
**Foundations of Trustworthy AI – Integrating Reasoning, Learning and Optimization**
**TAILOR**
**Grant Agreement Number 952215**

# Progress Report on SRIR Development in Collaboration with the ICT-48 CSA, PPP-AI, and AI4EU – v2

| | |
|---|---|
| **Document type (nature)** | Report |
| **Deliverable No** | D11.6 |
| **Work package number(s)** | WP2 |
| **Date** | Due M44, April 2024 => extended to M48 |
| **Responsible Beneficiary** | UNIBO |
| **Author(s)** | Roberta Calegari (UNIBO), Michela Milano (UNIBO) |
| **Publicity level** | Public |
| **Short description** | Description of the collaboration process with ICT-48 CSA, PPP-AI, and AI4EU that led to the finalisation of the SRIR |

| **History** | | | |
|---|---|---|---|
| **Revision** | **Date** | **Modification** | **Author** |
| Version 1 | 2024-09-10 | - | Roberta Calegari, Michela Milano |

| **Document Review** | | |
|---|---|---|
| **Reviewer** | **Partner ID / Acronym** | **Date of report approval** |
| Marc Schoenauer | #3, INRIA | 2024-09-10 |
| Annelot Bosman | #7, ULEI | 2024-09-10 |

## Table of Contents

## Summary of the report

This report outlines the processes and strategies employed to engage strategic European AI initiatives in the formulation and finalisation of the TAILOR Strategic Research and Innovation Roadmap (SRIR version 2). This final version of the SRIR aims to contribute to a cohesive ecosystem by fostering collaboration and coordination among various initiatives, ensuring the roadmaps cover diverse aspects while maintaining alignment.

Summary of the Report
The report details the following key elements:

- **Revisiting Previous Deliverables**: the report recalls the engagement process described in Deliverable 11.5 and how its concepts have been practically implemented to collect feedback for the current and final version of the SRIR. Additionally, gaps highlighted in the first version of the SRIR are reported, explaining the actions taken to address those comments.
- **Development of the Final SRIR**: it discusses the evolution from the first version of the SRIR (Deliverable D2.1) to the final version (Deliverable D2.5).
  The first version was primarily based on insights from academic research partners within the TAILOR project. The second version of the SRIR addresses the comments and feedback from reviewers of the first version. It expands the roadmap by incorporating necessary research topics and priorities from industrial domains. It operationalizes these collected research topics into a guiding roadmap, combining and mapping them with respective priorities.
- **Final process for engaging the relevant community**: the report describes the process implemented for collecting final messages from the TAILOR consortium and external consultations, highlighting the ultimate direction and vision for fostering trustworthy AI in Europe. It details the engagement procedure, ensuring that all key initiatives were consulted, resulting in an SRIR that represents a holistic outcome based on input from a significant part of the AI and Robotics community. The actions taken in the collection of the SRIR are linked to the content found in the final version of the SRIR.

The comprehensive approach taken ensures that the final SRIR not only reflects academic insights but also incorporates industrial needs and priorities, thereby providing a holistic view of the landscape for research in Trustworthy AI. This alignment is crucial for advancing AI research and innovation in a coordinated manner across Europe.

# 1. Purpose of the SRIR and the engagement process

TAILOR is a Network of Excellence dedicated to creating an ecosystem and delivering robust research results in Trustworthy AI, focusing on Learning, Optimization, and Reasoning. Within Work Package 2, TAILOR is tasked with developing an SRIR. This roadmap aims to shape the main research directions of the project and extend its influence beyond the project's duration, providing guidelines to the European AI community and the European Commission for future work programs. The roadmap's ambition is to outline the foundations of Trustworthy AI from 2020 to 2030.

## Objectives of the SRIR

The SRIR has three primary objectives, aligned with those of TAILOR:

- **OB1:** Strengthen and expand a pan-European network of excellent researchers on the Foundations of Trustworthy AI.
- **OB2:** Enhance the scientific foundations for Trustworthy AI and translate them into technical requirements broadly adopted by industry.
- **OB3:** Foster collaborations between academic, industrial, governmental, and community stakeholders on the Foundations of Trustworthy AI.

## Contributions from Scientific Work Packages

The TAILOR project serves as the core source of content for the SRIR. The scientific work packages play a crucial role in defining current and long-term research directions by analysing gaps and needs from both research and industrial perspectives. By working internally, TAILOR aims to enhance the SRIR by defining actions and priorities and tightly integrating Learning, Optimization, and Reasoning (LOR) with Trustworthy AI (TAI) dimensions. This integration includes investigating how LOR can enforce TAI aspects such as fairness, privacy, explainability, robustness, and sustainability, and making LOR techniques trustworthy.
In particular, each work package contributed by focusing on specific aspects of Trustworthy AI that are the foundation of the SRIR.

**WP3: Trustworthy AI**

- Focus: investigating the fundamental principles and methodologies to ensure AI systems are fair, transparent, accountable, and robust.
- Contribution to the SRIR: WP3 provides the foundational research necessary to define the core attributes of Trustworthy AI. It develops theoretical frameworks, methodologies, and evaluation metrics to assess AI systems' adherence to ethical and legal standards. The insights from WP3 identify critical gaps in current AI technologies and propose new research directions to enhance the trustworthiness of AI.

**WP4: Integrating AI Paradigms and Representations**

- Focus: Developing methods to integrate various AI paradigms (e.g., symbolic reasoning, machine learning) and representations to create more comprehensive AI systems.
- Contribution: WP4 advances the integration of different AI approaches, combining their strengths to create more versatile and effective AI systems. This integration is crucial for building robust, adaptable, and trustworthy AI systems capable of handling

complex, real-world problems. The research from WP4 informs the SRIR on hybrid AI systems, identifying areas where combined methodologies can improve trustworthiness.

**WP5: Deciding and Learning How to Act**

- Focus: enhancing AI systems' decision-making and learning capabilities, particularly in dynamic and uncertain environments.
- Contribution: WP5 contributes by developing advanced algorithms and techniques for decision-making and learning that are both effective and reliable. This work package focuses on creating AI systems that can learn from experience and are capable of planning putting into action what they learned, adapting to new situations, and making sound decisions while maintaining trustworthiness. These advancements guide the SRIR in identifying best practices and future research areas in AI decision-making as well as grounding the principle of trustworthiness of AI.

**WP6: Learning and Reasoning in Social Contexts**

- Focus: understanding and improving how AI systems learn and reason within social contexts, including interactions with humans and other AI systems.
- Contribution: WP6 explores how AI can be designed to understand and navigate social dynamics, ensuring that AI systems are socially aware and can collaborate effectively with humans and other AI systems[1]. This work package contributes to the SRIR by highlighting the importance of social context in AI development and proposing research directions that prioritize social interaction and ethical considerations.

**WP7: Automated AI**

- Focus: developing automated AI techniques that can create and optimize AI systems with minimal human intervention.
- Contribution: WP7 advances the field of automated machine learning (AutoML) and other automated AI techniques, which are essential for scaling AI development and deployment. This work package ensures that automated AI systems are designed to be trustworthy by incorporating trustworthiness criteria into the automation processes. The research from WP7 informs the SRIR on how to develop scalable, reliable, and trustworthy automated AI solutions.

# Importance of Partner Engagement

Engagement of partners is essential to provide a coherent and holistic SRIR for several reasons summarised in the following.

*Diverse expertise and perspectives*: TAILOR encompasses a wide range of expertise and perspectives from various stakeholders, including academic researchers, industry experts, and governmental bodies. This diversity ensures that the SRIR addresses the comprehensive needs and challenges of Trustworthy AI across different domains.

[1] Socially aware AI systems refer to those capable of understanding human behaviors, and social dynamics, enabling more effective collaboration with both humans and other AI technologies in shared tasks or environments. See for instance: Cheng, Lu, Kush R. Varshney, and Huan Liu. "Socially responsible AI algorithms: Issues, purposes, and challenges." *Journal of Artificial Intelligence Research* 71 (2021): 1137-1181

Comprehensive analysis of gaps and needs: by involving partners, the project can conduct a more thorough analysis of the current gaps and future needs in AI research and application. This comprehensive analysis is critical for identifying priorities and directing efforts toward the most impactful areas.

*Alignment with industrial and societal requirements: e*ngaging industrial partners ensures that the research directions and technological developments proposed in the SRIR are aligned with real-world applications and industry requirements. This alignment is crucial for the practical implementation and adoption of Trustworthy AI solutions.

*Fostering collaboration and synergy:* collaboration between different stakeholders fosters synergy, enabling the pooling of resources, knowledge, and expertise. This collaborative approach enhances the quality and relevance of the SRIR, making it a robust guide for future research and innovation activities.

*Ensuring inclusivity and representation*: partner engagement ensures that the SRIR is inclusive and represents the views and interests of all relevant parties. This inclusivity is important for gaining broad acceptance and support for the roadmap across the European AI community.

*Holistic approach to Trustworthy AI*: by integrating inputs from diverse partners, the SRIR can adopt a holistic approach to Trustworthy AI. This approach ensures that all aspects of AI, from technical requirements to ethical considerations, are addressed in a balanced and comprehensive manner.

*Facilitating continuous improvement*: partner engagement allows for continuous feedback and updates to the SRIR. This dynamic process ensures that the roadmap remains relevant and up-to-date with the latest advancements and emerging trends in AI research and technology.

## Engaging the European AI Ecosystem

To pervasively spread the concept of Trustworthy AI across Europe, both in academia and industry, it is crucial to connect the SRIR development with other initiatives in the European AI ecosystem. TAILOR identifies the main initiatives for liaison on the SRIR topic and defines the process to engage them in the SRIR's definition and continuous update.

Engaging the broader AI, Data, and Robotics community is essential to ensure that the SRIR reflects the diverse views and needs of the entire ecosystem.

*Diverse perspectives*: the AI ecosystem comprises various stakeholders, including researchers, industry practitioners, policymakers, and end-users. Each group offers unique insights and priorities that are crucial for creating a comprehensive and balanced SRIR. By incorporating these diverse perspectives, the SRIR can address a wider range of challenges and opportunities, ensuring that it is relevant and useful to all stakeholders.

*Identifying gaps and needs:* engaging with the broader community helps identifying gaps in current research and applications that might not be evident from a single perspective. This comprehensive analysis is necessary for setting research priorities and directing efforts

where they are most needed. It ensures that the roadmap is not just theoretically sound but also *practically relevant*, addressing real-world challenges faced by the industry and society.

*Enhancing credibility and adoption:* a SRIR developed through inclusive engagement is more likely to gain acceptance and support from the community. When stakeholders see their concerns and inputs reflected in the roadmap, they are more likely to trust and adopt its recommendations. This buy-in is essential for the successful implementation of the roadmap's strategies and achieving its long-term goals.

*Promoting collaboration:* engagement fosters collaboration between different initiatives and projects within the AI ecosystem. This collaborative approach can lead to synergies, where the combined efforts and resources of various stakeholders lead to more significant advancements. It helps build a cohesive and unified AI community, where stakeholders work together towards common goals, rather than in isolated silos.

*Adapting to emerging trends:* the AI field is rapidly evolving, with new technologies and challenges emerging continuously. Ongoing engagement allows the SRIR to remain dynamic and responsive to these changes. Regular updates and feedback from the community ensure that the roadmap stays current and aligned with the latest developments in AI research and industry practices.

## Deliverable Organisation

The deliverable is organised into three main parts:

1) Gap Analysis: The first part addresses the gaps identified in the previous version of the TAILOR SRIR, based on reviewers' comments. This section reviews the existing roadmap and highlights areas requiring further attention or improvement.
2) Internal Consultation: The second part outlines the internal consultation process conducted to address the identified gaps. Particular attention is given to the input from the scientific task leaders within the TAILOR consortium. This section also includes contributions from all partners to ensure that the core research directions and priorities of TAILOR are well-represented in the SRIR.
3) External Engagement: The final section details the engagement with other EU initiatives. It presents the preliminary questions addressed to external communities and describes the planned actions, meetings, and tools used to facilitate this interaction. This ensures that external perspectives are incorporated into the SRIR.

Each section is linked to the outcomes collected during the process, leading to the final content of the SRIR.

# 2. Building the SRIR: Methodology

## Review of the First Version of the SRIR and Key Observations

The first version of the Strategic Research and Innovation Roadmap (SRIR v1) was published on April 13, 2022, providing an overview of topics relevant to achieving Trustworthy AI systems. The review comments highlighted three major observations:

**FEEDBACK 1**
**Merging Trustworthy AI (TAI) and Learning, Optimization, and Reasoning (LOR):**
**Observation**: the SRIR v1 indicated that the dimensions of Trustworthy AI and the Learning, Optimization, and Reasoning aspects of the TAILOR project were somewhat disconnected.
**Importance**: it is essential to integrate these dimensions more closely, identifying how LOR can be utilized to achieve Trustworthy AI features and vice versa. The reviewer endorsed the selected research areas, where data-driven and knowledge-driven principles come together, resulting in *hybrid AI* and *neural-symbolic computing*. **These principles must converge to advance toward ethical and legally compliant systems.**
**Action**: addressing this gap involves engaging the TAILOR consortium to identify these connections and define research topics aimed at strengthening the integration of TAI and LOR. This integration is crucial for the overall success of TAILOR and could influence future European Commission funding directions. The outcomes of the consortium's feedback will serve as the foundation for engaging external initiatives.

**FEEDBACK 2**
**Priorities and Actions:**
**Observation**: while SRIR v1 identified important research directions for innovation and foundational research across all scientific work packages, it did not sufficiently *prioritize these directions or outline specific actions for achieving them*.
**Importance**: to create a truly strategic roadmap, it is necessary to prioritize research topics and define short- and long-term objectives and actions. The SRIR v1 provided an overview of the research landscape, but SRIR v2 should guide researchers and policymakers in navigating this landscape.
**Action**: engaging the TAILOR consortium to establish priorities and actionable steps is essential. This extension of the SRIR with prioritised actions was planned and is necessary to move from listing important research topics to planning and programming research and development in Trustworthy AI. The outcomes of the consortium's feedback will serve as the foundation for engaging external initiatives.

**FEEDBACK 3**
**Connections with Other Initiatives:**
**Observation**: the SRIR v1 was produced by consulting the TAILOR consortium. However, Trustworthy AI and LOR could impact and be impacted by other European strategic initiatives related to AI and other disciplines.
**Importance**: to ensure a comprehensive and cohesive roadmap, it is necessary to identify relevant initiatives to be involved in the SRIR definition process and establish a timeline for their involvement.

**Action**: identifying and engaging with these initiatives will broaden the SRIR's scope, incorporating insights and aligning with broader European strategic objectives in AI and related fields.

## Methodology

Based on the feedback described above and the summary of actions discussed in detail in D11.5, the following consultations were conducted to address the identified gaps and build a synergistic roadmap for Trustworthy AI by involving the entire interested community.

**Second internal consultation**: this step addressed FEEDBACK 1 and FEEDBACK 2. It sets the groundwork for validation, communication, and collection of feedback for the next phase of external consultation.

**External consultation:** this step addressed FEEDBACK 3 and consolidated the insights from FEEDBACK 1 and FEEDBACK 2. Key initiatives consulted during this phase included all relevant EC-funded initiatives within the AI, Data, and Robotics community, both academic and industrial. In particular Industrial Consultation and Inclusion of EU Initiatives Connected to TAILOR have been organised as follows (full methodology depicted in Figure 1).

**Industrial Consultation:**
To gather insights from the industry, we organised Team Development Workshops. All the details of these workshops can be found in D8.3. The outcomes from these workshops are linked to the final version of the SRIR in the following chapter.

**Inclusion of EU Initiatives Connected to TAILOR:**
The following key initiatives were involved to ensure a comprehensive and inclusive SRIR:
ICT48 Networks of Excellence and the associated CSA (VISION)
AI-on-Demand Platform and Follow-Up Projects (Six ICT49 Projects)
AI, Data and Robotics PPP and Related CSA
HPC and Cybersecurity Initiatives

These initiatives are core for external consultation because they represent a broad spectrum of expertise, resources, and stakeholders within the European AI ecosystem. Their involvement ensures that the SRIR is comprehensive, relevant, and aligned with the current and future needs of the AI, Data, and Robotics communities.

**ICT48 Networks of Excellence and the associated CSA (VISION)**
*Expertise and Collaboration*: These networks consist of top AI research institutions across Europe, fostering collaboration and advancing state-of-the-art research in AI.
*Integration of Efforts*: Engaging these networks ensures that the SRIR aligns with the latest academic research and leverages collective expertise to address complex AI challenges.

**AI-on-Demand Platform and Follow-Up Projects (Six ICT49 Projects)**

*Accessibility and Utilization*: The AI-on-Demand platform provides resources and tools that make AI technologies accessible to a wide range of users, from researchers to businesses.
*Innovation and Development*: The follow-up projects (ICT49) focus on developing innovative AI solutions and applications. Their insights help ensure the SRIR addresses practical, real-world applications and technological advancements.

**AI, Data and Robotics PPP and Related CSA**

*Public-Private Partnership*: This initiative brings together public and private stakeholders to drive AI innovation and adoption in Europe.
*Strategic Alignment*: Engaging this partnership ensures that the SRIR is aligned with broader strategic goals, including economic growth, societal benefits, and competitive advantages in AI and robotics.

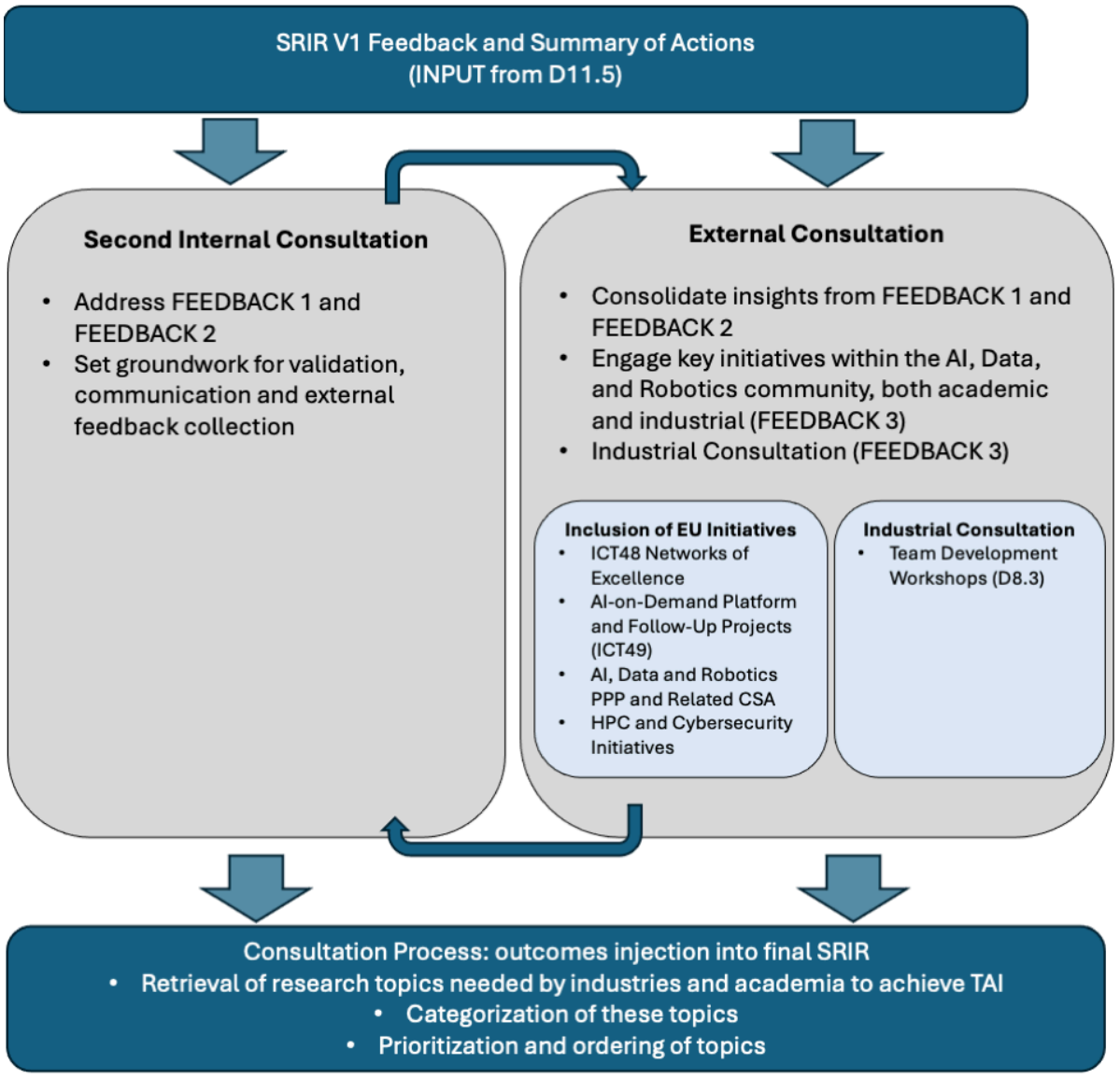



*Figure 1 – SRIR Development in Collaboration with the ICT-48 CSA, PPP-AI, and AI4EU: Process & Methodology*

**HPC and Cybersecurity Initiatives**
*High-Performance Computing (HPC)*: HPC is crucial for processing large datasets and complex AI models, making it a key component for advanced AI research and development.
*Cybersecurity*: Ensuring AI systems are secure is essential for trustworthiness. Engaging cybersecurity initiatives ensures that the SRIR incorporates best practices for protecting AI systems against threats.

Involving these initiatives ensures that the SRIR is informed by diverse perspectives, covering academic research, industrial needs, public-private collaboration, and technological infrastructure. By consulting these key initiatives, the SRIR is more likely to address the current and future needs of the entire AI ecosystem, ensuring its recommendations are practical and impactful. Aligning the SRIR with these initiatives helps to create a cohesive strategy for AI research and innovation in Europe, maximising the potential for synergies and coordinated efforts. Overall, by engaging these initiatives, the TAILOR project ensures that the SRIR is not only a robust academic document but also a practical roadmap that guides AI development and deployment in Europe, fostering an environment where Trustworthy AI can thrive.

Through internal and external consultations and multiple iterations, we achieved the following:

1) Retrieval of research topics needed by industries.
2) Categorisation of these topics.
3) Prioritisation and ordering of topics.

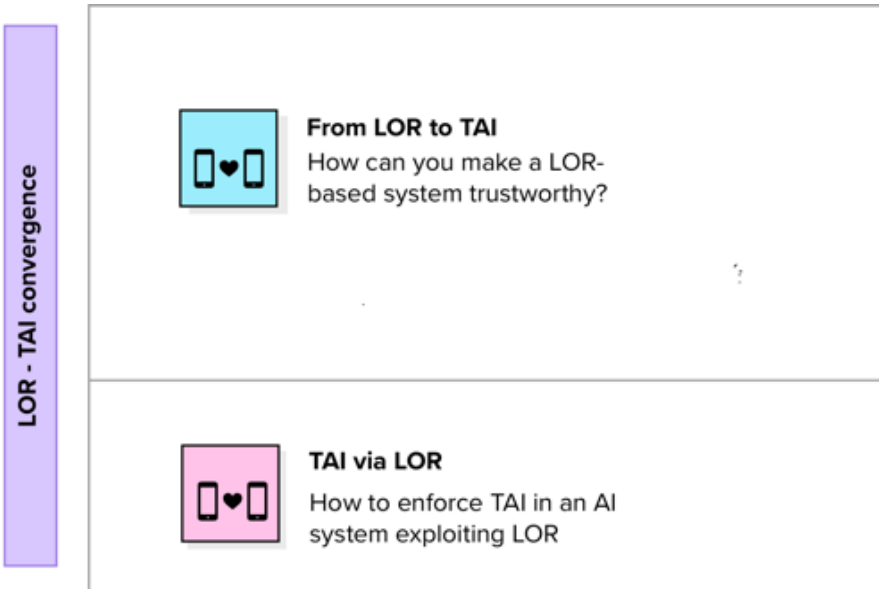
# 3. Process and outcomes

The development of the SRIR followed the above described structured methodology involving various workshops and working groups to gather insights, needs, and recommendations from different stakeholders. This approach covered both research and innovation perspectives, ensuring a comprehensive and collaborative process. The selected tool for facilitating this collaboration was Mural[2], which enabled interactive participation and feedback collection. All the murals are available in the Appendix, for illustration of work methodology. The virtual post-it notes are not intended for reading.

**First internal consultation**

During the TAILOR recurring meeting, the methodology for building the SRIR was discussed and validated with the consortium. Initial pillars were identified to formulate the research questions and gather necessary feedback.

The following key questions were identified for the consultation.




The first two questions focused on gathering feedback on how to converge LOR and TAI. The first question, "How can you make a LOR-based system trustworthy?" sought to understand the methods to ensure trustworthiness in systems based on Learning, Optimization, and Reasoning. The second question, "How to enforce TAI in an AI system exploiting LOR?" aimed to explore how Trustworthy AI principles can be integrated into AI systems utilising LOR techniques. The feedback collected for these questions addressed the gap identified in FEEDBACK 1. These insights are now part of the SRIR section "2. Categorization in topic sectors" and have evolved into the final recommendations by TAILOR.

[2] https://mural.co/

**Trustworthiness and certifications**

How can we generate trust? Do we need certifications?

The third question focused on certifications and trustworthiness, specifically on how to generate trust and what is needed to achieve it. The feedback collected here also addressed the gap identified in FEEDBACK 1 and contributed to the sections of the SRIR related to final recommendations.

**Mentoring, training**

Future of AI for public research is in hybridization: Big Tech do not have experts in all disciplines, while public research does.
* would you agree?
* which disciplines are the most promising ones

A fourth question addressed mentoring and training, exploring how education should evolve accordingly. The feedback gathered for this question further addressed the gap identified in FEEDBACK 1 and informed the sections of the SRIR related to final recommendations.

**Actions, Priorities**

Which are the main urgent actions that need to be done for obtaining trustworthy AI systems, to assess the risks of AI misuse.

The final question focused on concrete actions and priorities to fill the gap identified in FEEDBACK 2. The consolidated outcomes can be found in the SRIR sections "3. AI Development Layers" and "4. Prioritisation based on Showcases". These outcomes formed the basis for the final workshop organised by TAILOR (in Vaals, the Netherlands, April 2024) where all the consortium worked on the final prioritisation of all the topics identified and their actions.

For each of these questions, different feedback was asked according to 4 different perspectives: what is needed, challenges and risks, other initiatives to involve, and recommendations.

**Key Questions and Summary of Answers**

**Question 1: How can you make a LOR-based system trustworthy?**

Convergence of Learning, Optimization, and Reasoning (LOR) with Trustworthy AI (TAI): The first question aimed to gather feedback on strategies to enhance the trustworthiness of systems based on LOR principles.

Participants highlighted several approaches to make LOR-based systems trustworthy. One key aspect is ensuring fairness by implementing algorithms that prevent biased decision-making. Privacy was also a significant concern, with suggestions focusing on robust data protection measures to safeguard user information. To make the system's decisions transparent and understandable, methods to enhance explainability were discussed. Additionally, participants emphasised the importance of robustness, ensuring the system can perform reliably under various conditions. Finally, sustainability was noted, with recommendations to create systems that are energy-efficient and have minimal environmental impact.

**Question 2: How to enforce TAI in an AI system exploiting LOR?**

Enforcing Trustworthy AI in AI Systems Using LOR: The second question sought to understand how Trustworthy AI principles could be integrated into AI systems utilizing LOR techniques.

The discussion on enforcing Trustworthy AI in AI systems exploiting LOR revolved around several key strategies. Participants emphasized the importance of integrating ethical guidelines directly into the design and development phases of AI systems. Ensuring regulatory compliance was also critical, with systems needing to adhere to existing legal frameworks and standards. Continuous monitoring and evaluation were identified as necessary for the ongoing assessment of a system's trustworthiness. There was also a strong focus on user-centric design, with systems needing to be developed with a clear understanding of user needs and societal impacts. Lastly, participants highlighted the importance of interdisciplinary collaboration, encouraging AI developers to work closely with ethicists and legal experts to create well-rounded, trustworthy systems.

**Question 3: How to generate trust? What is needed?**

Feedback highlighted the need for well-defined certification processes that ensure AI systems meet specific trustworthiness criteria. Participants suggested that trust could be built through transparency in AI operations and decision-making processes. Establishing rigorous testing protocols and regular audits was seen as crucial. Additionally, the creation of industry standards and guidelines was emphasized to help harmonize practices and expectations across different sectors.

**Question 4: How should education and training evolve to support Trustworthy AI?**

Participants underscored the importance of updating curricula to include courses on ethics, transparency, and privacy in AI. Practical training programs and workshops that focus on real-world applications of Trustworthy AI were recommended. Collaboration between

academia and industry was seen as essential to ensure that educational programs are relevant and up-to-date. Moreover, ongoing professional development and certification programs for existing AI professionals were suggested to keep them informed about the latest developments in Trustworthy AI.

**Question 5: What are the concrete actions and priorities needed to fill the identified gaps?**

The responses emphasized the need for a clear roadmap outlining short- and long-term objectives. Prioritization of research topics was seen as crucial, with participants suggesting that focus should be given to areas with the highest potential impact. Establishing partnerships between academia, industry, and government was recommended to facilitate the translation of research into practical applications. Furthermore, securing funding and resources for priority areas was highlighted as a key action.

**Workshop to collect AI4Europe and AioD Perspective**

A workshop with the AI4Europe consortium was organised on Friday, May 19, 2023, with the participation of 24 members and 36 individuals having access to the collaborative mural document. The main objective of this workshop was to cover the research part of the SRIR. Through discussions and collaborative activities, participants shared their expertise, identified research needs, and proposed recommendations to shape the research direction of TAI aligned with the research direction of the platform.

**Workshop to collect ICT49 Perspective**

On Friday, May 23, 2023, a dedicated workshop was organised under the ICT-49 initiative, with 9 participants and 36 individuals having access to the mural document. The primary focus of this workshop was to cover the innovation part of the SRIR, specifically from the perspective of industries and SMEs. The workshop aimed to understand the requirements for enabling TAI in project exploitation, including desires, challenges, and the long-term direction of the platform.

**Workshop to collect ICT49 Perspective – Focus on exploitation**

A dedicated meeting was organised on Monday, June 19, 2023, with a specific focus on a TAI exploitation of the results of all the projects. The purpose of this meeting was to delve into the strategies and mechanisms required for effectively exploiting the research and innovation outcomes of the platform. This meeting allowed for detailed discussions and refinement of the concepts related to exploitation.

**Second internal consultation and ICT48**

Given the importance of trustworthiness in AI, two workshops were organised to address this specific aspect. The first workshop, organised on Friday, May 23, 2023, involved the ICT-49 TAI working group, while the second workshop took place during the TAILOR Network of Excellence (NoE) conference on June 6, 2023. The outcomes of these discussions were presented at the VISION event to gather a joint view of the NoEs and at the EGI conference.

**ICT-49 final event in Athens**

Discussions and finalisation of concepts during the ICT-49 event (June 29-30, 2023, in Athens) in a dedicated session on the SRIR (120 participants, https://www.ai4europe.eu/news-and-events/news/aiod-community-forum-and-ict-49-final-event-unveils-success-stories-and ).

**Industrial needs**

To understand the needs of the industry across various application domains, TAILOR organised seven Team Development Workshops (TDWs). Detailed information can be found in D8.3. These workshops examined the role of AI within their specific domains and identified the core AI research topics required for the next 5-10 years. The recommendations from the TDW reports ("input for the roadmap") were incorporated into the existing list of research topics from SRIR v1, expanding it into a comprehensive long list. TDWs aimed to identify the most promising and emerging AI topics within sectors such as the Public Sector, Future Mobility, and Healthcare. FBK conducted an analysis of the TDW reports to pinpoint potential use cases that could be translated into showcases. The analysis also explored topics that could foster research-industry collaboration. Several transversal subjects emerged as critical enablers for realising Trustworthy AI. These include expertise and education, trustworthiness/confidence measurability, privacy, personal data and GDPR, ethical use of data, human factors, and standardisation and certification.

**AI, Data and Robotics PPP and Related CSA**

Furthermore, involvement with the PPP-ADRA to foster collaboration and maximise the impact of the SRIR (ADR Forum 8-9 November 2023, Versailles, 150 participants).

**HPC and Cybersecurity Initiatives**

Several initiatives and meetings were organised to collect feedback on HPC and cybersecurity. In April 2023, a PRACE meeting focused on the use of AI and HPC, with an emphasis on green AI applications. A consultation meeting with a supercomputing and quantum computing center was held in January 2024. Additionally, in April 2024, an event was organised in the Emilia Romagna region to discuss the AI Act and its impact on businesses and cybersecurity. Regarding cybersecurity, several meetings with Siemens were conducted to gather more specific feedback on this aspect.

**Final Feedback collection**

A first round of summarising and clustering all the feedback received was conducted during the 3rd TAILOR conference, held on June 5-6, 2023, in Siena, Italy (https://tailor-network.eu/events/3rd-tailor-conference/ ).

The final actions for collecting feedback included several workshops and meetings. Key events included the TAILOR workshop “TAILOR SRIR – 2nd version” (https://tailor-network.eu/events/workshop-tailor-srir-2nd-version/ ) held in Vaals, which was attended by 40 participants, and a meeting with PREPAI for connection purposes in April 2023.

At the Vaals workshop, the primary goal was to work on the second version of the TAILOR SRIR. Participants tested the proposed sector map and showcases, working in groups to select essential research topics for these showcases. Following this workshop and the processing of its outcomes, the map with development layers was introduced. These layers help to order research topics according to their development stage, facilitating the progression of research and development towards innovations.

Additionally, at the 2024 Lisbon conference, a special workshop titled “What’s next in European AI?” focused on testing and developing the final steps of SRIR v2 (https://tailor-network.eu/events/4th-tailor-conference/ ). During this workshop, the larger audience aimed to select the top five most important research topics per showcase, guided by the sector map and development layers. The aggregated results from these workshops and consultations are included in the final version of the SRIR.

# 4. Conclusions

The first version of the Strategic Research and Innovation Roadmap (SRIR v.1) sketched the landscape of research topics for AI innovation considering the essential principles of *Trustworthy AI* (TAI) and methodologies based on *Learning, Optimization and Reasoning* (LOR). The second version of the SRIR (SRIR v.2) aims to add orderings and prioritisation to the first version in order to map out a road in the research landscape.

The development of the SRIR v.2 for TAILOR represents a significant achievement in advancing the framework for Trustworthy AI. This effort brought together a diverse range of stakeholders from academia, industry, and various AI-related initiatives across Europe. By leveraging collaborative tools and methodologies, the SRIR process ensured a thorough and inclusive approach to identifying and prioritising the research and innovation needs for the coming years.

A key strength of the SRIR development process was its commitment to inclusivity and collaboration. The various workshops and consultations allowed for the integration of a wide array of perspectives, ensuring that the final roadmap is not only comprehensive but also representative of the needs and aspirations of the broader AI community. This approach has helped in identifying critical research areas, highlighting gaps, and proposing actionable steps to address these challenges.

The SRIR v.2 emphasises the importance of interdisciplinary collaboration and continuous learning. By recommending updates to educational curricula and promoting ongoing professional development, the roadmap ensures that future AI professionals are well-equipped to navigate the ethical and technical complexities of AI. This forward-looking approach is essential for maintaining the relevance and impact of AI research and innovation.

The introduction of the AI development layers is another significant milestone. This tool provides a clear framework for ordering research topics according to their maturity levels, facilitating a structured progression from foundational research to practical innovations. This systematic approach will help align research efforts with industry needs, ultimately accelerating the deployment of Trustworthy AI solutions.

All the outcomes, details and recommendations are discussed in the final SRIR.

# 5. Appendix: Mural Files

TAILOR SRIR

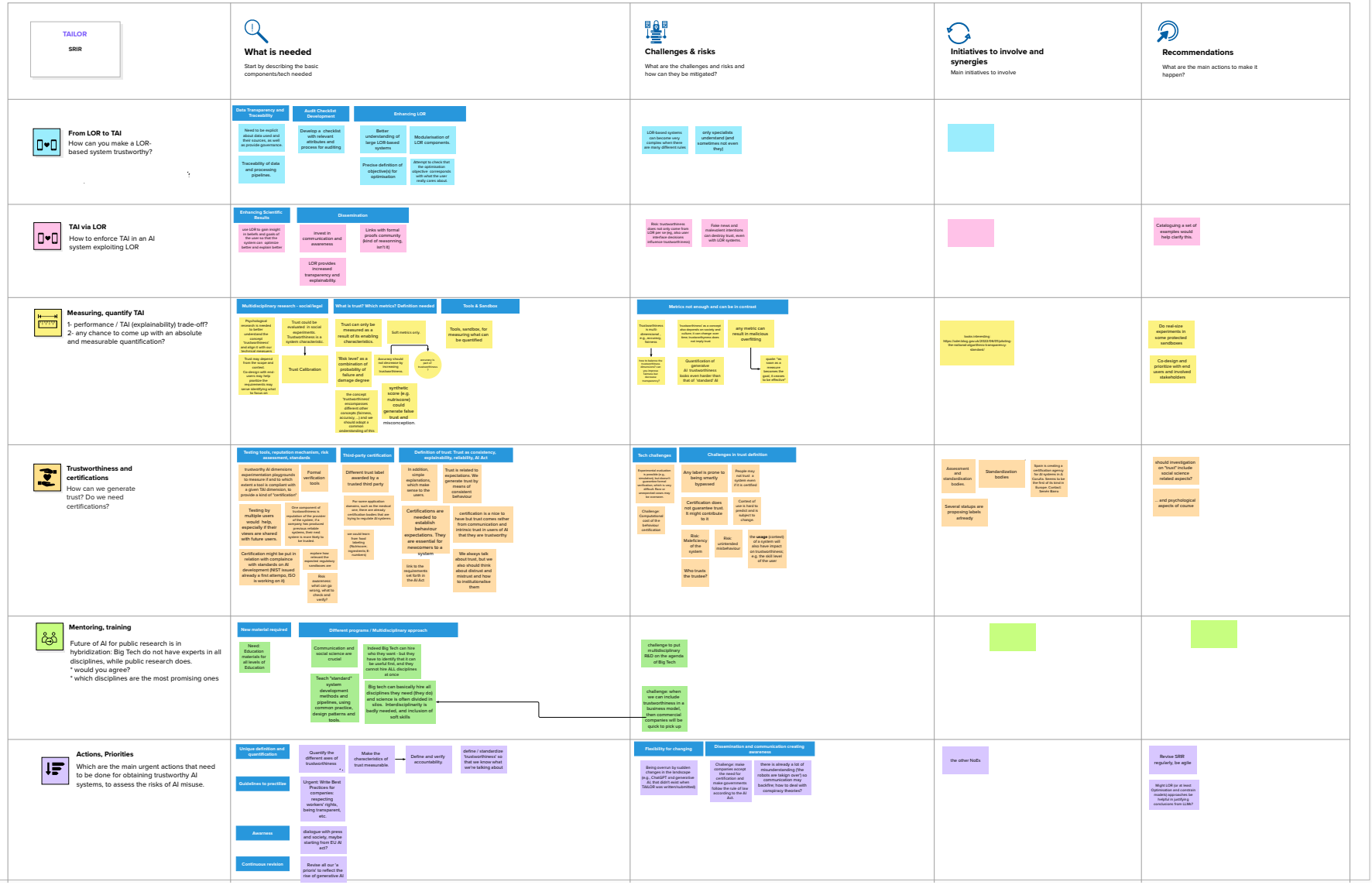
# TAILOR SRIR

**OBJECTIVE 1**
Providing mechanisms for strengthening and enlarging the Pan-European network of researchers

**OBJECTIVE 2**
Defining paths for advancing research foundations in AI providing a reference platform for AI researchers to share results, training on a topic, reuse effective AI methods

**OBJECTIVE 3**
Identifying directions for fostering collaborations between academic, industrial and community stakeholders of AI

## TAILOR SRIR
Feeds / Pillars



# ICT49 - AI4Europe - TAILOR SRIR Sustainability

**OBJECTIVE 1**
Providing mechanisms for strengthening and enlarging the Pan-European network of researchers

**OBJECTIVE 2**
Defining paths for advancing research foundations in AI providing a reference platform for AI researchers to share results, training on a topic, reuse effective AI methods

**OBJECTIVE 3**
Identifying directions for fostering collaborations between academic, industrial and community stakeholders of AI

## ICT49 - AI4Europe - TAILORAI4Europe SRIR

Sustainability Pillars

## How do we foster TAI and sustainability?

ICT49

SRIR and SUSTAINABILITY

### What is needed

Start by describing the basic components/tech needed

### Challenges & risks

What are the challenges and risks and how can they be mitigated?

### Initiatives to involve and synergies

Main initiatives to involve

### Recommendations for making it sustainable

What are the main actions to make it happen? How to make it sustainable (incentives, ...)

**Sustainability**
What features would be beneficial in a platform to support ICT-49 projects beyond their completion and make them exploitable?

- Spiros: Marketplace features like the ability to sell assets, or expose monetization mechanisms (subscription, per asset / usage, etc) - AIOD to support the advertisement / visibility the commercial (private products of its community (e.g.ICT-49 assets). - A horizontal, common approach for a ticket support mechanism that ICT-49 use to sustain their products - Common approach for documentation / help on the assets
- A dedicated space to showcase all completed ICT-49 projects with additional functionalities to promote KEAs for each project. Search functionalities that allow visitors to search for specific assets based on the domain, TRL of each asset etc. Connection with open code repositories (e.g. GitHub) where people can download packages of services along with installation instructions.
- From Exploitation WG: what is needed to exploit results ? • Clearly identified needs from SMEs and researchers • A set of fully developed and RE-USABLE AI assets and AI services • Core processes and tools (preferably OSS) to experiment, build and deploy the resulting AI solutions • A secure Marketplace with IPR enforcement for scalable & meaningful enterprise adoption
- connection with the other platforms dev in the ICT49 projects, avoiding duplication
- adopt the taxonomy of services in line with L-BEST to help understanding the services offered
- spiros: Security aspects and privacy - currently the security and privacy is left on the users that publish the assets - that could introduce security risks that further decrease the trust of potential stakeholders / customers and thus is a challenge for sustainability
- Difficulty for dihs in engaging smes and convincing them in using their AI assets. Not clear added value for SMEs in using the AIoD
- no resources & incentives for the DIHs to use the platform
- a one place for all the stakeholders is not very clear and useful. the risk of resulting not fit for anyone
- lack of clarity of functionalities and lack of competences
- the AIoD is only a technology platform
- 1. Strategy clarity 2.Resources to execute 3. Fund to sustain 4. Lack of engagement
- Oftentimes, it is difficult to have people from a completed project keep working on its post-project exploitation without additional funding
- to ensure post-project sustainability of the network of DIHs focusing on AI, synergies and collaborationns opprtuinites wiith networks such as ADRA and DTA could be explored, considering the possibiilty of establishing a topic group within it
- Sister project Bonseyes has developed a number of the functionalities needed for the marketplace module of AIoD platform, namely tools, workflows, licensing framework and user journey, to implement a indisutry marketplace This incide a non profit sustainability vehicle (association)
- Making the platform more "commercial-ready" (security, front-end,scalability) to increase trust and interest of external companies / customers
- Identify and dlegagte platform build out to experienced team and in B2B2C platforms Create proper sustainability vehcle with proper funding and staffing Have a realistic and tested business model for the platform
- Make the platform a central space for exploitation of ICT-49 assets, thus assisting projects in the oftentimes awkard task of defining exploitation pathways for RIAs. Create commercial packages for the more ready-to-market assets, thus incentivising projects to seek solutions with high TRLs

**Trustworthiness**

- Provide set of transparent and immutable services that can be validated and audited on platform, based on contenerized assets and tools
- spiros: - a navigational overview of current trustworthy AI efforts, so users can easily adopt / use work been done in all different projects that is more relevant (e.g. bias, security, robotics, domain-specific, etc). - A guided approach of users on trustworthiness (not just a list of relevant projects that would confuse the user)
- We need a repository with current successful practices and pathways for TAI so that future projects have an available know-how and instead of reinventing the wheel they can build on existing knowledge. A TAI certification supported by current working groups and the EC would also be very helpful
- forcing every asset to be Trustworthy
- providing tools, methodologies to help with implementing TAI solutions
- Spiros:Too many approaches and too high level projects, while most of them afraid to take a step against real problems (like copyright issues of generative AI). Many overlaps between projects (that remain high-level and not practical for average end-user)
- Lack of a setp-by-step methodology depending on the domain. For example, commercial AI systems built for an enterprise in a low-risk environment should not have to follow the same steps compared to a more high-risk domain (e.g., healthcare/sensitive data)
- lack of awareness on the topic migh be a risk
- where there is too much, it is like there isn't anything...
- competition between different approaches and tools does not help users
- IP leakage to China
- Networks such as ADRA
- ICT-49 TAI, ETAPAS, AI act, Adra-e, BDVA
- ICT49 TAI WG to leverage the work done
- Other funded projects on TAI (e.g. ETAPAS)
- Spiros: A top down approach for AIOD is needed to organise the current efforts and take concrete steps for the users (single guide). In the future create blueprints of such efforts (like the L-services)
- it should be integrated (somehow) with the regulatory sandboxes that will be established
- An assessment could be created that would first help projects realise the potential risks of their solution depending on the domain and the nature of the solution. Depending on the resulting score, only the necessary steps to ensure TAI should be followed.

**Certifications, Measuring, quantify TAI**
How can we generate trust? Do we need certifications?

- Reputational mechanism on the platform for assets with objective KPIs (difficult for people)
- Spiros Yes, but this needs to be done outside AIOD at a higher level
- We do need a TAI certification for sure. At the moment, there is no clear way to evaluate a projects effort for TAI before the final review. It is also hard for potential clients to realise the trustworthiness of a solution without a clear certification
- Ensure a clear alignment with the latest European Regulations such as the AI Act
- certification is required by the upcoming AI Act, this is something outside the legal and ESO approach
- see above comment on Regulatory Sandboxes
- force compliance with future Eu standards on AI when using the platform
- Subjective bias, discrimination
- Spiros: These need to come first with the objective of protecting citizens and their IP, privacy, not the industry that would first prioritize low cost at the expense of society and environment
- As mentioned before taking into consideration the AI Act ii and trustortnwss in technologies. The o ly risk is the one to fall into "ETHICS washing"
- At the moment, it seems that the community is quite far away from developing a standardised certification due to the multidisciplinarity of the problem and the amount of people required to collaborate. We should start smaller and generate small-scale certifications on a project-by-project basis that we can later consolidate
- Same as above
- Strong support from the legal system and politicians towards the direction of protecting the citizens- it wont work otherwise
- again, it is impossible (and i think wrong) that one platform answers to all the needs connected to AI. certification is a very structured process (also connected to legal implication) that probably cannot fit here
- Focus on each individual project. Create a first version of TAI certification for projects that are ethical supported by the ICT-49 working group. Each individual certification would generate successful practices and leasons learnt which would facilitate a more horizontal certification.

**Actions, Priorities**
Which are the main urgent actions that need to be done for the ICT-49 support?

- Strategic ROADMAP needed
- Spiros: EXploitation opportunities to ICT_49 offered by AIOD. Currently AIOD lacks completely of exploitation channels.Exploitation channels could include: pricing mechanism (as stated in comment above, subscription per usage / time, .etc) per asset, marketplace features (advanced / market friendly search for products, vendor dashboards, payment gateways, analytics dashboards, rating and reviews) opportunities for commercial advertisment of the users' assets to the entire network of AIOD, SEO optimisation to promote assets, events to promote assets of users,
- Spiros: The current features of AIOD are underwhelming compared to relevant effots.No real users except Horizon Europe actors so far. next urgent efforts need to focus on attracting the intended target audience
- Dedicated conferences/workshops focused on TAI. Present demos of ICT-49 projects and explain how this is the right way to build current and future AI solutions. Have opposite examples of how each of those solutions would affect ethics/human rights and so on if it did not follow the TAI approach. Have speakers from outside the EU (e.g., USA) where no regulations exist, to remind us on the dangers and rules to having a clear pathway and rules to ensure TAI.
- i think there should be a choice on on the target users of the platform, the one stop shop for everything is not working
- clear understanding of the vision, new configuration, purpose etc
- Spiros:AIOD doesn't have clear objectives (or they are implemented as such) in terms of exploitation
- In a global level, it seems that the general public cares more about the potential/ possibilities of AI and less about how bad/non-ethical practices could affect society
- the current vision is too technical to fit all the outcomes of the ICT49 projects
- AIoD has to show a clear added value for SMEs and DIHs. So the risk is an oversupply of competitive offering.
- Pre-PAI is fundamental to understand how different potential users perceive the platform and need
- Spiros: An urgent action is needed to agree on the goals of AIOD in terms of exploitation channels
- Dedicated conferences/ workshops around TAI with the support of ICT-49 projects along with project officers, representatives from the respective DGs etc. The EC needs to show more support on this by being more vocal on the importance of TAI
- select the real user of the platform
- clear understanding of the real value for the users and communication about it

# AI4Europe SRIR Sustainability

**OBJECTIVE 1**
Providing mechanisms for strengthening and enlarging the Pan-European network of researchers

**OBJECTIVE 2**
Defining paths for advancing research foundations in AI providing a reference platform for AI researchers to share results, training on a topic, reuse effective AI methods

**OBJECTIVE 3**
Identifying directions for fostering collaborations between academic, industrial and community stakeholders of AI

**AI4Europe SRIR**
Sustainability Pillars

## How do we foster sustainability?

AI4EUROPE
SRIR and SUSTAINABILITY

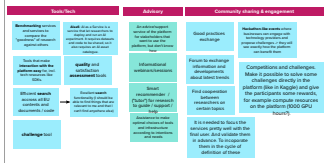
### What is needed
Start by describing the basic components/tech needed

### Challenges & risks
What are the challenges and risks and how can they be mitigated?

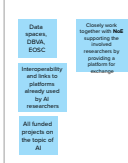
### Initiatives to involve and synergies
Main initiatives to involve

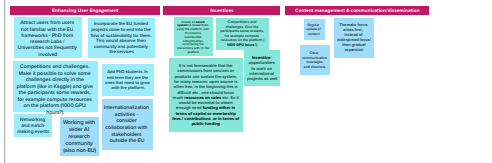
### Recommendations for making it sustainable
What are the main actions to make it happen? How to make it sustainable (incentives, ...)

**Services for sustainability**
Examples:
- Matchmaking
- LLM
- Success stories
- Training and mentoring

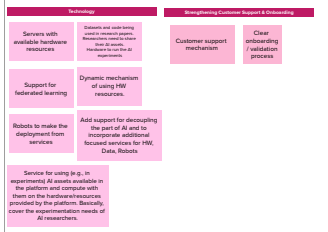


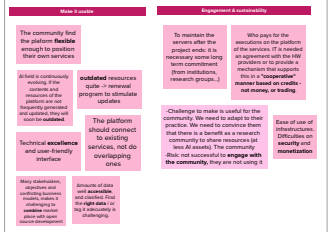


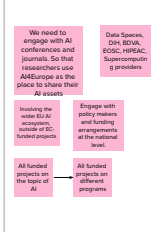


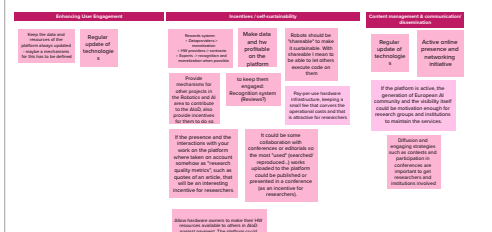



**Resources for AI (hardware, data, robots, people,...)**
- how do we bring the resources on the platform?
- Strategies to keep them engaged and retained on the platform

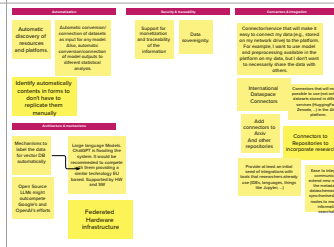


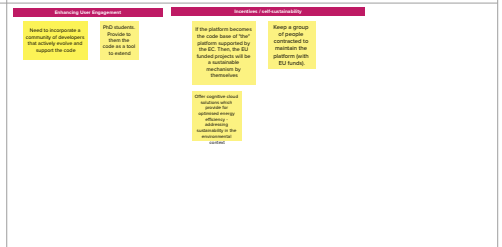






**Technologies & architecture**
Software system to interconnect, support, interoperate (e.g. api, connectors, interfaces, LLM...)








**Trustworthiness and certifications**
How can we generate trust?
Examples: *experimentation playgrounds to measure if and to which extent a tool is compliant with a given TAI dimension, to provide a kind of "certification"*