All-in-one middleware for industrial Human-robotic interaction

arise-middleware.eu

Coordinator

Consortium partners

Co-funded by the European Union

Document History

Ver.	Date	Description	Author	Partner
0.5	07.11.2025	ToC	Mariela Urra Schiaffino, Tuukka Puonti, Kelly Yen	Demos Helsinki
1.0	08.12.2025	First version for internal review	Mariela Urra Schiaffino, Tuukka Puonti, Kelly Yen	Demos Helsinki
1.5	18.12.2025	Commented version by internal reviewer	Mireya de Diego, Elena Mossali	CARTIF, INTELLIMECH
2.0	22.12.2025	Final version	Mariela Urra Schiaffino, Tuukka Puonti, Kelly Yen	Demos Helsinki

List of author(s), contributors(s) and reviewer(s)

Name	Partner (Acronym)	Contribution/Comment
Mariela Urra Schiaffino	Demos Helsinki	Main editor
Tuukka Puonti	Demos Helsinki	Main editor
Kelly Yen	Demos Helsinki	Main editor
Jussi Asikainen	Demos Helsinki	Assistant editor
Falguni Mazumder	Demos Helsinki	Assistant editor
Mireya De Diego	CARTIF	Internal reviewer
Elena Mossali	INTELLIMECH	Internal reviewer

Co-funded by the European Union

Publication Details

Grant Agreement Number 101135784

Acronym ARISE

Full Title	Agile, human-centric, and Real-tIme enabled SourcE technologies advancing industrial HRI in Europe
Topic	HORIZON-CL4-2023-DIGITAL-EMERGING-01-02
Funding Scheme	HORIZON Innovation Actions
Start Date	01/01/2024
Duration	42 months
Project URL	www.arise-middleware.eu
Project Coordinator	CARTIF
Deliverable	D3.2 Periodic release and report on the SSH Framework for human-centric and ethical HRI
Work Package	WP3 - Open SSH Framework for human-centric and agile HRI
Delivery Month (DoA)	M24
Version	1.0
Actual Delivery Date	31.12.2025
Type	Report
Dissemination Level	Public
Lead Beneficiary	Demos Helsinki
Keywords	SSH, human-centric, ethics, human-robot interaction

D3.2 Periodic release and report on the SSH Framework for Human-Centric and Ethical HRI

Co-funded by the European Union

Table of Contents

Co-funded by
the European Union

List of Figures

List of Tables

Co-funded by the European Union

Acronyms & Abbreviations

ACR.	Description
AI	Artificial Intelligence
ALTAI	Assessment List for Trustworthy AI
CEO	Chief Executive Office
COO	Chief Operating Officer
C-Suite	Chief-Suite

CoCo	Coexistence-and-cooperation
CTO	Chief Technology Officer
CV	Computer Vision
DPO	Data Protection Office
EIA	Ethics Impact Assessment
EHDS	European Health Data Space
ETAPAS	Ethical technology adoption in public administration services
EU	European Union
EU-OSHA	European agency for safety and health at work
GDPR	General Data Protection Regulation
HCAI	Human-Centered AI
HLEG-AI	High Level Expert Group on Artificial Intelligence
HR	Human Resources
HRI	Human-Robot Interaction
IEEE	Institute of Electrical and Electronics Engineers
IEC	International Electrotechnical Commission
ISO	International Organization for Standardization
KPI	Key Performance Indicator
LEADOR	Leadership in ethical AI development and operational research
MDR	Medical Device Regulation (EU)
NASA	National Aeronautics and Space Administration
NPS	Net Promoter Score
OECD	Organisation for Economic Co-operation and Development
PoC	Proof of Concept
PREM	Patient Reported Experience Measure
PROM	Patient Reported Outcome Measures
QA	Quality Assurance
R&I	Research & Innovation
REELER	Responsible ethical learning with robotics
SATORI	Stakeholders acting together on the ethical impact assessment of research and innovation
SGI	Smart graph interface

Co-funded by
the European Union

SHERPA	Shaping the ethical dimensions of smart information systems - a European perspective
SIENNA	Stakeholder-informed ethics for new technologies with high socio-economic and human rights impact
SME	Small and Medium-sized Enterprises
SPD	Situated participatory design
SRL	Societal Readiness Level
SSH	Social Sciences and Humanities
SUS	System Usability Scale
TEF	Test and Experimentation Facility
TRL	Technology Readiness Level
UEQ-S	User Experience Questionnaire (Short)
UI	User Interface
UMUX-lite	Usability Metric for User Experience lite
UX	User Experience
VIRT-EU	Values and ethics in responsible technology in the EU

1 Introduction

1.1 Purpose and Scope

This deliverable provides an update from ***D3.1 Initial Open SSH Framework for human-centric and ethical HRI* (the SSH Framework)**. Demos Helsinki, the social sciences and humanities (SSH) partner of the ARISE project, leads the development of the framework, which is ongoing throughout the ARISE project. Beyond iterating on the foundations and key elements of the SSH Framework, the purpose of the document is also to to improve and expand upon the initial version by leveraging learnings from the ARISE project's ongoing developments and engagements with the Human-Robot Interactions (HRI) industry. Furthermore, this deliverable discusses the contextual understandings that make up the foundations of the framework, articulates pathways towards human-centered and ethical technological innovation, details an overview of the various possible audiences of this work, and formulates practical and evidenced-based recommendations for HRI development and policy. The deliverable concludes by outlining the key next steps and expected evolution for future iterations of the SSH framework.

1.2 Intended Audiences

This deliverable is public. Therefore, the intended audience are all parties interested in SSH aspects of technological innovation and in the context of HRI. These include the consortium partners, researchers, practitioners, HRI industry actors, technologists, the European Commission, and decision-makers interested in the topic within the scope of the ARISE project. The framework will be developed and applied during the ARISE project and iterated further with ARISE partners and open-call beneficiary projects.

1.3 How to Use This Framework

This deliverable describes the foundations and key elements of the second version of the ARISE SSH framework for human-centric and ethical HRI. It presents an overview of undertaken and planned research activities, establishing a methodology to guide trustworthy and ethical technology development. This document was not designed to be read through sequentially. Instead, readers should use the table of contents to navigate to specific sections of the framework that are relevant to their needs and goals. **Section 2, *Approach and Methodology***, describes the contextual research, key terminologies, and general objectives and approaches on which the framework is built. **Section 3 *The ARISE framework for human-centric and ethical HRI*** describes the actual key components of the SSH framework, including the binding and non-binding requirements for human-centric HRI, modes of implementing the framework, and an SSH toolbox of approaches. **Section 4 *Iterations and* continuous learning** discusses how the framework has evolved since its initial release in 2024, and its expected evolution in future iterations.

The SSH Framework deliverable is complemented by a **web tool**[1], a digital resource designed to offer simplified navigation through the framework's sections. Currently under development, the web tool serves as an accessible entry point for understanding and utilising the framework, tailored specifically for a practical implementation. Its content is scheduled to be updated to align with this second version of the framework following its release.

[1] The ARISE SSH framework webtool can be currently accessed in https://arise-ssh.info/

Co-funded by
the European Union

2 Approach and methodology

2.1 Ethics, Society and Human-Centricity in HRI Development

Robots are increasingly entering human environments from workplaces to homes to care settings, making interactions between human and robots more frequent, varied and consequential. Thus, **human-robot interaction (HRI)** is no longer solely a technical domain but a social and ethical field that directly shapes human agency, trust and well-being. In these shared work spaces, robots act as co-actors rather than passive tools, meaning their design and behaviour must be aligned with the societal values, cultural contexts and diversity of people who interact with or are affected by them.

Ethical considerations are central to HRI because they provide a structured way to **anticipate impacts, balance stakeholder interests** and **prevent harm** as robotic systems enter human domains. Ethical considerations must be both embedded in design and attentive to systems behaviour in practice, including unintended consequences. Important dimensions of ethics such as autonomy, accountability, transparency, fairness, privacy and dignity guide this work, especially as robots increasingly mediate power and decision-making in everyday contexts[2]. In practice, ethical reflections should aim to assist human flourishing rather than constrain or replace it. They must be understood as an ongoing, context-sensitive process rather than a static checklist[3].

The development of robotics is fundamentally a social process, shaped by and shaping the norms, institutions and collective expectations of societies in which these systems operate[4]. As robots become embedded in workplaces and public services, they influence organisational structures, roles and power dynamics. Hence, **ethical reflections addressing systemic and institutional impacts are fundamentally important**. Building and maintaining public trust depends on how well these societal considerations are integrated and how well they demonstrate that robotic systems support social well-being, legitimacy and inclusiveness.

Human-centricity positions human needs, experiences and values at the core of technology development to create an environment where robotic systems genuinely support the people who use or are affected by them. This principle promotes well-being, empowerment and meaningful work by directing automation toward **enhancing human capabilities** rather than diminishing them. Importantly, human-centricity extends beyond design phases to shape organisational cultures and governance practices, thereby integrating respect for human expertise and agency throughout the entire socio-technical system.

[2] Alaieri, F., & Vellino, A. (2016). Ethical Decision Making in Robots: Autonomy, Trust and Responsibility. Social Robotics, 159–168. https://doi.org/10.1007/978-3-319-47437-3_16

[3] For more details on the foundational definition of ethics management underpinning this approach see section 2.3.4.1 Ethics management in emerging HRI technologies.

[4] Šabanović, S. (2010). Robots in Society, Society in Robots. International Journal of Social Robotics, 2(4), 439–450. https://doi.org/10.1007/s12369-010-0066-7

Co-funded by the European Union

2.1.1 Requirements for Industry 5.0

From Industry 4.0 to Industry 5.0

Industry 5.0 represents a value-based reorientation rather than a straightforward continuation of Industry 4.0's technology trajectory. While Industry 4.0 prioritised automation and digitalisation to increase the effectiveness of industrial processes, it inadvertently neglected the human costs associated with the processes of optimization[5]. In response, Industry 5.0 represents a profound shift in perspective by placing **human, social and environmental priorities at the center of technology development,** aiming to rebalance innovation toward societal benefit and human well-being[6]. Conceptualised by the European Commission, Industry 5.0 addresses the limitations of Industry 4.0 by offering a strategic blueprint for more responsible industrial transformation centered on human needs and societal values. The concept integrates European economic priorities regarding the digital transition and the European Green Deal, resting on three fundamental pillars: **sustainability, resilience and human-centricity.**[7,8] .

Sustainability

The sustainability pillar integrates both the environmental and social dimensions. Environmentally, sustainability calls for resource-efficient and circular design practices that minimise ecological impact across the lifecycle of technologies. Socially, it emphasises well-being, equality and long-term employability, aiming at technological change that supports resilient and inclusive societies. Thus, sustainability relies on ethical foresight that actively anticipates and mitigates potential social and ecological harms.

Resilience

Resilience is understood as the capacity of socio-technical systems to adapt to and thrive amidst economic, environmental and societal disruptions. Achieving resilience depends on organisational learning and the adaptability of workers. This creates systems that are enabled to respond effectively to unexpected challenges. At the design level, resilience is supported by principles that include redundancy, flexibility and inclusiveness, which, in turn, help in creating technologies and organisations that remain robust, responsive and capable of sustaining essential functions under changing conditions.

Human-Centricity

Human-centricity serves both as a design principle and organisational commitment that expands from a narrow user-centered design goal. It requires the intentional inclusion of stakeholders, such as workers, users and affected communities, throughout all phases of

[5] Nahavandi, S. (2019). Industry 5.0—A Human-Centric Solution. *Sustainability*, *11*(16), 4371. https://doi.org/10.3390/su11164371

[6] Padovano, A., Cardamone, M., Woschank, M., & Pacher, C. (2024). Exploring Human-Centricity in Industry 5.0: Empirical Insights from a Social Media Discourse. Procedia Computer Science, 232, 1859–1868. https://doi.org/10.1016/j.procs.2024.02.008

[7] European Commission: Directorate-General for Research and Innovation. (2024). ERA industrial technologies roadmap on human-centric research and innovation for the manufacturing sector. Publications Office of the European Union. https://data.europa.eu/doi/10.2777/0266.

[8] European Commission. (2024). Industry 5.0. Research-And-Innovation.ec.europa.eu. https://research-and-innovation.ec.europa.eu/research-area/industrial-research-and-innovation/industry-50_en

technological development. This process is underpinned by systematic engagement of influential actors and practices of care for vulnerable groups, ensuring that diverse perspectives are meaningfully integrated.

In the industrial context, a human-centric approach helps to ensure that collaborations between humans and machines enhance people's well-being at three systemic levels: **1) Micro level** (individual/industrial processes): empowering individuals and optimising human-robot collaborations to support meaningful and safe work, **2) Meso level** (organisational): shaping organisational culture and governance, using technology and organizational structures to enable the paradigm shift, and **3) Macro level** (societal): aligning technological development with broader societal values and expectations, emphasizing principles such as trustworthiness and human oversight[9].

To achieve impact at these levels, human-centricity should be implemented through two primary dimensions: the **technological dimension** and the **organizational level**. In the **technological dimension**, the aim is to design, develop and implement technologies with human needs, values and experiences as the starting point—rather than pursuing technological advancements solely for their state-of-the-art potential—as well as to enhance interactions between humans and machines. Identifying technologies that influence human-centricity is also key to this initiative. For instance, robotics is an area where the development process can focus on enhancing both physical and psychological safety, aligning with the core ideas of human-centricity. At the **organizational dimension**, the human-centric approach emphasizes worker well-being, specifically valuing employee expertise and contribution. It fosters inclusivity and prioritises continuous learning, reshaping existing skills rather than replacing them.

Importantly, human-centricity resists direct translation into technical specifications. Instead, it functions as a guiding principle rooted in intangible values, shaping responsible approaches to the ideation, design, deployment and governance of advanced industrial technologies.

Integration

The success of Industry 5.0 relies on the integration of human-centricity, sustainability and resilience into a coherent and mutually reinforcing system. A holistic approach functions as a design vision and a practical commitment that should be supported by cross-sector collaboration and participatory innovation. When these domains work in synergy, improvements in one area strengthen others, generating cross-over effects that advance multiple domains simultaneously.

While the ARISE framework acknowledges the integration of all three pillars, **human-centricity** emerges as the primary focal point given the project's specific emphasis on HRI.

[9] Callari, T. C., Vecellio Segate, R., Hubbard, E.-M., Daly, A., & Lohse, N. (2024). An Ethical Framework for Human-robot Collaboration for the Future People-centric Manufacturing: a Collaborative Endeavour With European Subject-matter Experts in Ethics. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.4927452

Co-funded by the European Union

Consequently, the SSH framework prioritizes the human element as the central driver for technological design of industrial robots[10].

2.2 Mapping of Existing Approaches

In order to establish a foundation for the ARISE SSH framework, it is important to examine the range of existing ethical frameworks and methods that have been developed within the European research landscape. Over the past decade, numerous initiatives have sought to address ethical questions arising from advances in AI and robotics. These efforts provide a conceptual and a practical guide for integrating ethical considerations into system design and development. The following subsections present a selection of these frameworks and highlight their key features, along with methodological approaches and relevance for the context of human-robot interaction.

Table 1. Similar projects to ARISE from SSH perspective and their key characteristics.

Project Name	Focus Area	Funder	Website	Methods	Main deliverables
SIENNA	Ethics (by design)	Horizon Europe	SIENNA	Literature review	Frameworks, Impact Assessment
SHERPA	Ethics	Horizon Europe	SHERPA	Delphi study	Frameworks, Guidelines, Impact assessment
REELER	Human-Proximity Model	Horizon Europe	REELER	Ethnographic study	Roadmap, Human-Proximity Model, Toolbox
VIRT-EU	Impact Assessment	Horizon Europe	VIRT-EU	Ethnographic study	PESIA framework, Ethics Primer and Stack
AEQUITAS	Ethics	Horizon Europe	AEQUITAS	Experiments	Frameworks, Tools
Robotics4EU	Ethics	Horizon Europe	Robotics4EU	Societal Readiness Plan	Maturity Assessment Model, Chatbot, Robotics Compass
ETAPAS	Ethics	Horizon Europe	ETAPAS	Literature Review	RDT Framework, Code of Conduct
SATORI	Impact Assessment	European Union Seventh Framework Pro- gramme (FP7/2007-2013)	SATORI	Foresight	Ethical Impact Framework
ALTAI	Compliance Assessment	EU Commission	ALTAI	Survey In-depth interviews Feedback reporting	Compliance assessment framework

[10] For more details on the foundational definition of human-centricity within the ARISE context see section 2.3.4.2 Human-centricity in industrial HRI

2.2.1 Relevant Frameworks, Methods and Tools

While numerous ethical frameworks and tools have been developed within European research projects, as outlined in Table 1, this section focuses on a selected set of approaches identified as particularly relevant for ARISE. The selection is based on their conceptual strength, practical orientation and suitability for supporting ethical and human-centric robotics development across different lifecycle stages. The following frameworks provide important reference points for the ethical grounding and methodological choices adopted in ARISE.

2.2.1.1 SIENNA

A comprehensive ethical framework was developed in the SIENNA project in its deliverable report 4.7, entitled "An ethical framework for the development and use of AI and robotics"[11]. It proposes a set of methods and procedures for developing, deploying and using AI and robotics systems in a way that adheres to ethical principles. Taking into account earlier proposals of the Ethics Guidelines for Trustworthy AI presented by the High-Level Expert Group on AI (HLEG-AI)[12] and IEEE's Ethically Aligned Design framework[13], SIENNA introduces six high-level ethical principles for incorporating ethics into research and development of AI & robotics. These high-level principles are based on the ethics guidelines by the SHERPA project, namely human agency, liberty and dignity; privacy and data governance; transparency; diversity, non-discrimination and fairness; individual, societal and environmental well-being; accountability.

For the development of robotics, SIENNA assumes the existence of shared practices or phases described as the "V-model" (Fig. 1), which can then be accompanied by guidelines for the incorporation of ethical considerations. This V-model approach is paired with a checklist, where the different phases of robotics development are measured against the six ethical principles mentioned before. An overview of the V-model checklist is illustrated in Figure 2.

[11] Brey, P., Jansen, P., Maas, J., Lundgren, B., & Resseguier, A. (2021). SIENNA D4.7: An ethical framework for the development and use of AI and robotics technologies (1.1). Zenodo. https://doi.org/10.5281/zenodo.7266848
[12] European Commission. (2019). Ethics guidelines for trustworthy AI | Shaping Europe's digital future. European Commission. https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai
[13] IEEE. (2019). Ethically Aligned Design: A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems. https://standards.ieee.org/wp-content/uploads/import/documents/other/ead_v2.pdf

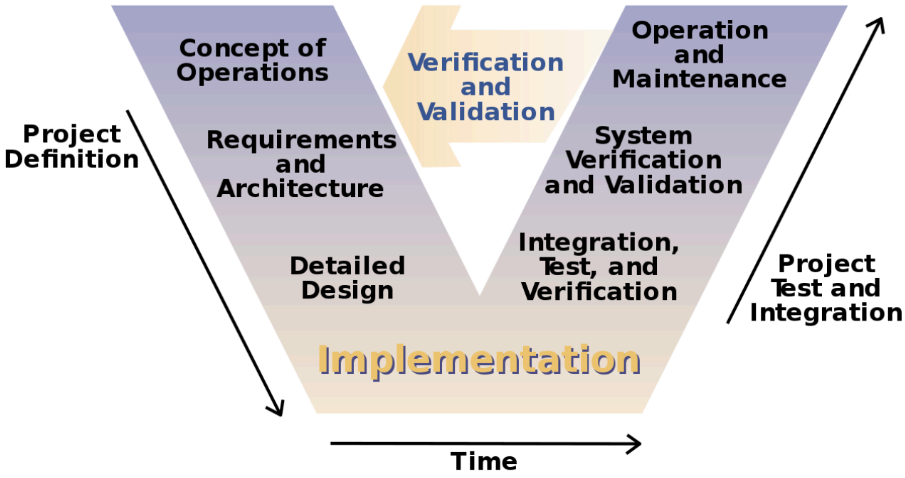


Figure 1. The V-model, a graphical representation of a systems development lifecycle, as presented in SIENNA. Source: SIENNA D4.7: An ethical framework for the development and use of AI and robotics technologies.

2.2.1.2 Ethics by Design

Ethics by Design, presented in the 'Ethics By Design and Ethics of Use Approaches for AI' framework[14], outlines a proactive human-centered approach to technology development. It aims to make people think about and address potential ethical concerns while they are developing a system. Ethics by Design is a general model, which can be implemented alongside various development processes such as V-Model described above or Agile[15].

The Ethics by Design model builds on six iterative phases:

1. **Specification of objectives**: Determining what the system is for and what it should be capable of doing)
2. **Specification of requirements:** Developing technical and non-technical requirements for building the system
3. **High-level design: Developing** high-level architecture
4. **Data collection and preparation:** Collecting, verifying, cleaning, and integrating data
5. **Detailed design and development:** Construction of a fully working system
6. **Testing and evaluation:** Testing and evaluating the system in context of the ethical requirements.

[14] European Commission. (2021). Ethics By Design and Ethics of Use Approaches for Artificial Intelligence. https://ec.europa.eu/info/funding-tenders/opportunities/docs/2021-2027/horizon/guidance/ethics-by-design-and-ethics-of-use-approaches-for-artificial-intelligence_he_en.pdf

[15] Beck, K., Beedle, M., van Bennekum, A., Cockburn, A., Cunningham, W., Fowler, M., Grenning, J., Highsmith, J., Hunt, A., Jeffries, R., Kern, J., Marick, B., Martin, R. C., Mellor, S., Schwaber, K., Sutherland, J., & Thomas, D. (2001). Manifesto for agile software development. Agile Manifesto. https://agilemanifesto.org/

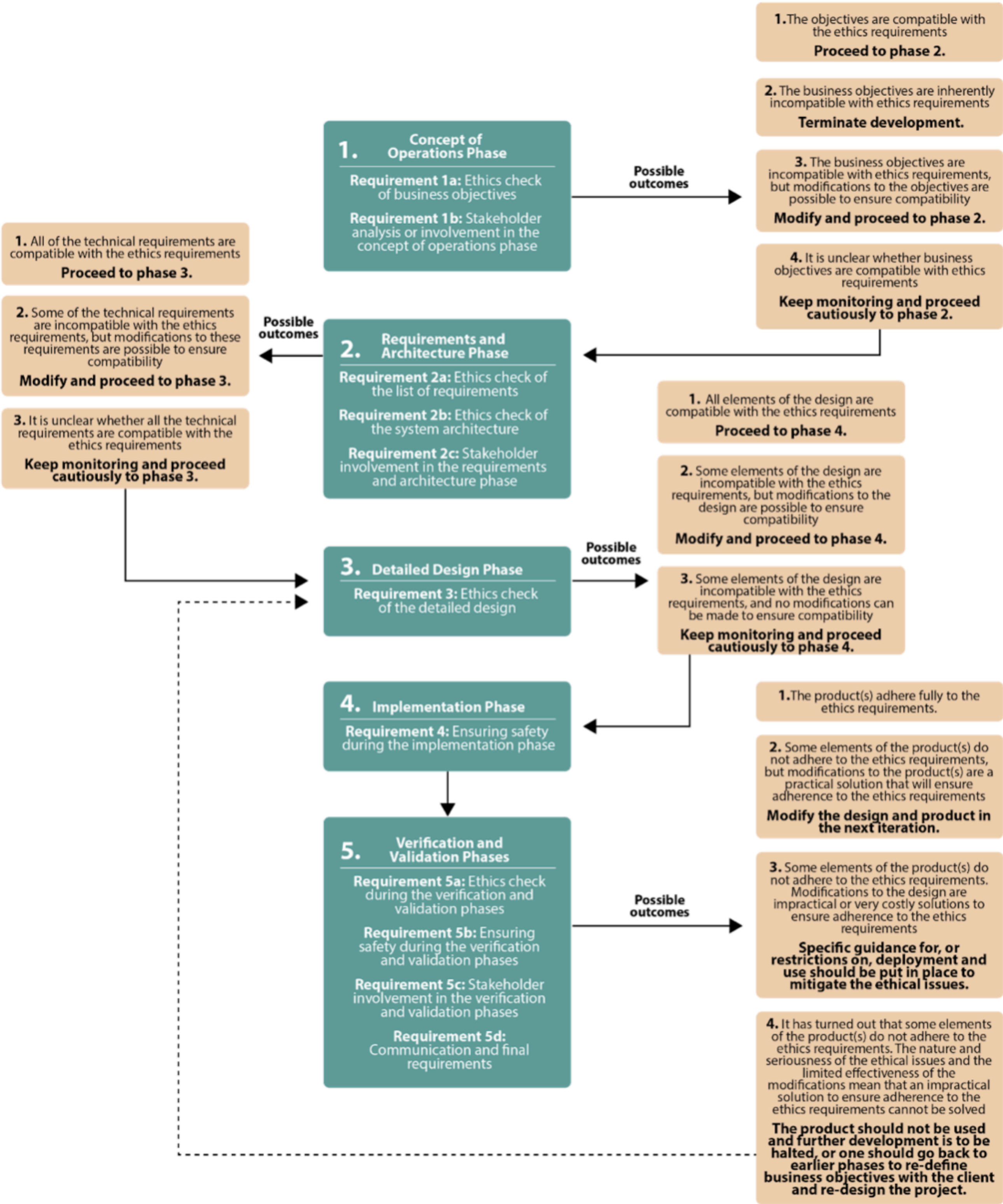


Figure 2. The V-model checklist, adapted from the SIENNA framework. Adapted from SIENNA D4.7: An ethical framework for the development and use of AI and robotics technologies.

Importantly, these are not linear steps that follow another, but rather steps that should be implemented into appropriate phases of the chosen development process. Figure 3 illustrates the six phases included in the Ethics by Design framework. As one of the most influential approaches in contemporary technology ethics, Ethics by Design represents a state-of-the-art method for embedding ethical considerations throughout system lifecycles.

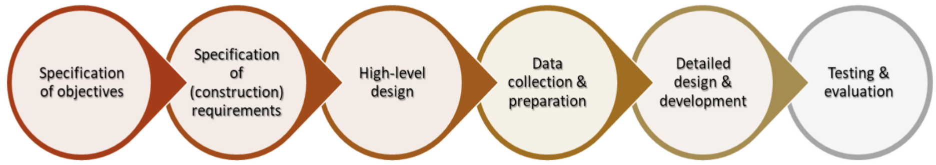


Figure 3. The six phases included in the Ethics by Design framework. Source: "Ethics By Design and Ethics of Use Approaches for Artificial Intelligence".

2.2.1.3 Impact Assessment & SATORI

The ethical impact assessment (EIA) mentioned in the SIENNA project's website[16] derives from the SATORI project. The SATORI project provides a comprehensive EIA framework and structured methodology[17] that evaluates the ethical implications of research and innovation (R&I) projects. The framework consists of six stages:

1. **Threshold analysis**: Determines if ethical issues are present that require a full EIA
2. **EIA plan**: Developed if threshold analysis identifies ethical issues (including team, budget, stakeholder engagement)
3. **Ethical impact identification**: Identifies potential ethical impacts using literature reviews, foresight methods (e.g., horizon scanning, Delphi), and stakeholder input
4. **Ethical impact evaluation**: Evaluates the severity, likelihood, and importance of ethical impacts
5. **Remedial actions**: Develops recommendations and design interventions to mitigate negative impacts and enhance benefits
6. **Review and audit**: Ensures accountability through regular reviews and audits of the EIA process.

2.2.1.4 Human-Proximity Model

Within the REELER project, ethical concerns emerging from the disconnect between robotics developers and directly affected stakeholders were coined as "proximity gap". This entails a physical and cultural distance, for example in the form of values and motives, between those parties. One solution provided by REELER was to close these situation-specific gaps by seeking alignment. They suggest the introduction of alignment experts, whose tasks would involve spanning the space between the robotic bubble with their core experts and the

[16] Ethical impact assessment - SIENNA. (2024, February 15). Sienna-Project.eu. https://www.sienna-project.eu/w/si/about-sienna/eia
[17] Satori Project. (2017). English version Ethics assessment for research and innovation -Part 2: Ethical impact assessment framework. https://satoriproject.eu/media/CWA17145-23d2017.pdf

affected stakeholders' communities to identify their separate motives, and to communicate them to one another in a move toward alignment and collaborative learning. Figure 4. Illustrates the human-proximity model[18], which is included in the broader awareness-raising toolbox by the REELER project.

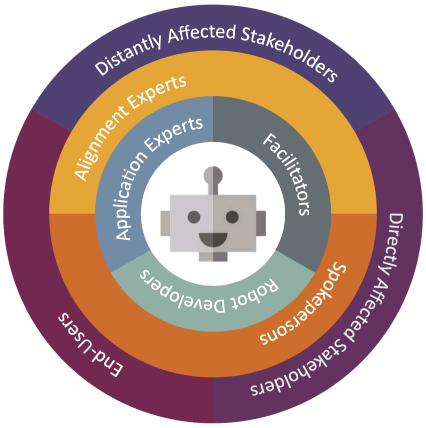


Figure 4. The Human-Proximity Model by the REELER project.

2.2.1.5 ALTAI

The Assessment List for Trustworthy Artificial Intelligence (ALTAI) was developed by the European Commission's High-Level Expert Group on AI (HLEG-AI) to operationalise the Ethics Guidelines for Trustworthy AI[19]. It provides a self-assessment tool for organisations and developers to evaluate whether their AI systems meet the seven key requirements of trustworthy AI: 1) human agency and oversight; 2) technical robustness and safety; 3) privacy and data governance; 4) transparency; 5) diversity, non-discrimination, and fairness; 6) societal and environmental well-being; and 7) accountability.

Rather than prescribing fixed metrics, ALTAI offers a structured reflection process, encouraging developers and organisations to identify, assess, and mitigate ethical risks throughout the lifecycle of AI systems. The tool can be adapted for robotics and HRI contexts by guiding designers to consider how robotic systems respect human autonomy, remain transparent in interaction, and uphold safety and inclusivity in real-world settings. ALTAI thus functions as both a practical checklist and a governance framework.

2.2.2 State of the Art of Existing Approaches

The review of existing initiatives indicates that the European landscape of ethical frameworks, methods and tools for AI and robotics is extensive and diverse. In particular, increased policy attention and sustained research funding over recent framework programmes have

[18] Human-proximity model. (2020). REELER project. https://reelertoolbox.ab-acus.com/
[19] European Commission. (2020, July 17). Assessment List for Trustworthy Artificial Intelligence (ALTAI) for self-assessment | Shaping Europe's digital future. Digital-Strategy.ec.europa.eu. https://digital-strategy.ec.europa.eu/en/library/assessment-list-trustworthy-artificial-intelligence-altai-self-assessment

Co-funded by the European Union

contributed to a growing body of work addressing ethical aspects of AI and robotic systems. Many of the projects analysed in the previous section have been supported through European research and innovation programmes, reflecting a broader strategic emphasis on responsible and human-centric technological development.

Across this body of work, a variety of ethical frameworks explicitly refine earlier approaches. For example, many frameworks draw on established principle-based guidance, adapting it to specific application contexts (e.g. SIENNA framework uses SHERPA's and IEEE's ethically aligned design approaches and refines them). Despite differences in scope and emphasis, **there is a strong convergence around a core set of ethical concerns,** including **human agency and autonomy, dignity, privacy and data protection, fairness and non-discrimination, well-being, transparency** and **accountability.** In addition to these high-level principles, a range of practical methods and tools have been developed to support its implementation.

Ongoing research, emerging initiatives and evolving regulatory discussions continue to shape the field in Europe and internationally. Monitoring developments and critically assessing their relevance for human-robot interaction forms an integral part of the work on the ARISE SSH framework. In addition, the question of applying these frameworks and tools has to be evaluated closely. Many of them analyse the question "what?" closely, but do not give an adequate answer to "how?". Through this assessment, the ARISE project aims to develop a unified framework that enriches these existing approaches. The following section describes the methodology through which this work is carried out.

2.3 Methodology

2.3.1 Objectives of this Framework

The primary objective of the ARISE SSH framework is to **integrate key ethical, legal, and social implications of HRI development into a unified and actionable framework.** Although centred on HRI, the approach extends to the broader fields of robotics and AI to facilitate coherent application across technological domains. This strategy addresses two critical limitations in the current landscape: first, the **fragmentation** where ethical, legal, and social dimensions are treated separately from technical expertise; and second, the **abstract nature of existing frameworks**, which often lack the specificity required for industrial application. This targeted approach is critical for smaller ecosystem actors—including SMEs, research labs, and pilot teams—facing the challenge of aligning with upcoming **Industry 5.0** requirements.

Embedding ethical, societal and environmental priorities into current industrial systems remains difficult, particularly within the dominant paradigm of mass production[20]. Several systemic constraints include the inherent rigidity, uneven organizational capacities, limited institutional support and the dual need to integrate new technological components and

[20] Nasir, V., Hosseini, A., Binfield, L., Hasani, N., Ghotb, S., Diederichs, V., Hansen, E. (2025). Human-centric Industry 5.0 manufacturing: a multi-level framework from design to consumption within Society 5.0. International Journal of Sustainable Engineering, 18(1). https://doi.org/10.1080/19397038.2025.2551000

reconfigure work structures (roles, workflows, leadership and decision-making). The introduction of these new practices also requires broader participation of different actors within technology organisations, from managers and engineers to operators and designers, which creates uncertainty, a lack of ownership and resource constraints that small actors may struggle to meet. Nevertheless, **SMEs and other smaller actors can play a pivotal role** as their agility and capacity for customized production places them in an ideal position to prototype more human-centric models, allowing them to catalyse wider industrial change from within.

However, addressing these challenges requires a shift that extends beyond technology development alone. Stakeholders across the industrial ecosystems (manufacturers, workers, clients, regulators, communities) often differ in their stance on the role ethics should play in the industrial transformation[21]. The integration depends on a shared language and understanding of ethics and human-centricity that clarifies evolving responsibilities, enables inclusive dialogue and co-creation and aligns projects with societal objectives. Establishing this common ground of principles and practices is key to supporting the application and acceptance of ethical practices across the future industry.

In response to the current challenges, this framework pursues **5 practical objectives:**

1. **Unify ethics management and human-centricity** into a coherent approach that moves past checklist compliance toward emerging best practices, aligning legal considerations with shared societal values.
2. **Contextualise guidance for industrial HRI**, translating broad principles into project steps for design, PoCs, pilots, and scaling phases, initiating and maintaining broader-level discussions to raise awareness across disciplines.
3. **Reduce complexity and lower the entry barrier** for SMEs and experiment pilots with methods, templates, and examples that can be adopted incrementally, starting from the earliest stages of technology maturity.
4. **Embed participatory practices** so ethical considerations inform real design and deployment decisions.
5. **Enable measurable progress** by establishing indicators and accountability mechanisms that allow teams to track and demonstrate their advancement from basic compliance to ethical excellence and societal acceptance.

The ARISE SSH framework is the primary output of ARISE's SSH & ethics workstream, integrating state-of-the-arts knowledge, experiential insights, and mentoring experiences into a single, iterative reference that HRI developers can use to plan, execute, document and improve ethics management and human-centricity alignment.

Iterative development and validation

The contents of this framework were developed iteratively, combining insights from prior research with contextual evidence collected through ethnographic studies, and then validated through the support and mentoring programme for ARISE's first open call beneficiaries. The

[21]Nasir, V., Hosseini, A., Binfield, L., Hasani, N., Ghotb, S., Diederichs, V., Hansen, E. (2025). Human-centric Industry 5.0 manufacturing: a multi-level framework from design to consumption within Society 5.0. International Journal of Sustainable Engineering, 18(1). https://doi.org/10.1080/19397038.2025.2551000

first version of this framework was tested during the design and deployment of the activities within the dedicated support service. The findings from this validation (see section 2.3.3 Insights informing this framework) informed this second iteration while further improvements are expected to be incorporated in the final version, as the framework continues to be applied throughout the remainder of the first open call and throughout the second open call. Feedback from the ARISE TEFs and consortium partners was also gathered and continues to inform its development to better integrate the framework's concepts and activities within the structures of HRI development. The final version of the framework is planned to be released by the end of the project (month 42).

2.3.2 Research Methodology: Integrating Contextual and Ethnographic Knowledge

The SSH framework is grounded in **transdisciplinary, qualitative research** that integrates participatory, ethnographic, and human-centric design methodologies to analyse HRI practices and processes, and to surface practitioners' values, aspirations, and needs from an SSH perspective. The research is **ongoing throughout the project**, continuously informing updates as new evidence emerges. It examines current and future practices in which SSH and, in particular, how requirements concerning ethics and human centricity can be embedded. A key research focus is the **day-to-day, in-context work of HRI development.** To this end, the research employs a collaborative approach that brings together SSH experts, ARISE consortium partners, HRI technology developers, and end users participating in the ARISE open-call programmes.

In practice, the research activities aimed at generating contextual knowledge have comprised expert interviews, workshops, a survey and notably onsite ethnographic visits to all four TEF partners. The research scope also encompasses the ARISE open calls, involving direct engagement with beneficiary projects through a dedicated mentoring service focused on ethics management and human-centricity. These interactions serve as the primary mechanism for testing and validating both the SSH framework and its associated mentoring model. Finally, the coordination of external ethics assessments, conducted by the ARISE Ethics Committee, provided a complementary perspective, shedding light on formal compliance challenges and generating further critical learnings. The specific activities are chronologically detailed below:

- **Spring 2024:** Alongside an initial literature review (February-May), SSH partners conducted semi-structured interviews with ARISE HRI professionals and external technology ethics experts, to scope current use of ethics frameworks, tools, and SSH practices in HRI development. Interviews were conducted between May and August.
- **Autumn 2024:** A series of workshops with technical partners examined Ethics by Design principles and how they are, or could be, reflected within ARISE and beyond.
- **Autumn 2024**: Two SSH experts visited all four ARISE TEFs to conduct ethnographic fieldwork—combing observations, interviews, workshops and visual methods—analysing day-to-day practices, workflows, and decision processes, and surfacing practitioners' values, aspirations, and needs related to ethics and human-centricity. These 2-3 day visits helped bridge disciplinary perspectives

between humanities and technology. Field notes were systematically analysed to distill insights, and the photographic material was curated into a short and an extended series (Fig. 5), complementing the findings with non-verbal evidence and spatial context. The ethnographies identified what works well, distilled good practices, and tested ways to embed existing ethics frameworks and processes into HRI development. Additionally, the first iteration of the ARISE SSH framework was released by the end of the year.

Figure 5. Selected images from the ethnographic photo series. Credits: M. Urra Sch.

- **Spring - Autumn 2025:** The ARISE mentoring support service for the first open-call beneficiaries kicked off in May 2025. The programme provides 12 months of funding and mentoring to 13 SME-led HRI projects across seven countries in industrial and healthcare contexts. Each project pairs a technology provider with an industrial end user who supplies the real-world environment for testing and validation. A dedicated support service on ethics and human-centricity, led by ARISE SSH experts, guides teams in applying the first release of the SSH framework, both through targeted sections in project deliverables and via support activities (workshops, clinics, 1:1 calls). Feedback is collected continuously to refine both the framework and the support service. Additionally, the SSH experts conducted a survey to gather perspectives on ethics from a diverse range of professionals operating within the ARISE robotics ecosystem.

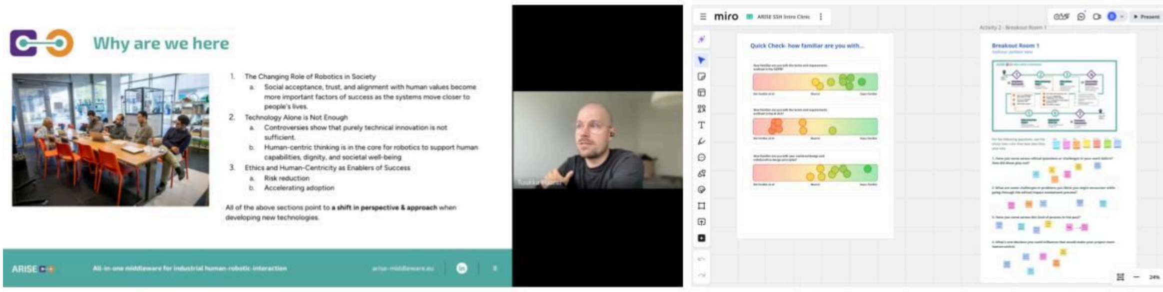


Figure 6. Visual documentation of online workshops and Miro board .

- **Autumn 2025**: Insights from all research activities informed the guidelines, tools, specifications, and recommendations in this second version of ARISE SSH framework, ensuring that they are relevant, understandable, and fit to be used by real HRI development projects.

The **qualitative data collected** in all activities comprises literature review insights, interview transcriptions, field notes, photo series, workshop notes, workshop outputs (e.g. Miro boards), feedback forms, and survey responses, as well as the tacit knowledge gained through immersive engagement in the research context. This data was systematically analysed using content analysis to derive the preliminary findings reported in the next section. Furthermore, consolidated syntheses of these findings are scheduled for dissemination via academic publications in 2026.

2.3.3 Insights Informing this Framework

As elaborated upon in the previous section, the ARISE SSH framework builds on key learnings from contextual and ethnographic research. The following subsections synthesize key learnings from these engagements into **4 themes** (see table 2 at the end of the section), with each theme covering challenges and opportunities for the ARISE project to develop more effective strategies for integrating ethics and human-centricity in HRI work.

2.3.3.1 Structured and Hierarchical Decision-Making

HRI development work typically takes place in **highly structured environments.** Teams have specific projects, tasks, and responsibilities allocated to them that align with their organization's higher-level goals. They work with processes and frameworks such as Scrum[22] and Agile[23] to ensure that their work meets internal and external deadlines while also making optimal use of their organization's resources. Bringing in additional novel considerations, such as **human centricity and ethics, into this highly structured work environment can be disruptive** to these efficiency-optimized workflows if not carefully introduced and integrated. This disruption may add changes to scopes and timelines by introducing requirements often perceived as not functional; it demands lateral and multi-actor coordination of inputs across

[22] Scrum.org. (2024). What is Scrum? Scrum.org. https://www.scrum.org/learning-series/what-is-scrum/
[23] Beck, K. (2001). Manifesto for Agile Software Development. Agilemanifesto.org. https://agilemanifesto.org

silos, beyond the usual decision-making paths; and surfaces trade-offs that may conflict with already established objectives and KPIs.

The highly structured nature of HRI development also means that **higher-level decision-making** (e.g., setting the project roadmap, outlining and integrating organizational goals, deciding on external partnerships, etc.) **may not be accessible to everyone working on the project**, **nor at every phase of the project**. This poses a challenge for integrating ethics and human-centricity as often, ethical concerns arise during the development process at a phase where making major changes to design requirements may not be feasible, or may lead to conflict with criteria that was set by decision makers at the organization (e.g., the CEO issues a directive to integrate AI models that poses threats to the well-being of users).

An additional challenge arises from **ambiguity in ownership and accountability**. Encouraging individuals to 'consider ethics' in their work is not enough if there is no shared understanding of who flags emerging concerns, how they are triaged, and **where and when they enter existing workflows**. Without defined roles, decision paths and documentation expectations, ethical issues remain diffuse and rarely convert into concrete actions.

Opportunities:

- Design tools for integrating a human-centric and ethical approach **into existing workflows and structures** that are common in the HRI industry, such as AGILE.
- Ensure that discussions around ethics **are not left in the abstract**, but grounded in the actual work that takes place by including practical considerations of who is accountable for tracking and integrating emerging ethical issues of a project.
- Look for ways to **integrate ethical considerations during key decision-making moments of a project**, including when higher-level agenda and roadmap setting takes place at the management and leadership level of the organization.
- Overcome the lack of ownership by **defining SSH roles and expectations**, and **specifying the needs for new capacities** within existing or new HRI roles.

2.3.3.2 Limited Capacity for Interdisciplinarity

Due to the highly technical and complex nature of robotics development, HRI workers are typically expected to utilize their specialized backgrounds to contribute to a **niche part of the robotic solution.** This means that HRI workers often complete their tasks with self-imposed blinders– working only with the information that's necessary for their roles. This specialized information is already highly complex and constantly updated, requiring them to be aware of new external developments. Trying to consistently understand and keep up with what other team members in the project or organisation are working on would require an enormous amount of time, energy, and background knowledge that they likely lack. This isn't to say that cross-team interactions never happen, but rather that **interactions with other teams are often carefully scoped and timed.**

While people in highly technical roles, such as engineers and developers, might have great interest in ensuring ethical and human-centric HRI development, they frequently lack time and resources to engage deeply with ethical regulations, human-centric methodologies, and general SSH topics (including the content of this framework). This creates a critical tension: as the ethnographic observations confirm, these individuals **hold significant agency over the granular decisions that define how technical components interact with users.** These technical choices shape the entire human-machine relationship, determining not only the functional safety of the system, but also the quality of the experience of interaction, both of which carry profound ethical and human-centric implications. Overcoming this challenge requires that technical roles be **supported by complementary expertise**, facilitating a **low-threshold engagement** with ethical and human-centric issues that fits within their existing constraints.

In addition, **recognising emerging ethical issues is difficult without a holistic and even anticipatory view** of the robotic system and **how it will be deployed and used in real settings.** In many cases, HRI components or modules are designed without a clear picture of their real-world application, leaving ethical impacts too abstract or invisible. This was evident during the TEF visits: TEFs primarily develop technological components, run experiments, and validate results at an experimental/validation level, serving its expertise as role models and support hubs for actors (e.g., open-call beneficiary projects) who will later apply these solutions in operational contexts. Consequently, individual specialists may not readily see downstream effects. For example, a mechanical engineer who's concerned with designing and fabricating the housing for a humanoid robot's upper arm likely wouldn't be able to intuitively connect their contribution to end user's well-being or perceptions of agency and fairness once the system is deployed.

Applying ethics and human-centric principles to emerging industrial technologies means working with **multi-level complexity**. Yet multidisciplinary work introduces new challenges. For instance, bringing together workers from diverse backgrounds to discuss ethics and human centricity can reveal **gaps between workers' overall perception of the project,** which became visible in the mentoring workshops. Multidisciplinary teams might be better at considering impacts ranging from the **immediate context of use** (e.g., factory practices, new skills, new actors) to the **societal sphere** (e.g., demographic shifts, work–life conditions, inclusion of women in the workforce). But at the same time multidisciplinarity may introduce unclear boundaries of responsibilities (e.g. who should be aware of evolving ethical concerns, who should participate in ethics-specific tasks) which do not fit the pre-existing conditions of HRI industry.

Ethics must be understood as an **empirical practice** situated at the intersection of values, choices, and real-world consequences. Because technology and SSH have traditionally sat in separate disciplines, applying ethics in HRI also requires introducing **new capacities** within teams, presenting an added challenge **specially for smaller or emerging actors** such as SMEs and experiment pilots.

In short, a challenge that may arise is the limited capacity of HRI workers to **think in interdisciplinary terms** because of the **lack of shared language** around human centricity in HRI development, the **limited contextual scope** of their contribution relative to the entire solution, and a **general lack of time, energy, and resources** for building new capacities.

Co-funded by
the European Union

Opportunities:

- Develop a **shared language and understanding** for ethics and human-centricity to facilitate regular and inclusive discussions across disciplinary boundaries. Produce contextually grounded essentials such as guiding principles, key stakeholder values, plausible negative impacts, and the positive-seeked outcomes, using narrative elements (including real scenarios), so they are easy to grasp across teams.
- Leveraging expertise to **translate complex SSH concepts into low-threshold, actionable guidance**, enabling technical teams to integrate ethical considerations without exceeding their resource constraints.
- Leverage the insights and learnings from **key positions that already play interdisciplinary roles** in technology development contexts, such as project managers and designers.
- Knowing that cross-team interactions may be more resource-intensive than intra-team interactions, suggest **interdisciplinary interventions** appropriately.
- Provide concrete **examples of integrating ethics that map to specific roles** within a technology organisation (i.e., here are the considerations for <ethical principle> that a software developer can apply).

2.3.3.3 Context Specific Needs and Challenges

Although HRI work is often discussed as a monolith, it is important to recognize that **different contexts of HRI development have different needs and challenges.** Even just in ARISE, the technology partners and beneficiary projects represent a wide range of organizations, from research oriented labs to product oriented startups and industrial facilities. Solutions range from social robots for health care to collaborative robots for industrial manufacturing settings like construction, storage or fabrication of very specialized components. Each of these sectors already follows very specific regulatory frameworks. As HRI and robotics are transversal to several sectors in the industry, sectors introducing these technologies might not be aware of additional norms outside their boundaries. Accordingly, when bringing ethical considerations into this diverse set of contexts, not all regulatory frameworks will be familiar, nor every tool, principle, and practice will be equally applicable or critical in every setting.

In ARISE, there have been repeated requests to 'apply' the SSH framework across both TEFs and open-call beneficiary projects. For beneficiary projects, a dedicated SSH and human-centricity support service has provided a clear entry point and limits to consolidate requirements, guidance, and documentation. For TEFs, the application has been more abstract, reflecting heterogeneous structures, workflows, and mandates across the facilities, a challenge that underscores the need for flexible, context-sensitive integration. Future open calls may further widen this spectrum of contexts.

Technology readiness levels (TRLs) may also vary. **Many ethics frameworks and tools assume well-known end users and relatively high TRLs.** In contrast, in the context of ARISE, much of what is being developed is not at the highest TRL levels but is rather seen as a "platform" or prototype of something to be iterated further. At the same time, **solutions being developed are assemblages of different technologies** (components, modules), **with a distributed network of working teams** bringing them together, raising compatibility, alignment and cross-organisational communication challenges. This reality poses difficult questions for considerations of ethics, responsibility and accountability: design choices affect the affordances of the technological solution up to a point, but also leave open a wide range of uses, of which the lower TRL levels might not have certainty. **How could ethics considerations be embedded in earlier TRL levels? Which design choices can preclude unethical uses at the lower TRL levels?** Addressing these questions is essential to ensure that ethics remains an integral part of the development process, even under uncertainty.

These points highlight a practical tension across TRLs: **at lower TRLs,** ethical impact and human-centric alignment can be **too abstract to operationalise; at higher TRLs, architectures and processes are already fixed,** making fundamental course corrections difficult.

Opportunities:

- **Consider an approach of specificity:** provide tools and examples of human-centric HRI development that are tangible and appropriate for specific robotics contexts.
- **Consider an approach of generalizability:** provide tools and frameworks that can be flexibly utilized across diverse contexts: TRLs, organization types, and solution types.
- Acknowledging that ethics management is hard to operationalise at low TRLs and hard to change at high TRLs, the implication is to **embed ethics early and iteratively, developing transferable documentation** (across different teams, organisations and contexts), **staged checkpoints and design handrails**, so societal considerations mature and become more tangible alongside the technology.

2.3.3.4 Leverage Existing Norms and Practices

While SSH and human-centricity may be relatively novel terms in the space of robotics, there are several existing norms and practices that can be leveraged to push for more ethical robotics development. One key aspect of HRI development is iteration: **teams often work in iterative cycles on a particular part of the project**, incorporating research, requirements, prototyping, development, testing, and user validation and feedback in a continuous cycle. Once a solution has been tested and validated in a controlled setting, the team may repeat similar cycles of iteration for testing in real-world settings. The ethnographic research indicates that this iterative, task-specific approach is intentionally exploratory: teams accept that some trials will fail, and that small, contained failures are a productive source of learning rather than a negative outcome.

Building on this mindset, ethics and SSH tools and frameworks can be packaged in **small, manageable actions** that can be embedded early and continuously. A familiarity with iteration also means that HRI developers may be more willing to explore, test out, and integrate novel and unfamiliar considerations of ethics and human-centricity if these considerations are introduced gradually and in digestible portions.

Another existing framework that ARISE can leverage is the concept of 'user' and **'user-centricity'.** While Industry 5.0 and the ARISE project promote ***human-centric* technology development, *user*-centric design** is already a familiar paradigm for many HRI developers, particularly those in design-related roles. Instead of re-inventing the wheel, ARISE can expand upon this existing design framework while introducing the more holistic values of human-centricity. However, in the context of HRI, the concept of *human-centricity* should still be broadened to address **impacts at a societal level**, beyond the immediate contexts of use, consumption, and the mass production paradigm, and even including environmental considerations.

An additional learning from the ethnographic work and mentoring process is the growing and **sustained interest** in holding technology development to higher ethical standards. Many HRI practitioners have explicitly voiced their desire to develop robotic solutions that prioritize the **safety, well-being, and agency of their users**. At the same time, there's growing systemic, policy-induced momentum from directives such as the **AI Act** and the **Commission's commitment to Industry 5.0.** ARISE's work in human centric HRI development comes at a crucial and opportunistic time for aligning with industry, cultural, and policy trajectories.

Opportunities

- Design around the **iterative and explorative nature of HRI** work to gradually integrate and test out human-centric/ethics framing tools early and iteratively.
- Expand upon **user-centricity tools and paradigms** to encompass a HRI-specific common understanding.
- Capitalize on ongoing **interest in and institutional support for human-centric technology development.**

Table 2. Themes and opportunities from ethnographic research.

Theme	Opportunities
Structured and hierarchical decision making	• Integrating human-centric and ethical approaches into existing workflows and structures. • Discussions around ethics contextually grounded. • Integrating ethics at key decision-making moments of a project. • Overcoming the lack of ownership by defining SSH roles and expectations.
Limited capacity for interdisciplinarity	• Developing a shared language for ethics and human-centricity. • Leveraging expertise to translate complex SSH concepts into low-threshold actionable guidelines. • Leveraging positions that already play interdisciplinary roles. • Plan interdisciplinary interventions appropriately. • Providing concrete examples relevant to specific roles.
Context specific needs and challenges	• Approach of specificity: Examples of tangible human-centric HRI. • Approach of generalizability: Adaptable tools to diverse contexts. • Awareness that ethics are hard to apply at low TRLs but hard to change at high TRLs.
Leverage existing norms and practices	• Design around the iterative and explorative nature of HRI work. • Capitalise on ongoing interest in and institutional support for human-centric technology development.

2.3.4 A Holistic Approach to Human-centric HRI

Drawing on the research and insights presented, this section establishes the rationale and the guiding narrative that underpin the framework.

The ARISE SSH framework integrates **ethics** (the management of ethical impacts) and **human-centricity** (the alignment to design principles that improve current conditions) as one **integrated practice**. It is a holistic approach to industrial HRI development that advances societal well-being while anticipating and minimizing harm. The approach integrates regulatory duties and value-driven choices into a coherent view, using **narrative sense-making** to bridge collaboration challenges across contexts and disciplines. Ethical considerations are introduced early and iteratively, with clear argumentation and actionable measures, so they can evolve and mature along with the technology.

The framework is defined by **five main characteristics:**

- **Holistic and integrated →** Integrates compliance with EU regulations and human-centric aims—such as agency, inclusion, well-being, and environmental responsibility—into a single practice, ensuring they are addressed together rather than in silos or parallel tracks.
- **Narrative sense-making →** Bridges disciplinary vocabularies and perspectives by using context-specific stories to connect abstract values to concrete design choices. It emphasizes communication to make ethical reasoning open to discussion, argumentation, and transferability across contexts.
- **Collaborative →** Designed to co-produce decisions with those who build and operate HRI systems, as well as those affected by them. It recognises the influence of ecosystem stakeholders, ensuring that needs, risks, and opportunities are surfaced in a timely manner and at the appropriate depth.
- **Iterative form the beginning →** Embeds ethics and human-centricity from the outset, iterating alongside each technical sprint. Initial low-threshold prompts evolve into sophisticated, transferable requirements that mature with the technology as it advances through higher TRLs.
- **Applicable across contexts →** Modular and context-sensitive, the framework bridges diverse solution types, organisation sizes, and development stages. It prioritises identifying phase-specific challenges and capturing transferable evidence, ensuring insights remain accessible across different projects and settings.

Building on this approach, the next section sets out two contextualised narratives for HRI. The first, **ethics management**, elaborates on how ethical considerations can be embedded and organised within HRI projects. The second, **human-centricity**, proposes the elements of a common understanding of what human-centricity signifies for industrial HRI. Together, these narratives provide a conceptual foundation for translating the framework's principles into practice.

2.3.4.1 Ethics Management in Emerging HRI Technologies

The transformative capacity of technology is historically evident. For better or worse, examples ranging from the steam engine to the smartphone evidence that technology can radically reshape individual lives and societal structures.

Consequently, ethics must be regarded as an integral component of emerging technologies. In this framework, **ethics management is defined as the process by which ethical issues are recognised, discussed, and embedded into organisational policy and practice.** However, establishing a functional approach requires navigating the complexity of defining *"ethics"* itself. While history offers several philosophical definitions, the context of applied ethics in emerging technology requires a pragmatic stance. Drawing on Reijers, Young & Coeckelbergh (2025)[24], we posit that for the purpose of its application, a precise academic definition is less critical than a broad, shared understanding of the concept's key elements.

This shared understanding starts with the **normative and anticipatory nature of ethics**; it concerns the distinction between how the world is now and how it should or must be, focusing mostly on the second. An ethical lens allows actors to distinguish between desirable and undesirable outcomes, judging technologies based on their potential impact rather than solely on their functionality. Furthermore, ethics defines the **limits of actions in relation to others**, emphasising rights and obligations serving the common good over personal gains. This relational aspect highlights the duty of care, particularly towards those in positions of vulnerability.

Equally critical is the **discursive dimension**, which sees ethics as a process of argumentation and communication. It is insufficient to only assert that a choice is 'good'; ethical management requires the capacity to explain the reasoning behind design choices and subject them to open discussion. Finally, ethics is **inherently contextual**. It is not an abstract theoretical exercise but an empirical practice rooted in real-world challenges, emerging where specific values, choices, and consequences intersect.

Within this scope, ethical management serves as an essential mechanism for realising a human-centric vision. It functions to influence regulations, drive design, and offer principles that guide behaviour toward outcomes explicitly aligned with societal values. In doing so, ethics management becomes a cornerstone of Industry 5.0, ensuring technological advancement is fully integrated with societal well-being rather than operating in isolation.[25]

'Applying' ethics: Practices of ethics management

Moving from the definition towards the practice of ethics, the challenge for HRI development within ARISE is based on the operational question: *How can projects anticipate impacts and navigate ethical trade-offs while prioritising human well-being?* In the context of technology

[24]Reijers, W., Young, M. T., & Coeckelbergh, M. (2025). Introduction to the Ethics of Emerging Technologies. In Palgrave philosophy today. https://doi.org/10.1007/978-3-031-85887-1

[25] Callari, T. C., Vecellio Segate, R., Hubbard, E.-M., Daly, A., & Lohse, N. (2024). An Ethical Framework for Human-robot Collaboration for the Future People-centric Manufacturing: a Collaborative Endeavour With European Subject-matter Experts in Ethics. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.4927452

development, we draw and expand from Reijers, Young & Coeckelbergh (2025)[26] to present three distinct mechanisms in which ethics is typically operationalised:

1. **Compliance with Norms and Regulations:** The adherence to binding legal frameworks as well as industry standards related to ethics management—whether mandatory or voluntary—is designed to manage and mitigate the negative implications of technology.
2. **Tools and methods:** This refers to the use of structured frameworks developed by various entities, most notably Ethical Impact Assessments (EIAs). These provide step-by-step methodologies that guide projects through concrete actions, translating ethical concerns into manageable project requirements.

Together, these mechanisms can be categorised as **'technical commitments'**. They represent tangible, measurable actions that can be directly integrated into existing management processes.

3. **Principles and codes of ethics:** This dimension represents a more abstract category of commitment, which this framework coins as **'commitment to excellence'** or commitments to care. These are intangible and public pledges to act with integrity, prioritizing societal well-being. A classic example of such a pledge is the 'Hippocratic oath', used in the medical field. Unlike the technical commitments that focus on risk mitigation, these commitments guide developers toward what they should actively pursue: contributing to the greater societal good.

The role of systemic conditions:
While HRI developers can implement the previous three mechanisms, their effectiveness is intrinsically linked to external systemic conditions. These are elements within the landscape that can facilitate or hinder the integration of ethics management—factors that projects cannot always control but must navigate. Recognising these conditions allows teams to proactively identify the root causes of advancements or implementation gaps, rather than treating systemic barriers as individual failures. Key systemic conditions include:

- **Institutional Support:** The availability of mandates, adequate resources, and access to specialized expertise (e.g., the EU Commission support to Industry 5.0 and consequent Horizon EU funds).
- **Organisational Capacity:** The internal skills, knowledge, and readiness within developer teams to handle ethical complexity.
- **Stakeholder Involvement:** The availability and willingness of key actors such as intended users, funders and work teams to engage. Without active participation, conducting actions to assess real needs and potential risks becomes impossible.

This structure (Fig.7) forms the foundation of the ARISE approach to HRI project's ethics management. However, it is crucial to recognise that the implementation of these elements does not guarantee that new developments will be flawlessly ethically exempt from the risk of

[26] Reijers, W., Young, M. T., & Coeckelbergh, M. (2025). Introduction to the Ethics of Emerging Technologies. In Palgrave philosophy today. https://doi.org/10.1007/978-3-031-85887-1

Co-funded by the European Union

producing harm. There are no 'magic bullets' in ethics management. Instead, it requires an attitude of continuous commitment to best practices and new or improved ways to implement responsible innovation.

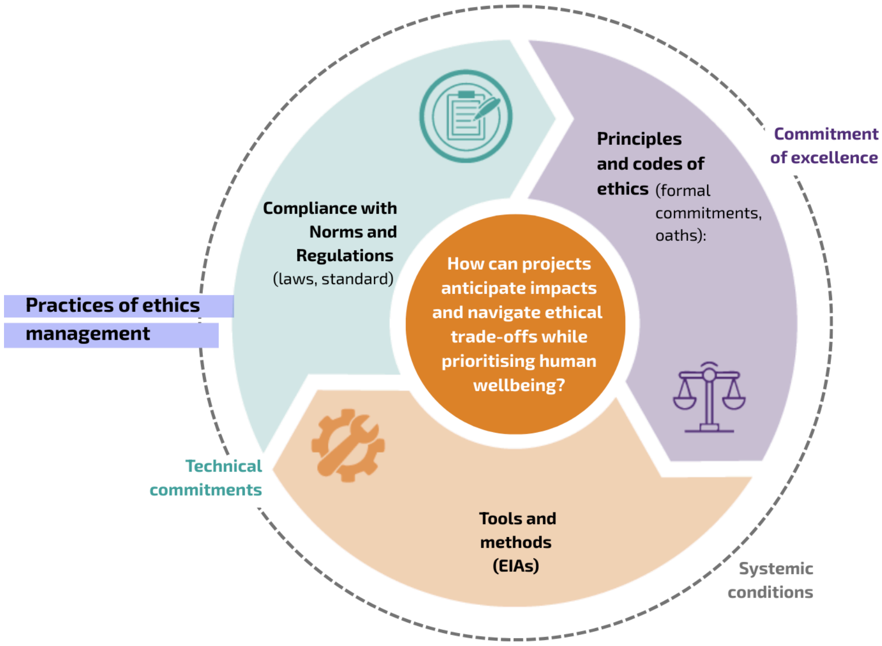


Figure 7. ARISE approach to HRI ethics management.

2.3.4.2 Human-centricity in Industrial HRI

Whereas ethics management is primarily concerned with the identification and mitigation of risks, human-centricity represents the proactive ambition to align technology development with human needs, values, and capabilities. Within this framework, human-centricity serves as one of the core value propositions of Industry 5.0, while ethics management provides the essential structure needed to uphold it. Yet, because human-centricity often remains an abstract ideal, the ARISE framework offers an operational narrative rooted specifically in the context of industrial HRI.

This narrative shifts the focus from merely collaborating with robots for efficiency and precision to a collaboration that enhances both the **quality of human experience** and **societal outcomes**. It posits that HRI technology is only truly human-centric when it creates value across three systemic levels: **the individual operator, the organisation, and the broader ecosystem.**

At the **immediate level of interaction (micro)**, the narrative begins with **dignity at work.** Here, the physical and psychological integrity of the operator is central; safety standards are rigorous, data privacy is guaranteed, and the environment is designed to minimize stress. Any type of data related to people is used only to guide seamless human-machine interactions, never to measure or predict individual performance. Within this secure space, **human agency**

and meaningful control are preserved, ensuring that technology augments human capabilities rather than reducing the worker to a passive observer. Operators retain the autonomy to make decisions, understand system actions, and intervene meaningfully at any point. This interaction is supported by systems designed for **learning and explicability**, where transparent logic makes acquiring new skills engaging, preventing deskilling, and instead supporting the continuous professional development of the individual.

Moving to the **organisational sphere (meso),** the narrative shapes the culture and composition of the workforce itself. Through **inclusivity and accessibility**, technology is used to actively lower barriers to entry, accommodating diverse languages, physical abilities, and experience levels, crucially facilitating the integration of underrepresented groups, such as women, into industrial roles. This inclusive reality is not imposed from the top down but is achieved through **collaboration and co-creation,** using embedded methods to ensure systems reflect the actual needs and values of the diverse stakeholders who work with them daily.

Finally, this human-centric logic expands to the **societal and environmental scale (macro)**. Here, HRI systems contribute to societal resilience by addressing critical demographic transitions. By assuming physically demanding or repetitive tasks, collaborative robots reduce the physiological burden on workers, thereby supporting an aging workforce, extending healthy working lives, and maintaining the viability of essential industries in the face of labor shortages. Simultaneously, this level embraces a **responsibility for environmental well-being**, recognising that human health is inseparable from environmental health. Therefore, industrial HRI must integrate energy efficiency, sustainable lifecycles, and waste reduction as core design criteria, ensuring that technological advancement contributes to a sustainable and resilient future for the broader ecosystem.

Ultimately, this narrative defines a new standard for industrial environments: one where technology serves to empower individuals, build resilient organisations, and protect the wider world. To turn this ambition into operational reality, the next section details the framework's core components and the practical mechanisms for their implementation.

3 The ARISE SSH Framework for human-centric and ethical HRI

This section introduces the main components of the ARISE SSH framework. The framework is structured into four interconnected parts: **Foundations**, **Requirements and Compliance**, **Practice,** and **the SSH Toolbox**. Sections are not intended to be used in a linear sequence. Instead, they can be consulted independently or in combination, depending on specific needs.

Foundations (3.1) introduces the ethical principles, roles, and responsibilities that establish human-centric and ethical HRI. **Requirements and Compliance** (3.2) summarises relevant legislation, guidelines, and standards, and provides practical compliance considerations in the context of ARISE. **Practice: Models of implementation** (3.3) outlines the core implementation models, including the Ethics Impact Assessment (EIA) and the SSH mentoring model. **The SSH Toolbox** (3.4) provides practical resources, methods, and templates to support the application of the SSH insights that are provided in the earlier chapters.

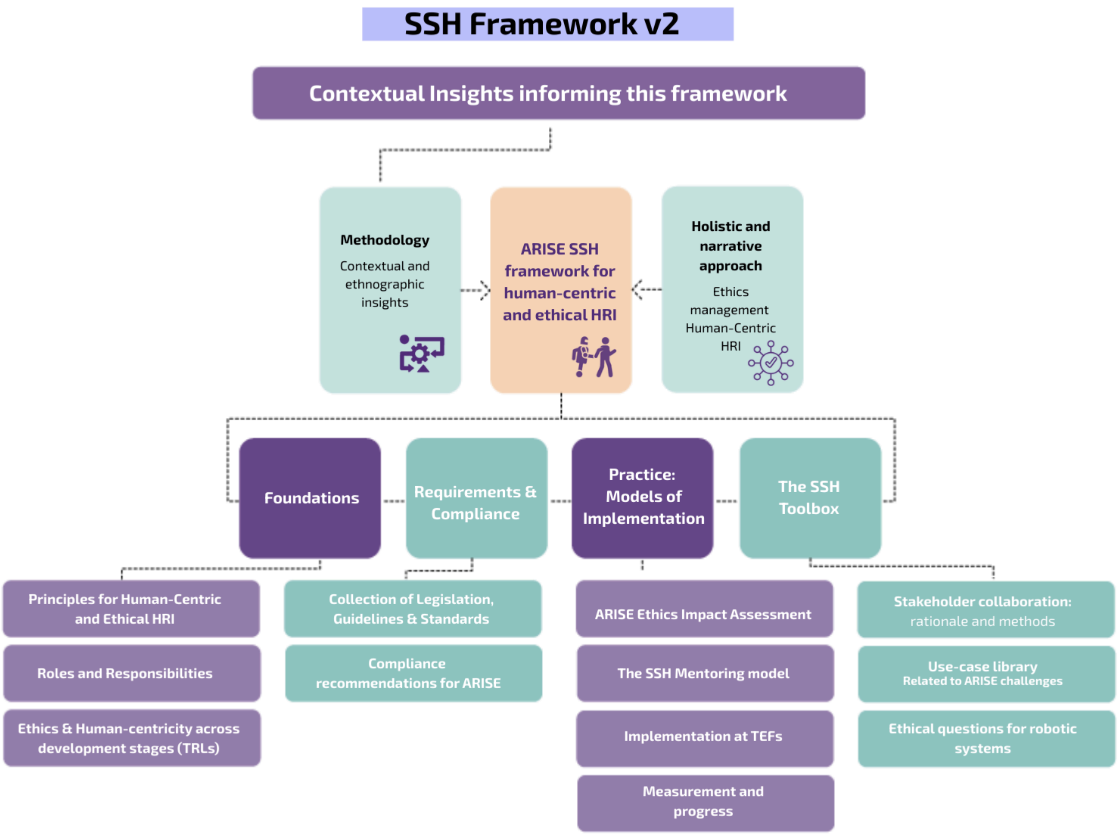


Figure 8. Elements of the ARISE framework for human-centric and ethical HRI.

3.1 Foundations

3.1.1 Principles for Human-Centric and Ethical HRI

Background
Ethical principles in the context of AI have been developed, adapted, and widely adopted across various initiatives. At the time of releasing this framework, the AI industry within the EU and related fields, including robotics, commonly based their ethical frameworks on the work of the High-Level Expert Group on AI (AI HLEG). One such framework is the **Ethics by Design framework**[27]**,** which derived 6 ethical principles for AI systems: **Respect for Human Agency, Privacy and Data Governance, Fairness, Individual, Social and Environmental Well-being, Transparency, and Accountability and Oversight**. These principles serve as the basis for guidelines that can be implemented during various stages of the technology development process in order to ensure that resulting AI systems protect fundamental rights.

In addition to the Ethics by Design principles, other research initiatives have developed ethical guidelines tailored to specific contexts that are closely aligned with the goals of ARISE. Notably, the **Ethical Framework for Human-Robot Collaboration in Manufacturing**[28] outlines a set of principles designed to guide human-robot collaboration in industrial settings, particularly within the vision of Industry 5.0. While these principles greatly overlap with those of the Ethics by Design Framework, they emphasize the unique ethical considerations relevant to robotics in industrial manufacturing settings. In particular, they focus on safeguarding the integrity of human workers by addressing aspects such as autonomy, continuous learning, and skills development.

8 Principles for ARISE
The ARISE ethical principles are derived from the two key frameworks described above: the Ethics by Design guidelines and the Ethical Framework for Human-Robot Collaboration in Manufacturing. Given the centrality of AI systems in HRI development and the importance of the industrial context in many ARISE use cases, these frameworks have been merged and adapted into the eight principles outlined below.

These principles will serve as the foundation for commitments to excellence. They will be piloted through the commitments of excellence within the open-call projects and applied more broadly across the TEFs and other ARISE activities. The eight principles are as follows:

1. **Respect for Human Agency** - Human beings must be respected to make their own decisions and carry out their own actions. Respect for human agency encapsulates three more specific principles, which define fundamental human rights: **autonomy, dignity,** and **freedom**. Respecting **autonomy** means allowing people to think for themselves, decide for themselves what is right and wrong, and choose how they want to live their lives. **Dignity** entails that every human being possesses an intrinsic

[27] See section 2.2 Mapping of Existing Approaches
[28] Callari, T. C., Vecellio Segate, R., Hubbard, E.-M., Daly, A., & Lohse, N. (2024). An Ethical Framework for Human-robot Collaboration for the Future People-centric Manufacturing: a Collaborative Endeavour With European Subject-matter Experts in Ethics. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.4927452

worth which should never be compromised. Hence, people should not be instrumentalized, objectified, or dehumanized, but must be treated with respect at all times, including when using or being subjected to AI-based and HRI systems. Respecting **freedom** requires that people are not constrained in taking those actions which they should be able to pursue as autonomous persons, such as freedom of movement, freedom of speech, and freedom of access to information.

2. **Privacy and Data Governance** - People's right to privacy and data protection should be respected at all times. AI systems must be built in a way that embeds the **principles of data minimisation** and **data protection** by design and by default as prescribed by the EU's General Data Protection Regulation (GDPR). Privacy rights must be safeguarded by data governance models that ensure data accuracy and representativeness; protect personal data and enable humans to actively manage their personal data and the way the system uses it. Data minimisation and data protection should never be leveraged to hide bias or avoid accountability, and these should be addressed without harming privacy rights. Importantly, ethical issues can arise not only when processing personal data but also when the AI system uses non-personal data (e.g., racial bias).

3. **Fairness** - People should be given **equal rights and opportunities** and should not be advantaged or disadvantaged undeservedly. This does not require identical outcomes, e.g., that people must have equal wealth or success in life. However, there should be no discrimination on the basis of the **fundamental aspects of one's own identity,** which are inalienable and cannot be taken away. Procedural fairness requires that the procedure was not designed in a way that disadvantages specific individuals or groups. Substantive fairness entails that the AI does not foster discrimination patterns that unduly burden individuals and/or groups for their specific vulnerability.

4. **Individual, Social and Environmental Well-Being** - Systems should contribute to, and not harm, individual, social, and environmental well-being. **Individual well-being** means people can live fulfilling lives where they are able to pursue their own needs and desires in mutual respect. **Social well-being** refers to the flourishing of societies whose basic institutions, such as healthcare and politics, function well and where sources of social conflict are minimized. **Environmental well-being** refers to the well-functioning of ecosystems, sustainability, and the minimization of environmental degradation. Possible research participants, end-users, affected individuals and communities, and relevant stakeholders should be identified at the very early stage to allow for a realistic assessment of how the AI and robotic systems could enhance or harm their well-being. This recognition should contemplate current uses and future deployments of technology at higher maturity levels (e.g., higher TRL). Documented choices should be made during development to support well-being and avoid harm.

5. **Transparency** - The purpose, inputs, and operations of AI and robotics systems should be knowable and understandable to their stakeholders. Transparency impacts all elements relevant to an AI system: the data, the system and the processes by which it

is designed and operated. S**takeholders must be able to understand** the main concepts behind these systems (how, and for what purpose these systems function and come to their decisions). IP rights, confidentiality, or trade secrets claims cannot prevent transparency as long as they can be preserved appropriately, for instance by way of selective transparency (e.g., confidentially to trustworthy third parties), technology, or confidentiality commitments.

6. **Accountability and Oversight** - Humans should be able to understand, supervise, and control the design and operation of AI and robotics systems, and the actors involved in their development or operation should take responsibility for the way that these applications function and for the resulting consequences. Accountability presupposes certain levels of transparency as well as oversight. To be held to account, developers or operators of AI systems must be able to explain how and why a system exhibits particular characteristics or results in certain outcomes. **Accountability depends on oversight**: to be able to take responsibility and act upon it, developers and operators of AI systems must understand and control the functioning and outcomes of the system. Hence, to ensure accountability, developers must be able to explain how and why a system exhibits particular characteristics.

7. **Worker Dignity and Equity** - In line with the core principles of Industry 5.0, dignity and equity are at the core of human-centric, inclusive workplaces. Human worker experience should be respected, and the role of human workers in HRI acknowledged and compensated accordingly. It's important to recognize that **workers are a heterogeneous group** in respect to their preferred types of tasks. Some find fulfillment in performing routine tasks. As robots often replace this type of work, organizations should provide their workers with **opportunities that align with their unique interests and talents**, rather than channeling them into pre-determined paths. Moreover, **robots should not be used to set the standards for human performance and behaviour**, as this poses a risk of *'instrumentalisation'* of humans.

8. **Human Resilience Through Continuous Learning and Support** - Human skills and their development should complement, not compete with advances in robot development. Adequate possibilities for **learning new skills** and **supporting worker psychological safety** in HRI should be fostered. This involves acknowledging the critical role of management. Leadership should be supported in practices that prioritise well-being and psychological safety in human-robot coordination. Workers should be taught the technical skills required to troubleshoot robots effectively to ensure the security and primacy of their roles in HRI environments . Providing opportunities for workers to bridge any skill gaps to engage meaningfully with robotic technology can help foster a sense of agency and empowerment in HRI settings.

3.1.2 Roles and Responsibilities.

To support the usability of this framework, this section clarifies which **internal roles** constitute its intended audience. At this stage, the framework's application has focused on two specific contexts within the ARISE project: the open-call mentoring programme (beneficiary projects) and the Test and Experimentation Facilities (TEFs). However, the roles defined here are designed to be transferable, anticipating broader adoption by other actors in the HRI ecosystem in later stages.

While applicable to robotics development across diverse organizational contexts, this framework is particularly tailored to the reality of small to medium-sized enterprises (SMEs) and TEFs. The following table lists out the internal roles that are commonly found at these HRI projects, illustrating how human-centric and ethical considerations translate to each.

It is relevant to note that these roles represent functions rather than rigid job titles. Depending on the project size and structure, these actors will be found in different compositions. For instance, in a TEF, these roles may be found in different departments while in an SME or pilot experiment, a single individual may hold multiple responsibilities. These examples are intended as a starting point—illustrative rather than exhaustive—to establish a shared understanding of the responsibilities incorporated into the application of the framework.

Table 3. Roles and responsibilities

Category	Examples of role titles	Examples of ethics and human-centric responsibilities
Technical Development	Software developer Robotics developer Electrical engineer Systems engineer Mechanical engineer Computer scientist Automation engineer Researcher Full stack developer Firmware engineer QA/testing engineer	• Data is stored and processed in accordance with GDPR principles • Physical safety of users interacting with robot • Comprehensive testing for possible safety failure points • Software and mechanical systems are designed in a way that allows for human oversight and override
Fabrication and Manufacturing	Machinist Technician	• Materials for manufacturing are sourced and used sustainably
Design	Industrial designer User researcher UI/UX designer	• Design user testing and feedback sessions to not only evaluate usability, but also end-user well-being • Stakeholder engagement plans; coordination of inclusive co-creation sessions
Project Management	Team lead Project manager Coordinator	• Development process include checkpoints for assessing emerging ethical concerns • User engagement sessions are appropriately timed to align with key decision making timelines

Business & Marketing	Sales engineer Marketing Business director	• Seeking out funding and market opportunities that align with human centric goals • Communications of results aligned with Industry 5.0 objectives
Administrative roles	HR Legal Recruitment Proposal management	• Strategically allocate and advocate for resources to ensure organizational capacity for incorporation of human-centric HRI development
Upper level decision-makers	CTO CEO COO Board member Founder or co-founder	• Setting a culture of prioritizing and valuing SSH perspectives of HRI work • Making human-centric HRI work a company priority • Ensuring that organizational goals align with human-centric values
SSH support	*HRI Researcher* *Social science and humanities researchers, experts and advisors*	• *Raising awareness of emerging ethical concerns that haven't been recognized* • *Providing advice and guidance to different roles in the organization when needed* • *Support capacity-building* • *Designing, adapting or deploying tools and processes for the organization to incorporate human-centric and ethical perspectives*

It is important to note that roles within the SSH Support category, listed at the end of the table, are rarely found as distinct, in-house positions at most SMEs due to resource constraints and the novelty of the requirements. However, the competence they represent—bridging technical development with societal and ethical considerations—provides immense value, enabling organizations to strategically embed human-centricity into their workflow. Where internal recruitment is not feasible, this critical capability can be effectively fulfilled through partnerships with external experts, a model which is detailed in section 3.4.1.1 (Stakeholder Mapping).

3.1.3 Ethics and human-centricity Across Technology Readiness Level

As elaborated upon in 2.3.3.3 (Context specific needs and challenges), human-centric and ethical concerns may vary significantly depending on the maturity of the technology. Recognizing the pitfalls of prescribing a one-size-fits-all framework for incorporating human-centricity in HRI development, this section introduces levels 3 - 9 of the **Technology Readiness Level (TRL)** scale to ground the SSH framework in specific TRL contexts. Levels 1 and 2 are not considered in this framework as they don't typically apply to the context of HRI development at SMEs. TRL was first developed by NASA in the 1970s, and has since been adopted and incorporated into the EU ecosystem starting with the 2020 EU Horizon Europe program[29]. At the time of this writing, TRL remains a widely recognized and familiar framework for technology development in the European context.

[29] Horizon Europe 2020 Work Programme 2014 - 2015 General Anexxes
https://ec.europa.eu/research/participants/data/ref/h2020/wp/2014_2015/annexes/h2020-wp1415-annex-g-trl_en.pdf

Co-funded by
the European Union

Note that projects rarely progress in a linear, sequential path along the TRL scale. Instead, HRI development is iterative- projects can move between different stages of TRL in iterative development cycles. This framing fails to precisely represent the full life-cycle of HRI development, but it can serve as a rough estimate of the current stage of a project at a specific point in time.

Table 4. Technology Readiness Level [30]

Technology Readiness Level	Description	Example in the context of HRI development
3 - Experimental proof of concept	Initial experiments validate the concept. Active research and development begins and a proof-of-concept model is constructed. Feasibility of separate technology components are being validated through analytical and laboratory studies. Reaching to this point, one can conclude that **the new technology is feasible from a scientific point of view.**	A novel computer vision (CV) model is developed and tested. The vision model is able to identify and interpret its surroundings, as validated with simple command line interactions.
4 - Technology validated in lab	**Components are integrated and tested in lab conditions as a prototype.** Basic technological components are integrated "ad-hoc" to establish that they will work together in a laboratory environment.	The output from the CV model feeds into and drives simple motions of the robotic arm. This CV + arm integration is tested in a lab setting.
5 - Technology validated in relevant environment	The integrated basic technological components are performing for the intended applications in a simulated and controlled environment. Configurations are being developed but can undergo fundamental changes. Reaching to this point, one can conclude that **the new technology is feasible from a technological point of view.**	CV model + arm integration is tested with varying environment conditions, validating whether CV interpretation of environment leads to expected motion.
6 - Technology demonstrated in relevant environment	Prototype demonstrated in a relevant but not fully operational setting. **A model or prototype representing a near desired configuration is being developed at a pilot scale**, generally smaller than full scale. The model or prototype is demonstrated in a real environment, so to confirm the engineering is feasible.	Robot arm degrees of motion, movement velocities, force of grip, and other technical parameters are being tested and validated. Robotic solution is able to identify and pick up the tools it's expected to retrieve in a simulated manufacturing environment.
7 - System prototype demonstration in operational environment	Near-final system tested in actual operational conditions. **The working model or the prototype is demonstrated in an operational environment, typically under industrial conditions and timings (i.e. field tests)**. At this stage, the final design is very close to completion.	The robot is able to fully operate in controlled field tests in an actual manufacturing environment.
8 - System complete and qualified	The final system is completed and meets all specifications. **TRL 8 commonly represents the end of technology development.** At this stage, operations are well understood, operational procedures are being developed, and final adjustments are being made. Technology is ready for implementation into an already existing technology or technology system.	The robot can be integrated into the manufacturing processes and environments it's intended for. The solution has passed all regulatory requirements and is ready for the market.

[30] Technology Readiness Level (TRL) | Aalto University. (2025, July 28). Aalto.fi. https://www.aalto.fi/en/services/technology-readiness-level-trl

9 - Actual system proven in operational environment	**Technology is in use and proven in real-world operations.** Actual application of the technology in its final form is being conducted under a full range of operational conditions. Sometimes referred to as "system operations", this stage is where technology is further refined and adopted.	The robot has been proven, validated, and is in use in the markets it's designed for

As a next step, the framework will further elaborate on how to incorporate ethical and human-centric requirements as technical maturity advances. A critical focus for future iterations, particularly relevant for the TEFs, is navigating requirements at lower TRLs, where ethical implications often remain abstract because the context of use is not yet fully defined. To address this, the framework advocates for embedding ethics early and iteratively. By establishing transferable documentation (adaptable across different teams, organisations, and contexts), staged checkpoints, and guiding parameters, projects can ensure that societal considerations mature and become more tangible alongside the technology. The future versions of this framework will look into providing specific strategies for these 'abstract' phases.

3.2 Requirements and Compliance

3.2.1 Collection of Relevant Legislations, Guidelines and Standards

As part of the initial ARISE SSH framework, we provide a collection of European regulations and industry standards relevant to SSH perspectives of the ARISE project. While the current listing does not include all possible perspectives, it aims to provide a comprehensive overview of the current approaches relevant to our framework and its use in the project.

Tables 5, 6, and 7 present a categorization of multiple industry standards and EU-level legislation related to HRI. The chosen standards and legislation are EU-centric but also include international standards in cases where they are more applicable for this framework, for example, ISO-standards. The tables are categorized into three sections: **EU level legislation, ethical guidelines for AI,** and **industry standards related to ethics and human centricity.** The reasoning used for these categories is to group items by importance and level of detail. Items in the **EU legislation** table are especially critical for all HRI workers to pay attention to, as there are legal obligations and penalties tied to most items. Items in the **Ethical Guidelines for AI** table provide a good starting point and overview for applying existing ethical standards to technologies that utilize AI. Items in the **Industry standards** table provide extensive, detailed documents that examine ethics, safety, and human-centricity in a particular context of technology development.

Note that EU-level legislation includes both binding and non-binding legislation, while the remaining items listed in these tables are all non-binding recommendations. In addition to the existing legislation, we monitor the field for upcoming and relevant legislation regarding HRI and will be updating the forthcoming iterations of the framework to include these.

Table 5. Binding and non-binding EU legislation

Binding EU Legislation
GDPR General Data Protection Regulation
EU AI Act - AI Liability Directive (proposed but not enacted) - General Purpose AII Code of Practice (first draft published in November 2024)
Machinery Directive
Non-binding EU Legislation
EU-OSHA Guidelines European Agency for Safety and Health at Work Guidelines

Table 6. Ethical guidelines for AI

Ethical Guidelines for AI
AI HLEG Deliverables - Ethics Guidelines for Trustworthy AI - ALTAI Assessment List for Trustworthy Artificial Intelligence (ALTAI) for self-assessment
Ethics By Design and Ethics of Use Approaches for AI
OECD AI Principles

Table 7. Industry standards

Industry standards for ethics and human-centricity
IEEE 7000-2021 IEEE Standard Model Process for Addressing Ethical Concerns during System Design
IEEE 7014-2024 IEEE Standard for Ethical Considerations in Emulated Empathy in Autonomous and Intelligent Systems
ISO 26000:2010 Guidance on social responsibility
ISO 27501:2019 The human-centred organization — Guidance for managers
ISO 9241-210:2019 Ergonomics of human-system interaction Part 210: Human-centred design for interactive systems
Industry Standards for AI and Autonomous Systems
ISO/IEC TR 24028:2020 Information technology — Artificial intelligence — Overview of trustworthiness in artificial intelligence
ISO/IEC 24027:2021 Information technology — Artificial intelligence (AI) — Bias in AI systems and AI aided decision making
ISO/IEC TR 24368:2021 Information technology — Artificial intelligence — Overview of ethical and societal concerns

ISO/IEC AWI TR 5469 Artificial intelligence — Functional safety and AI systems
ISO/IEC AWI TS 6254 Information technology — Artificial intelligence — Objectives and approaches for explainability and interpretability of ML models and AI systems
IEEE 7001-2021 IEEE Standard for Transparency of Autonomous Systems
IEEE 7010-2020 IEEE Recommended Practice for Assessing the Impact of Autonomous and Intelligent Systems on Human Well-Being
Industry Standards for Safety in Robotics
ISO 10218 - 1 & 2 Robots and robotic devices — Safety requirements for industrial robots Part 1: Robots & Part 2: Robot systems and integration
ISO 13482:2014 Robots and robotic devices — Safety requirements for personal care robots
ISO/TS 15066:2016 Robots and robotic devices — Collaborative robots
ISO/TR 20218-1:2018 Robotics — Safety design for industrial robot systems Part 1: End-effectors
ISO 13849-1:2023 and ISO 13849-2:2012 Safety of machinery — Safety-related parts of control systems Part 1: General principles for design & Part 2: Validation
ISO 12100:2010 Safety of machinery — General principles for design — Risk assessment and risk reduction
IEC 61508-1:2010 Functional safety of electrical/electronic/programmable electronic safety-related systems - Part 1: General requirements (see Functional Safety and IEC 61508)

3.2.2 Compliance Recommendations for ARISE

This section outlines one possible structured approach for navigating the landscape of collections, standards, and legal requirements for HRI projects. The following approach serves as a recommended starting point for projects in the ARISE HRI context to ensure sufficient compliance with legal and ethical standards. Specifically, this approach outlines steps for ensuring compliance with the EU level **AI Act**, **GDPR**, and AI-HLEG's recommendations for trustworthy AI (**ALTAI**). A full template for ensuring compliance with key standards has been incorporated to the 'Ethics and Human-Centricity Action Plan' explained in section 3.3.2.3 . A template can be found in the Appendix 2

After completing these steps, a project, experiment or TEF should be able to recognise:

- Whether they are legally compliant with the new **AI Act** and **GDPR**
- The **types of personal data** that are being collected and the procedures needed to ensure **safe and ethical data processing**
- Recommendations for how to **mitigate risks** related to any AI systems in use
- Which **requirements for trustworthy AI**, as recognised in the EU, is project the in compliance with

3.2.2.1 AI Act

While the AI Act officially entered into force on August 1, 2024, at the time of this writing, not all parts of the act have entered into force. Regardless, it is recommended for HRI developers to adopt a proactive approach in ensuring legal compliance with all parts of the Act:

1. Identify the role that your organization plays in the development of the AI system, relying on the definitions provided in Article 3 of the AI Act (provider, deployer, distributor, authorized representative, importer, distributor, operator).
    - If you are a **deployer** or **provider** of AI systems, refer to articles 4 and 50 to understand the literacy and transparency obligations, respectively, you are subject to.

2. Identify all parts of the technology solution and development lifecycle that utilize AI. For each instance of AI usage, categorize the level of risk according to articles 5, 6, and 7 of the AI Act (unacceptable, high, or neither).
    - Any instances of **unacceptable risk** are prohibited under the AI Act, unless they fall under the exemptions listed in Article 2
    - For all instances of **high risk,** refer to articles 8-17 for requirements that must be fulfilled under the AI Act.
    - AI systems that are not prohibited or classified as high risk are not regulated by the AI Act.

As an additional recommended step, use the information above to fill out the EU AI Act Compliance form. The online form should generate a list of recommendations for how your team can address the legal obligations you're subject to.

3.2.2.2 ALTAI Self-Assessment

The Assessment List for Trustworthy Artificial Intelligence (ALTAI) self-assessment is a tool developed by the High-level expert group on AI, presented in July 2020. ALTAI builds on the ethics guidelines for trustworthy AI to outline 7 key requirements for trustworthy AI.

To complete the full self-assessment, your team can use either the PDF file provided on the ALTAI website or the web-based tool. Working through the web-based self-assessment is strongly recommended as it generates a set of recommendations for ensuring compliance with each of the 7 requirements. However, note that at the time of this framework release, the ALTAI web tool has presented some technical difficulties, and the website may not be consistently accessible. The text-based ALTAI provides the same content and questions as the webtool, but won't result in concrete recommended next steps.

Alternatively, the questions listed in 3.4.3 (Ethical questions for robotic systems) build off of many of the same principles as the ALTAI assessment, but with the two additional ARISE principles (Worker dignity and equity, and Human resilience through continuous learnings). Completing either the ALTAI self-assessment or the 3.4.3 questions for your team's robotic and AI systems will likely lead to similar reflections and insights.

3.2.2.3 GDPR

The GDPR was made applicable to all member states in May of 2018. To ensure minimal compliance with the GDPR, work through the following questions with your team:

1. List out all of the features of your robotics solution that collect or process data about a data subject (an identified or identifiable person).
    a. For each of these features, determine whether the following special categories of data are being collected or processed: racial or ethnic origin, sexual orientation, political opinions, religious or philosophical beliefs, trade-union membership, genetic, biometric or health data, personal data related to criminal offences.
    b. For each of the features, also indicate the types of personal data that's being collected or processed. (name/surname, home address, IP address, cookie ID, advertising identifier of phone, medical data that could be a symbol that uniquely identifies a person)

2. If your solution requires the collection of special categories of data, elaborate on what exception under the GDPR (Article 9) allows you to collect this data.

3. If your solution requires the collection of personal data, please elaborate on how you will ensure the rights of the data subject. For instance, you can provide a link to a template of an informed consent form, or explain what other basis you have under the GDPR to process personal data.

4. Read through the "Who monitors how personal data is processed within a company" section of this webpage, and explain whether a Data Protection Officer (DPO) is or is not necessary.

3.2.2.4 Recognizing Additional Relevant Regulations and Standards

While checking compliance with the AI Act, ALTAI self-assessment, and GDPR are important first steps, going beyond minimal compliance requires projects to examine their specific contexts and needs to determine appropriate next steps. Projects should examine **country-specific regulations** for the locations where they are developing or deploying their solution. Projects related to **healthcare** need to be aware of the European Health Data Space (EHDS), ISO standards regarding clinical trials and usability testing, MDR regulation, and other regulatory and ethical guidelines for working in healthcare contexts. Projects developing solutions for **industrial and manufacturing contexts** should pay particular attention to standards such as ISO 10218 for general robot safety, ISO 15066 for collaborative robotics, ISO 13849 standards for functional safety and risk assessment, and the European Agency for Safety and Health at Work Guidelines (EU-OSHA).

Given the complexities and varied contexts of robotics development, it is impractical to develop a one-size-fits-all recommendation for how robotics developers should ensure regulatory and ethical compliance. Ultimately, it is the responsibility of each team to

Co-funded by
the European Union

understand the regulatory landscape they work in and recognize how to use relevant standards appropriately. As a reminder, the tables in section 3.2.1 can serve as a starting point for finding additional relevant standards, guidelines, and regulations.

3.3 Practice: models of implementation

This section presents different ways the ARISE SSH framework can be implemented for the diverse purposes of embedding SSH considerations into the workflows of HRI. Recognising that human-centricity is not a one-size-fits-all requirement, the framework offers distinct operational models tailored to specific contexts found within the context of ARISE.

The following subsections detail these practical applications:

- **ARISE Ethics Impact Assessment (EIA)**: The method for identifying risks and benefits, Section 3.3.1.
- **SSH Mentoring Model:** A structured support pathway designed specifically for SMEs and open-call beneficiaries to provide support and build internal capacity, Section 3.3.2.
- **Implementation at TEFs:** A tailored approach for Testing and Experimentation Facilities, Section 3.3.3.
- **The SSH Toolkit:** A consolidated set of resources to support independent implementation ,Section 3.3.4.

3.3.1 ARISE Ethics Impact Assessment (EIA)

The **ARISE Ethics Impact Assessment (EIA)** constitutes a key method for identifying and evaluating risks and benefits from an ethics perspective. It provides a structured process for HRI developers to assess technological components, implementation strategies, and deployments, ensuring that practical operations align with high-level ethical standards.

EIA objectives
The primary goal of the EIA is to provide the structure to identify and follow up on both **benefits (desirable impacts)** and **concerns (undesirable impacts)**. It serves as a consolidation point, integrating insights from diverse sources. In ARISE, these sources primarily include the external ethical reviews by the ARISE Ethics Committee, the ALTAI assessment (specifically for AI components), and internal project analysis. Beyond identification, the EIA also aims to capture appropriate **prevention strategies, mitigation actions, and assurance mechanisms.** Crucially, it aims to recognise impacts across systemic levels, ranging from the immediate context of use to broader societal implications. Tied into the ethical considerations recognised across systemic levels, the EIA also encourages projects to reflect on potential environmental impacts, validating that sustainability and resource use are assessed where relevant.

The ARISE EIA organises these insights using the framework's common notions: the **micro, meso, and macro levels.** This structure ensures that impacts are recognised not just at the level of the immediate robot-user interaction, but also regarding organisational culture and broader societal effects. The level of granularity and detail will depend on the project's own context and TRL. For instance, lower TRLs may focus on abstract societal risks, while higher TRLs will require detailed planning and dedicated resources to assess immediate contexts of use.

Importantly, the documentation produced by the EIA should be treated as a **transparency mechanism**, readily available for review by internal actors within projects, stakeholders, and external auditors (such as the Ethics Committee).

Relation to the Ethics and human-centricity action plan: It is important to distinguish the EIA process from the reporting document. The **EIA** is the analytical methodology used to identify impacts and risks. The **Ethics and Human-Centricity Action Plan**, introduced later in 3.3.2.3, is the living document where these findings are recorded. In practice, the tables generated by the EIA (Appendix 1) constitute the content of a section of the Action Plan. Therefore, conducting the EIA is the primary mechanism for populating and updating the project's Action Plan.

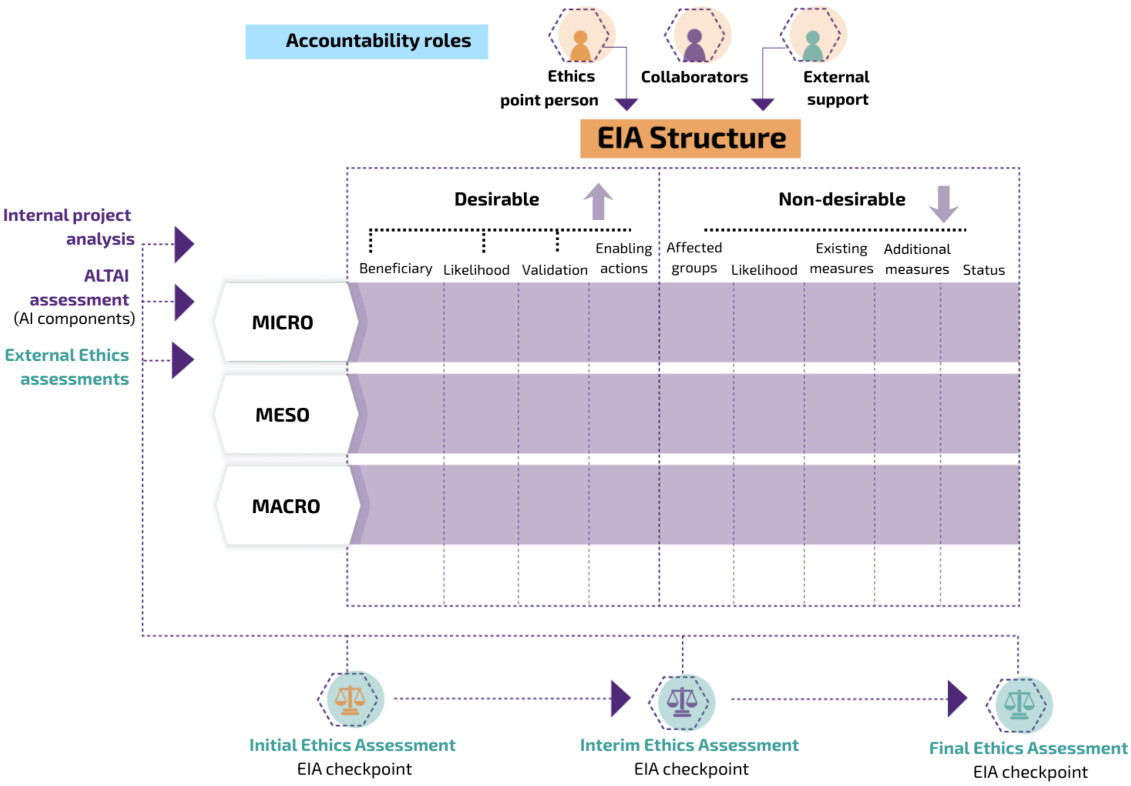


Figure 9: ARISE EIA structure.

3.3.1.1 Roles and Accountabilities

Completing the EIA requires diverse inputs from legal, technical, managerial domains and other domains. While one individual may not hold all the answers, it is essential that clear accountability is defined.

Establishing this accountability presents a distinct challenge, especially for smaller actors who often lack specialized internal roles regarding ethics or societal impact. Furthermore, because Industry 5.0 requirements are not mainstream in standard technical training, teams may feel under-equipped to complete the EIA with confidence. Therefore, the EIA is designed as a low-threshold exercise and capacity-building process to help teams acquire these skills iteratively. It is also advisable to complement the completion of the assessment with external expertise if needed. For instance, within ARISE, the EIA is completed in coordination with an external assessment of ethics experts.

Therefore, the first step to complete the EIA is to clearly **define roles**:

1. **Ethics point-person:** Each project must designate a specific lead responsible for the completion of the EIA. Their role is not necessarily to write every answer or possess all critical expertise, but to facilitate the necessary dialogue and schedule follow ups as needed. They act as the bridge, ensuring the right questions are asked of the right people, at opportune moments of the project's development.
2. **Collaborators:** Specific actors within the organisation can be designated to collect input for the EIA. Together, the collaborators give representation to the most relevant roles or departments within a specific project being assessed. They are expected to contribute with specific data and be available for dialogues when needed, ensuring that claims made in the assessment can be validated.
3. **External support:** The project clarifies whether there is external support or any type of mechanism supporting the identification of benefits and concerns. In ARISE this support corresponds to the SSH experts as project partners and mentors, as well as the external Ethics Committee.

Finally, regarding responsibilities, it is important to clarify that while any of these actors facilitates the process, they should not bear the ultimate accountability for the EIA's integrity or for critical safety decisions (e.g., halting on experimentation if mitigation strategies are insufficient), unless this is clearly stated. Instead, it is the responsibility of each project or organisation to explicitly define this chain of command, with critical decision-making power typically residing in higher management positions.

3.3.1.2 Assessment structure

The EIA makes a distinction between **desirable impacts** (which should be maximised and validated) and **undesirable impacts** (which must be mitigated or prevented). Projects are expected to map these across the three distinct systemic levels: **Micro**, **Meso** and **Macro**.

A. **Assessing desirable impacts:** Recognising the desirable impacts forms part of the value proposition of HRI projects and greatly relates to the human-centric values. They can signal directionalities for improvements beyond the avoidance of harm. Importantly, desirable impacts refer specifically to **human, societal** and **environmental benefits** and not just commercial or financial KPIs. Examples of desirable impacts are: *improved worker safety, empowerment and upskilling strategies, support for an inclusive workforce* and *energy and carbon emission reduction*. To identify desirable impacts, projects must recognise:

    ○ **Type of benefit:** *description of how the improvement is achieved*
    ○ **Primary beneficiary:** *who benefits?*
    ○ **Likelihood:** *how certain is that this benefit will materialise? Is this a hypothesis or a fact?*
    ○ **Validation:** *what information or testing is missing to confirm this benefit? (e.g., "Need to validate through co-creation with industrial workers' unions")*
    ○ **Enabling actions:** *what steps (actions, strategies) are required to ensure this positive outcome occurs?*

B. **Assessing undesirable impact**: This step is key to the project's risk management strategy, designed to identify and avoid actions that could lead to harm if left unaddressed. Documenting potential negative impacts, along with corresponding avoidance measures, serves to validate decision-making and provides protection against unforeseen liabilities. Specifically, projects must recognise:

    ○ **Type of risk:** *description of the undesirable impact and how it happens*
    ○ **Affected group(s):** *Who is impacted?*
    ○ **Likelihood:** *without intervention, what is the probability that the impact materialises? Is this a hypothesis or a fact?*
    ○ **Existing measures:** *what measures are currently in place?*
    ○ **Additional mitigation:** *what additional actions are planned in order to avoid or mitigate?*
    ○ **Post-mitigation status:** *after mitigation, is it resolved, is there a need to follow up? how frequently should follow ups occur? Is the level of remaining risk acceptable?*

The template tables with all sections of the EIA are available in the Appendix 1

3.3.1.3 Assessment Process

The process of filling out the EIA template will depend greatly on the specific intention, context, and maturity of the project. Therefore, in general terms, the framework does not offer a fixed set of instructions for every possible scenario. Instead, the EIA should be understood not as a one-time event, but as an evolving process of reflection and documentation, to be completed in accordance with the context.

EIA process within ARISE

In ARISE, for both open call projects and TEFs, the process is formalised around **three key milestones** involving validation by external expertise[31], and interacting with other models of implementation of the SSH framework.

The ARISE Ethics Committee: The external Ethics Committee consists of three independent experts. Their mandate is to review both TEFs and open call projects at the start, middle, and end, identifying concerns, recommending mitigations, and highlighting strengths[32].

The assessment cycle proceeds as follows:

1. **Initial Ethics Assessment (scoping ethical impact):** Its primary purpose is to identify risks early, before experiments begin, or at early phases. Based on available documentation (the open call proposal, or specially gathered material), the committee evaluates compliance with Horizon Europe standards, anticipates ethical concerns, and offers recommendations. **Outcome:** A consolidated report that serves as an initial foundation for the EIA and the project's ethical management.

2. **Interim Ethics Assessment (following up on progress and evolving challenges):** Set as a milestone, this review focuses on progress and adaptation. Based primarily on the first version of the Ethics and Human-Centricity Action Plan, it evaluates whether initial recommendations have been addressed and identifies new risks that may have emerged as the technology matured. **Outcome:** An interim report guiding the team to update their EIA and course-correction if necessary.

3. **Final Ethics Assessment (evaluating the final outcomes):** Assess the extent to which ethical requirements were met and concerns resolved. It relies on the final Ethics and Human-Centricity Action Plan and the Roadmap for Future Use. It critically assesses whether there are any issues that need to be critically considered if the technology solution continues its development or deployment. **Outcome:** A final report as validation of the project's ethical management trajectory.

Assessments serve as key checkpoints for refining the EIA. These inputs, combined with findings from other evaluations like ALTAI, should drive the continuous evolution of the document. Teams must therefore plan their own internal processes for monitoring and reporting updates. Ultimately, the ownership of the EIA process and how it is tailored to the project context should be managed by the project team.

[31] The process of external ethical assessments is detailed in ARISE D3.4 Ethics Summary Reports v1

[32] Detailed information regarding the selection and relationships to the Ethics Committee can be found in D3.2 ARISE's Ethics Summary Reports

Co-funded by the European Union

3.3.2 The SSH Mentoring Model

The SSH Mentoring Model defines how the ARISE SSH Framework is implemented in the context of the beneficiary projects of the ARISE open-call. This implementation is delivered through a specific support channel as part of the 12-month ARISE mentoring programme: **Support Service III: Human-Centricity by Design using the ARISE SSH Framework**, designed and conducted by the ARISE SSH partner within the project, Demos Helsinki.

The primary objective of this service is to ensure beneficiaries adopt a human-centric and ethical approach while meeting necessary regulatory and societal requirements. Delivered as a structured pathway, the service focuses on building internal capacity and facilitating the integration of expertise not only within the current project but as transferable knowledge for future initiatives. The model unfolds across distinct stages: scoping societal impacts, establishing ethics management practices, deepening the collective understanding of human-centricity, and preparing for alignment at higher Technology Readiness Levels (TRLs).
The SSH Mentoring Model is operationalised through two complementary components: **support activities** (the enablers) and the **alignment pathway** (the structural requirements). The relationship between these components and the overall structure and flow of the mentoring model is visualized in Figure 10 below. The following sub sections provide more detailed explanations for each component of the mentoring model.

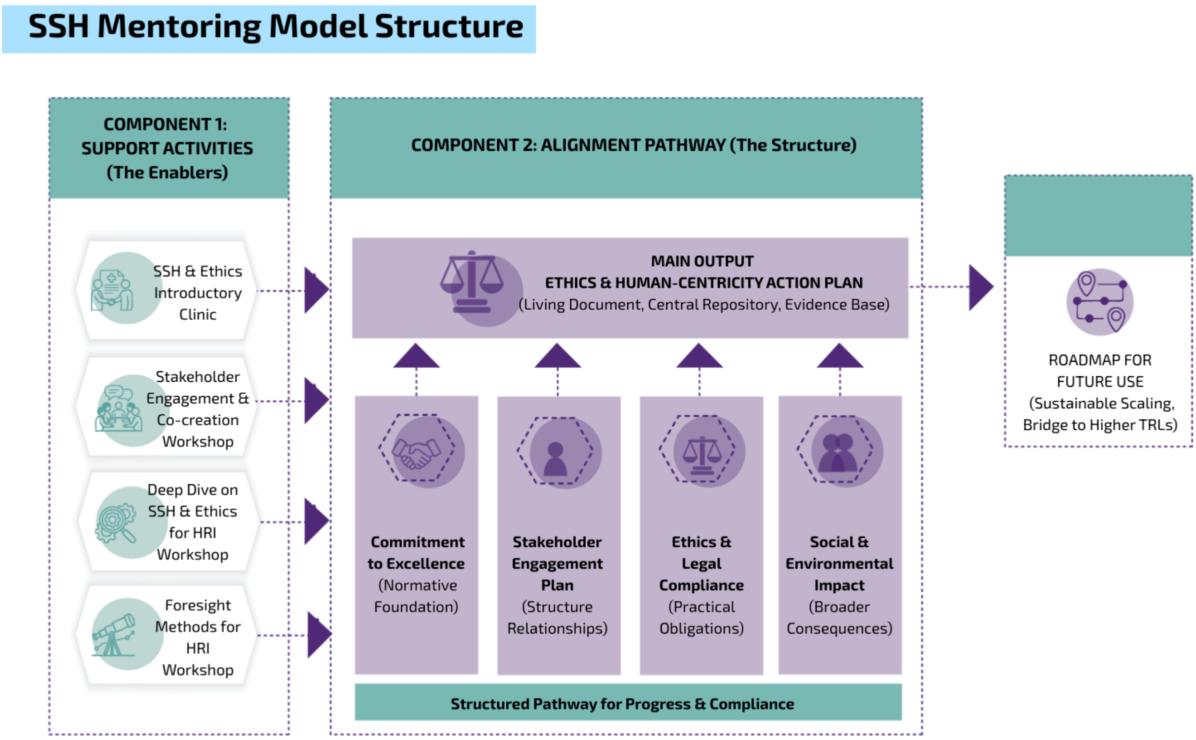


Figure 10. ARISE Mentoring Model structure

3.3.2.1 Component 1: Support Activities (the enablers)

The mentoring guidance is delivered through a blend of systematic assessments, participatory workshops, and individual support when needed. These sessions are designed specifically to equip beneficiary projects with the knowledge and tools necessary to align with the mentoring objectives and to complete and refine their Ethics and Human-Centricity Action Plan, described in the following section.

Importantly, the aforementioned **Ethical Impact Assessment (EIA),** including the external ethical assessments, provides a structured approach for identifying ethical issues.

Several **participatory activities** are designed to deepen collective understanding of ethics in HRI. The activities include:

- **SSH & ethics introductory clinic:** Introduce the rationale and structure of the SSH framework and its associated web tool, and to the process and methods that will be available as support throughout the project.
- **Stakeholder engagement & co-creation workshop:** Promotes good practices for identifying and engaging with key stakeholders, and supports the recognition of their needs and preferences in context and throughout the project lifecycle. It introduces practical tools and methods to facilitate meaningful engagement and co-creation.
- **Deep dive on SSH & Ethics for HRI workshop:** Explores key components of ethics management in depth, including foundational concepts, guiding principles, and regulatory frameworks, and explores the systemic conditions needed to support ethical practices throughout the TRLs.
- **Foresight methods for HRI workshop:** Introduces the foresight rationale and specific methods and tools to anticipate desired and undesired outcomes at different time scales, focusing on medium and long-term. Additionally, it aims to align industrial HRI development with desirable and shared future visions.

The mentoring activities are designed to support project teams in understanding and engage meaningfully with the themes and questions in the Ethics and Human-Centricity Action Plan and the entire Ethics Impact Assessment process. Through structured assessments, participatory workshops, and targeted individual support, the mentoring process establishes a shared conceptual foundation around human-centricity, stakeholder engagement, ethical risk, and future impacts in HRI. This enables participants to better contextualise the Ethics Action Plan's questions (see Appendix 2: Ethics and Human Centricity Action Plan v1), reflect on their relevance to their specific technological and organisational contexts, and develop analytical and reflexive capacity needed to revisit and refine their responses as the project evolves.

3.3.2.2 Component 2: Alignment Pathway (the structure)

While the support activities provide guidance, the alignment pathway defines the structure connecting the aims of the SSH framework to the forms of its application. In practice, this translates as the core domains where projects must demonstrate progress and compliance. This pathway is organised around **four strategic pillars**, as illustrated in Figure 10:

Pillar 1, Commitment to Excellence: This pillar establishes the normative foundation for the project. It invites teams to move beyond passive compliance and actively engage with an ethical stance. Under this pillar, teams are encouraged to articulate their adherence to the **Principles for Human-Centric and Ethical HRI** (section 3.1.1) in their own words, explicitly defining how each principle resonates with their specific technological aims and operational context. It supports the establishment of a common and specific language to guide decisions and trade-offs during development stages. Ultimately, this commitment serves to foster an open culture that welcomes broad discussions on ethical questions and alignment with societal goals.

Pillar 2: Stakeholder Engagement Plan: This pillar structures the project's relationship with the outside world. It proposes a four-step progression, guiding projects from initial considerations of inclusivity toward advanced participatory practices:

1. **Addressing Diversity and Inclusion:** Invites an open reflection on accessibility, representation and vulnerabilities to prevent unjustified exclusion at different phases.
2. **Stakeholder mapping:** Propose and utilize systematic methods to map stakeholder groups at the micro, meso and macro levels. It aims to recognize specific actors' needs and expectations as well as their level of influence and impact.
3. **Stakeholder engagement:** After recognition, projects are encouraged to select and apply relevant tools (e.g. surveys, interviews, focus groups) to gather insights and monitor how stakeholder needs evolve as the technology matures.
4. **Participatory practices:** Once the closer stakeholders have been recognised, projects can invite them to influence their trajectory through collaborative and co-creation methods.

Pillar 3: Ethics & Legal Compliance: The third pillar addresses legal compliance and standards regarding ethical innovation, focusing on the practical obligations that are currently in place at the EU level, local level and specific to their field of application. The aim is to demonstrate awareness and provide concrete evidence for how these obligations are met in practice. Following the recommendations of this framework (see section 3.2.2), the guidance is structured to be compliant towards:

1. **EU regulatory framework:**
   a. AI Act
   b. GDPR
2. **Ethics guidelines and standards:**
   a. ALTAI assessment
3. **Additional local and case-specific regulations and standards**

Pillar 4: Social and Environmental Impact: This pillar prompts projects to consider the broader societal and environmental impacts of their work by identifying both desirable and undesirable impacts. Projects are expected to draw on internal analysis and on external assessments—the external ethics assessment and ALTAI—to recognise concerns and potential consequences. Importantly, the recognition of concerns is not supposed to be static; instead, the structure prompts for follow-ups across development processes and translates into specific actions. The outcome aims for concrete strategies to mitigate negative impacts and enhance positive ones. This pillar is closely related to the **ARISE Ethics Impact Assessment (EIA)** detailed in the previous section 3.3.1. In practice, projects fulfill the requirements of this pillar by applying the EIA methodology to systematically map impacts, define mitigation strategies, and plan for the maximisation of positive outcomes.

3.3.2.3 The Ethics and Human-Centricity Action Plan

To operationalise the four strategic pillars, beneficiary projects are required to maintain a dedicated living document known as the **Ethics and Human-Centricity Action Plan**. This document serves a dual critical function within the mentoring model. Primarily, it acts as the central repository where the outcomes of support activities are crystallised into concrete project strategies. For instance, insights generated during a collaborative workshop are directly applied to populate the corresponding section of the plan.

Beyond its role as an internal repository, the document forms the primary evidence base for formal assessments conducted by the ethics mentors and by the external ethics committee. Structured to mirror the pillars of the alignment pathway, the Action Plan is explicitly designed to function as a live document. Rather than being a static submission, it remains open to continuous updates at any stage, allowing teams to refine their strategies as the project evolves and new ethical or technical insights emerge.

It is important to note that the pathway to completing the Action Plan is **iterative**. Projects are not expected to finalise all sections in a single step. Instead, the model follows a cyclical approach over the 12-month mentoring period, where specific pillars come into focus depending on the project's maturity stage.[33]

The template of the Ethics and Human-Centricity action plan is available in Appendix 2

3.3.2.4 Roadmap for Future Use

The final outcome of the SSH mentoring model prepares for sustainable scaling. While the Action Plan manages ethics during the controlled ARISE phase, the Ethics Roadmap ensures these practices survive the transition to the market.

As projects reach the end of the deployment and experimentation phase, they will draft this roadmap to capture risks and requirements that were out of scope for the pilot, but are critical for real-world integration. This document serves as the strategic bridge between the experimental phase and full industrial deployment at higher TRLs, ensuring the continuity of human-centric practices.

3.3.3 Implementation at TEFs

Future iterations of the SSH framework will focus on tailoring its application to the specific context of the ARISE TEFs. Unlike the open-call beneficiaries, who develop context-specific solutions, TEFs act as providers of expertise, technology modules, and standards. Their

[33] The completion of this pathway is expected to introduce improvements and a final form for the second open call.

developments are often "open-ended", designed to be versatile rather than situated in a single context.

Applying human-centric requirements at the TEF level presents a distinct challenge. Because TEFs focus on developing modular capabilities (missions, tasks, and skills) rather than fully deployed solutions, they inherently lack the 'situated context' where ethical concerns typically surface. For technology developers, translating abstract SSH principles into these decontextualised technical components has proven difficult, primarily because the end-user and environment remain hypothetical, and there is often no shared language to bridge the disciplines. Furthermore, as detailed in section 2.3.3, technology-focused roles often lack the specific resources and interdisciplinary capacities required to bridge this translation gap effectively.

The next stages will look into the integration of SSH at the general ARISE framework, using the TEFs as an entry point. This requires close, interdisciplinary collaboration between SSH experts and technical partners to embed ethical requirements directly into the granular components of robotic behavior.

This process can be conducted by making translations in two directions:

- From task to SSH, evaluating specific existing robotic actions (task, skills) to attach relevant non-technical requirements. For example, for the standard "transport object from A to B" skill, the translation defines inherent requirements regardless of context: presence of safety mechanisms, clear and inclusive communication signals, and data minimisation protocols during navigation.
- From SSH to task, translating high-level human-centric concepts into specific technical capabilities. For example, the abstract value of "empowering the operator" is translated into specific interaction modalities that support dignity and autonomy. Implementing manual hand-guiding teaching allows the operator to teach the robot physically, retaining their expertise rather than relying on complex coding.

This integration is closely linked to the ARISE Task 4.5 Starter Kit for Open, Human-Centric, and Agile HRI. The insights gained from mapping ethics to robotic skills at the TEF level can be consolidated into the Starter Kit. This ensures that methods, tools, and best practices for responsible HRI are shared as both technical and non-technical components across the TEFs and the wider ARISE ecosystem.

Finally, the implementation at the TEF level must also align with the measurements of progress detailed in the following section. As TEFs validate the technical maturity (TRL) of these modules, they will simultaneously begin to benchmark their potential for societal alignment using the Societal Readiness Level (SRL) indicators.

Co-funded by
the European Union

3.3.4 Measurement and Progress

Future stages of the framework will focus on how progress regarding ethics management and human-centricity is measured across different settings. However, measuring these dimensions remains a significant challenge: unlike technical performance, ethical impact is often qualitative, context-dependent, and latent. To address this, future work will aim to translate these 'abstract' values into tangible indicators. A key mechanism for this will be the introduction of Societal Readiness Levels (SRL),detailed in the next section, which will serve as a standard to benchmark synchronised maturity alongside technical development.

3.3.4.1 Technology Readiness Level & Societal Readiness Level

A forward-looking approach to human-robot interaction requires advancing technical maturity but also confirmation that new systems are socially and ethically prepared for real-world deployment. Alongside the widely used technology readiness level (TRL) scale[34], discussed in section 3.1.3, the **societal readiness level (SRL)** model offers a complementary lens for assessing how well a technology fits with societal expectations[35]. SRL focuses on, for example, value alignment, legitimacy, responsible use, and the preparedness of organisations and communities to adopt new technologies in ways that aim to enhance societal well-being.

Bringing TRL and SRL into dialogue provides a way to track progress that considers both technological and human-centric development. Ideally, each step up in TRL should be accompanied by a corresponding rise in SRL. As technical capabilities advance, so should the clarity of the technology’s purpose and the integration of societal aspects.

Coupling TRL with SRL offers a possible practical tool for steering development trajectories. Early prototypes could be paired with exploratory engagement and values clarification, while mid-level development stages could integrate participatory testing and an ethical impact assessment. Later stages of technology development can focus on organisational preparedness and policy alignment to clarify possible long-term societal implications.

Table 8. Levels or Technology Readiness Level & Societal Readiness Level

Level of progression	Technology Readiness Level	Societal Readiness Level
1	**Basic principles observed**	Identifying problems and identifying societal readiness.
2	**Technology concept formulated**	Formulation of problem, proposed solution(s) and potential impact, expected societal readiness; identifying relevant stakeholders for the project.
3	**Experimental proof of concept**	Initial testing of proposed solution(s) together with relevant stakeholders.

[34]Technology Readiness Level (TRL) | Aalto University. (2025, July 28). Aalto.fi. https://www.aalto.fi/en/services/technology-readiness-level-trl
[35]Societal Readiness Levels (SRL) defined according to Innovation Fund Denmark. (n.d.). https://innovationsfonden.dk/sites/default/files/2019-03/societal_readiness_levels_-_srl.pdf

4	**Technology validated in a lab**	Problem validated through pilot testing in a relevant environment to substantiate proposed impact and societal readiness.
5	**Technology validated in a relevant environment**	Proposed solution(s) validated, now by relevant stakeholders in the area.
6	**Technology demonstrated in a relevant environment**	Solution(s) demonstrated in a relevant environment and in co-operation with relevant stakeholders to gain initial feedback on potential impact.
7	**System prototype demonstration in operational environment**	Refinement of project and/or solution and, if needed, retesting in a relevant environment with relevant stakeholders.
8	**System complete and qualified**	Proposed solution(s) as well as a plan for societal adaptation complete and qualified.
9	**Actual system proven in an operational environment**	Actual project solution(s) proven in a relevant environment.

A practical way to integrate TRL and SRL within ARISE could be to use the two frameworks as a flexible guide for mentoring and collaboration with the beneficiary projects and TEF partners. For example, early development activities might be supported with lightweight values clarification and stakeholder reflection, while more mature prototypes could benefit from structured participatory testing, societal impact reviews or organisational preparedness discussion, depending on the context and stage of the project. These elements need not be tied rigidly to specific TRL or SRL thresholds, but rather use the elements to illustrate how technical and societal readiness could be advanced in parallel. As the project progresses, the most suitable ways of operationalising this dual-readiness approach in mentoring and TEF interactions will be analysed more deeply in next stages.

3.4 The SSH Toolbox

The SSH Toolbox gathers practical resources to support the consideration of social, ethical, and human-centric aspects throughout HRI development. It includes methods and reasoning for co-creation and collaboration, a use-case library illustrating how ethical issues emerge in HRI contexts, a consolidated set of ethical questions connected with the ARISE HRI principles, and a collection of templates to help structure analysis and documentation of ethical concerns and impact. The tools and methods that are presented in this section are intended to be used flexibly and consulted as needed to support project participants in applying SSH perspectives in practice.

As part of the ongoing efforts to ensure the accessibility and usability of this framework, the ARISE SSH partner developed a web tool that serves as a digital guide for referencing and integrating the contents of this section. Following the release of this deliverable, the web tool will be updated to include the most recent changes to the ARISE SSH Framework and Toolbox. The web tool is accessible at https://arise-ssh.info.

3.4.1 Stakeholder collaboration

The ARISE SSH framework ensures alignment with the principles of human-centricity and Industry 5.0, both of which emphasise meaningful human interaction, collaboration, and co-creation. Growing evidence suggests the consideration of the 'human' in human-centricic industries should expand from the narrow focus on the 'operators' to include other actors such as designers, engineers, managers, decision-makers, end-users, consumers, and wider societal actors who influence or are influenced by technological developments[36]. This broader framing situates the ARISE SSH framework within principles of social well-being, inclusion, and responsible integration of technologies.

Why collaboration?
Research insights point out that if ethical concerns and human inputs remain isolated from technological advancements, HRI technologies could easily reinforce rigid or even exploitative work conditions and other existing vulnerabilities over time[37]. In response, many propositions suggest that the first step for integration is the establishment of a **shared understanding of the potential implications** of developments by all the parties involved. This step is critical, as what ethics mean and imply in practice often significantly differs between engineers, operators, managers, designers, safety specialists, and social actors. Therefore, there is a need for "socialisation" of ethical terms, common concepts and collective knowledge of ethical issues to which social actors would habitually refer.[38]

Participatory approaches grounded in inclusive dialogue among key stakeholders are essential to promote trust-building, familiarity and agency. They provide critical insights into the necessary trade-offs between economic efficiency and ethics alignment (e.g. balancing productivity with well-being and accountability) prevalent in Industry 5.0[37]. Within the context of technology development and design, collaborative sessions that integrate ethical reflections on a regular basis help organisations anticipate operational and organisational risks while supporting individuals' physical and psychological safety, preparedness, confidence, and capability building for adapting to evolving human–robot interactions[37]. Furthermore, because many ethical impacts materialise beyond the immediate context of test and use, some kind of collaboration must also extend to involve the wider ecosystem of stakeholders.

The ARISE SSH framework proposes that the emerging HRI industry must develop a more **tangible narrative around human-centricity that is shared among all relevant ecosystem actors.** This is necessary to clearly recognise when actions are truly human-centric and the distinction between what is desirable and what is simply acceptable. Early, lightweight practices, starting with stakeholder mapping, joint ethics check-ins, and different types of dialogues, can seed common ground that matures over time.

---

[36] Nasir, V., Hosseini, A., Binfield, L., Hasani, N., Ghotb, S., Diederichs, V., … Hansen, E. (2025). Human-centric Industry 5.0 manufacturing: a multi-level framework from design to consumption within Society 5.0. International Journal of Sustainable Engineering, 18(1). https://doi.org/10.1080/19397038.2025.2551000
[37] Reijers, W., Young, M. T., & Coeckelbergh, M. (2025). Introduction to the Ethics of Emerging Technologies. In Palgrave philosophy today. https://doi.org/10.1007/978-3-031-85887-1
[38] Callari, T. C., Vecellio Segate, R., Hubbard, E.-M., Daly, A., & Lohse, N. (2024). An Ethical Framework for Human-robot Collaboration for the Future People-centric Manufacturing: a Collaborative Endeavour With European Subject-matter Experts in Ethics. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.4927452

Levels of engagement

Different forms of participation grant stakeholders varying degrees of influence in shaping research and innovation processes. Following the taxonomy outlined in the proEthics framework[39], engagement can range from ***inform*** and ***consult***, where communication is largely one-directional, to more reciprocal modes such as ***involve*** and ***collaborate,*** and ultimately ***co-creation***, where decisions are shaped jointly with shared ownership of outcomes. These distinctions matter ethically: lighter forms of engagement may be efficient but risk overlooking lived experiences and reinforcing existing power asymmetries, while deeper forms require investing resources, sharing control and accommodating diverse viewpoints.

Selecting an appropriate level of participation is a methodological and ethical decision, which are both shaped by context, stakeholder composition, and the purpose of the activity. Terminology also plays an important role, where labels such as "stakeholder", "citizen", or "expert" determine who is invited into the process and whose knowledge is recognised. A shared vocabulary around participation helps organisations design processes that connect with human-centric values and support trust-building.

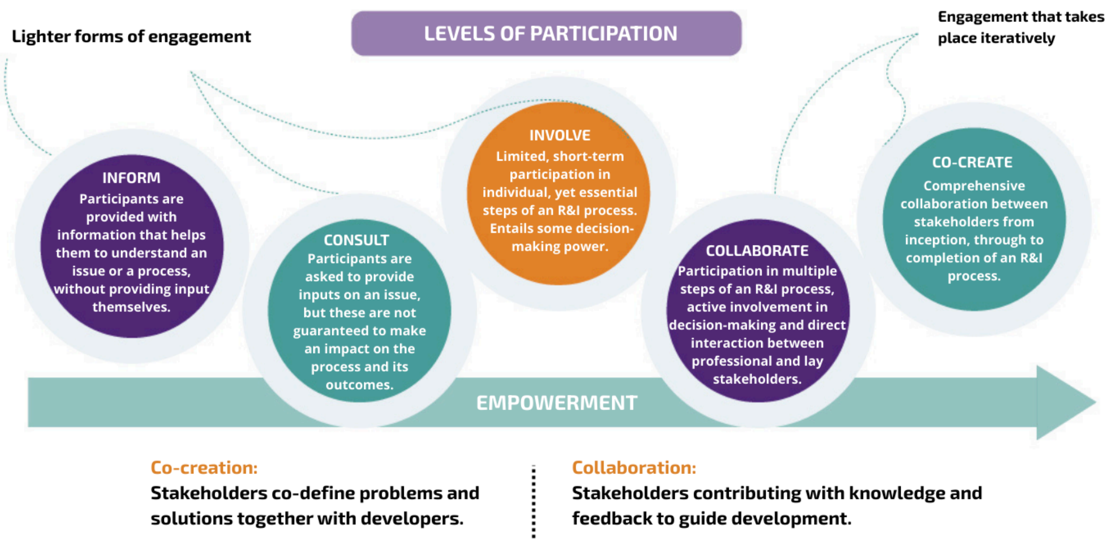


Figure 11. Levels of Stakeholder Engagement adapted from proEthics, Ethics Framework and Guidelines- A guide for research funding organizations implementing participatory activities.

3.4.1.1 Stakeholder mapping

Consistent with the preceding analysis, HRI projects require a broad view of stakeholders—from the immediate ecosystem to the periphery—to strategically determine where intensive collaboration is needed in place of lighter inclusion. The practical method for identifying these actors and defining engagement boundaries is **stakeholder mapping.**

[39] Wiarda, Giannelos, Schuerz, Reber, & Doorn. (2023). Ethics Framework and Guidelines for Participatory Processes in the Activities of Research Funding Organizations. Zenodo. https://doi.org/10.5281/zenodo.8089673

Conventionally, stakeholder mapping involves a systematic listing of all relevant stakeholders, followed by organizing them into diagrams, models, and other frameworks that support analysis. When evaluating the ethical and human-centric implications of robotics solutions, it's crucial to be aware of not only internal capacity and responsibility, but also the external context in which stakeholder needs are embedded in. The table below provides a starting point for mapping out the relevant external stakeholders for an HRI project.

Table 9. Stakeholder roles & descriptions

Category	Examples of actors	Description
Funding	Investors Research institutions EU Commission	Funders often have influence over the requirements, roadmap, and goals of the project. SMEs may be responsible for continuously informing funders about the progression of projects, and adhering to the terms set when funding was acquired.
Policy	Local, regional, national, and EU policy makers	Legislation outlines the legally binding obligations SMEs must adhere to throughout project development. Alignment with existing legislation and paying attention to upcoming implementation of new directives can be crucial to the success of a project.
Direct users	Primary end-user of solution	End users are the primary actors interacting with the robotic solution most intimately and consistently. Committing to human-centricity means committing to prioritizing their short and long-term well-being through deep engagement.
Indirect users	Managers of end-users, Family members of end-users	Indirect users include those closely related to direct users, or actively involved in the contexts where the robotic solution is used. Their understanding of the needs and challenges of the problem context can complement and fill in blind spots from engagements with the direct users.
The public	The immediate surrounding community where the robotic solution is deployed	The public refers to all actors who are embedded in the community and social context where a robotic solution is being introduced (e.g. local organisations, civil society groups) . Although these actors may not constitute a target user group, their opinions of and reactions to a robotic solution can set social and cultural norms that ultimately influence the adoption success of the solution.
Collaborators	Other SMEs Independent HRI workers Research labs Living labs	Technology firms often form partnerships with other firms to collaboratively develop a project. Understanding the goals, challenges, and capacities of these collaborations is crucial for the success of projects where more than one technology firm is involved.
Technology ecosystem	Developers and maintainers of open source software	Robotic technologies are rarely developed completely in-house, and often involve leveraging existing libraries, standards, and modules to piece together the resulting solution. Understanding the needs and processes of actors in the technology ecosystem can lead to more streamlined and better informed integration of open source or third-party components.
SSH support	External advisors External ethics committee	Experts with insight into the social and ethical implications of a developing technology can provide support during crucial decision making stages of the project. External advisors, for instance, can help recognize and mitigate emerging ethical issues before they have severe impacts on the usability and marketability of the solution.

After recognizing the possible external actors that may influence or be impacted by the HRI solution, HRI teams can use a variety of methods and frameworks to understand the relationships, needs, and dynamics of external stakeholders and strategically plan out their project's engagement approaches. Examples of stakeholder mapping methods include:

- **Micro, Meso, Macro Proximity Mapping:** Map proximity to use context. Intuitive, useful for first approaches (Fig. 12)
- **Power-Interest Grid:** Map influence vs. engagement. Helps plan communication and participation levels (Fig. 13)
- **Influence-Impact Grid:** Focuses on who is most affected by project outcomes.
- **Value Exchange Mapping:** Shows what each stakeholder gives and gains (e.g., time, safety, data, trust)
- **Stakeholder Relationship Mapping:** Visualize how stakeholders are connected to one another
- **Stakeholder Radar:** Offers a detailed look at stakeholders' engagement levels and roles by two attributes (level of engagement and domain)

The following section exemplifies the use of the **proximity mapping** and **power-interest grid** methods using a fictional project.

Introducing AURORA
AURORA is a project led by NextGen Components, a robotics and AI engineering firm, specializing in high-precision technology for the aerospace industry. **AURORA's goal: a collaborative robot system that lets skilled operators set up complex welds using voice and simple hand cues, cutting setup time while keeping humans firmly in control.** NextGen's team in AURORA is composed of 5 engineers, plus one colleague with a collaborative & UX design background. They are led by a project manager and closely supported by the CEO. For AURORA, NextGen Components partners with AstraWeld, a welding facility producing high-end, custom aerospace components. NextGen builds the AI-enabled interfaces and control logic; AstraWeld provides the production environment, expert operators and test lines to validate usability, safety and quality.

Table 10. An introduction to AURORA, a fictional ARISE project.

As seen in Figure 12, the **Micro** level focuses on actors directly interacting with the robotics development. In AURORA, it includes operators, trial participants, line supervisors, and welding engineers (the target end user). SSH perspectives at this level include concerns for physical safety, psychological well-being, and ergonomic optimization for direct interactions.

The **Meso** level refers to actors involved in the organisational context. Meso-level stakeholders include managers, C-suite level decision makers, engineering and design teams, legal experts and advisors. The SSH perspectives include concerns for worker autonomy and agency, safeguarding data privacy, supporting managers, and managing the workload between workers and robots.

The **Macro** level concerns actors in external contexts, such as policymakers, investors, funding institutions, and local communities. The SSH concerns include awareness of societal demand and needs, awareness of vulnerable and under-represented groups, and structural support for safe, ethical workplaces.

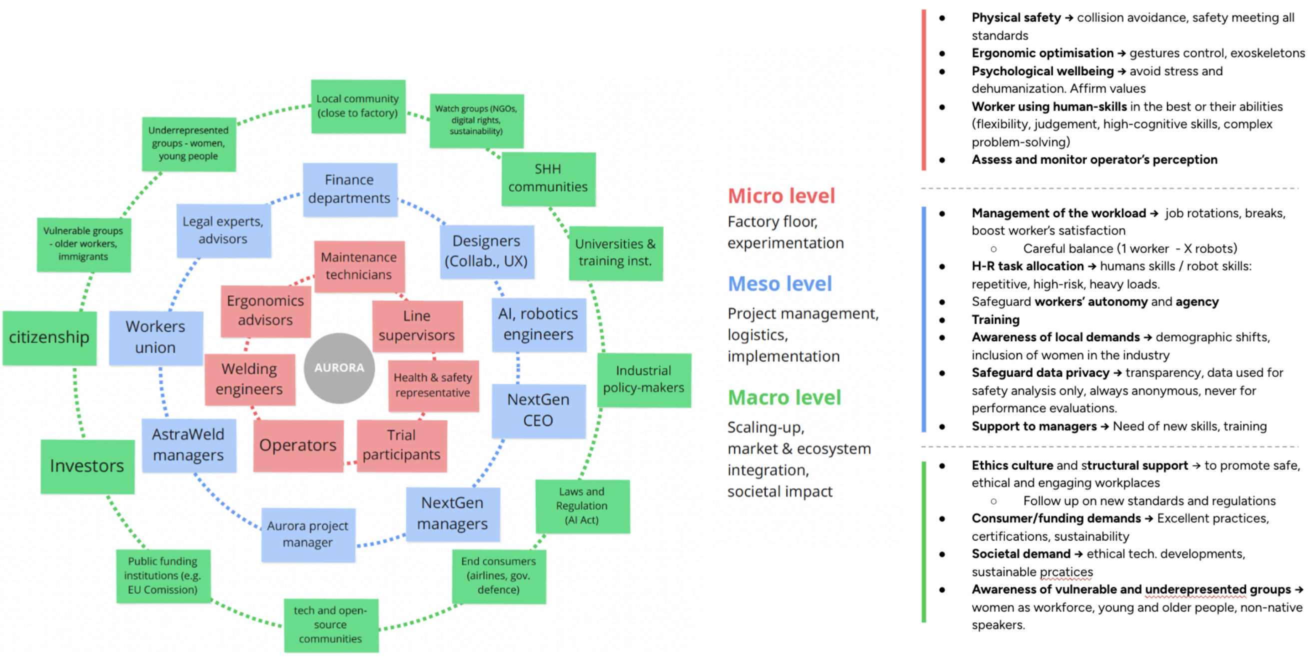


Figure 12. Micro, Meso, Macro mapping for the fictional project AURORA.

Taken together, the Micro, Meso, and Macro mapping framework prompts teams to reflect on the needs and concerns of actors involved at every scale of the technology solution, from interaction level concerns with testers and users to broad, societal-level implications.

The **Power-Interest Grid** supports a more strategic understanding of stakeholder relationships by mapping actors according to their degree of influence over the project's direction (**power**) and their level of engagement or stake in its outcomes (**interest**). This method helps determine which stakeholders should be closely involved, kept informed or periodically consulted. For instance, in AURORA (Fig. 13), frontline actors such as operators, trial participants and welding engineers fall into the quadrant of high interest and high power, meaning that these groups should be kept involved, consulted and informed during the development process. Whereas AstraWeld and NextGen managers fall into low interest and high power and thus, should be kept informed and consulted when necessary.

The Power-Interest Grid functions as a prioritisation tool that helps project teams balance formal authority with experiential insight and allocate engagement resources efficiently while improving and safeguarding inclusivity.

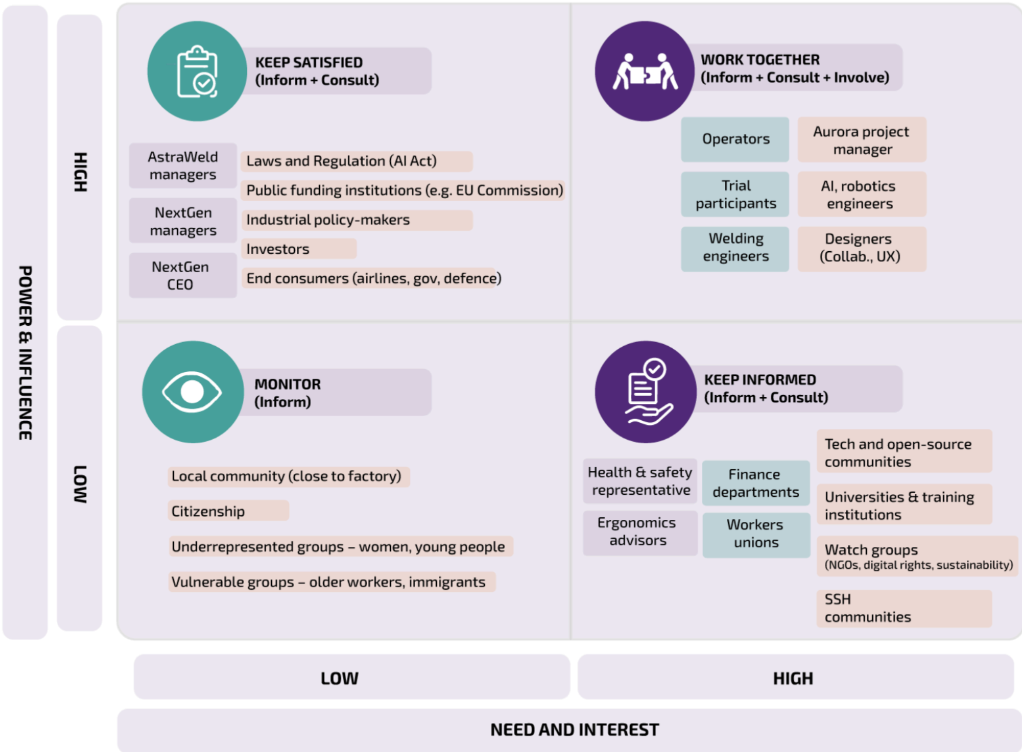


Figure 13. Power-Interest Grid for the fictional project AURORA.

3.4.1.2 Stakeholder Engagement: Pre-engagement Considerations

Prior to stakeholder engagement, teams must explicitly define the values and metrics to be evaluated. While traditional strategies often focus on usability and efficiency, robotics developers should prioritise broader human-centric metrics such as user trust, perceived agency, and impacts on physical or emotional well-being. The ethical principles outlined in this framework serve as a vital starting point for identifying these "beyond-usability" indicators, ensuring that engagement sessions validate the meaningful incorporation of specific values into the system.

To integrate these values strategically, teams can leverage theoretically grounded approaches such as **Value Sensitive Design**[40], which accounts for human values in a principled manner suitable for HRI contexts. Additionally, the **Ethics by Design**[41] approach offers a framework to

[40]What Is VSD :: Value Sensitive Design @ Khoury College. (n.d.). Vsd.ccs.neu.edu. https://vsd.ccs.neu.edu/introduction/what-is-vsd/

[41]European Commission. (2021). Ethics By Design and Ethics of Use Approaches for Artificial Intelligence. https://ec.europa.eu/info/funding-tenders/opportunities/docs/2021-2027/horizon/guidance/ethics-by-design-and-ethics-of-use-approaches-for-artificial-intelligence_he_en.pdf

structure development processes, ensuring that ethical considerations are proactively incorporated into the product lifecycle.

3.4.1.3 Stakeholder Engagement Methods

Following pre-engagement reflection, HRI development teams must select specific methods to operationalise their strategy. This selection requires a careful assessment of relevant stakeholders, project context, and internal capacities to ensure that the chosen methods are appropriate. Tailoring strategies to these specific needs is essential for generating meaningful, actionable insights rather than generic feedback.

Qualitative engagement approaches: Qualitative methods are best suited for generating rich, deep insights, often from smaller participant groups, revealing the 'why' behind behaviors and preferences. Key methods include:

- Semi-structured interviews
- Focus groups
- Participatory workshops
- Observational studies in real settings
- Case analysis of contextual use and acceptance
- Questionaries

Quantitative engagement methods: When teams need to validate hypotheses across a larger participant group or integrate human-centric metrics into verifiable project checkpoints (e.g., KPIs), quantitative methods may be more appropriate. The following list categorises these methods by the primary dimension they validate:

- **Usability**: Standardised tools to measure ease of use include System Usability Scale (SUS), User Experience Questionnaire short (UEQ-S) and Usability Metric for User Experience (UMUX-lite).
- **Experience**: Metrics capturing the user's internal state include Patient reported outcome measures (PROMS), Patient reported experience measures (PREMs), and self-reported scales for trust, comfort, and well-being.
- **Performance**: Objective measures of human–robot interaction efficiency, such as setup time reduction, task accuracy, error rates, and learning curve.
- **Satisfaction and acceptance**: Indicators of adoption likelihood include Net Promoter Score (NPS), Acceptance scores.

3.4.1.4 Making Sense of Engagement Data

To make sense of both qualitative and quantitative data generated from stakeholder engagement sessions, HRI teams can use a variety of tools such as

- **User journey mapping:** Visualizes a specific stakeholder group's interaction with the solution from the very initial contact, through their engagement, to their ultimate goal
- **Personas:** A fictional but research-based representation of real stakeholders that capture their core needs, values, and motivations

- **Scenarios:** A detailed narrative that describes how a user interacts with the solution and captures a user's motivations and behaviors

Relying on design-based insight generation methods, HRI development teams can more effectively translate the disjointed mass of data into actionable insights.

3.4.1.5 Additional Collaboration Toolkits and Approaches

In addition to the tools, processes, and methods described above, Table 11 provides an overview of the design processes and approaches that HRI teams can rely on to structure their engagement and feedback strategies.

Table 11. General design approaches

Name of design approach	Description
Design Thinking	Empathise, define, ideate, prototype, test.
Participatory Design	Users participate directly in shaping tools and solutions.
User-Centred Design	Focus on real human needs, requirements, and usability.
Double Diamond	Structured phases for divergent and convergent decision-making.
Value-Sensitive Design	An approach to the design of technology that accounts for human values in a principled and comprehensive manner throughout the design process.

Table 12 lists out ethics and human-centric design tools which can be applied to general technology development, while Tables 13, 14 and 15 list out design toolkits that were developed for particular robotics and technology contexts. Rely on the following tables to further explore how your team can leverage design processes to strategically integrate stakeholder feedback into HRI development workflows.

Table 12. Ethics and Human Centered Design Toolkits

Name of tool or approach	What is it used for?
Human-Centered Design Toolkit	Tools supporting human-centric thinking throughout the design process.
Ethics for Designers Toolkit	Tools to explore ethical questions across project stages.
The Tarot Cards of Tech	Card deck of questions prompting tech developers to contemplate the disruption, usage, equity, access, and future implications of technologies they build.

Table 13. Design toolkits for AI

Name of tool or approach	What is it used for?
HCAI Toolkit	Tools for embedding human-centricity in AI system design.

Co-funded by the European Union

Table 14. Design toolkits for robots

Name of tool or approach	What is it used for?
Safety around robots	Step by step toolkit for improving the safety of collaborative robots, designed for actors developing, integrating, manufacturing, and using collaborative robots
LEADOR	End-to-end methodology for co-designing and evaluating social robots.
LEGO® Serious Play® + Design Thinking Workshops	Methods enabling participants to conceptualise meaningful robotic solutions.
Smart Graph Interface (SGI)	Framework for strategic decisions in human–robot collaborative workplaces.
Co-Creation Sessions of Assistive Robots	Participatory sessions to refine assistive robot design.
Social Robot Co-Design Canvases	Canvases facilitating shared decision-making across expert domains.
Situated Participatory Design (SPD)	Designing interactions directly in the robot's deployment environment.
Participatory Design for Explainable Robots	Methods enabling robots to articulate their decision-making.
HRI Toolkits	Suites of tools for designing and evaluating HRC systems.

Table 15. Design toolkits for industrial robots

Name of tool or approach	What is it used for?
SHAREWORK Toolkit	Modular system for safe collaboration without physical barriers.
CoCo System	System enabling dynamic switching between coexistence and cooperation modes.

3.4.2 Use-case library

The ARISE use-case library is a practical **repository designed to ground abstract ethical principles in the specific context of HRI.** Rather than offering generic solutions, the library presents detailed, situated descriptions of HRI workflows, illustrating the ethical concerns inherent to these specific operational contexts. This allows teams to move beyond theoretical discussions and examine how ethical tensions emerge from specific combinations of tasks, technologies, and human factors.

The use-case library follows a consistent structure designed to clarify the objectives, workflow, and ethical considerations of each illustrative use case. Each use case begins with a description of the purpose of the robotic application, discussing the rationale for its use and the technologies selected to support the task. This is followed by a general workflow that describes how the robot and the human actor(s) operate within the broader sequence of activities.

Co-funded by
the European Union

A table of roles and responsibilities then maps the human and non-human actors involved. For example, a human worker, a robot, a sensor, and a PC (system). The roles' functions within the workflow are then specified in the table's cells. Finally, each use case includes a set of questions that are organised based on the ARISE ethical principles (see section 3.1.1). Each principle has dedicated questions for each use case. The complete set of use cases and their descriptions is illustrated below in Tables 16 to 31.

3.4.2. Challenge 1 Use Case

Dismantling and assembly of high-value products: Battery Disassembly

The primary objective of this application case is to optimize and expedite the disassembling process of batteries by utilizing a collaborative approach between humans and robots. This strategy integrates mixed reality and artificial vision technologies to accurately detect and localize screws within the battery, enhancing efficiency and accelerating the disassembly procedure. The actors of the use case are: Human workers, PC (System), Robot/Camera, and MR HoloLens glasses.

The general flow of the work task is as follows: the dimensional camera captures GRB and depth image of the object and forwards them to the PC for analysis. The PC then proceeds with coordinate calculations and transformations to both robot reference space and world coordinates. These coordinates are then sent forward to the robot, which guides the human. The PC's role is to oversee the interaction between the human and the robot. Meanwhile, the collaborative robot selects screws for the disassembly process. The actors of this use case, and their roles, are summarized in Table 16.

Table 16. The actors and their roles in the battery disassembly use case.

Human workers	PC (System)	Robot/Camera	MR HoloLens glasses
Supervising the disassembly process. Establishing the framework and making critical decisions, such as selecting screws to retrieve.	Calculating and transforming coordinates. Coordinating robot actions. Relaying data to AR glasses for enriched visual experiences.	Providing accurate data. Perform physical actions.	Provide real time visual information. Track user interactions for an immersive user experience.

Table 17. Associated questions for each ARISE principle for challenge 1

ARISE Principle	Associated Question
Respect for Human Agency	How is workers' informed consent ensured before the HRI solution is deployed?
Privacy and Data Governance	How is it ensured that the camera embedded in the system collects data solely for the purpose of the HRI solution?
Fairness	How does the system account for differences among workers to ensure guidance from the robot and MR interface is equally usable and does not disadvantage certain workers?

Individual, Social and Environmental Well-being	What potential health and safety risks to workers are assessed and what measures are implemented to address them?
Transparency	How transparent is the system to human workers regarding how screws are selected, how coordinates are calculated and how guidance decisions are made by the PC-robot system?
Accountability and Oversight	How are accountability and liability for the HRI system's performance established, especially in the case of system failures?
Worker Dignity and Equity	Are there measures in place to ensure that all team members benefit from the time savings and reduced physical strain provided by the collaborative robot?
Human Resilience through continuous learning and support	What ongoing training and support are provided to help workers adapt to using MR glasses and working alongside a robot?

3.4.2.2 Challenge 2 Use Case

Complex product picking in industrial warehouses: Fruit picking

The main purpose of this application case is to make certain tasks that would be done by a human worker, like picking fruits, can be done by a collaborative robot, making HRI possible and perform the work more efficiently. The actors of the use case are: Human workers, MR HoloLens glasses, system, Collaborative robot and camera. The users engage with the mixed reality glasses through intuitive gestures, voice commands and gaze. The MR glasses then send forward this input to the user by providing a live camera view, object information and system feedback. The glasses also forward relevant data to the PC via a Wi-Fi connection (WebSockets protocol). The PC then processes the data and sends the processed data back to the user. The PC's role is to act as a central hub and facilitate the communication between the user, the glasses and the robot, sending commands to the robot based on the input information received from the user. The robot provides updates on its actions and its environment to the PC. The camera's role is to capture visual data and send it to the PC for adjustments of camera settings. The actors of this case, and their roles, are shown in Table 18.

Table 18. The actors and their roles in the fruit picking use case.

Human workers	MR HoloLens glasses	System	Collaborative robot and camera
Supervise operations Provide instructions to the collaborative robot via MR HoloLens glasses	Provide interface between human workers and the system Provides location of objects as a digital sphere so that the user is aware of the pick-up position of the robot	Receive instructions Provide real-time visual feedback	Robot sorts based on instructions from human workers Camera captures visual data for object recognition and situational awareness Aids decision-making in the system

Table 19. Associated questions for each ARISE principle for challenge 2

ARISE Principle	Associated Question
Respect for Human Agency	To what extent can workers decide when, how and to what degree the collaborative robot performs picking tasks, rather than merely supervising or reacting to automated behaviour?
Privacy and Data Governance	If the HRI solution uses data from a third party, how is API integration compliance ensured, and is a compliance form provided?
Fairness	How does the project ensure that visually impaired, far-sighter or near-sighted workers can effectively use MR HoloLens glasses?
Individual, Social and Environmental Well-being	What protocols are in place for safe physical interaction, including detailed fail-safe mechanisms?
Transparency	Are the participants informed about the purposes, risks, and benefits of their involvement in the research activities?
Accountability and Oversight	What processes are in place to ensure compliance with ethical, legal and organizational standards throughout the research activities?
Worker Dignity and Equity	How is training for the HRI solution made accessible and equitable for all employees?
Human Resilience through continuous learning and support	How are workers encouraged to provide feedback about the technology and their training needs?

3.4.2.3 Challenge 3 Use Case

Flexible collaborative robots: Assembly and packing robotic assistant

In this application case, the robot should recognize and manage the heavy product, safeguarding the operator from injuries and safety problems, while the operator should monitor the robot's activity and intervene, for example, if inaccuracies occur or unexpected forces are requested to perform the task.

A product is selected to be packed from the company management system's database by an operator. After this, the robot receives the product code information and obtains the necessary information to recognise the product in the space (CAD and images). The robot also receives the information on the related tray. Subsequently, the operator prepares the packing area by placing the first tray on a pallet. Through physical interface and voice input commands, the operator advises the robot so that it can recognise the object and its pose. Then the robot picks up the object and places it in one of the free slots. Because the pre-formed tray constrains the product's positioning and requires a specific force to achieve the correct coupling, the positioning sequence in the slots is pre-programmed. In case of incorrect coupling in the product's final slot, the operator can intervene by physically acting on the robot for a successful positioning of the product in its slot. In this phase, the robot bears the

Co-funded by
the European Union

product's weight and cooperates with the human to achieve precise positioning. Once the tray is filled, the operator will perform a quality check and set up another tray until the desired number of trays for the current pallet is reached. At any time, the operator can stop, start or give high-level commands to the robot. During the whole process, a camera monitors the operator's posture, and they are alerted if their ergonomic indices fall below a threshold. The actors of this use case, and their roles, are summarized in Table 20.

Table 20. The actors and their roles in the assembly and packing robotic assistant use case.

Collaborative robot and operator	Human workers
Execute force operations	Supervise the packing process
Execute fine manual operations	Handle exceptions
Share some tasks with human operators	Manage the empty pre-formed tray layer or empty boxes
	Manage orders from company management system
	Coordinate shipments

Table 21. Associated questions for each ARISE principle for challenge 3

ARISE Principle	Associated Question
Respect for Human Agency	How does the system design and work process ensure that operators can understand and deliberately influence the robot's actions?
Privacy and Data Governance	Will research data be anonymized or pseudonymised? If so, what techniques will be implemented?
Fairness	How are the camera features designed to consider diverse backgrounds, facial features, languages, and accents of participants?
Individual, Social and Environmental Well-being	What sustainability practices are adopted for dismantling, assembly and the environmental impact of the use case?
Transparency	How transparent is the system to operators regarding when and why the robot requests specific forces, follows pre-programmed positioning sequences or triggers ergonomic alerts during packing process?
Accountability and Oversight	How will the implementation of health and safety procedures, data collection measures and informed consent processes be monitored and evaluated?
Worker Dignity and Equity	How does the system ensure equitable safety measures for all operators, particularly in interventions involving physical interactions with the robot?
Human Resilience through continuous learning and support	What mechanisms are in place for operators to provide feedback on the robotic system and the overall workflow?

Co-funded by the European Union

3.4.2.4 Challenge 4 Use Case

Smart programming: Robotic Kitting Assistant

One of the standard and crucial phases in producing assembled parts is kitting. A specific product is selected from the company's management system, obtaining its Bill of Materials (BoM) with the various components and their location within the warehouse. Following a lean paradigm, all the items constituting the kit should be retrieved from a warehouse, bringing only what is needed to the assembly station. This application case considers both automated warehouses, involving a robotic cell, and non-automated warehouses, consisting of shelving, which humans should manage.

The flow of the process starts when the operator selects a product to be assembled from the company management system, and the robot receives a list of objects, their quantity, and their location within the company. The needed items could be stored in an automated or non-automated warehouse. Then the operator reads the object list from the robot's touchscreen. The robot can follow the operator to the automated warehouse in follow-me mode or independently. Once in the warehouse, the operator takes items from the shelves and places them inside the bin, which will constitute the assembly kit. The operator then updates the list of items to be picked via the screen or with natural language. While this operation is happening, the robot will communicate with the warehouse, requesting the bins with the remaining components. Once the selected bin is retrieved, the robot should pick the right number of items and move them to the assembly kit bin, simultaneously updating the list of items. Once all the objects of the list have been collected, the operator performs a quality check and sends the robot to the assembly station that requested the kit. The actors of this use case, and their roles, are summarized in Table 22.

Table 22. The actors and their roles in the robotic kitting assistant use case

Autonomous mobile robot and collaborative robot	Human workers
Offer a simple way to codify different objects	Teach the robot on what it has to do
Take input naturally from operators	Supervise the kitting process
Autonomously create the kit	Handle exceptions
Notify lack of necessary items	Provide additional instructions if it is necessary
Navigate autonomously inside the warehouse	Task and process scheduling

Table 23. Associated questions for each ARISE principle for challenge 4

ARISE Principle	Associated Question
Respect for Human Agency	What are the detailed procedures for obtaining informed consent, and can templates of the Informed Consent Forms and Information Sheets in language intelligible to participants be provided?

Privacy and Data Governance	How does the use case demonstrate compliance with relevant data protection regulations, such as GDPR, and what security measures are in place to protect collected data from unauthorized access?
Fairness	How is inclusivity ensured in addressing the challenges posed by diverse practitioner characteristics?
Individual, Social and Environmental Well-being	How does the use case ensure compliance with relevant local, national or European health and safety guidelines and regulations?
Transparency	How are the details of informed consent, data processing and other relevant procedures communicated in an understandable and accessible manner to all participants?
Accountability and Oversight	How are accountability and liability for the HRI system's performance established, especially in the case of system failures?
Worker Dignity and Equity	How does the HRI solution address potential biases, especially in recognition technologies used in interfaces like touchscreens and natural language processing?
Human Resilience through continuous learning and support	How does the organization promote a culture where continuous improvement and innovation are encouraged, and learning from mistakes is seen as part of growth?

3.4.2.5 Challenge 5 Use Case

Enhancing robot functionality through multimodal HRI interactions: Interactive learning session

This use case demonstrates the robot's ability to adapt technology to the user's needs and environment, emphasising a human-centred approach in HRI. It showcases how robots can support personal growth and enrichment activities, making technology an empowering tool rather than a constraint.

The focus of this example use case emphasises the patient-robot flow. The robot first assesses the user's (patient) skill level and specific interests in using speech recognition. Next, the robot proceeds to personalise the session based on the user's responses and observable cues. After this, the robot provides step-by-step instructions for the session, utilising multimodal communication such as gestures and verbal instructions. The robot then monitors the user's progress during the session and adjusts the pace or repeats instructions as needed. In addition, the robot uses facial expression recognition and activity recognition to observe signs of frustration and confusion. If these signs are observed, the robot will offer encouragement, suggest breaks, or modify the teaching method to keep the session enjoyable. During the session, the robot will also provide real-time feedback and tips for improvement, encouraging exploration and creativity while simultaneously adapting the session to incorporate the user's ideas and preferences. After the session, the robot reviews the user's achievements and areas for improvement. Based on the user's engagement and feedback, the robot suggests future sessions or related hobbies. The personalised profile of the user is stored for future user interactions with the robot. The actors of this use case, and their roles, are shown in Table 24.

Co-funded by
the European Union

Table 24. The actors and their roles in the interactive learning session use case.

Care staff	Patients	Interactive robot	Care environment
Scheduling physical activities via a mobile-based interface	Different needs, limitations and preferences	Range of sensing and interaction capabilities	Physical building layout Organizational constraints

Table 25. Associated questions for each ARISE principle for challenge 5

ARISE Principle	Associated Question
Respect for Human Agency	How will the informed consent of participants (staff, employee, volunteers, patients, etc.) be obtained to ensure ethical requirements are met?
Privacy and Data Governance	What technical and organizational measures are implemented for further processing of previously collected personal data to safeguard the rights of data subjects?
Fairness	How are the voice command and camera features designed to consider diverse backgrounds, cognitive abilities and technical capabilities?
Individual, Social and Environmental Well-being	What potential health and safety risks to workers are assessed and what measures are implemented to address them?
Transparency	How transparent is the system to users regarding what personal data are collected, how these data are interpreted to personalise the session and how the personalised user profile is used over time?
Accountability and Oversight	What processes are in place to ensure compliance with ethical, legal and organizational standards throughout the research activities?
Worker Dignity and Equity	What mechanisms are in place to continuously monitor and correct biases in algorithms, especially those related to facial expression and activity recognition?
Human Resilience through continuous learning and support	What features does the robot have to create a psychologically safe learning environment where users feel free to take risks and express creativity?

3.4.2.6 Challenge 6 Use Case

Developing robotic systems able to autonomously perform fetch and carry tasks in healthcare environments: Asking for help while carrying medical samples

This case focuses on using assistive robots for the safe and efficient transportation of medical samples between patient wards and laboratories within healthcare environments. In case of the robot being blocked (for instance, because of a hospital bed blocking the passage in a corridor), the robot looks for someone and asks for help.

The main flow consists of two interactions that need to be considered, the first one being fetch & carry flow. In this scenario, a staff member calls the robot and requests it to bring a medical sample to a known location such as a laboratory. Next, a nurse places the sample on the robot and then lets the robot go. The second main flow is a help request flow, which follows the first flow. In this case, the robot has faced an interruption in its navigation. This could be due to, for example, an object blocking the passage or a door being closed. This results in the robot looking for help to overcome these obstacles in navigation. The actors of this use case, and their roles, are illustrated in Table 26.

Table 26. The actors and their roles in the asking for help while carrying medical samples use case.

Hospital staff members	Mobile manipulator (robot)
Send requests for the robot for samples	Carry samples
Load the samples on to the robot	Navigate obstacles
Assist robot when interruptions in navigation happens	Ask for help when navigation is interrupted

Table 27. Associated questions for each ARISE principle for challenge 6

ARISE Principle	Associated Question
Respect for Human Agency	What are the procedures and criteria for identifying and recruiting research participants (e.g., number of participants, inclusion/exclusion criteria, direct/indirect incentives, risks, and benefits for participants?)
Privacy and Data Governance	How does the use case ensure that the data processed is relevant and limited to the purposes of the project, in compliance with the data minimization principle?
Fairness	How are the voice command and camera features designed to consider diverse backgrounds, facial features, languages, and accents of participants?
Individual, Social and Environmental Well-being	What protocols are in place for safe physical interaction, including detailed fail-safe mechanisms?
Transparency	Are the participants informed about the purposes, risks, and benefits of their involvement in the research activities?
Accountability and Oversight	How will the implementation of health and safety procedures, data collection measures and informed consent processes be monitored and evaluated?
Worker Dignity and Equity	Is the robot’s interface designed to be easily usable by all staff, regardless of their technical skill levels or physical capabilities?
Human Resilience through continuous learning and support	What channels are available for staff to report issues or suggest improvements regarding the robot’s operations?

3.4.2.7 Challenge 7 Use Case

Leveraging HRI to improve the efficiency of workers in high precision flexible tasks: PCB Desoldering

Reworking of printed circuit boards (PCBs) that failed to t pass the quality control check is considered a high-precision task. Defective PCBs should be recovered (i.e., defective

Co-funded by the European Union

components must be identified and replaced by operational ones). The process of replacing components is intricate and time-consuming, in the sense that an operator must (i) recognize the defective components on the board by the information received from the quality control station, (ii) pick the functional component among the pool of components (it may be time-consuming, due to similarity of components, for example, resistors with identical shape and different resistance values), (iii) select the suitable tools among the available tools, depending on the type of components and the level of operator expertise, (iv) solder and desolder components taking into account to avoid applying excessive heat to components and board, to orient components correctly, to inspect new soldered joints, and etc., which causes high levels of physical and mental stress on them.

The main flow starts with the automated guided vehicle carrying defective PCBs from the control station room to the reworking station room. During this, the UR5e robot is preparing the rework station by the data received from the control unit. Preparation of the rework station includes identification of the proper tools and functional components from the available pool of tools and components via computer vision. The robot also picks and places all the needed objects on the rework station. The identification of correct tools happens through BGR and depth images that are sent from the robot to the PC, where the PC runs object detection algorithms to identify the correct tools and components. In addition, their positions are calculated and the automated guided vehicle is provided with a map of the building so that it can navigate across the shop floor. The actors of this use case and their roles are summarized in Table 28.

Table 28. The actors and their roles in the PCB desoldering use case.

Human workers	UR5e Robot	AGV
Soldering/desoldering	Pick the functional component among the pool of components Select the suitable tools among the available tools Set up working space	Carry defective PCBs from the control station room to the reworking station

Table 29. Associated questions for each ARISE principle for challenge 7

ARISE Principle	Associated Question
Respect for Human Agency	How are participants, such as employees, being given opportunities to train their skill in HRI?
Privacy and Data Governance	What are the detailed procedures for obtaining informed consent regarding data processing?
Fairness	How is inclusivity ensured in addressing the challenges posed by diverse practitioner characteristics?

Individual, Social and Environmental Well-being	What sustainability practices are adopted for dismantling, assembly and the environmental impact of the use case?
Transparency	What information is made available to operators about the system's assumptions, confidence levels, and limitations when supporting high-precision reworking tasks such as component identification and soldering assistance?
Accountability and Oversight	How are accountability and liability for the HRI system's performance established, especially in the case of system failures?
Worker Dignity and Equity	Are there safeguards to prevent certain operators from consistently receiving more challenging or stressful rework tasks based on the system's performance assessment?
Human Resilience through continuous learning and support	Are there mechanisms for operators to report issues, provide suggestions and contribute to continuous improvement efforts without fear of reprisal or negative impacts on their employment?

3.4.2.8 Challenge 8 Use Case

Leveraging HRI for improving ergonomics in high precision tasks: Workpiece holder

The robot holds the workpiece in the appropriate position and orientation while an operator is working on it. Through machine vision, the physical features of the operator including their height and neck length are provided to the robot for calculating the right position and orientation of the workpiece, prioritizing the ergonomics and comfort of the operator. In addition, for the further adjustment of orientation, the operator sends a speech command to the robot.

The main flow of the process is initiated by capturing RGB and depth images of the operator and sending these images to the PC. The PC then extracts the relevant physical features of the operator, and the appropriate ergonomics and comfort of the operator are defined. In the next step, the desired joint angles of the robot are obtained and sent to the robot for holding the workpiece. The operator's speech is captured and used for voice commands for the robot to make required adjustments. The actors of this use case, and their roles, are shown in Table 30.

Table 30. The actors and their roles in the workpiece holder use case.

Operator	UR5e robot
Send move commands to the robot verbally	Adjusts the workpiece orientation Estimate the pose of the workpiece according to physical parameters of operators through images.

Table 31. Associated questions for each ARISE principle for challenge 8

ARISE Principle	Associated Question
Respect for Human Agency	How does the system ensure that operators retain meaningful control over the positioning and adjustment of the workpiece, including the ability to override, correct, or stop robot actions?
Privacy and Data Governance	How are the operator's physical and biometric data handled in terms of data minimisation, storage duration, access control, and purpose limitation?
Fairness	How are the voice commands designed to consider diverse backgrounds, and non-standard accents?
Individual, Social and Environmental Well-being	How does the use case ensure compliance with relevant local, national or European health and safety guidelines and regulations?
Transparency	How are the details of informed consent, data processing and other relevant procedures communicated in an understandable and accessible manner to all participants?
Accountability and Oversight	How will the implementation of health and safety procedures, data collection measures and informed consent processes be monitored and evaluated?
Worker Dignity and Equity	How can the system be audited for bias in the recognition and adjustment algorithms to ensure fair treatment of all operators?
Human Resilience through continuous learning and support	What training and support mechanisms are provided to operators to effectively interact with the robot, especially for those less familiar with advanced technology.

3.4.3 Ethical questions for robotic systems

While 3.4.2. demonstrates how to operationalize the 8 ARISE principles in particular robotic development contexts, this section restructures those considerations into a list of typically encountered ethical questions in HRI development. These questions have been organised according to the 8 ARISE ethical principles, which, as a reminder, build off of the principles presented in the Ethics by Design framework, the Ethics Guidelines for the Trustworthy AI and the SIENNA and SHERPA projects. Many of the questions were borrowed or modified from the Assessment List for Trustworthy AI (ALTAI), which provides examples of general-level ethical questions related to AI systems. These questions were adapted or refined to better reflect the specific challenges and contextual dependencies of human-robot collaboration.

The resulting list of questions aims to remain connected to well-established ethical approaches while being directly usable within ARISE's HRI context. The formulation of these examples draws on the original work by Callari et al. 2024[42]. When working through these

[42] Callari, T. C., Vecellio Segate, R., Hubbard, E.-M., Daly, A., & Lohse, N. (2024). An Ethical Framework for Human-robot Collaboration for the Future People-centric Manufacturing: a Collaborative Endeavour With European Subject-matter Experts in Ethics. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.4927452

questions with your team, treat them as a discussion guide for genuine reflection on the principles presented, rather than a checklist.

3.4.3.1 Respect for human agency

- Did you align the AI system with the relevant standards (e.g. ISO25, IEEE26) or widely adopted protocols for (daily) data management and governance?
- Is the robot designed to physically or socially interact with humans, guide their actions or take autonomous decisions that influence human behaviour, safety or task performance in shared spaces?
- Could the robot's embodied form, movement, gaze, voice or social behaviour create confusion about whether users are interacting with a machine or a human, or lead to mistaken assumptions about the robot's capabilities or intentions?
- Could the robot cause physical, social or emotional over-reliance?
- Could the robot's proximity, movement patterns, timing, gestures or social cues unintentionally pressure or steer the user's decisions, actions or pace of work in undesirable ways?
- Does the robot engage in social behaviours that influence user expectations, emotions or social dynamics, including interactions between multiple people?
- Does repeated embodied interaction with the robot risk creating emotional attachment, dependency, addictive use or behaviour manipulation due to the robot's physical presence, responsiveness or social characteristics?

3.4.3.2 Privacy and data governance

- Did you align the AI system with the relevant standards (e.g. ISO25, IEEE26) or widely adopted protocols for (daily) data management and governance?
- Did you consider how the robot's continuous sensing (e.g. cameras, microphones, proximity sensors, touch sensors) affects bodily, spatial and informational privacy as well as the physical, mental and moral integrity of individuals in shared environments?
- Depending on the use case, did you establish mechanisms for users to easily flag privacy concerns related to the robot's sensing, recording, storage or sharing of data captured during physical human-robot interactions?
- Is the robot trained on, or does it process, personal data such as biometric information, spatial/location data, or environmental context that may directly or indirectly identify individuals?
- Did you put in place any of the following measures under the General Data Protection Regulation (GDPR), or a non-European equivalent?
    a. Data Protection Impact Assessment (DPIA)
    b. Designate a Data Protection Officer (DPO) and include them at an early stage in the development, procurement or use phase of the AI system;
    c. Oversight mechanisms for data processing (including limiting access to qualified personnel, mechanisms for logging data access and making modifications);
    d. Measures to achieve privacy-by-design and default (e.g. encryption, pseudonymisation, aggregation, anonymisation);

   e. Data minimisation, in particular personal data (including special categories of data);

- Did you implement the right to withdraw consent, the right to object and the right to be forgotten?
- Did you consider the privacy and data protection implications of data collected, generated or processed over the course of the AI system's life cycle?
- Did you consider the privacy and data protection implications of the AI system's non-personal training data or other processed non-personal data?
- Did you align the AI system with the relevant standards (e.g. ISO25, IEEE26) or widely adopted protocols for (daily) data management and governance?

3.4.3.3 Fairness

- Did you align the AI system with the relevant standards (e.g. ISO25, IEEE26) or widely adopted protocols for (daily) data management and governance?
- Did you establish a strategy to avoid creating or reinforcing bias in the robot's perception, movement, behaviour or task allocation - for example, biases in detecting different bodies, mobility aids, gestures, clothing or interaction style?
- Did you consider diversity and representativeness in the data used to train the robot's sensing and interaction capabilities, making sure it can reliably perceive and respond to users of different body shapes, skin tones, mobility patterns, communication styles and cultural norms?
- Did you implement training and awareness initiatives to help robot designers and developers understand how embodied systems may introduce bias?
- Did you ensure mechanisms that allow for the flagging of issues related to bias, discrimination or poor performance of the AI system?
- Is your definition of fairness commonly used and implemented in any phase of the process of setting up the AI system?
- Did you ensure the robot can adapt its speed, proximity, communication modality, gesture recognition and interactive behaviours to accommodate diverse human preferences, abilities and comfort levels?
- Did you assess whether the robot's physical interaction capabilities and interfaces are accessible to people with disabilities or those at risk of exclusion?
- Did you ensure that Universal Design principles are taken into account during every step of the planning and development process, if applicable?
- Did you take into account how the robot's physical presence, behaviour and task role may differently impact various end-users?
- Did you include diverse stakeholders in the design and testing of the robot's embodied interaction, ensuring the system works equitably across real-world contexts?

3.4.3.4 Individual, social and environmental well-being

- Does the AI system impact human work and work arrangements?
- Did you pave the way for the introduction of the AI system in your organisation by informing and consulting with impacted workers and their representatives (trade unions, work councils, etc) in advance?
- Did you adopt measures to ensure that the impacts of the robot on human work, such

as physical workload changes, new coordination demand and altered responsibilities, are well understood?
- Could the AI/robotic system create the risk of de-skilling the workforce?
- Does the system promote or require new (digital) skills?
- Could the AI/robotic system have a negative impact on society or democracy?
- Are there potential negative environmental impacts related to the robot's manufacturing, resource use, energy consumption, disposal or embodied sensor infrastructure?
- Where possible, did you establish mechanisms to evaluate the environmental impact of the AI system's development, deployment and/or use (for example, the amount of energy used and carbon emissions)?

3.4.3.5 Transparency

- Did you put in place measures that address the traceability of the AI system during its entire lifecycle?
- Did you explain the decision(s) of the AI system to the user?
- Did you continuously survey the users if they understand the decision(s) of the AI system?
- In cases of interactive AI systems (e.g. chatbots, robo-lawyers), do you communicate to users that they are interacting with an AI system instead of a human?
- Did you establish mechanisms to inform users about the purpose, criteria and limitations of the decision(s) generated by the AI system?

3.4.3.6 Accountability and Oversight

- Did you establish mechanisms that facilitate the AI system's auditability (e.g. traceability of the development process, the sourcing of training and the logging of the AI system's processes, outcomes, positive and negative impact)?
- Did you ensure that the AI system can be audited by independent third parties?
- Did you foresee any kind of external guidance or third-party auditing processes to oversee ethical concerns and accountability measures?
- Did you organise risk training and, if so, does this also inform about the potential legal framework applicable to the AI system?
- Did you consider establishing an ethics review mechanism that explicitly addresses robotics-specific concerns such as physical risks, proxemics, role expectations, emotional influence and long-term embodied interaction?
- Did you establish a process to discuss and continuously monitor and assess the AI system's adherence to this Assessment List for Trustworthy AI?
- Did you establish a process for third parties (e.g. suppliers, end-users, subjects, distributors/vendors or workers) to report potential vulnerabilities, risks or biases in the AI system?
- For applications that can adversely affect individuals, have redress by design mechanisms been put in place?
- Have the humans (human-in-the-loop, human-on-the-loop, human-in-command) been given specific training on how to exercise oversight?
- Did you establish any detection and response mechanisms for undesirable adverse

Co-funded by the European Union

effects of the AI system for the end-user or subject?
- Did you ensure a 'stop button' or procedure to safely abort an operation when needed?
- Did you take any specific oversight and control measures to reflect the self-learning or autonomous nature of the AI system?

3.4.3.7 Worker Dignity and Equity

- How is it ensured that human workers are treated with dignity and respect in workplaces where robots operate, including through fair task allocation, respectful robot behaviours and avoiding dynamics where robots undermine workers' status or roles?
- Are there measures in place to address any form of discrimination or harassment in the workplace, particularly in scenarios involving human-robotic/AI system interactions?
- How does the organisation's compensation structure reflect the value of human workers' contribution, especially in comparison to the efficiencies gained through automation?
- In what ways are the unique contributions of human workers in technically integrated workplaces recognised?
- How are diverse workers' needs in settings that include both human and robotic/AI workers accommodated?
- How is the significance of human roles communicated within the organisation and in external communities?
- What mechanisms are in place for workers to provide feedback or raise concerns about their working conditions, particularly regarding their interactions with robotic/AI systems?
- How is training for new technologies, especially those involving collaboration with robotic/AI systems, made accessible and equitable for all employees?

3.4.3.8 Human Resilience Through Continuous Learning

- How are learning opportunities designed to complement and enhance human skills alongside advancements in robotics/AI systems?
- What strategies are in place to ensure that the skills development offered keeps pace with the evolution of technology in the workplace?
- What types of learning resources are available to workers to help them adapt to and collaborate effectively with robotic/AI systems?
- What measures are taken to support the psychological safety of workers interacting with robotic/AI systems?
- How are potential anxieties or concerns associated with increased automation and human-machine interactions addressed?
- How is feedback from employees about the effectiveness of the learning and support programs provided and used?
- How are learning and development opportunities integrated into workers' regular work schedules without causing disruption or undue stress?
- What career pathways are available for workers in a hybrid environment of humans and robotic/AI systems?

- How are career pathways communicated to employees, and how do they reflect the potential for long-term career growth despite technological changes?
- What long-term support systems are in place for employees to continue developing their skills over time?
- How is the culture of continuous learning and resilience among its workforce promoted?
- What incentives and rewards are provided to encourage ongoing personal or professional development?

4 Iterations and continuous learning

4.1 Improvements from the first iteration

The ARISE SSH framework has undergone significant evolution since its initial release in month 24 (December 2024). First, this second iteration is structured around more clearly delineated and articulated **roles** for both i**nternal and external stakeholders**. Recognizing that tangible outcomes require a clear, shared understanding of *who* is involved, *what* each actor is responsible for, and *how* actors are expected to carry out their responsibilities, this second iteration of the framework is built on actor-appropriate perspectives. One of the explicit aims of this second version is to **bridge theory-grounded, high-level principles with contextually-grounded, actionable goals.**

In the twelve months between the publications of the first and second versions, the ARISE SSH partners had the opportunity to further analyze and **incorporate contextual knowledge** from the ethnographic field visits and mentorship activities. This second iteration of the SSH framework more clearly acknowledges the critical challenges, points of friction, and underlying assumptions that the emerging HRI industry may encounter when applying the framework, while also highlighting the key opportunities and starting points the industry can leverage in its existing development workflows. In doing so, HRI projects can more effectively understand, adapt, and implement the framework to meet the requirements of Industry 5.0.

While the first version established the key foundational elements for ethical HRI, this second iteration focuses on making those elements **tangible and actionable**. The framework has been restructured to consolidate practical measures, drawing directly from the experience of deploying these concepts in real-world settings. The open call mentoring process was particularly pivotal, providing an opportunity to assess how ethical requirements resonate with SMEs and how they can be adapted to pre-existing industrial structures. This hands-on engagement allowed the **validation of previously theoretical challenges** and **tested concrete implementation solutions.** This evolution is made tangible in Section 3.3 (Practice: Models of Implementation), which introduces distinct operational models, including the EIA, the SSH Mentoring Model, and the SSH Toolkit. Additionally, the section outlines specific **implementation strategies** for TEFs and defines **metrics for measuring progress**. By moving toward this delineated structure, the framework becomes more readily usable for in-situ HRI development, translating theoretical principles into applied practice and actionable next steps.

4.2 Expected evolution

The development of the ARISE SSH Framework is an ongoing, iterative, and collaborative process. While the framework has evolved to better complement the needs and contexts of the HRI industry, there are still several key limitations. Some of the objectives have not been fully achieved and some elements need to be further developed, which will require a deeper involvement of the consortium partners.

The ongoing mentorship program with the open call beneficiaries has served as the primary testing ground for validating the initial SSH framework. While this engagement has yielded rich insights, the ARISE SSH partner has yet to formulate a systematic approach for evaluating impact within the current resource constraints. Although some positive feedback was received, the general reliability of the data remains limited by a small sample size and inconsistent gathering methods. Furthermore, this engagement has highlighted the critical need to align SSH methodologies with other areas of the project, such as business management. Such integration is justified not only by operational overlaps, such as stakeholder engagement, but also by the potential for financial advantages that require further investigation.

Additionally, the application of the framework at the TEF level requires further exploration. This version of the framework delineates the core challenges and proposes a preliminary operational approach (section 3.3.3), but successful implementation will require deeper collaboration with other consortium partners. Specifically, closer collaboration with WP4 (Set up, operation, and maintenance of testing and experimentation facilities) is essential to bridge the interdisciplinary gaps encountered during this phase. Furthermore, this collaboration will allow us to move beyond general project-level guidance and explore specific implementation at the technical 'middleware level'. Strengthening this link is critical to ensuring the framework does not remain isolated but becomes fully embedded in the wider TEF ecosystem.

In terms of the framework's internal evolution, further validation is needed to identify overlooked assumptions, misaligned recommendations, and pathways for better integrating the framework into HRI development workflows. Moving forward, the next iterations will explicitly address the limitations and opportunities listed in section 2.3.3 Insights Informing This Framework and will look into introducing a dedicated toolkit section on foresight methods. Future guidance will help beneficiaries incorporate learnings from their Ethics Action Plans into dynamic project roadmaps. Consequently, the next stages of mentorship will increasingly focus on preparing projects to mitigate emerging ethical concerns as their technology matures, placing a clearer emphasis on connecting Societal Readiness Levels (SRL) to Technology Readiness Levels (TRL). This practical connection allows for context-aware progression, ensuring ethical understanding advances appropriately alongside technical development.

Finally, significant effort will be directed toward operational usability and ecosystem integration. This development involves articulating pathways that better align with specific project settings, such as aligning business objectives with Industry 5.0 and human-centric aims, thereby demonstrating that ethical compliance supports internal success. Further investigation and iteration is needed to outline strategies for negotiating between social, societal, technological, and business agendas. To make these concepts accessible, navigation will be improved through a more user-friendly interface that integrates the written framework with the corresponding web tool. Ultimately, the framework's content will be consolidated into the ARISE Starter Kit (Task 4.5). This transition is essential for moving the framework from the open call experimentation phase into the permanent infrastructure of the TEFs and the broader ARISE HRI ecosystem.

Co-funded by
the European Union

4.3 Relevance for HRI governance and policy

The ARISE SSH framework supports current and emerging governance of human-robot interaction by facilitating the practical interpretation and implementation of existing regulatory and policy instruments instead of proposing new regulation. In the current state of European policy regulation, the framework provides structured guidance for translating high-level ethical and legal requirements into operational practices relevant to embodied and interactive robotics systems.

In particular, the framework can assist organisations in operationalising compliance with instruments such as the EU AI Act and data protection regulation. By assisting in the integration of ethical principles, lifecycle and technology readiness considerations and concrete practices, the framework can be used as an evaluation tool to assess the ethical state of HRI systems in relation to the current and anticipated policy processes (see section 3.2). This also enables policymakers to identify areas where existing regulatory guidance may require further clarification or contextual adaptation for robotic systems.

The framework is aligned with the Industry 5.0 perspective, which emphasises human-centricity, sustainability and resilience in industrial development. By focusing on ethical, social, and organisational dimensions alongside technical progress, it supports policy approaches that promote the responsible adoption of robotics in ways that enhance human capabilities, work practices, and organisational preparedness.

The connection of high-level ethical guidelines and practical design principles allows the ARISE SSH framework to serve as a vital bridge between ethical theory, policy objectives, and engineering practice. Its accessibility makes it particularly relevant for SMEs and robotics developers, supporting the consistent uptake of human-centric principles across the entire HRI development lifecycle. Future developments of the framework will continue making ethical considerations in HRI explicit, contributing to broader policy goals related to Trustworthy AI implementation and enhancing public trust in robotic technologies for developing a sustainable industry in line with European values.

References

Alaieri, F., & Vellino, A. (2016). Ethical Decision Making in Robots: Autonomy, Trust and Responsibility. Social Robotics, 159–168. https://doi.org/10.1007/978-3-319-47437-3_16

ARISE SSH. (2025). arise-ssh.info. https://arise-ssh.info/

Beck, K., Beedle, M., van Bennekum, A., Cockburn, A., Cunningham, W., Fowler, M., Grenning, J., Highsmith, J., Hunt, A., Jeffries, R., Kern, J., Marick, B., Martin, R. C., Mellor, S., Schwaber, K., Sutherland, J., & Thomas, D. (2001). Manifesto for agile software development. Agile Manifesto. https://agilemanifesto.org/

Brey, P., Jansen, P., Maas, J., Lundgren, B., & Resseguier, A. (2021). SIENNA D4.7: An ethical framework for the development and use of AI and robotics technologies (1.1). Zenodo. https://doi.org/10.5281/zenodo.7266848

Callari, T. C., Vecellio Segate, R., Hubbard, E.-M., Daly, A., & Lohse, N. (2024). An Ethical Framework for Human-robot Collaboration for the Future People-centric Manufacturing: a Collaborative Endeavour With European Subject-matter Experts in Ethics. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.4927452

Ethical impact assessment - SIENNA. (2024, February 15). Sienna-Project.eu. https://www.sienna-project.eu/w/si/about-sienna/eia

European Commission: Directorate-General for Research and Innovation. (2024). ERA industrial technologies roadmap on human-centric research and innovation for the manufacturing sector. Publications Office of the European Union. https://data.europa.eu/doi/10.2777/0266.

European Commission. (2019). Ethics guidelines for trustworthy AI | Shaping Europe's digital future. European Commission. https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai

European Commission. (2020, July 17). Assessment List for Trustworthy Artificial Intelligence (ALTAI) for self-assessment | Shaping Europe's digital future. Digital-Strategy.ec.europa.eu. https://digital-strategy.ec.europa.eu/en/library/assessment-list-trustworthy-artificial-intelligence-altai-self-assessment

European Commission. (2021). Ethics By Design and Ethics of Use Approaches for Artificial Intelligence. https://ec.europa.eu/info/funding-tenders/opportunities/docs/2021-2027/horizon/guidance/ethics-by-design-and-ethics-of-use-approaches-for-artificial-intelligence_he_en.pdf

European Commission. (2024). Industry 5.0. Research-And-Innovation.ec.europa.eu. https://research-and-innovation.ec.europa.eu/research-area/industrial-research-and-innovation/industry-50_en

European Commission (2014). Horizon Europe 2020 Work Programme 2014 - 2015 General Anexxes. https://ec.europa.eu/research/participants/data/ref/h2020/wp/2014_2015/annexes/h2020-wp1415-annex-g-trl_en.pdf

IEEE. (2019). Ethically Aligned Design: A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems. https://standards.ieee.org/wp-content/uploads/import/documents/other/ead_v2.pdf

Nahavandi, S. (2019). Industry 5.0—A Human-Centric Solution. Sustainability, 11(16), 4371. https://doi.org/10.3390/su11164371

Nasir, V., Hosseini, A., Binfield, L., Hasani, N., Ghotb, S., Diederichs, V., Hansen, E. (2025). Human-centric Industry 5.0 manufacturing: a multi-level framework from design to consumption within Society 5.0. International Journal of Sustainable Engineering, 18(1). https://doi.org/10.1080/19397038.2025.2551000

Padovano, A., Cardamone, M., Woschank, M., & Pacher, C. (2024). Exploring Human-Centricity in Industry 5.0: Empirical Insights from a Social Media Discourse. Procedia Computer Science, 232, 1859–1868. https://doi.org/10.1016/j.procs.2024.02.008

Reelertoolbox. (2020). Human-proximity model. REELER project. https://reelertoolbox.ab-acus.com/

Reijers, W., Young, M. T., & Coeckelbergh, M. (2025). Introduction to the Ethics of Emerging Technologies. In Palgrave philosophy today. https://doi.org/10.1007/978-3-031-85887-1

Šabanović, S. (2010). Robots in Society, Society in Robots. International Journal of Social Robotics, 2(4), 439–450. https://doi.org/10.1007/s12369-010-0066-7

Satori Project (2017). English version Ethics assessment for research and innovation -Part 2: Ethical impact assessment framework. https://satoriproject.eu/media/CWA17145-23d2017.pdf

Scrum.org. (2024). What is Scrum? Scrum.org. https://www.scrum.org/learning-series/what-is-scrum/

Societal Readiness Levels (SRL) defined according to Innovation Fund Denmark. (n.d.). https://innovationsfonden.dk/sites/default/files/2019-03/societal_readiness_levels_-_srl.pdf

Technology Readiness Level (TRL) | Aalto University. (2025, July 28). Aalto.fi. https://www.aalto.fi/en/services/technology-readiness-level-trl

What Is VSD :: Value Sensitive Design @ Khoury College. (n.d.). Vsd.ccs.neu.edu. https://vsd.ccs.neu.edu/introduction/what-is-vsd/

Wiarda, Giannelos, Schuerz, Reber, & Doorn. (2023). Ethics Framework and Guidelines for Participatory Processes in the Activities of Research Funding Organizations. Zenodo. https://doi.org/10.5281/zenodo.8089673

Appendix

Appendix 1: ARISE EIA template tables

Table 1.1: EIA accountability roles

EIA roles	Definition	Name(s), position	Contact details
Ethics point person	The designated lead responsible for facilitating the EIA process. Acting as a bridge, they coordinate input from relevant experts and manage the assessment timeline, rather than providing all technical answers themselves.		
Collaborator(s)	Subject matter experts from across the project who support the Point-Person by contributing specialized knowledge, providing data evidence, and validating the accuracy of the assessment responses		
External support	External experts responsible for providing independent guidance, mentoring, and validation to support the identification of ethical benefits and concerns.	*E.g. ARISE SSH experts, external Ethics Committee*	

Table 1.2: EIA Desirable impacts

Systemic level	Type of benefit	Primary beneficiary	Likelihood	Validation	Enabling actions
Example: Micro	*Reduction of physical stress. The robot takes over high-load lifting (20kg+)*	*Factory operators*	*Hypothesis, not yet tested in real users*	*Validation through ergonomic assessment e.g. RULA*	*Implementing "hand-guiding" features, Operators can adjust the drop-off height to their specific body size.*
Example: Meso	*Lowering skill entry-levels. No-code platforms allow operators without programming background to configure the robots*	*HR department, entry-level workers, workers without technology background*	*High, same interface validated in previous experiments*	*Comparison of the onboarding training time required*	*User Experience (UX) design to improve the interface tablet; visual, step-by-step tutorials.*
Example: Macro	*Energy efficiency, carbon footprint reduction when compared to*	*Environment, local community*	*Hypothesis: Estimated 15% reduction in energy usage.*	*Comparative energy measurement over 3 months*	*Lifecycle Assessment (LCA) monitoring, and transparent reporting of energy flows*

	previous machinery

Table 1.3: EIA Undesirable impacts

Systemic level	Type of risk	Affected group(s)	Likelihood	Existing measures	Additional mitigation	Post-mitigation status
Example: Micro	*Operators may over-rely on the robot's safety sensors, decreasing attention to the workspace and increasing the risk of missing non-robot hazards.*	*Factory operators*	*High: Observed and described in literature*	*Standard ISO 13849*	*Implementing "hand-guiding" features, Operators can adjust the drop-off height to their specific body size.*	*Acceptable: Risk is managed, but requires monthly monitoring by the safety officer*
Example: Meso	*De-skilling or loss of professional identity: Skilled welders feel replaced by a machine that "anyone can run." This leads to low morale or resistance to new machines.*	*Welding area, welders*	*Very high: Workers'unions have already voiced concerns*	*Not currently*	*Involve senior welders in robot skill design and learning*	*Monitor regularly: Requires close dialogue with union representatives*
Example: Macro	*Demographic exclusion. New HRI system requires high levels of digital fluency creating an entry barrier for local older workers*	*Local community*	*High: The region's demographic data shows 60% of the available workforce is over 45*	*Standard training*	*Partner with local vocational schools to provide courses targeted for older adults*	*Track and give visibility to the ratio of local hires of workers.*

Appendix 2: Ethics and Human Centricity Action Plan v2

Introduction

Beneficiary projects of the ARISE open calls are required to develop an Ethics and Human Centricity Action Plan—an ethics-specific document designed to guide the management of ethical issues throughout their participation in ARISE and beyond. The plan is intended to remain a live document, open to updates at any stage of the project. It serves as both a guiding tool for projects—providing a structured template to gather the necessary knowledge and plan actions that ensure compliance with ethics regulations and alignment with human-centric principles—and as a basis for assessment by the Ethics Mentors and the external Ethics Committee. The template for the Ethics Action Plan includes the following sections:

1. **Commitment to Excellence:** Projects are expected to establish a strong ethical foundation and clearly communicate this commitment internally and externally. This includes affirming a human-centric approach, reflecting on societal and environmental impacts, and adhering to the ARISE ethical principles.
2. **Stakeholder Engagement Plan**: Projects should outline strategies for involving stakeholders such as users, test participants, or industry associations. This includes methods for identifying and tracking stakeholder needs, creating a stakeholder map, and describing participatory practices.
3. **Compliance with Legal and Ethical Guidelines:** Projects must demonstrate compliance with relevant regulations and standards, including the EU AI Act and GDPR, complete the ALTAI assessment, and address other applicable rules.
4. **Social and Environmental Impact:** Projects should consider the societal and environmental impacts of their work by identifying and addressing both desirable and undesirable outcomes, and planning developments accordingly. This effort is supported by the process of completing the ARISE Ethics Impact Assessment (EIA), which includes a review by the external committee.

Note: The Ethics & Human-Centricity Action Plan is a comprehensive, iterative document. Instead of being completed all at once, it is designed to be addressed and reviewed at different stages of the mentoring programme.

Section 1: Commitment of Excellence

1.1 Commitment to a human-centric approach

Guiding questions:

- *How does your team define human-centricity in the context of HRI?*
- *How does your project commit to human-centric development of HRI (e.g. putting people before machines, recognising stakeholder's needs, considering cross-cultural perspectives, supporting gender balance, etc) ?*

1.2 Commitment to ARISE principles

Ethical principles are widely recognized as important tools to guide responsible technological development. Their aim is to guide the ethical requirements in order to address concerns and align to societal values. The ARISE project has defined eight ethical principles. The first six are adapted directly from the Ethics by Design framework published by the European Commission in 2021, reflecting established industry and policy standards. The final two have been specifically developed to address the unique HRI challenges in ARISE.

ARISE principle	How does your project align and commit to this principle?
Respect for Human Agency	
Privacy and Data Governance	
Fairness	
Individual, Social and Environmental Well-being	
Transparency	
Accountability and Oversight	
Worker Dignity and Equity	
Human Resilience Through Continuous Learning and Support	

Section 2: Stakeholder Engagement Plan

2.1 Diversity and Inclusion Mapping

Describe how the PoC addresses user needs and diversity, gender balance and inclusiveness and cultural sensitivity.

2.2 Stakeholders

Identify and list the key stakeholders who are affected by, involved in, or can influence your project (e.g. intended users, affected communities, working unions, partners, policymakers, etc). For each stakeholder group, briefly describe their role, and potential interest, need or concerns related to your technology.

2.3 User Needs

1) How are you collecting (or planning to collect) the needs and feedback of intended users or other relevant stakeholders outside your team? Please describe the

methods you use (e.g., co-creation workshops with operators, testing solutions in real industrial contexts).

2) How is the input you receive from these users and stakeholders influencing your design and development decisions? Please explain how you plan to adapt your approach if your understanding of user needs evolves throughout the process.

2.4 Participation practices

Do you have plans to involve human participants in testing or validating your solution? If so:

1) Describe your method for selecting and recruiting human participants
2) Describe the activities the participants will be involved in
3) Describe how you will ensure the health and safety of the participants involved in the experiment.
4) List out the information that will be collected from each participant, including all types of feedback and data collected during the experiment, as well as any personal data.
5) Will data collected from your participants be anonymized? If so, describe your anonymization process.
6) What information does your participants need to be made aware of before participating in the experiments?
7) Describe your procedures for ensuring informed consent of participation.

If participants are not involved in the testing or validating of your solution, describe how your team plans to ensure your solution's human readiness and centricity during the development process. relevant

Section 3: Compliance to Legal and Ethics Guidelines

3.1 EU Legal Compliance

3.1.1 AI Act:

1) How would you characterize your organization according to the EU AI Act? You can find definitions for each of the roles here. Note that you may select multiple types
    a) Provider
    b) Deployer
    c) Distributor
    d) Importer
    e) Authorised representative
    f) Product manufacturer

2) Please list out, in detail, all of the features of your solution that use AI. If possible, describe the type of AI algorithm being used, for instance "A generative AI model to generate text for chatbot feature". Read through this summary of the EU AI Act, and for each of the AI related features listed above, decide if it poses unacceptable risk (Chapter 2, article 5), high risk (Chapter 3), limited risk, or minimal risk.

AI Feature	Level of risk (According to EU AI Act)

3) Please fill out the compliance checker form, and list out your results, including:
   a) Transparency obligations
   b) AI literacy obligations
   c) Provider obligations
   d) Risk related obligations
   e) Indications that you're developing an AI technology that's prohibited by the AI Act

3.1.2 GDPR (General Data Protection Regulation):

1. In column 1, list out all of the features of your robotics solution that collects or processes data about a data subject (an identified or identifiable person)
2. In column 2, list out whether the following special categories of data is being collected or processed (racial or ethnic origin, sexual orientation, political opinions, religious or philosophical beliefs, trade-union membership, genetic, biometric or health data, personal data related to criminal offences)
3. In column 3, indicate whether personal data is being collected or processed. If so, indicate the types of personal data that's being collected or processed. (name/surname, home address, IP address, cookie ID, advertising identifier of phone, medical data that could be a symbol that uniquely identifies a person)

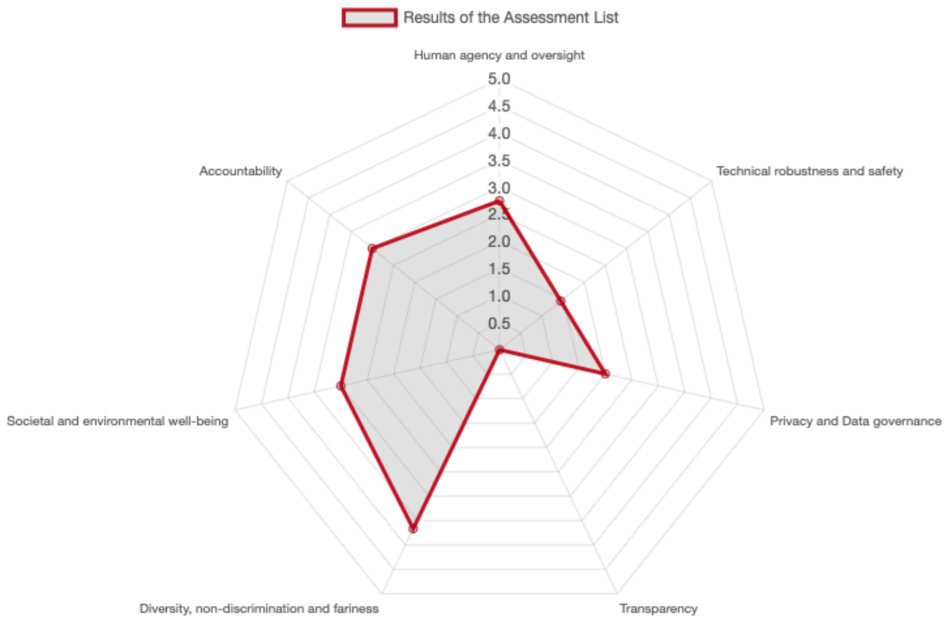
Feature that collects or processes data about a data subject	Special categories of data being collected	Personal data being collected

4. If your solution requires the collection of special categories of data, elaborate on what exception under the GDPR (Article 9) allows you to collect this data.
5. If your solution requires the collection of personal data, please elaborate on how you will ensure the rights of the data subject. For instance, you can provide a link to a template of an informed consent form, or explain what other basis you have under the GDPR to process personal data.
6. Read through the "Who monitors how personal data is processed within a company" section of this webpage, and explain whether a Data Protection Officer (DPO) is or isn't necessary.

3.2 Adherence to Ethical Guidelines and Standards

3.2.1 ALTAI:

1. Visit this website and create an account. Complete the self assessment for all 7 sections of the ALTAI. Each section contains several multiple choice and yes/no answers. Once completed, insert a screenshot of your self assessment results such as the one seen below.



2. Copy and paste all of the recommendations for each of the sections from the results of the self assessment into the chart below:

Principle	Recommendations
Human agency and oversight	
Technical Robustness and Safety	
Privacy and Data Governance	
Transparency	
Diversity, Non-Discrimination and Fairness	
Societal and Environmental Wellbeing	
Accountability	

3.2.2 Additional standards

Please list out any additional standards your team has used to guide the development of your AI solution (options: ethics by design and ethics of use approaches for AI, OECD AI Principles, ISO Overview of Trustworthiness in AI, other)

Section 4: Social and Environmental impact

1. Check the Ethics Impact Assessment (EIA) section of the SSH Framework for Human-Centric and Ethical HRI. Conduct the assessment drawing from your project's external ethics assessments, the ALTAI assessment and the project's own analysis. Include the EIA tables in the section below.

2. Describe briefly how your team will continue to recognize and address ethical concerns that emerge throughout the project.