines. The analysis revealed significant gaps: nearly half of the respondents were unsure whether their organizations follow formal AI standards. Furthermore, AI familiarity influenced risk perception, with more experienced users being more aware of algorithmic and resource-related challenges.

Building on these findings, the deliverable outlines the design of a self-assessment tool which is a chatbot built using a Retrieval-Augmented Generation (RAG) architecture. Two models are proposed and tested: a basic Naive RAG and an advanced Graph RAG, the latter enabling more complex, structured responses based on regulatory knowledge graphs. The chatbot helps users assess their AI systems against the EU AI Act by offering personalized compliance support through an intuitive conversational interface.

User acceptance of this tool was evaluated using the Quantitative Effect and Technology Acceptance Model (QETAM), which showed moderate acceptance, with trust and ease of use as key influencing factors. To ensure practical relevance, the chatbot will be also tested against real industrial AI use cases in AIMS5.0 and legal expert-defined checklists from our legal and ethical experts in the project.

We conclude the report by highlighting the need for broader stakeholder inclusion, expansion of the regulatory coverage, and continued refinement of the chatbot to ensure higher trust, usability, and compliance effectiveness. It also showcases dissemination activities, including academic publications and workshops, underlining the project's contribution to the responsible and ethical use of AI in future manufacturing ecosystems.

This Work has been supported by the following recent peer-reviewed publications:

1) Á. Kővári, Y. Ghafourian, C. Hegedűs, B. Abu Naim, K. Mezei, P. Varga, and M. Tauber, "Let's have a Chat with the EU AI Act," in Proc. IEEE/IFIP Network Operations and Management Symposium (NOMS 2025), 2025.
2) Y. Ghafourian, F. Lindner, O. Kattan, K. Karathanasopoulou, M. Tauber, and G. Dimitrakopoulos, "Understanding Stakeholders of Industrial AI: Insights from a Persona-Based Questionnaire," in *Proc. IEEE/IFIP Network Operations and Management Symposium (NOMS 2025)*, 2025.

Table of Contents

1. Introduction 6
1.1 Objective and scope of the document 6
1.2 Structure of the deliverable report 6
2. Understanding Stakeholders of Industrial AI: Insights from a Persona-Based Questionnaire 8
2.1 Design methodology of the persona-based questionnaire 8
2.1.1 Background 8
2.1.2 Related Literature 9
2.1.3 Survey ..10
2.2 Analysis of the results 10
2.2.1 Descriptive Statistics 10
2.2.2 Exploratory Analysis 10
2.2.3 Chatbot Acceptance in Specific 13
2.3 Discussions and Conclusions 16
2.3.1 Acknowledgements 17
3. The Self Assessment tool: Let's talk to the EU AI Act 18
3.1 The Problem Space 18
3.2 The Regulatory Landscape 19
3.2.1 About the AI Act 20
3.3 The Chatbot 20
3.3.1 System Design and Architecture 20
3.3.2 Naive RAG 21
3.3.3 Graph RAG 23
3.4 Methodology for evaluation & fine tunning 26
3.4.1 EU AI Act Checklist 26
3.4.2 Use Case Stories 26
3.5 Future Direction and Conclusion 27
4. Outreach to the Community 29
4.1 Papers 29
4.2 Posters or oral presentations at relevant events 29
4.3 Dedicated AIMS5.0 WP7 workshops 30
4.3.1 Workshop with the PAL project at Zittau/Görlitz at 09th.April 2025 30
5. Conclusions 31
6. Appendix 32
6.1 Annex A: Abbreviations 32
6.2 Annex B: Overview of WP 7 - related AI-Standards 33
6.2.1 AI: ethical, social, governance, oversight and human aspects impact: 33
6.2.2 Trustworthiness, Safety, Reliability, Robustness, Controllability, and Risk management: 33
6.2.3 EC Standardization Requests for the AI-Act 34
6.2.4 Domain-specific AI-related/human-centered Standards (Examples) 35
7. References 36

List of Figures

Figure 1: How do you expect AI to add value to your use case/company? 12
Figure 2: What challenges and risks do you see when using AI for your specific use case at work? 12
Figure 3: Which measures do you consider most suitable to overcome these challenges and/or mitigate the risks? 13
Figure 4: PU, PT, PEU, SI, and FC Measures to Address AI/ML Acceptance with No. of Responses~ >20% (Multiple Choice; $N=59$) 14
Figure 5: Correlation Vector of Key Indicators with No. of Responses ≥ 20% (Multiple Choice; N = 59) 15
Figure 6: Overview of the Naïve RAG 21
Figure 7: Simple question/answer pair from the interface of our developed chatbot 22
Figure 8: Graph-based RAG 24
Figure 9: Global-search 25
Figure 10: Local-search 25

List of tables

Table 1: “What is your level of familiarity with Artificial Intelligence (AI) / Machine Learning (ML) technologies?” (multiple choice, N = 59) 11
Table 2: Abbreviations 32

1. Introduction

1.1 Objective and scope of the document

(This Deliverable is supported by 2 peer-reviewed articles accepted for publication in Proc. IEEE/IFIP Network Operations and Management Symposium (NOMS 2025), 2025)

This deliverable, D7.2, is all about helping industrial AI systems become more people-friendly, ethically responsible, and legally compliant. It focuses on Tasks T7.1 and T7.2 of the AIMS5.0 project and outlines how we've worked with industrial partners - developers, managers, researchers—to understand their concerns and needs around AI in the workplace.

In Task T7.1, we took a human-centered approach by developing "personas that reflect the roles and experiences of AI practitioners across industry. This helped us get a clearer picture of the real challenges people face when it comes to adoption of AI and trust.

Task T7.2 built on that work by looking at the broader context—laws, ethical norms, and existing guidelines. This involved reviewing key documents like the GDPR, the EU AI Act, and ethical principles from leading institutions. One of the key outcomes is a self-assessment chatbot designed to help AI users and developers easily understand how their systems align with these evolving standards. This chatbot currently integrates the EU AI Act and we are working towards integrating more ethical guidelines and standards into it.

Together, these tasks feed into a larger effort to support responsible, transparent, ethical and human-centric AI deployment in Europe’s digital workspaces.

1.2 Structure of the deliverable report

This deliverable is structured as follows:

Section 2 dives into how we built and analyzed the persona-based questionnaire, what we learned about the people working with AI, and what this tells us about how they view AI risks and responsibilities.

Section 3 introduces the self-assessment chatbot and explains how it works, how it was built, and how it helps navigate complex regulations like the EU AI Act.

Section 4 summarizes how we’ve disseminated and advertised this work so far—through papers, workshops, and community outreach.

Section 5 wraps things up with our key takeaways from Tasks 7.1 and 7.2, and gives a preview of what’s coming next in Task 7.3.

Sections 6 and 7 include extra reference material and background information.

1.3 Progress since D7.1

Since the publication of deliverable D7.1, significant advancements have been made toward achieving the objectives outlined in the initial phase of the project. D7.1 focused on compiling and evaluating methods and sources for establishing guidelines applicable across all industrial AI use-cases, laying the foundational groundwork for subsequent tasks.

Building upon this foundation, deliverable D7.2 has introduced developments, particularly with the introduction of a user-centric self-assessment tool—an AI-powered chatbot designed to facilitate compliance with the European Union's AI Act and related frameworks. This chatbot is built on a Retrieval-Augmented Generation (RAG) architecture, transitioning from preliminary theoretical groundwork into tangible application.

Moreover, comprehensive stakeholder engagement has expanded significantly through a detailed persona-based survey. The interviews conducted in D7.1 with the stakeholders of industrial AI have laid the ground for this survey. With participation from 61 respondents across European AI research and development projects, the survey has deepened our understanding of AI familiarity, risk perception, and guideline adherence among professionals in industrial settings. This quantitative and qualitative assessment has identified critical gaps, particularly highlighting that nearly half of the respondents were uncertain about adherence to formal AI standards within their organizations.

In parallel, deliverable D7.2 has refined and further developed the initial personas and role categorizations first established in D7.1. The integration of these refined personas into the self-assessment tool ensures that user interactions with the chatbot are more precisely aligned with the practical and regulatory needs of diverse AI users.

Lastly, deliverable D7.2 marks significant methodological progress through the implementation and comparative testing of two distinct chatbot models: a basic Naive RAG model and an advanced Graph RAG model. The introduction and evaluation of these models represent a considerable methodological advancement, enhancing the chatbot's capability to deliver structured, accurate, and contextually relevant responses based on complex regulatory knowledge graphs.

Overall, the progress from D7.1 to D7.2 reflects a transition from foundational research and stakeholder mapping toward practical implementation and iterative testing of tools designed to support ethical and compliant AI applications within industrial contexts.

2. Understanding Stakeholders of Industrial AI: Insights from a Persona-Based Questionnaire

2.1 Design methodology of the persona-based questionnaire

The rapid evolution of industrial artificial intelligence (AI) has given rise to a diverse landscape of developers and users across various sectors. Understanding their personas—including their backgrounds, expertise, and work environments—is crucial for fostering innovation and ensuring compliance with AI regulations. This study presents insights from a persona-based questionnaire aimed at mapping the industrial AI ecosystem and informing the development of a user-centered self-assessment compliance tool. Through a structured survey of 61 participants from European AI research projects, the study identifies key attributes of AI professionals, their industry affiliations, and their approaches to AI adoption and compliance. Findings reveal gaps in awareness and application of AI guidelines, with half of the respondents uncertain whether their organizations follow formal AI standards. Moreover, AI familiarity influences risk perception, with experienced users more likely to identify algorithmic and resource-related challenges. The study integrates the Quantitative Effect and Technology Acceptance Model (QETAM) to assess chatbot acceptance for AI compliance support, ensuring the proposed tool aligns with user needs.

2.1.1 Background

In recent years, the rapid advancement of industrial artificial intelligence (AI) has created a dynamic ecosystem comprising developers and users engaged in creating and deploying AI applications across various sectors. Understanding the personas of these individuals - including their backgrounds, domains of work, and the types of applications they develop—is critical to fostering innovation and compliance in this field. Our study seeks to explore these personas, providing insights into the sectors they represent and the AI application domains they focus on. This knowledge is not only essential for mapping the industrial AI landscape but also for tailoring tools and frameworks that support their work effectively.

Building on this understanding we are developing a self-assessment tool to support compliance with AI regulations and guidelines, designed with a user-centered approach informed by the personas of AI developers and users from our questionnaire. This self-assessment tool is a chatbot powered by large language models (LLMs) and built on the Retrieval-Augmented Generation (RAG) framework. The tool is intended to assist AI developers and users in navigating the complex landscape of regulatory requirements ensuring compliance with existing regulations, guidelines, and standards. By analyzing the personas of our target users, which includes understanding their demographic backgrounds, organizational contexts, roles, and AI expertise, we can tailor the tool to align closely with their needs and expectations.

Understanding personas helps us identify the individual needs and goals, the breadth of AI use cases, organizational sizes, and economic activities, allowing the tool to cater to a wide range of scenarios rather than being narrowly focused. For instance, by knowing whether users work in small startups or large corporations, the tool can adjust the scope of its guidance, offering tailored compliance pathways for resource-constrained environments or robust systems for enterprises with complex regulatory obligations. Similarly, insights into the types of AI being used and the areas of application ensure that the tool covers diverse industries, from healthcare to manufacturing, while addressing sector-specific compliance challenges. Furthermore, understanding the economic and geographical contexts of their organizations helps us address region-specific regulatory nuances and compliance requirements. To ensure its practical utility and widespread acceptance, we leverage the Quantitative Effect and Technology Acceptance

Model (QETAM) [1], a novel framework that integrates the strengths of the Technology Acceptance Model (TAM) [2] with Failure Mode and Effects Analysis (FMEA). By combining these detailed persona insights with direct feedback on the usability and acceptability of the tool, our research aims to create a solution that meets the expectations of AI developers and users while fostering responsible and compliant AI development practices.

2.1.2 Related Literature

2.1.2.1 Attitudes Towards AI in Industry

While workers' attitudes towards automation and digitalization are of interest in general [3], we conducted in previous work also structured interviews with 25 industry experts from large European research and development projects to gain insights into the adoption of AI in industrial settings in specific [4]. These interviews aimed to identify diverse user needs, challenges, and benefits related to AI implementation. According to our findings, organizations could be categorized into three different categories: early adopters/avant-garde, followers, and independent learners. Early adopters are mainly large companies that have the need for specific AI applications beyond the state of the art. Therefore, they are also actively involved in AI research and development activities. Followers and independent learners are mostly small and medium-sized companies (SMEs) with less experience in AI. While all of them share similar concerns and expectations, early adopters seem to have already more concrete mitigation measures in place to handle potential risks due to AI.

Similarly, Kutz et al. [5] assessed success factors and challenges of AI implementation among German manufacturers. They find that both success factors and challenges are most often related to data, development and operational processes, and stakeholder engagement. Additionally, Somer and Thalmann [6] reviewed the literature concerning different risk management approaches when using AI. Among others, they identify algorithm regulation, human oversight, governance, certification, and auditing as possible measures.

2.1.2.2 Technology Acceptance

The original Technology Acceptance Model (TAM), introduced by Davis [2], was developed to model users' acceptance of information systems and technologies. TAM aims to identify the key determinants of computer acceptance, enabling the prediction of user behavior across various end-user computing technologies and diverse user groups. The foundational TAM framework includes two primary beliefs: Perceived Usefulness (PU) and Perceived Ease of Use (PEU), both of which are shaped by external factors. TAM has been widely applied in different contexts. For example, Alfadda and Mahdi [7] employed TAM to analyze user attitudes toward using Zoom for language learning and to investigate the effects of gender and experience on usage. Additionally, TAM has been utilized in healthcare to study the adoption of health IT [8]. However, researchers noted limitations in its application, such as infrequent measurement of certain variables and insufficient theoretical explanation of its constructs.

In another study, Chatterjee et al. [9] used a combined TAM-TOE framework to explore the social, environmental, and technological factors influencing the adoption of embedded artificial intelligence in digital manufacturing organizations. TAM also shares similarities with the Unified Theory of Acceptance and Use of Technology (UTAUT), which incorporates core constructs like perceived ease of use and perceived usefulness. UTAUT was built on TAM and seven other theoretical models, focusing on factors such as performance expectancy, effort expectancy, social influence, and facilitating conditions to explain technology acceptance. UTAUT has been applied in various domains, such as identifying factors affecting students' acceptance of MOOCs in Saudi Arabian higher education [10]. Similarly, Puspitasari et al. [11]

used UTAUT to explore variables influencing the adoption of the Integrated Licensing Services Information System.

2.1.3 Survey

To assess the different industry stakeholders' knowledge, attitude and usage of AI at work, we designed and distributed a standardized online questionnaire. It covered information about participants' personal background (socio-demographics) and work background (organization, role), their familiarity and proficiency with AI, the implementation of AI at work, the used guidelines and standards, and the expected benefits and challenges thereof. In a second part, participants' specific acceptance of a chatbot that supports implementation and compliance processes with AI was assessed by using an adaptation of the technology acceptance model (see Section II.B). For this purpose, the participants were confronted with a scenario, where they can use such a company-wide chatbot that is able to answer any of the users' questions concerning compliant and targeted applications of AI for their specific use cases. Afterwards, they were asked about their acceptance of such a chatbot using the adapted technology acceptance model.

The questionnaire was administered to the partners of three European-funded R\&D projects (AIMS5.0, AI REDGIO.0, and Arrowhead fPVN) concerning AI and was answered by 61 participants, of which two were excluded from the analysis due to missing answers.

2.2 Analysis of the results

2.2.1 Descriptive Statistics

The participants in the survey were mostly male (75%, n=44), on average 39 years old (N=59), and held a Ph.D. (44%, n=26) or Master's degree (41%, n=24). Their professional role was situated in research (61%), engineering (44%), management (36%), and/or technology (17%) (multiple choice, N=59). The vast majority of participants has already used AI (Table 1). Participants mostly report that they feel somewhat knowledgeable with AI (27%, n=16) and familiar in applying it at work (29%, n=17). Some of them have already led projects and published in this field (25%, n=15).

The organizations, where the respondents work at, are primarily located in Germany (32%, n=19), Austria (15%, n=9), the Netherlands (14%, n=8), Poland (10%, n=6), and Greece (9%, n=5). 59% of them are profit-oriented and 48% research organizations (multiple answers per participant were possible). 61% have ≥ 250 employees, 24% < 250, 12% <50, and 3% <10. They operate in the sectors of information and communication as well as professional, scientific and technical activities (both 36%), in the manufacturing (32%), education (31%), agriculture, forestry and fishing as well as semiconductor (both 5%) (multiple choice, N=59).

Table 1: "What is your level of familiarity with Artificial Intelligence (AI) / Machine Learning (ML) technologies?" (multiple choice, N = 59)

Familiarity Level	Percentage
More than 10 years of experience in developing and deploying machine learning models.	14%
Having led several projects from conception to implementation and published research in this field.	25%
I possess a deep understanding of AI/ML algorithms, having built custom models and frameworks for various applications in my career.	17%
Extensive experience with AI/ML, including hands-on work with neural networks, deep learning, and natural language processing.	19%
I have strong skills in AI/ML and have used these skills to solve complex problems.	24%
I am familiar with various machine learning frameworks and libraries and have applied them to practical problems in my work.	29%
I have a foundational understanding of AI/ML and have completed a few projects, but I am still expanding my expertise in this area.	27%
I am a beginner/have a theoretical understanding of AI/ML and am in the process of gaining practical experience.	20%
I am using it but have only little or no theoretical understanding.	20%
I have no direct experience with AI/ML technologies but am interested in learning more about them.	5%

In our questionnaire, we have also asked participants in a free form about the specific purposes for which they use AI/ML. Responses highlighted six main themes (N=51): 1) AI and ML usage: leveraging artificial intelligence and machine learning in various applications, 2) vision and Recognition Applications: Keywords like "detection" and "classification" in the use of AI for image-based applications. 3) specific Techniques: use of advanced AI methods such as Deep Learning, 4) industry Applications: use of AI in "Manufacturing", "maintenance", and "production", 5) data Utilization and use of AI to make data-driven applications, and 6) research and Development: being involved in "research" to innovate and expand AI/ML capabilities, and 7) other. The most common uses included leveraging AI to enhance productivity, applying computer vision for tasks like detection and classification, and adopting advanced techniques for both industry-specific and research-driven purposes.

The expected benefits of AI/ML application concern reducing costs, increasing quality, improving efficiency, or saving resources, among others (Figure 1). Roughly one fourth (24%, n=14) applies standardized guidelines or checklists when using AI. Almost half of respondents (49%, n=29), however, cannot tell if guidelines are applied or not. The guidelines applied are internally developed (100%), less frequently official standards (36%) and external guidelines (29%) are used (multiple choice, N=14). Mostly, legal regulations (64%) and/or ethical aspects (57%) are considered (multiple choice, N=14).



Figure 1: How do you expect AI to add value to your use case/company?

Despite (or because of) this, compliance-related risks (37%) using AI/ML at work are less frequently mentioned by the participants than data- (71%) or people-related ones (68%) (multiple choice, N=59; Figure 2).

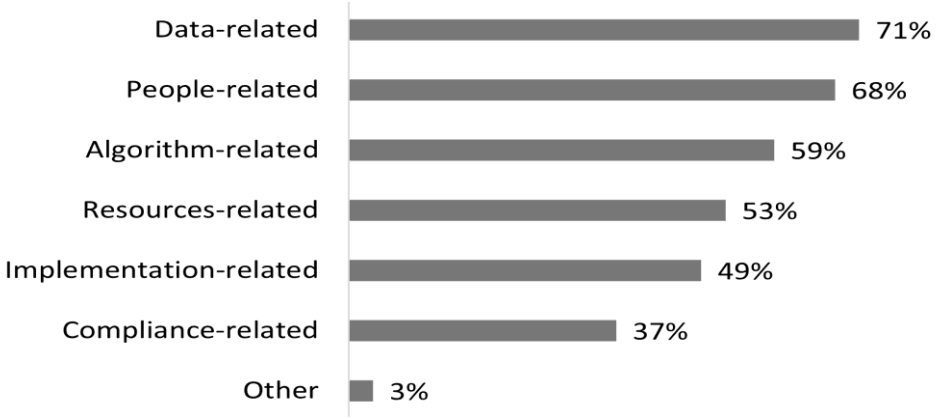

Figure 2: What challenges and risks do you see when using AI for your specific use case at work?

According to the respondents (multiple choice, N=59), the most urgent measures to address such issues are thus also people- (know-how, training, responsibilities, motivation, fears) and data-related (availability, quality, biases). Too high expectations of AI in practice need to be managed properly (29%) and the benefits (and limits) need to be clear (15%). Furthermore, stakeholders of industrial AI require the qualifications and competencies to successfully implement and use AI at work (24%). The end-users need to be integrated into the development processes early on (24%) and trust in AI needs to be developed (22%). Further measures concern the use of suitable datasets (22%), standardized hardware and software (17%), and combining AI with traditional approaches (15%; Figure 3).

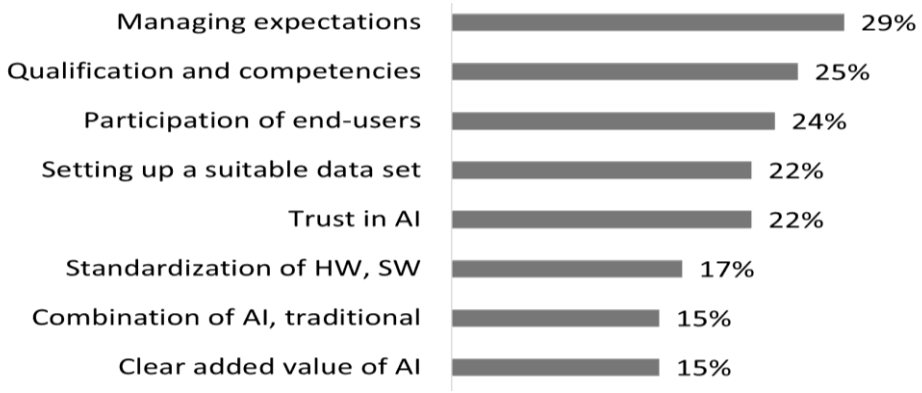

Figure 3: Which measures do you consider most suitable to overcome these challenges and/or mitigate the risks?

2.2.2 Exploratory Analysis

To be able to draw further conclusions concerning different individual and organizational needs, characteristics, requirements, and challenges concerning AI applications and guidelines, we tested several exploratory hypotheses related to meaningful associations between different variables using chi-square tests of independence.

First, we wanted to know whether there are associations between the respondents' level of familiarity with AI, the challenges they see, and the countermeasures they find relevant. From our survey data we can only state with statistical significance that respondents with higher familiarity with AI considerably more often mentioned resource-related issues as challenging ($\chi^2(1, N = 59) = 5.19, p = 0.02^*$). Also, algorithm-related challenges were mentioned more often by them than by those with less AI familiarity, although not statistically significant ($\chi^2(1, N = 59) = 3.64,\ p = 0.06$).

Second, we were also interested if there were differences in the application of AI guidelines between organizations' characteristics. But neither the size of the organization (≥ 250 vs. less employees; $\chi^2(1, N = 30) = 0.01, p = 0.92$), nor the organization's country of operation (Germany, Austria, or Netherlands; $\chi^2(2, N = 20) = 1.63, p = 0.44$ was associated with whether specific AI guidelines were followed or not. There was just a difference between profit and non-profit organizations, whereas profit-oriented organizations seemed to apply guidelines more often, although not statistically significant ($\chi^2(1, N = 30) = 3.77, p = 0.05$).

2.2.3 Chatbot Acceptance in Specific

The acceptance of chatbots for AI compliance verification depends on several key factors, which are addressed by the QETAM model, those are perceived trust, usefulness, ease of use, social influence, and facilitating conditions.

The results in Figure 4 indicate a moderate acceptance of the chatbot, with Perceived Trust (PT) scoring the highest average (6.48) and Social Influence (SI) the lowest (5.33). This suggests that users generally trust the chatbot's reliability, security, and fairness but that external factors such as peer influence, industry adoption, and marketing (5.33) have minimal impact on their decision to use it.

Facilitating Conditions (FC) (6.28) and Perceived Usefulness (PU) (5.65) also received relatively positive scores, indicating that users find the chatbot somewhat helpful and compatible with their workflows, though there is room for improvement in usability and efficiency. However,

Perceived Ease of Use (PEU) (5.92) had mixed ratings, implying that users might experience some difficulties navigating or fully utilizing the chatbot's features.

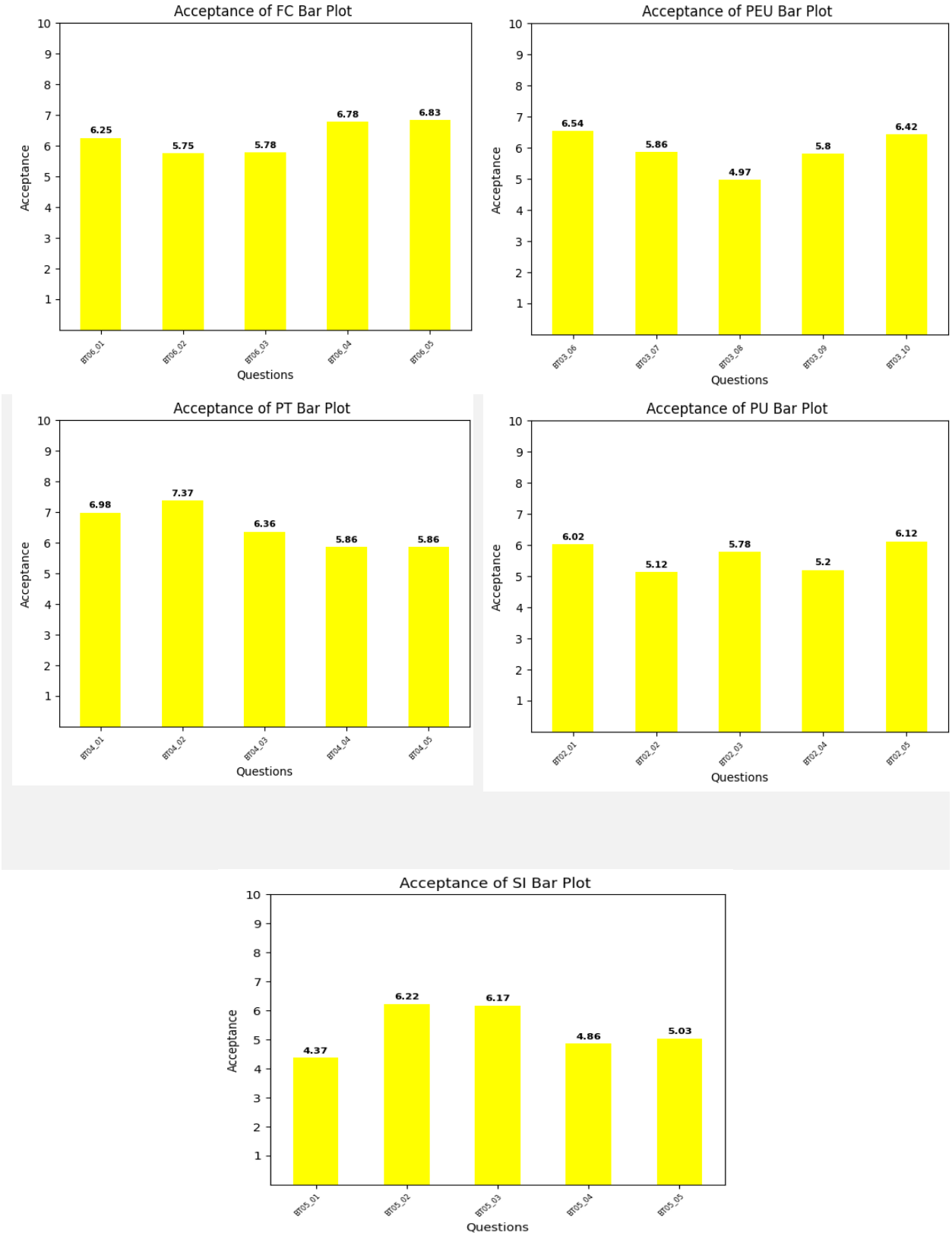

Figure 4: PU, PT, PEU, SI, and FC Measures to Address AI/ML Acceptance with No. of Responses~ >20% (Multiple Choice; $N=59$)

The lower Social Influence score (5.33) suggests that the chatbot lacks strong encouragement or visibility within the industry, which could be a barrier to widespread adoption. To improve acceptance, enhancing usability and increasing social influence will be crucial. Additionally, focusing on intuitive design and simplifying navigation can help address challenges in ease of use, making the chatbot more accessible and user-friendly.

The total BIU value is 7254.2933, classified within the Low (5001-20000) category, which suggests a moderate level of acceptance of the chatbot for AI compliance verification. While users acknowledge its potential, several factors may contribute to its limited adoption. Perceived trust appears to be a concern, as respondents may have reservations about the chatbot's reliability, security, and ability to provide unbiased and up-to-date recommendations.

Similarly, perceived usefulness remains moderate, indicating that while the chatbot is considered helpful, it may not yet be seen as a highly efficient or indispensable tool for ensuring compliance with AI guidelines and standards. Additionally, perceived ease of use suggests that while the chatbot is not necessarily difficult to navigate, it may lack intuitive design features that facilitate seamless adoption. The influence of external factors, such as colleague recommendations, industry adoption, and marketing efforts, appears to be limited, potentially affecting user confidence in the tool's relevance and credibility. Furthermore, facilitating conditions, including technical support, training materials, and integration into existing workflows, may not be sufficiently developed to encourage widespread use. To enhance adoption and increase the BIU score, improvements should focus on building trust through transparency and security, demonstrating clear efficiency gains, refining usability, increasing industry endorsement, and strengthening support resources to align more closely with user needs and expectations.

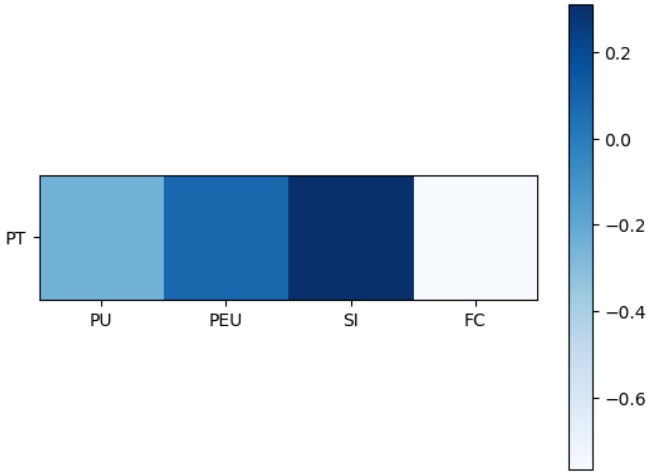

Figure 5: Correlation Vector of Key Indicators with No. of Responses ≥ 20% (Multiple Choice; N = 59)

Figure 5 presents the analysis of the correlation vector, revealing that the Social Influence (SI) and Perceived Trust (PT) constructs demonstrate the strongest correlation. This finding highlights the pivotal role of social reinforcement and trust in fostering user adoption. Furthermore, the significant association between Perceived Ease of Use (PEU) and Facilitating Conditions (FC) underscores the importance of a supportive environment in enabling effective and seamless usage.

2.3 Discussions and Conclusions

Our results show that the main areas where AI is currently applied in industry and research and development activities is in line with previous research and findings and also the desired benefits are similar to expectations [12]. Furthermore, the findings regarding AI guideline awareness raise important questions about organizational readiness for upcoming AI regulations. The fact that 50% of respondents are uncertain about their organization's AI guidelines suggests a potential gap between policy development and communication within organizations. This knowledge gap could become particularly critical as the EU AI Act comes into force, requiring organizations to implement robust compliance frameworks. The higher awareness of AI guidelines in profit-oriented organizations might indicate that commercial pressures and potential liability concerns are driving more structured approaches to AI governance in the private sector.

Regarding concerns when applying AI, notably, respondents with higher familiarity with AI are more often concerned about resource- and algorithm-related risks than others. This can mean that novel users of AI might underestimate resource- and algorithm-related challenges when planning AI applications. Therefore, a prospective chatbot can take this into account, by proposing relevant information to such users. Likewise, these concerns by more knowledgeable users need to be addressed by the chatbot as well.

These findings also point to the potential value of AI compliance support tools, such as the proposed chatbot. The fact that even experienced AI practitioners express significant concerns about resource and algorithmic challenges suggests that automated compliance support tools could fill a crucial gap, helping organizations navigate both technical and regulatory complexities. However, the success of such tools will likely depend on their ability to address the specific concerns of users at different levels of AI expertise, while also raising awareness of compliance requirements that might currently be underappreciated.

This study provides key insights into the personas of industrial AI stakeholders, their expertise, challenges, and approaches to compliance. The findings highlight significant gaps in AI guideline awareness and application, with nearly half of the respondents unsure whether their organizations follow formal AI standards. Moreover, the level of AI familiarity influences risk perception, with more experienced users recognizing algorithmic and resource-related challenges more frequently than less familiar users. Overall, our sample is represented mostly by engineers and managers from European AI projects with accordingly considerable knowledge of AI and ML. As the intended group of end-users of such a chatbot is, however, broader, it is necessary to consider further stakeholder groups such as administration or shop floor workers. Furthermore, by leveraging the Quantitative Effect and Technology Acceptance Model (QETAM), the study evaluates the acceptance of a chatbot designed to support AI compliance. the QETAM model's analysis reveals that the chatbot's moderate acceptance is primarily influenced by perceived trust and social influence, which are strongly correlated. Enhancing user trust through improved reliability and security measures, along with bolstering social advocacy, is essential for increased adoption. Additionally, the significant association between perceived ease of use and facilitating conditions underscores the need for a supportive environment, including comprehensive training and seamless integration into existing workflows, to promote effective utilization. Addressing these factors is crucial for elevating the chatbot's acceptance and effectiveness in AI compliance verification.

However, while our findings are valid for the sample analyzed, the number of respondents does not yet allow for sound generalization of the results to the entire population of interest. Therefore, future work will need to evaluate more responses from a more diverse sample of employees in the industry.

2.3.1 **Acknowledgements**

We would also like to thank Helmut Berrer, Miriam Groß, and Michaela Zalesak from Economica for their valuable advice in designing our questionnaire.

3. The Self Assessment tool: Let's talk about the EU AI Act

As artificial intelligence (AI) regulations evolve and the regulatory landscape develops and becomes more complex, ensuring compliance with ethical guidelines and legal frameworks remains a challenge for AI developers. This study introduces an AI-driven self-assessment chatbot designed to assist users in navigating the European Union AI Act and related standards. Leveraging a Retrieval-Augmented Generation (RAG) framework, the chatbot enables real-time, context-aware compliance verification by retrieving relevant regulatory texts and providing tailored guidance. By integrating both public and proprietary standards, it streamlines regulatory adherence, reduces complexity, and fosters responsible AI development. Theis section explores the chatbot's architecture, comparing naive and graph-based RAG models, and discusses its potential impact on AI governance.

3.1 The Problem Space

The rapid evolution of artificial intelligence (AI) technologies has enabled transformative applications across industries that are empowered by AI components and services. Nevertheless, it has also introduced significant challenges related to ethics, legal compliance, and governance. As global and regional regulations, such as the European Union's AI Act [13], gain traction, developers and publishers of industrial AI systems face increasing complexity in navigating diverse guidelines, standards, and legal frameworks. This complexity is compounded by the need to ensure ethical soundness, regulatory adherence, and conformance to national and international standards. Although there are tools based on Large Language Models (LLMs) that help the users to navigate these documents, such as ChatGPT, or tools provided by standardization organizations, these tools are either fundamental inquiry tools or very generic and do not cover copyrighted standards.

To address this challenge, we are building an innovative AI-driven self-assessment tool—a chatbot designed to assist AI developers and users in evaluating the compliance of their applications with ethical guidelines, legal regulations, and industry standards. Leveraging a Retrieval-Augmented Generation (RAG) framework, and hosted on cloud infrastructure, the chatbot provides real-time, context-aware guidance, allowing users to assess their AI systems efficiently. By facilitating the compliance verification process, this solution not only encourages a consistent approach to ethics and compliance but also reduces the time and expertise required to navigate complex regulatory landscapes. Furthermore, the chatbot enables developers and AI project managers to access not only public regulations but also company-owned standards and guidelines—whether these are proprietary regulations developed internally or industry standards the organization has acquired. The integration of these standards will allow users to browse through hundreds of regulations more easily and allows for organizations to maintain the recency of their standards by focusing the chatbot's local repository on the most up-to-date versions. By collecting all necessary documents in one system, the chatbot empowers developers and organizations to build AI systems that are not only innovative but also responsible and legally sound.

The users of this self-assessment tool will be able to interact with the chatbot and identify the relevant and applicable sections from regulatory frameworks that pertain to their AI-empowered applications. By engaging in a conversational interface, users can pose specific questions about their system's design, deployment, or use-case scenarios and receive tailored, easy-to-understand guidance drawn from a comprehensive repository of legal and ethical guidelines. This is enabled by the descriptive power of LLMs, which allow the chatbot to interpret user queries in nuanced ways, extract precise information from complex documents, and present it in a user-friendly manner.

This Section introduces the architecture, functionalities, and implications of our chatbot which is designed as a self-assessment tool for AI ethics and compliance. To the best of our knowledge, it is the first attempt to utilize LLMs and RAG to create a tool that navigates the regulatory landscape and facilitates compliance for the AI community. Additionally, we highlight key ethical considerations, user-centric design principles, and the tool's potential to foster responsible AI development.

3.2 The Regulatory Landscape

The regulatory landscape surrounding AI is rapidly evolving as governments and organizations worldwide work to address the ethical, legal, and societal challenges posed by AI technologies. This landscape is diverse and multifaceted, encompassing a range of document categories that vary in scope, nature, and focus areas. These categories include high-level frameworks and guidelines, such as the OECD AI Principles [14] and UNESCO's recommendations [15], which provide non-binding ethical and governance principles. Binding legislation and statutory laws, like the EU's AI Act [13] and the General Data Protection Regulation (GDPR) [16], the Digital Markets Act (DMA) [17], the Digital Services Act [18], and the Data Governance Act (DGA) [19], establish enforceable rights and obligations, focusing on areas such as data privacy, risk management, and accountability [20].

Technical standards, including those developed by ISO/IEC and IEEE, define operational specifications for AI systems, ensuring performance, interoperability, and safety [21]. Sector-specific regulations, such as FDA guidelines for AI in medical devices [22] or autonomous vehicle safety standards or the Cyber Resilience Act (CRA) [23], address unique challenges in critical domains (see Appendix 6.2). Additionally, national AI strategies (e.g., Hungary's AI strategy[1]) and policies set long-term goals for innovation and international competitiveness, while voluntary codes and best practices, often developed by industry leaders, foster ethical development without legal mandates.

With regulatory documents spanning ethical guidelines, binding legislation, technical standards, and sector-specific rules—each varying by jurisdiction, focus area, and binding nature—identifying the most relevant requirements can be a daunting task. The self-assessment tool proposed in this document addresses this challenge by enabling stakeholders to efficiently map their AI applications to the applicable regulations and standards. This not only saves time but also ensures comprehensive compliance, fostering responsible AI development while adapting to the dynamic nature of the regulatory ecosystem. By bridging the gap between complex regulatory frameworks and practical implementation, the self-assessment tool aims to become an essential asset in aligning AI innovations with ethical, legal, and technical benchmarks.

In their highly related work, Huang et al. [24] explored the ethical challenges arising from AI's rapid integration into areas like healthcare, finance, and governance. They highlight key concerns at individual (privacy, safety), societal (bias, accountability, democracy), and environmental levels (resource use, pollution). Reviewing 146 global guidelines, the authors identify core ethical principles: transparency, fairness, responsibility, non-maleficence, and privacy. They discuss approaches to mitigate these challenges, including technological solutions (e.g., explainable AI, fair algorithms), ethical frameworks (virtue ethics, deontology), and legal regulations (e.g., GDPR). Methods for evaluating AI ethics, such as the Moral Turing Test and international standards, are also examined. Their work stresses the need for global consensus,

[1] https://bitly.cx/HD59x

interdisciplinary efforts, and integrated ethical frameworks to ensure AI benefits society while mitigating harm.

Initially, Abu Naim et al. [4] proposed a self-assessment tool to facilitate the adoption of AI in industrial digital workspaces by addressing trust, ethical, and regulatory challenges. They outlined a four-step methodology: (1) defining personas based on expert interviews to categorize AI adopters into early adopters, followers, and independent learners; (2) mapping relevant AI regulations and ethical standards to ensure compliance; (3) fine-tuning LLMs to provide tailored guidance on AI implementation; and (4) developing an AI-powered self-assessment tool integrating profiling, memory, planning, and action modules. The tool aims to bridge the gap between AI innovation and legal compliance, enhancing user trust and promoting sustainable AI adoption in alignment with Industry 5.0 goals.

3.2.1 About the AI Act

The EU AI Act [13] is a pioneering regulatory framework that categorizes AI systems by risk level—ranging from minimal to unacceptable—and imposes corresponding obligations. High-risk AI, such as in healthcare or law enforcement, faces strict requirements for transparency, accountability, and data governance. Additionally, the EU AI Act is notable for its emphasis on fostering trust in AI technologies while maintaining a balance between innovation and regulation. By mandating practices like human oversight, bias mitigation, and robust documentation, the AI Act seeks to ensure AI systems operate ethically and safely. Beyond risk mitigation, the Act fosters trust in AI through mandates on human oversight, bias mitigation, and documentation. Its extraterritorial scope, akin to the GDPR, extends to entities outside the EU impacting EU citizens. Rooted in European values of fundamental rights and fairness, the Act sets a global precedent for ethical AI governance [25] .

3.3 The Chatbot

3.3.1 System Design and Architecture

Modern, widely adopted LLMs like GPT-4 [26] are well-suited for general-purpose use and text comprehension, as they are trained on an extensive and broad knowledge base. However, this is often not sufficient when accurate and up-to-date information is required. Modern language models often grapple with challenges like limited access to the (i) latest information, (ii) insufficient expertise in specific domains or tasks, and (iii) a tendency to generate inaccurate or fabricated responses due to their probabilistic design.

The architecture we have designed seeks to address or mitigate these issues through the implementation of a Retrieval-Augmented Generation (RAG) framework. This approach ensures that, during the process of answer generation, the LLM is provided with (i) up-to-date and (ii) domain-specific information relevant to the query. By doing so, it significantly reduces (iii) the likelihood of hallucinations and enhances the accuracy of the generated responses.

In general, the implementation of a RAG system consists of two components. In a preparatory phase, referred to as Indexing, we digitize the regulatory documents (like the EU AI Act). During this phase, information is extracted from the documents, and a database is constructed. Later, in the RAG phase, elements from this database that resemble the information implied by the user's intent are retrieved and provided to the context window of the LLM. The LLM then interprets this information and generates an appropriate response.

Various types of RAG models are being developed, each exhibiting varying levels of performance depending on the nature of the data and the query derived from it. Initially, we implemented

the most basic form, referred to as the Naive RAG [27], to explore the capabilities of a basic approach to the problem and in order to have a benchmark for comparing the results of other RAG types.

In the next phase of the project, we will implement a more structured, hierarchical model called Graph RAG [28], to assess whether it can show a significant improvement compared to the naive RAG.

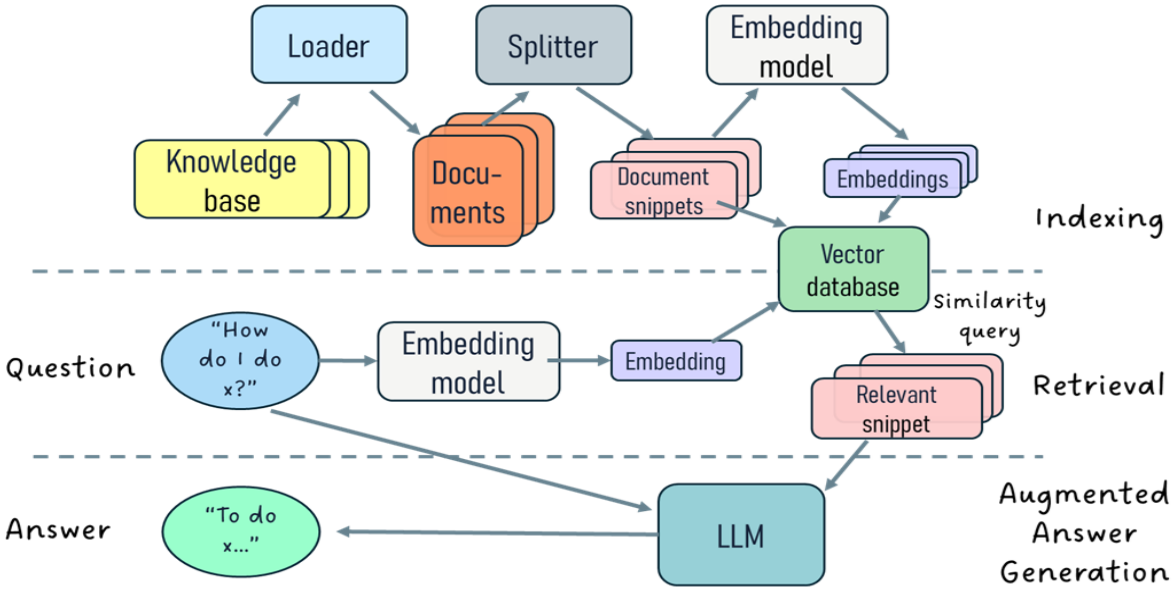

Figure 6: Overview of the Naïve RAG

3.3.2 Naive RAG

3.3.2.1 Indexing

The input to the indexing phase is a document-like file. Using various document processing techniques, like Optical Character Recognition (OCR), we extract text, tables, and visual content, making it possible to perform multimodal document processing. This approach contrasts with traditional indexing methods that only handle textual content, ensuring that valuable information derived from table structures and visual elements is preserved. The extracted information elements are then summarized and transformed into vectors using an embedding model and stored to perform semantic similarity search during retrieval. Figure 6 illustrates the entire process in the upper Indexing section.

3.3.2.2 Retrieval-Augmented Generation

In the context of RAG, the user submits a question or provides instruction in textual form. This input is similarly converted into a vector using the same embedding model utilized during the indexing phase. A similarity search is then performed against the vector database containing vector representations of the indexed document segments. The retrieved document segments are subsequently provided to the LLM, along with the user's intent, to facilitate response generation.

3.3.2.3 Functionality

The system implementing the naive RAG that we developed:

- Is capable of processing and integrating multimodal documents into the knowledge base regardless of their length, based on the Table of Contents of the documents.

- During chat, it interprets the user intent, searches for similarities in the knowledge base based on semantic analysis, and formulates a response by interpreting the resulting document segments, including source attribution.

The operation of a naive RAG is based on the semantic similarity between the user query and the information extracted from the processed documents, therefore, it is quite effective when querying specific expressions within a relatively simple context.

Manual evaluation of answer authenticity: To assess the reliability of the chatbot's responses, we plan to validate its outputs against labeled ground-truth answers. One approach involves developing a structured questionnaire based on the EU AI Act. This questionnaire will help determine whether a given use case falls under the AI Act's definition of AI, assess its risk classification (e.g., high-risk, low-risk), and identify the applicable regulatory requirements. Our legal experts will provide authoritative answers to this questionnaire based on the AI Act. We will then compare the chatbot's responses with those provided by the legal experts to measure accuracy and alignment with the official regulatory framework.

Industry use-case-based evaluation: Another evaluation approach involves collaboration with AI developers from the industry who have built AI-powered applications [29]. These developers will provide detailed descriptions of their use cases, which will be reviewed by our legal experts. The experts will identify the most relevant regulatory documents that need to be considered for each use case. The chatbot's recommendations will then be compared against the expert-identified regulations to assess its effectiveness in real-world scenarios.

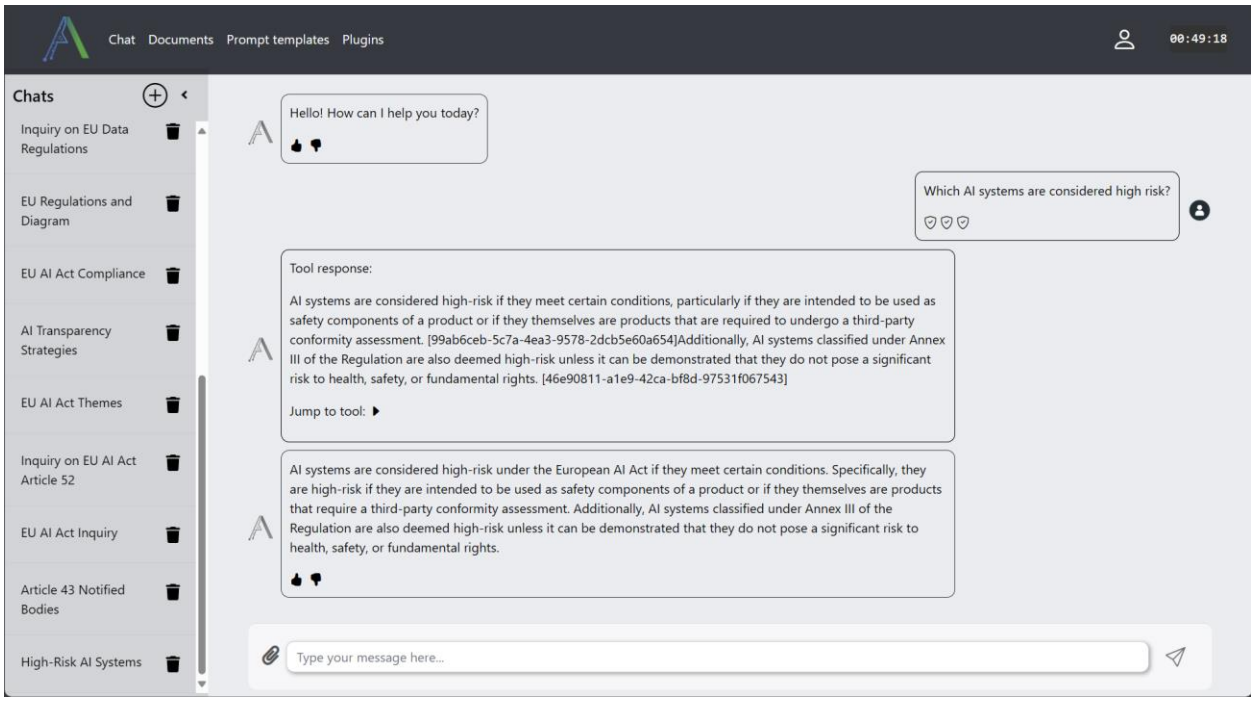

Figure 7: Simple question/answer pair from the interface of our developed chatbot

The example shown in the Figure 7 clearly demonstrates that the naive RAG successfully identifies semantic similarities. In response to the simple conceptual question "Which AI systems are considered high risk?", it correctly identified the corresponding definitional section of the AI Act, accurately interpreted both the user query and the text segment containing the answer, and then generated the response.\looseness=-1
While this approach can yield satisfactory results in certain cases, it may not be suitable for more complex scenarios [30], i.e.:

- When answering a question requires uncovering the relationships between disparate pieces of information based on their shared attributes in order to provide new, synthesized insights.

- When a holistic understanding of summarized semantic concepts in large datasets or documents is required.

3.3.3 **Graph RAG**

In contrast to the naive RAG, the Graph RAG does not search for semantic similarities in vector databases. Instead, it constructs knowledge graphs from related and unrelated data and their relationships from the source documents, which could enable a deeper level of interpretation of the source documents and with more complex questions, such as:

- List the 8 areas included in the annex referenced in the definition of high-risk AI systems.
- Summarize the 5 most important topics of the AI Act!

Where semantic matching is not feasible, significant advancements can be achieved in answer generation [31].

While Naive RAG retrieves information chunks based on keyword similarity, which can lead to missing essential information chunks, the knowledge graph structure used by Graph RAG retrieves information not just based on semantic similarity but also on relationships between entities, ensuring more relevant and contextually rich retrieval.

Naive RAG processes documents as a collection of independent chunks, which may result in fragmented and less coherent responses. In contrast, Graph RAG enables reasoning over connected nodes, helping the model understand how different pieces of information relate, leading to more coherent and accurate answers. This capability also enables multi-hop retrieval, meaning the system can follow connections between related nodes, allowing it to access information that leads to a more insightful response—information that a traditional RAG model might have overlooked.

Naive RAG provides answers without offering much insight into why a particular piece of text was retrieved. In contrast, Graph RAG can trace responses back to structured knowledge, ensuring greater transparency and explainability in the answers.

3.3.3.1 Indexing

According to the indexing pipeline of Graph RAG model, as Figure 8 depicts, the processed documents must first be divided into chunks of a specified token size to ensure they are of manageable dimensions. This is necessary because, in the next step, an LLM is used to identify the entities, their relationships, and claims present in the source document required to construct the knowledge graphs.

First, all entities in the text should be identified, including their name, type, and description. Following this, all relationships between clearly related entities should be identified, including the source and target entities and a description of their relationship. Next, claims should be extracted and linked to the detected entities, such as the subject, object, type, description, source text span, start- and end-dates. This process may require multiple iterations of LLM processing, and the definition of extraction prompts should take into account domain-specific features to ensure higher-quality results.

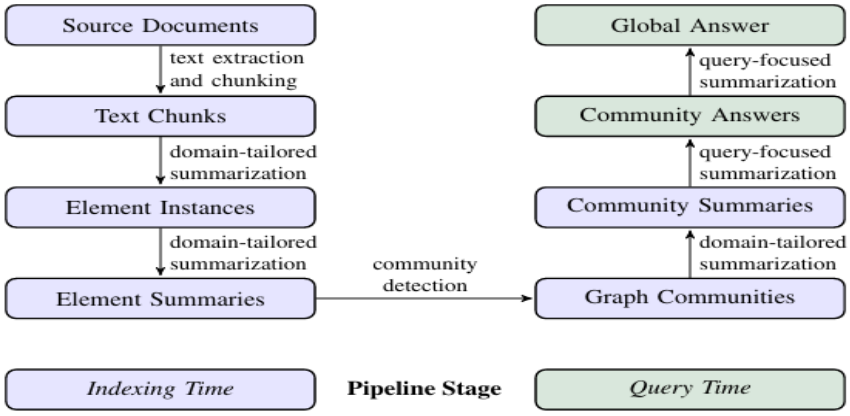

Figure 8: Graph-based RAG

In the second step, the instance-level summaries created must be transformed into a single block of text for each graph element (i.e., entity node, relationship edge, and claim-covariant). This requires an additional summarization round using the LLM for the corresponding instance groups.

The resulting index can be modeled as a homogeneous, undirected, weighted graph, where the nodes represent entities, the edges represent the relationships between them, and the weights on the edges denote the count of these relationships. Given such a graph, a range of community detection algorithms may be employed to partition the graph into communities of nodes with stronger interconnections within the community than with nodes outside of it [31].

As the final step of the indexing phase, summaries must be created for the communities formed in the previous step, starting from the bottom of the hierarchy as follows:

- Leaf-level communities: The summaries of connected elements should be ordered in descending order on the basis of the degree of the source and target nodes. Then, iteratively, the context window is expanded by adding descriptions of the source, target, connected covariants, and edge.
- Higher-level communities: If the summary of each element fits within the context window, the same procedure as the previous case should be applied. If not, the lower-level communities should be sorted in descending order based on the sum of the token counts of their element summaries and replace the summaries of the community elements (longer) with the community summary (smaller) [31]

The summaries of the communities are standalone and contain comprehensive semantic information of all the content of the elements lower in the hierarchy. If the user query asks for more detailed information, the relevant information can be found following the connection to the element to a deeper level of the hierarchy.

3.3.3.2 Retrieval-Augmented Generation

Once the knowledge graph has been constructed, graph RAG ensures inquiry in two ways, depending on the type of information requested by the user query. If it concerns the holistic understanding of the whole data corpus, then global search in Figure 9 should be applied. If the user query is focused on interpreting specific entities and the relationships between them, then local search in Figure 10 should be used.

- Global Search: the community summaries on a certain level should be random shuffled and divided into smaller chunks (to precede information loss), then intermediate answers are generated based on each chunk and these answers are rated by helpfulness between 0-100. In the final step these answers are sorted in descending order and injected into the context window until the token limit is reached to generate the final answer.

- Local Search: firstly, entities are identified from the knowledge base based on semantic similarity to the user query. Subsequently, the entities, relationships, covariates, document segments, and community reports associated with the selected entities must be retrieved and prioritized. Based on the most relevant results, the final response can then be generated.

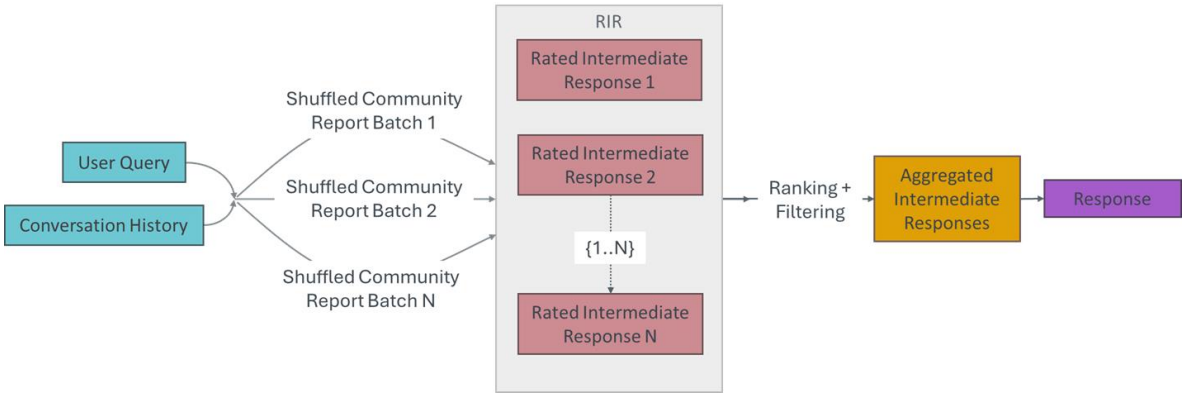

Figure 9: Global-search

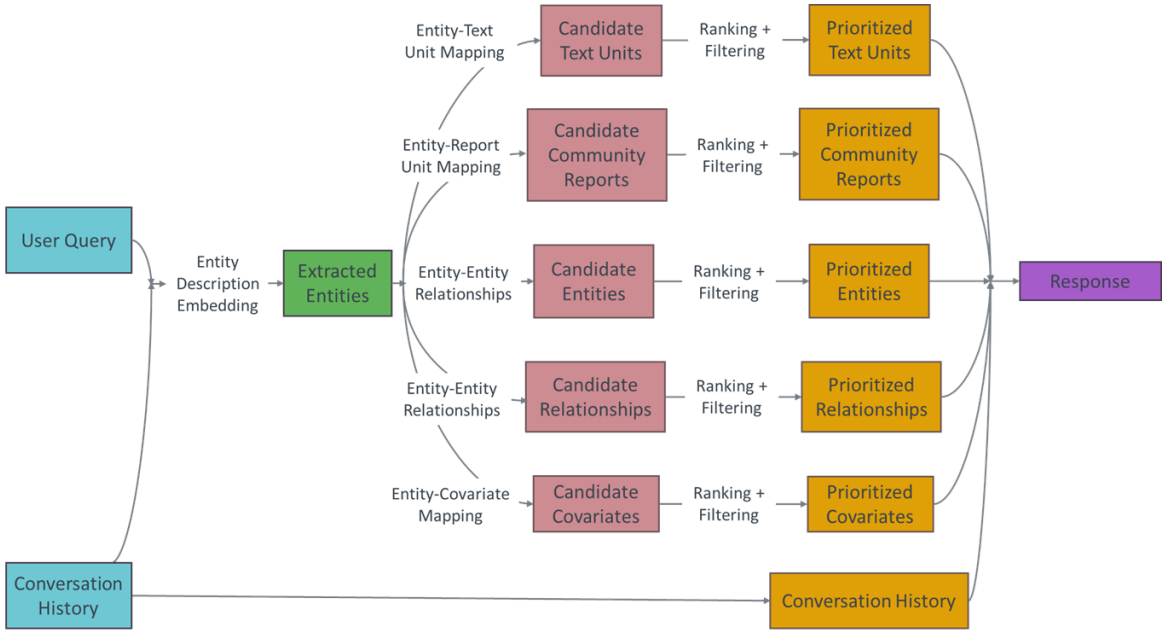

Figure 10: Local-search

3.3.3.3 Expectations

Graph RAG enables substantial improvements in both comprehensiveness and diversity when compared to naive RAG models. Furthermore, in datasets or task-specific domains where global queries are more common, Graph RAG demonstrates a clear increase over naive RAG models, while achieving competitive results compared to other global models, all with reduced token consumption [31]. We aim to further examine these aspects to generate increasingly accurate and effective responses within our self-assessment tool.

All of this with the aim of assisting AI-related developers and engineers in their work, so that they are aware of the legal and ethical regulations that arise during development in relation to questions that arise, for example, from our industrial partner, who is supporting and strengthening human and manufacturing cycles through the development of AMRs (Autonomous Mobile Robots) and exoskeletons:

- "What are the key safety standards and guidelines for AMRs with human detection systems to ensure safe interaction between robots and human workers in a factory setting?"
- "Are there ethical guidelines on the use of cameras in the workplace, specifically for human detection systems in AMRs, to prevent surveillance or misuse of data?"

3.4 Methodology for evaluation & fine tunning

3.4.1 EU AI Act Checklist

The evaluate our self-assessment tool, we have developed a checklist with our legal experts in the project. This checklist comprises 50 questions that will examine at first if an application with embedded AI falls into the definition of AI in the AI Act and later determines with category of AI (prohibited, high risk, etc.) the application falls into and accordingly provides the respective articles for it.

The answers to this checklist are provided by an AI Act expert. We aim to use this checklist to evaluate our self-assessment tool. Since we have the ground truth answer to this checklist, we plan to ask the chatbot with the same questions of the checklist and then compare the answers given by the chatbot to the answers provided by our expert.

3.4.2 Use Case Stories

To complement the structured checklist, we will also test the chatbot against realistic industrial AI use cases in AIMS5.0 to evaluate how well it performs in practical scenarios. To this aim, we have provided a set of questions to the developers of our industrial use cases and their owners detailing the kind of AI that is being developed in the use cases. Additionally, the following questions have been asked:

If you had access to a standard's expert, what questions would you ask or what knowledge are you interested in getting in your use case regarding standards guidelines and regulations?
Ethical Compliance Checklist:

- *Does the project store or process sensitive personal data (e.g. address, salary, gender ...) or private information (social networks, hobbies ...) of the employee?*
- *Does the project enable or expand performance and behavior monitoring (e.g. cameras at workstations, reporting, time recording...)?*
- *Does the project have an impact on the employee's working hours/work content and/or work organization (work intensification, underload/overload, shifting of work, new or different tasks, monotony of tasks, transfer to another position if necessary...)?*
- *Does the project have an impact on occupational health and safety (working under full protection, heavy lifting, noise pollution, psychological risks, etc.)?*
- *Is the use of artificial intelligence (AI) and/or algorithms that can be used for the use case and may have an impact on employees, planned?*

Use cases have provided this information and the answer to these questions. Then we will provide this information per use case to the self-assessment tool and at the same time ask our experts in the legal and ethical team to see if the answers provided for each use cases are accurate and complete and determine what is missing and what needs to be improved.

3.5 Future Direction and Conclusion

This deliverable introduced an AI-driven self-assessment chatbot designed to support AI developers and stakeholders in navigating the evolving regulatory landscape. The current instance of our self-assessment tool covers the EU AI Act in particular as it is the first binding regulation in the EU. By leveraging a Retrieval-Augmented Generation (RAG) framework, the chatbot enhances compliance verification by retrieving and interpreting relevant legal documents, reducing complexity, and ensuring adherence to ethical and regulatory standards.

Our implementation of both Naive RAG and Graph RAG architecture demonstrates the potential for improving AI governance tools. While the Naive RAG provides effective semantic matching for straightforward inquiries, the Graph RAG introduces a structured knowledge representation that facilitates deeper regulatory interpretation and more complex query resolution. The chatbot's ability to integrate proprietary and public regulations further enhances its value for organizations managing AI compliance.

Looking ahead, we anticipate three key developments in our work. 1) Regulatory Coverage: As regulatory landscapes evolve, we aim to expand our chatbot's scope to encompass a wider range of regulations. This will not only ensure that user queries are addressed in compliance with binding regulations but also highlight the most critical documents users should be aware of. By integrating a more comprehensive set of legal frameworks, our chatbot will provide more effective guidance on compliance requirements. 2) Checklists as a complementary input: We are developing legal and ethical checklists based on existing regulations and guidelines. These checklists will be presented to users at the beginning of their interaction with the self-assessment tool. Users' responses to these checklists will serve as a complement to the prompts they submit to the chatbot, enhancing the chatbot's ability to provide precise and contextually relevant guidance. This approach ensures that the tool delivers more comprehensive and

tailored support, helping users navigate complex regulatory environments with greater accuracy and confidence. 3) Expand our implementation of the self-assessment tool by integrating it with Graph RAG, enabling precise answers to more complex prompts.

4. Outreach to the Community

4.1 Papers

#	Title	Journal / Workshop / Conference	Authors
1	Engineering Secure, Trustworthy, and Ethically Sound AI-Based Computer Systems	Magazine (ECRIM News Software Security Special Theme)	Yasin Ghafourian, Markus Tauber, Germar Schneider, Andrea Bannert, Olga Kattan, Erwin Schoitsch
2	Industry 5.0: Research Areas and Challenges With Artificial Intelligence and Human Acceptance	Journal (IEEE Industrial Electronics Magazine)	George Dimitrakopoulos, Pal Varga, Thomas Gutt, Germar Schneider; Hans Ehm, Alfred Hoess, Markus Tauber, Konstantina Karathanasopoulou, Anna Lackner, Jerker Delsing
3	Let's have a Chat with the EU AI Act	Conference (Proceedings of the 6th Workshop on Management for Industry 5.0 (MFI5.0), at IFIP/IEEE NOMS 2025 - Network Operations and Management 2025 Symposium)	Ádám Kővári, Yasin Ghafourian, Csaba Hegedűs, Belal Abu Naim, Kitti Mezei, Pál Varga, and Markus Tauber
4	Understanding Stakeholders of Industrial AI: Insights from a Persona-Based Questionnaire	Conference (Proceedings of the 6th Workshop on Management for Industry 5.0 (MFI5.0), at IFIP/IEEE NOMS 2025 - Network Operations and Management 2025 Symposium)	Yasin Ghafourian, Fabian Lindner, Olga Kattan, Konstantina Karathanasopoulou
5	MULTI-PARTNER PROJECT: ARTIFICIAL INTELLIGENCE IN MANUFACTURING LEADING TO SUSTAINABILITY AND THE CONSIDERATION OF HUMAN ASPECTS (AIMS5.0)	Proceedings of the Design, Automation and Test in Europe (DATE) Conference 2025	Anouar Nechi, Yasin Ghafourian, Belal Abu Naim, Thomas Gutt, Georgios Dimitrakopoulos, Amira Moualhi, Mladen Berekovic, Pal Varga and Markus Tauber

4.2 Posters or oral presentations at relevant events

#	Title	Name of the event
1	AIMS5.0 WP7 Poster presentation	CPS Summer School in Sardinia, September 2024
2	AIMS5.0 WP7 Poster presentation	European Conference on EDGE AI Technologies and Applications, October 2024
3	AIMS5.0 WP7 Poster presentation	Deep Tech Workshop in Budapest, November 2024
4	AIMS5.0 WP7 Poster presentation	ECS Summer school of 3 industry associations 2024

4.3 Dedicated AIMS5.0 WP7 workshops

4.3.1 **Workshop with the PAL project at Zittau/Görlitz at 09th.April 2025**

We have organized an inter-project workshop between PAL and AIMS-WP7 on workspaces, acceptance, and AI in an industrial context. The workshop took place in person on April 9 in Zittau, Germany. This event provided an opportunity for participants to present their work, engage in discussions, and exchange ideas.

5. Conclusions

Deliverable D7.2 marks a significant milestone in the AIMS5.0 project by presenting the outcomes of Tasks T7.1, T7.2, and laying the results obtained for T7.3—all centered around fostering human acceptance, trust, and ethical integration of AI in digital workspaces aligned with Industry 5.0 principles.

In **Task 7.1**, we developed detailed stakeholder personas to better understand the diverse backgrounds, expertise levels, and industrial contexts of AI practitioners. This persona-driven insight was critical for shaping a user-centric approach to AI compliance.

Task 7.2 extended this work by translating those insights into the development and early validation of a self-assessment tool—an AI-powered chatbot—designed to help users navigate complex regulatory environments, particularly the EU AI Act. Built using Retrieval-Augmented Generation (RAG), both Naive and Graph-based models were explored, enabling intuitive, context-aware compliance support.

The tool's evaluation, informed by the Quantitative Effect and Technology Acceptance Model (QETAM), revealed moderate user acceptance, with perceived trust and ease of use being key adoption drivers. These findings provide essential input for **Task 7.3**, which focuses on further enhancing the chatbot through integration with legal expert-defined checklists and real-world industrial use cases.

While results show promising potential, they also highlight the importance of expanding outreach to additional stakeholder groups beyond the AI-savvy participants of our survey. Moreover, increasing the regulatory breadth and improving chatbot usability and reliability will be critical next steps.

In sum, this deliverable combines empirical insights with technical innovation to advance responsible and effective AI integration in manufacturing. It not only reflects the interdisciplinary goals of AIMS5.0 but also sets the stage for future improvements in regulatory compliance tools, ensuring alignment with evolving standards and user expectations across Europe's industrial ecosystem.

6. Appendix

6.1 Annex A: Abbreviations

Table 2: Abbreviations

Abbreviation	Meaning
AIMS5.0	Artificial Intelligence in Manufacturing leading to Sustainability and Industry5.0
TAM	Technology Acceptance Model
QETAM	Quantitative Effect and Technology Acceptance Model
RAG	Retrieval-Augmented Generation
LLM	Large Language Model
FMEA	Failure Mode and Effects Analysis
UTAUT	Unified Theory of Acceptance and Use of Technology
TOE	Technology-Organization-Environment (framework)
AI Act	European Union Artificial Intelligence Act
GDPR	General Data Protection Regulation
DMA	Digital Markets Act
DSA	Digital Services Act
DGA	Data Governance Act
CRA	Cyber Resilience Act
OECD	Organization for Economic Co-operation and Development
UNESCO	United Nations Educational, Scientific and Cultural Organization
FDA	U.S. Food and Drug Administration
PU	Perceived Usefulness
PEU	Perceived Ease of Use
PT	Perceived Trust
SI	Social Influence
FC	Facilitating Conditions
BIU	Behavioral Intention to Use
SMEs	Small and Medium-sized Enterprises

6.2 Annex B: Overview of WP 7 - related AI-Standards

Appendix 6.2 provides some additional background information to chapter 3.2 “The Regulatory Landscape”, addressing standards which are relevant in context of the topics addressed by WP 7 (public acceptance of AI – “Trustworthiness”, safety and security), the AI-Act, and specific man-machine collaboration (“teaming”) standards (AI, Robotics, Smart manufacturing). Of particular interest are evolving standards (“under development”), because in these cases we have on the one hand a “look ahead” on upcoming standards and requirements, on the other hand appropriate results and experiences from a research project could be brought forward to standardization committees and groups.

6.2.1 AI: ethical, social, governance, oversight and human aspects impact

Published:

- ISO/IEC TR 24368:2022 Information technology — Artificial Intelligence (AI) — Overview of ethical and societal concerns
- ISO/IEC 5259-5:2025 Artificial intelligence — Data quality for analytics and machine learning (ML) — Part 5: Data quality governance framework
- *ISO/IEC TS 12791:2024 Information technology — Artificial intelligence — Treatment of unwanted bias in classification and regression machine learning tasks*
- *ISO/IEC TR 24027:2021 Information technology — Artificial Intelligence (AI) — Bias in AI systems and AI aided decision making*
- ISO/IEC 38507:2022 Information technology — Governance of IT — Governance implications of the use of artificial intelligence by organizations

Under development:

- ISO/IEC CD TS 22443 Information technology — Artificial intelligence— Guidance on addressing societal concerns and ethical considerations.
- ISO/IEC AWI TS 25571 - Artificial Intelligence — Example template for documenting ethical issues of an AI system
- ISO/IEC AWI TS 25568 - Information technology — Artificial Intelligence — Guidance on addressing risks in generative AI systems
- ISO/IEC CD 42105 (under development) - Information technology — AI— Guidance for human oversight of AI system
- **NEW PWI: Governance and management of human oversight of AI systems (a PWI is not listed on the official ISO webpage as “standard under development”, this is only starting with AWI), it is only known to insiders (members of the WGs) (in this case AIT).**

6.2.2 Trustworthiness, Safety, Reliability, Robustness, Controllability, and Risk management

Published:

- ISO/IEC TR 5469:2024 Artificial intelligence — Functional safety and AI systems
- ISO/IEC TS 8200:2024 Information technology — Artificial intelligence — Controllability of automated artificial intelligence systems
- *ISO/IEC TS 12791:2024 Information technology — Artificial intelligence — Treatment of unwanted bias in classification and regression machine learning tasks*
- ISO/IEC 23894:2023 Information technology — Artificial Intelligence — Guidance on Risk Management
- *ISO/IEC TR 24027:2021 Information technology — Artificial Intelligence (AI) — Bias in AI systems and AI aided decision making*

- ISO/IEC TR 24028:2020 Information technology — Artificial Intelligence — Overview of trustworthiness in artificial intelligence
- ISO/IEC TR 24029-1:2021 Artificial Intelligence — Assessment of the robustness of neural networks — Part 1: Overview
- ISO/IEC 24029-2: 2023 Artificial Intelligence — Assessment of the robustness of neural networks — Part 2: Methodology for the use of formal methods

Under development (highlighted in red: started last year):

- ISO/IEC CD TR 21221 Information technology - Artificial intelligence - Beneficial AI systems
- ISO/IEC AWI TS 22440-1 - Artificial intelligence — Functional safety and AI systems — Part 1: Requirements
- ISO/IEC AWI TS 22440-2 - Artificial intelligence — Functional safety and AI systems — Part 2: Guidance
- ISO/IEC AWI TS 22440-3 - Artificial intelligence — Functional safety and AI systems — Part 3: Examples of application
- ISO/IEC AWI TS 25223 Information technology - Artificial intelligence - Guidance and requirements for uncertainty quantification in AI systems
- ISO/IEC AWI 24029-3 Artificial intelligence (AI) — Assessment of the robustness of neural networks — Part 3: Methodology for the use of statistical methods
- ISO/IEC AWI TS 25570 - Information Technology — Artificial Intelligence — Reliability assessment of AI systems
- ISO/IEC AWI TR 42109 Information technology — Artificial intelligence —Use cases of human-machine teaming

6.2.3 EC Standardization Requests for the AI-Act

The EC AI-Act which is particularly regulating the human-centred, society and public interests to protect the citizens against violation of their rights and misuse of powerful AI applications. Harmonized standards are the basis for suppliers and users to prove “presumption of conformity”. The EC has sent 10 requests for standardisation to CEN/CENELEC for that purpose, which are now handled in CEN/CLC JTC 21, “Artificial Intelligence”. Standardization members of AIMS5.0 are active at least on national level in the mirror Committees, e.g., AIT in ASI K001 AG42.

The **European Commission’s AI Act** served as a foundational guide, outlining ten critical requirements for **high-risk AI systems**. These include:

1. Risk management system for AI systems [Art. 9 AI-Act]
2. Governance and quality of datasets used to build AI systems [Art. 10 Data and data governance]
3. Record keeping - built-in logging capabilities in AI systems [Art. 11 Techn. doc., Art. 12 record-keeping]
4. Transparency and information to the users of AI systems [Art. 13 Transparency and user info]
5. Human oversight of AI systems [Art. 14 Human oversight]
6. Accuracy specifications for AI systems [Art. 15 Accuracy, robustness and cybersecurity]
7. Robustness specifications for AI systems [Art. 15 Accuracy, robustness and cybersecurity]

8. Cybersecurity specifications for AI systems [Art. 15 Accuracy, robustness and cybersecurity]
9. Quality management system for providers of AI system [Art. 17 QM for providers]
10. Conformity assessment for AI systems [Art. 19 + Art. 43 Conformity Assessment]

These requirements highlight the need for harmonized standards to support compliance and ensure the safe and responsible deployment of AI technologies across Europe.

6.2.4 Domain-specific AI-related/human-centered Standards (Examples)

Automotive:

ISO/TS 5083 - Road vehicles — Safety for automated driving systems — Design, verification and validation (under publication, will replace ISO/TR 4804, Road vehicles — Safety and cybersecurity for automated driving systems — Design, verification and validation).

ISO/PAS 8800 - Road vehicles – Safety and artificial intelligence.

Robotics (new examples):

ISO/CD 21423 - Robotics - Autonomous mobile robots for industrial environments -- Communications and Interoperability (under development).

ISO_NP 25785-1_(ed1) - Safety requirements for industrial mobile robots with actively controlled stability (legged, wheeled, or other forms of locomotion) Part 1: robots (under development).

ISO/CD 18646-7 Robotics — Performance criteria and related test methods for service robots, Part 7: Travelling around humans (under development).

7. References

[1] K. Karathanasopoulou, E. Tsaousi, D. Mitsiogianni, G. Dimitrakopoulos, and D. Georgiadis, "Technology Acceptance Modelling for Investigating the Uptake of Electric, Connected, Autonomous, and Shared Mobility Technologies," in *Proceedings of the 11th International Conference on Vehicle Technology and Intelligent Transport Systems (VEHITS)*, 2025.

[2] F. D. Davis, "Perceived usefulness, perceived ease of use, and user acceptance of information technology," *MIS quarterly*, pp. 319–340, 1989.

[3] F. Lindner, K. A. Przybysz, G. Schneider, and S. Keil, "Workers' Attitudes Towards Digital Transformation and Perceived Well-Being At Work: A Case Study," *ACC JOURNAL*, vol. 30, no. 2, pp. 47–62, 2024, doi: 10.2478/acc-2024-0006.

[4] B. A. Naim *et al.*, "A Self-assessment Tool to Encourage the Uptake of Artificial Intelligence in Digital Workspaces," in *NOMS 2024-2024 IEEE Network Operations and Management Symposium*, 2024, pp. 1–5.

[5] J. Kutz, J. Neuhüttler, J. Spilski, and T. Lachmann, "Implementation of AI Technologies in manufacturing - success factors and challenges," *The Human Side of Service Engineering*, vol. 62, no. 62, 2022, doi: 10.54941/ahfe1002565.

[6] P. Somer and S. Thalmann, "Risk Management of AI in Industry: A Literature Review," in *The 15th Mediterranean Conference on Information Systems (MCIS) and the 6th Middle East & North Africa Conference on Digital Information Systems (MENACIS)*, 2023.

[7] H. A. Alfadda and H. S. Mahdi, "Measuring students' use of zoom application in language course based on the technology acceptance model (TAM)," *Journal of Psycholinguistic Research*, vol. 50, no. 4, pp. 883–900, 2021.

[8] S. A. Kamal, M. Shafiq, and P. Kakria, "Investigating acceptance of telemedicine services through an extended technology acceptance model (TAM)," *Technology in Society*, vol. 60, p. 101212, 2020.

[9] S. Chatterjee, N. P. Rana, Y. K. Dwivedi, and A. M. Baabdullah, "Understanding AI adoption in manufacturing and production firms using an integrated TAM-TOE model," *Technological Forecasting and Social Change*, vol. 170, p. 120880, 2021.

[10] M. Altalhi, "Toward a model for acceptance of MOOCs in higher education: the modified UTAUT model for Saudi Arabia," *Education and Information Technologies*, vol. 26, no. 2, pp. 1589–1605, 2021.

[11] N. Puspitasari, M. B. Firdaus, C. A. Haris, and H. J. Setyadi, "An application of the UTAUT model for analysis of adoption of integrated license service information system," *Procedia Computer Science*, vol. 161, pp. 57–65, 2019.

[12] E. A. Armutak and M. Fendri, *Unlocking Value from Artificial Intelligence in Manufacturing*.

[13] European Commission, *The EU Artificial Intelligence Act*.

[14] Organisation for Economic Co-operation and Development, *OECD AI Principles*. [Online]. Available: https://oecd.ai/en/ai-principles

[15] S. United Nations Educational and Cultural Organization, *Recommendation on the Ethics of Artificial Intelligence*. [Online]. Available: https://www.unesco.org/en/articles/recommendation-ethics-artificial-intelligence

[16] European Union, *General Data Protection Regulation (GDPR)*. [Online]. Available: https://gdpr-info.eu/

[17] European Union, *Digital Markets Act (DMA)*. [Online]. Available: https://digital-markets-act.ec.europa.eu/

[18] European Union, *Digital Services Act (DSA)*. [Online]. Available: https://digital-strategy.ec.europa.eu/en/policies/digital-services-act-package

[19] European Union, *Data Governance Act*. [Online]. Available: https://digital-strategy.ec.europa.eu/en/policies/data-governance-act

[20] C. Novelli, F. Casolari, P. Hacker, G. Spedicato, and L. Floridi, "Generative AI in EU law: Liability, privacy, intellectual property, and cybersecurity," *Computer Law & Security Review*, vol. 55, p. 106066, 2024.

[21] M. Cantero Gamito and C. T. Marsden, "Artificial intelligence co-regulation? The role of standards in the EU AI Act," *International journal of law and information technology*, vol. 32, no. 1, eaae011, 2024.

[22] M. Aboy, T. Minssen, and E. Vayena, "Navigating the EU AI Act: implications for regulated digital medical products," *npj Digital Medicine*, vol. 7, no. 1, p. 237, 2024.

[23] European Union, *Cyber Resilience Act*. [Online]. Available: https://digital-strategy.ec.europa.eu/en/policies/cyber-resilience-act

[24] C. Huang, Z. Zhang, B. Mao, and X. Yao, "An Overview of Artificial Intelligence Ethics," *IEEE Transactions on Artificial Intelligence*, vol. 4, no. 4, pp. 799–819, 2023, doi: 10.1109/TAI.2022.3194503.

[25] G. de Gregorio and P. Dunn, "The European risk-based approaches: Connecting constitutional dots in the digital age," *Common Market Law Review*, vol. 59, no. 2, 2022.

[26] OpenAI, *ChatGPT: Optimizing Language Models for Dialogue*.

[27] Ivan Ilin, *Advanced RAG Techniques: an Illustrated Overview*.

[28] Microsoft Research, *GraphRAG: A Graph-Based Retrieval-Augmented Generation System*.

[29] R. AlMahasneh, G. Hollósi, D. Ficzere, M. Bancsics, C. Lukovszki, and P. Varga, "Uncovering Common AI Challenges Across Industrial Domains in the Transition to Industry 5.0," in *2024 20th International Conference on Network and Service Management (CNSM)*, 2024, pp. 1–7.

[30] J. Larson and S. Truitt, *GraphRAG: Unlocking LLM discovery on narrative private data*.

[31] Darren Edge *et al., From Local to Global: A Graph RAG Approach to Query-Focused Summarization*.