Handling with AI-enhanced Robotic Technologies for flexible ManUfacturing

D1.3

Human-AI teaming interaction model

Deliverable ID:	D1.3
Project Acronym:	HARTU
Grant:	101092100
Call:	HORIZON-CL4-2022-TWIN-TRANSITION-01
Project Coordinator:	TEKNIKER
Work Package:	WP1
Deliverable Type:	Report
Responsible Partner:	DBL
Contributors:	All partners
Edition date:	18 December 2024
Version:	03
Status:	Final
Classification:	PU

This project has received funding from the European Union's Horizon Europe - Research and Innovation program under the grant agreement No 101092100. This report reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

HARTU Consortium

HARTU "Handling with AI-enhanced Robotic Technologies for flexible manufactUring" (Contract No. 101092100) is a collaborative project within the Horizon Europe – Research and Innovation program (HORIZON-CL4-2022-TWIN-TRANSITION-01-04). The consortium members are:

1	Tekniker	FUNDACION TEKNIKER (TEK) 20600 Gipuzkoa | Spain	Contact: Iñaki Maurtua inaki.maurtua@tekniker.es
2	dfki Deutsches Forschungszentrum für Künstliche Intelligenz German Research Center for Artificial Intelligence	DEUTSCHES FORSCHUNGSZENTRUM FUEF KUENSTLICHE INTELLIGENZ GMBH (DFKI) 67663 Kaiserslautern | Germany	Contact: Vinzenz Bargsten vinzenz.bargsten@dfki.de
3	aimen TECHNOLOGY CENTRE	ASOCIACIÓN DE INVESTIGACIÓN METALÚRGICA DEL NOROESTE (AIMEN) 36418 Pontevedra| Spain	Contact: Jawad Masood jawad.masood@aimen.es
4	eng	ENGINEERING INGEGNERIA INFORMATICA S.P.A. (ENG) 00144 Rome| Italy	Contact: Riccardo Zanetti riccardo.zanetti@eng.it
5	TOFAŞ TÜRK OTOMOBİL FABRİKASI A.Ş.	TOFAS TURK OTOMOBIL FABRIKASI ANONIM SIRKETI (TOFAS) 34394 Istanbul | Turkey	Contact: Nuri Ertekin nuri.ertekin@tofas.com.tr
6	PHILIPS	PHILIPS CONSUMER LIFESTYLE BV (PCL) 5656 AG Eindhoven | Netherlands	Contact: Erik Koehorst erik.koehorst@philips.com
7	ULMA	ULMA MANUTENCION S. COOP. (ULMA) 20560 Gipuzkoa | Spain	Contact: Leire Zubia lzubia@ulmahandling.com
8	deepblue	DEEP BLUE Srl (DBL) 00193 ROME | Italy	Contact: Erica Vannucci erica.vannucci@dblue.it
9	FMI ImProvia	FMI HTS DRACHTEN B.V. (FMI) NL-4622 RD Bergen Op Zoom, Netherlands	Contact: Floris goet floris.goet@fmi-improvia.com
10	TECNOALIMENTI	TECNOALIMENTI S.C.p.A (TCA) 20124 Milano | Italy	Contact: Marianna Faraldi m.faraldi@tecnoalimenti.com
11	Dipartimento Meccanica Matematica Management MUR Dipartimento di Eccellenza 2018-2022 2023-2027	POLITECNICO DI BARI (POLIBA) 70126 Bari | Italy	Contact: Giuseppe Carbone giuseppe.carbone@poliba.it
12	OMNIGRASP	OMNIGRASP S.r.l. (OMNI) 70124 Bari | Italy	Contact: Vito Cacucciolo vito.cacucciolo@omnigrasp.com
13	ITRI Industrial Technology Research Institute	INDUSTRIAL TECHNOLOGY RESEARCH INSTITUTE INCORPORATED (ITRI) 310401 Hsinchu | Taiwan	Contact: Curtis Kuan curtis.kuan@itri.org.tw

Funded by the European Union

14	INFAR INFAR INDUSTRIAL CO., LTD.	INFAR INDUSTRIAL Co., Ltd (INFAR) 504 Chang-hua County | Taiwan	Contact: Simon Chen simon@infar.com.tw

Funded by the European Union

Document history

Date	Version	Status	Author	Description
[11/10/2024]	01	Draft	DBL	Introduction, state of the art, methods, model description
[04/12/2024]	02	Draft	DBL	Model application of an HARTU UC
[18/12/2024]	03	Final	DBL	Reviewed

Funded by the European Union

Executive Summary

This document describes the development of the **Human-AI Teaming Interaction Model (HATIM).**

HATIM is a descriptive model of Human-AI Interaction, developed to support the creation of teams paring Human and Autonomy (AI agents able to operate autonomously) in manufacturing.

The model is composed of a selection of variables that, according to the literature, can describe the Human Operator, the Autonomous Agent and the resulting Team. These have been selected and organised in macro blocs of Operator, AI-Agent, Mediator, Environment and Task, to describe the connection between them and how they affect teaming. A particular innovation of the model is a novel approach to the definition and description of the human and AI roles within the team, based on concepts from narrative and storytelling.

The document includes the theoretical basis of the model in the form of a literature review, a description of the model and its components, a guide on the application of the model as part of a User Centred Design effort, and a trial application of the model in the context of the HARTU project.

Funded by the European Union

Funded by
the European Union

Table of contents

Funded by the European Union

List of figures

Funded by the European Union

List of tables

Acronyms

List of the acronyms	
HARTU	Handling with AI-enhanced Robotic Technologies for flexible manufactUring
HATs	*Human–Autonomy Teams*
HATIM	Human-Autonomy Teaming Interaction Model
SA	Situational awareness
SSA	Shared situational awareness
SSM	Shared Mental Model
WL	Workload

Funded by the European Union

1. Introduction

1.1. Objective and scope of the document

As the manufacturing industry increasingly integrates Artificial Intelligence (AI) and advanced automation into its processes, the concept of Human-AI teaming has emerged as a critical area of study. This integration goes beyond mere automation, fostering a collaborative partnership between human operators and AI systems to achieve shared objectives. In this context, understanding the foundational elements that contribute to effective Human-AI teaming is essential. "*The potential of autonomous agents working with humans opens up an interesting question, which involves both articulating a clear definition of Human–Autonomy Teams (HATs) as well as identifying the factors that make these teams successful"* (O'Neill et al., 2022).

This document seeks to explore the state of the art regarding research in Human-AI teaming through a state-of-the-art review of available literature and understand the work and process needed to successfully design and implement teaming in manufacturing. A Human-AI Teaming Interaction Model (HATIM) was developed, informed by a comprehensive literature review and DBL's deep understanding of user needs and the human role within the envisioned TO-BE scenarios. Partners' inputs were collected through focus group sessions to ensure that the model would respond to the internal acceptability criteria and needs for the development of the TO-BE scenarios.

Funded by the European Union

2. State of the Art

2.1. Introduction

Rapid advancement in AI technology creates the novel opportunity to pair human operators with machines that have unprecedented levels of Autonomy and Agency, more akin to other human workers than traditional automation. This necessitates a transition from the concept of human operators using tools, albeit very advanced ones, to human operators engaging with Agents in a collaborative effort, solving problems as a team.

This document presents a collection of factors that, according to the literature, should contribute to the successful design and implementation of Human-Autonomy teams. It is worth noting that, here, the term "Human-Autonomy teaming" is used as a generic term to address the collaboration between humans and any AI-enabled system, be them physical or virtual. Therefore, it also encompasses other similar terms such as human-robot teaming, Human-AI teaming, human-agent teaming, and human-machine teaming.

Human-Autonomy Teaming (HAT) is a growing field that offers great potential in various domains, including industrial processes, healthcare, decision-making, and more. Industries can leverage HAT to optimise complex production workflows, streamline logistics, and minimise errors, leading to improved productivity and reduced costs. The synergy of human expertise and machine capabilities enables real-time decision-making and adaptive responses in dynamic manufacturing environments, ultimately fostering innovation and competitiveness within the industrial sector.

However, the successful design and implementation of HAT require careful consideration. In the existing body of research, several factors have been introduced as the main challenges to be faced for implementing effective Human-Autonomy Teams. Hagos & Rawat (2022) mention Heterogeneity, Communication, Coordination, and Adaptability as the challenges associated with the current HAT approaches. The National Academies of Sciences, Engineering, and Medicine (2022) believes the core characteristics of effective Human-Machine teams are Team Heterogeneity, Shared Cognition, Communication and Coordination, and Social Intelligence. Transparency (Akash et al., 2019), Shared Awareness (Forster et al., 2016), and Trust and Cooperation (Zanatto, 2019) are also among the primary problems to address for improving HAT.

2.1.1. Defining Human–Autonomy Teaming

The concept of Human–Autonomy Teaming (HAT) is not new and originated in the 1990s, but the term has recently been adopted. To define HAT, it is important to consider the difference between Human–Autonomy Teams and Human–Automation Teams, the latter have a degree of interdependence among humans and automation, but agents are not autonomous. To qualify as a HAT, certain criteria automation and teamwork must be met, and these criteria represent boundaries for our treatment of what constitutes a HAT versus what does not. Teams involve two or more members working interdependently toward a common goal, with interdependence in relation to both task activities and achieving outcomes (Campion et al., 1993; Salas et al., 1992).

Furthermore, to be considered a HAT the Agent must have a sufficient level of autonomy. As mentioned above, the transition from automation to autonomy is weak and in constant evolution, still it is possible to refer to the Level of Autonomy scale proposed by Parasuraman et al. (2000) to find a solid starting point.

Funded by the European Union

Partial levels of agent autonomy are present at level 5 and above (spontaneously recommend and execute actions unless they are vetoed) while high levels of agent autonomy can be seen above level 7 (require no human intervention).

Literature on human-human teaming argues that “real teams” are bounded; that is, those involved understand which members are on the team and which members are not (O’Neill et al., 2022). Therefore, autonomous agents must be divided into, and recognized by humans, as distinct entities that represent unique roles on the team that would otherwise have to be filled by a human in the role. If they are not recognized by humans as team members, there is no HAT (O’Neill et al., 2022).

To summarise, according to O’Neill et al. (2022) HAT can be defined as “interdependence in activity and outcomes involving one or more humans and one or more autonomous agents, wherein each human and autonomous agent is recognized as a unique team member occupying a distinct role on the team, and in which the members strive to achieve a common goal as a collective.” The autonomy aspect of HAT implies an autonomous agent, therefore different from automation.

2.1.2. Defining the concept of Roles in HATs

Vladimir Propp’s narrative theory, articulated in his seminal work *Morphology of the Folktale*, identifies archetypal character roles based on their functional contributions to a story (Propp et al., 1968). Propp categorizes these roles into distinct types, including heroes, helpers, and antagonists, emphasizing that “the roles of characters may be described in terms of their functions in the course of the action”. This perspective is particularly useful for understanding the interaction dynamics in HAT, where both human operators and Autonomous agent adopt complementary roles to achieve complex tasks within collaborative environments. In the context of HAT, autonomous agent can embody various roles, such as the "Helper," which performs supportive functions like task automation and data analysis, or the "Supervisor," which oversees quality and compliance by providing real-time feedback. Other roles, such as the "Optimizer," resonate with Propp's concept of strategic assistance, focusing on enhancing operational efficiency and suggesting improvements to workflows. As Propp noted, "the actual means of the realization of functions can vary... but the function, as such, is a constant" (Propp et al., 1968). This underscores the importance of defining roles based on their functional contributions rather than their specific identities.

Conversely, human operators within HATs may assume roles such as "Collaborator" when engaging closely with AI systems or "Supervisor" when maintaining overarching strategic oversight of the collaborative process. Propp’s assertion that “the sequence of functions is always identical” (Propp et al., 1968) highlights the importance of these dynamic roles, as they reflect the adaptable nature of task allocation in HATs. The application of Propp's framework to define roles in HATs contributes to developing a structured approach for adaptive task sharing (ATS) models, which aim to enhance productivity and flexibility in manufacturing settings (Schmidbauer et al., 2021). These models propose that the distribution of tasks between human and AI agents should be fluid and responsive to real-time conditions, thereby fostering human agency and improving cognitive ergonomics. Roles within HATs are characterized by a shared understanding of competencies and responsibilities, rather than being statically assigned. This collaborative framework enables dynamic task selection based on situational needs and preferences, thus promoting transparency and flexibility in production processes.

Funded by the European Union

2.2. Methods

The literature has been selected through keyword queries, encompassing topics such as artificial intelligence and human-machine interaction. The search efforts concentrated on several key domains, including human factors, artificial intelligence, and computer science, while also encompassing related areas like Human-Robotic Interaction, Human-Robotic Teaming, Human-Automation Teaming, Human-autonomy teaming, Autonomous Systems, and Hybrid Intelligent Systems. Emphasis was placed on sources published within the past five years, although valuable works predating 2018 were also considered for inclusion.

2.3. Key Factors, Challenges, and Concerns

2.3.1. Trust

The successful exploitation of HAT is majorly dependent on the trust between humans and AI-enabled systems (Hagos & Rawat, 2022). Trust serves as the cornerstone for facilitating interaction between humans and machines (Pinto et al., 2022). Trust is a multifaceted concept, it is the inner belief that the AI will act in the user's best interest, mitigating risk throughout. It is a vital element for the acceptance and utilisation of automation, as highlighted by various studies (Hancock et al., 2011; Hoff & Bashir, 2015; Malle & Ullman, 2021). Its definition includes inherent trust tendencies (Dzindolet et al., 2003) and trust that's adjusted over time (Ezer et al., 2019; Lee & See, 2004). Adjusting or evolving trust is related to experience and it’s built over time as the user internalises notions on when AI can be dependable and when it cannot (McDermott et al., 2018). A mismatch between trust and objective performance can lead to either over-reliance or avoidance (Parasuraman & Riley, 1997; Robinette et al., 2016). Lee & See (2004) emphasise how evolving trust is shaped by continuous evaluation of an entity's actions, methods, and ultimate goals. In this context, it doesn't matter if the entity is a human or a machine; the user deduces intent, and method based on visible actions and creates an internal model.

There are various dimensions that contribute to the formation of trust in the HAT context. Lee & See (2004) state Utility, Predictability, and Intent as the main factors to be considered for establishing trust. In a manufacturing context, utility refers to the machine's ability to perform tasks efficiently, predictability pertains to the consistency of machine performance, and intent relates to the machine's programmed objectives. Pinto et al. (2022) suggest four dimensions that affect human trust when interacting with an intelligent agent: Risk Perception, Competence, Benevolence, and Reciprocity. Risk Perception refers to an individual's assessment of the potential dangers or uncertainties associated with relying on an intelligent agent. In the context of HAT, risk perception might involve evaluating the likelihood of a machine malfunctioning, making errors, or not performing as expected, as well as weighing the potential negative outcomes and their consequences. Competency pertains to the perceived ability and skill of the intelligent agent to effectively and efficiently perform its designated tasks. If a machine is perceived as competent, users are more likely to trust it to perform its tasks without constant supervision or intervention. Benevolence relates to the perceived intention of the machine to act in the best interests of the user. It's the belief that the system is designed and operates in a manner that prioritises the user's needs and safety, without causing harm or danger. Reciprocity denotes the belief that if a user invests time, effort, or trust in a machine, the machine will "reciprocate" by performing its tasks reliably and effectively. It's the expectation of a mutual, beneficial relationship between the user and the machine, where both parties contribute to the success of the endeavour.

Funded by the European Union

The National Academies of Sciences, Engineering, and Medicine (2022) highlights the importance of Social Interactions, e.g., reputation, and the formal/informal communication that can affect that reputation, for instance, gossip. Chiou & Lee (2023) illustrated in a model (Figure 1) the social decision situations that are ingrained in a goal environment. They demonstrate how trust is created by iterative interactions between a human agent (AH) and an automated agent (AA) and the outcomes of those interactions over a span of time.

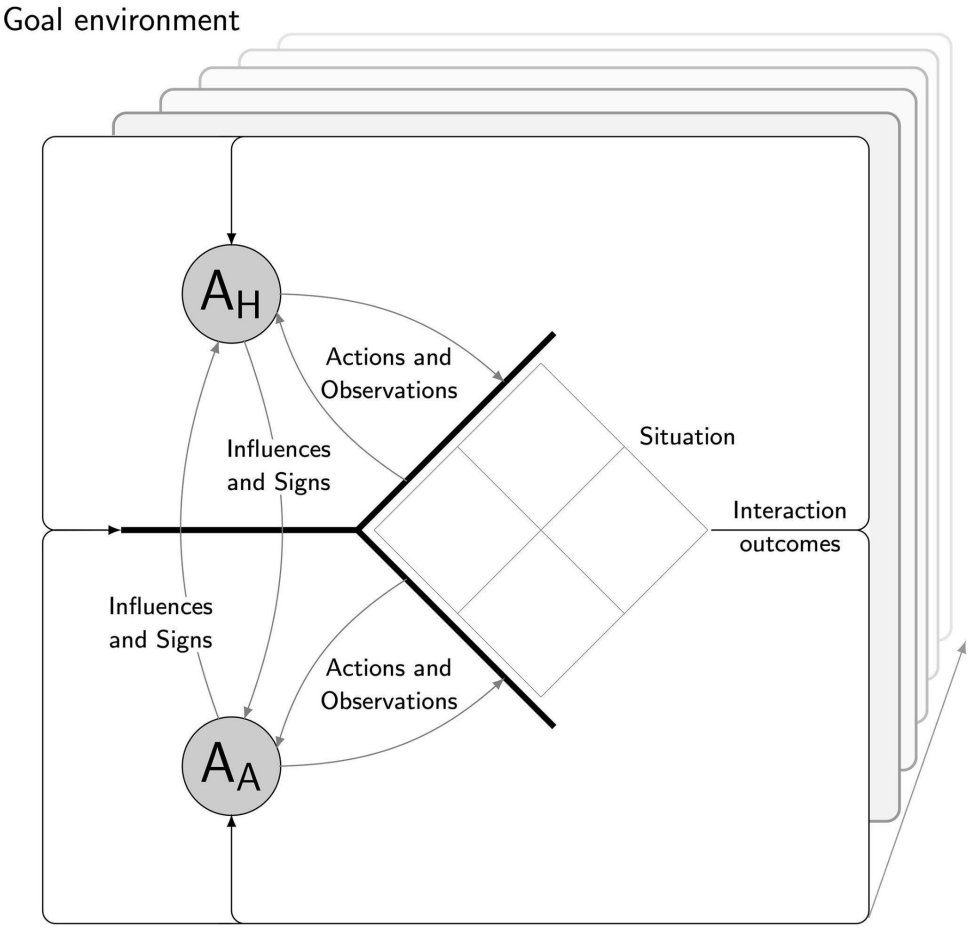


Figure 1 – The model illustrates the development of trust in a goal environment through ongoing interactions between a human agent (AH) and an automated agent (AA), as well as the results of these interactions over a period of time (Chiou & Lee, 2023).

The collaboration between humans and intelligent machines in a team may impose risks on the operator or the system itself, and as mentioned before, the perception of such risk can influence trust in HAT. To overcome such fear and create trust, (Gustavsson et al., 2022) suggest considering three key factors: knowledge, control, and self-preservation. To overcome a lack of knowledge, they advise visualising machine's tasks, illustrating human's tasks, and demonstrating safety systems in place. To overcome the lack of control, they propose giving the operators proper training and simulations as well as providing them with the possibility to determine, modify and stop the machine's activities. To calm the sense of self-preservation, they recommend pacifying the basic instincts by reinforcing desirable traits and avoiding intimidating implications, plus, performing risk assessment and risk visualisation to inform the operators about how to avoid/overcome the potential risks.

Funded by the European Union

Table 1. Summary of key findings on the role of trust in HAT.

Categories	*Key Findings*
Nature of Trust	• Multifaceted concept in automation acceptance • People have inherent trust tendencies • Trust adjusts over time • Misjudging trust can lead to over-reliance or avoidance
Role of Trust in HAT	• Essential for the acceptance and utilisation of automation in human-machine teams. • Determines the likelihood of system usage: low trust can decrease usage, while excessive trust can lead to unquestioning approval of system operations. • Serves as the foundation for interaction between humans and AI • Directly impacts the efficiency and effectiveness of human-machine collaboration. • Knowledge: Visualise machine tasks, illustrate human tasks, and demonstrate safety systems in place. • Control: Provide operators with proper training and simulations. Allow operators to determine, modify, and halt machine activities. • Self-Preservation: Reinforce desirable traits and avoid intimidating implications. • Conduct risk assessments and visualise risks to inform operators on potential dangers.
Factors influencing Trust in HAT	• Utility, Predictability, Intent • Risk Perception, Competency, Benevolence, Reciprocity • Social interactions, like reputation • Iterative interactions between human and automated agents • Familiarity: Becoming familiar with a complex system allows you to know whether it can be trusted or not, whether it is an autonomous system or not.

Funded by the European Union

2.3.2. Anthropomorphism & Design Features

The appearance and characteristics of intelligent machines is an important factor that contributes to the quality of interaction between humans and machines and eventually influence the effectiveness and performance of HM teams. There is, in fact, an inherent human inclination to anthropomorphize various elements, from animals and planets to geometric shapes (Burghardt, 2017; Eddy et al., 1993; Guthrie, 1995; Milstein, 2011). This tendency has been harnessed in marketing products (Portal et al., 2018; Tuškej & Podnar, 2018) and in designing technology, especially robots (Duffy, 2003)

To elucidate the anthropomorphisation of non-human entities, Epley et al. (2007) introduced a Three-Factor Theory. According to this theory, perceiving human qualities in non-humans arises from:

Elicited agent knowledge: Given that our understanding of non-human entities is limited compared to humans, anthropomorphic interpretations become a straightforward and efficient approach.

Effectance motivation: An inherent human desire to understand, control, and predict, amplifying our tendency to anthropomorphize.

Sociality motivation: The human need for social bonds and the drive to combat isolation compels us to view non-human entities as human companions.

In line with Epley's theory, research indicates that the more human-like a robot appears, the more positively it is received and collaborated with (Goetz et al., 2003). People are also more inclined to heed its guidance (Kiesler et al., 2008) and even express empathy towards it (Riek et al., 2009). Focusing on human-robot teams as an example, if the intended task is demanding, either physically or mentally, operators tend to respond more favourably to a robot with human-like qualities compared to one that is more machine-like (Panagou et al., 2024). Anthropomorphic characteristics are linked to roles in which robots are anticipated to engage in regular social interactions with humans. On the other hand, robots displaying machine-like attributes are better suited for contexts that prioritise specific tasks, like functioning as lab or workplace assistants, as well as military robots (Panagou et al., 2024).

Panagou et al. (2024) studied the effect of robot design features on human capabilities and the resulting impacts on operators in the workplace from different dimensions. The diagram in Figure 2 shows the result of their studies as to how the interconnection of different robot design features and the HF capabilities influences humans in various aspects.

Merely providing a robot with a human-like appearance may not suffice to enhance human-robot interactions, as observed by Kahn Jr et al. (2007). Research focusing on verbal communication has explored the role of accents and natural voices in increasing acceptance (Eyssel et al., 2012; Markowitz, 2017; Tamagawa et al., 2011). In contrast, investigations into non-verbal communication, as conducted by (Breazeal et al., 2005), have emphasised the significance of gestures and gaze. Studies have shown that robots engaging in gaze shifting during social interactions tend to be more engaging (Admoni & Scassellati, 2017), and robots exhibiting joint attention are perceived as more competent (Huang & Thomaz, 2010). Interestingly, even when robots make mistakes, it can increase the perception of their human-likeness (Mirnig et al., 2017).

Ragni et al. (2016) conducted a study comparing an erroneous robot with a flawless one during a reasoning task. Surprisingly, the erroneous robot elicited more positive emotions. Moreover, robots that display

Funded by the European Union

typical human behaviours, such as cheating, are also perceived as more human-like (Short et al., 2010). In their study, Short and colleagues examined three robot behaviours during a "rock-paper-scissors" game with a human: a control condition with a robot behaving normally, a verbal-cheat condition in which the robot lied about winning the game, and an action-cheat condition in which the robot not only lied but also changed its gesture to match the winning one. Interestingly, the verbal-cheat behaviour was interpreted as a mechanical malfunction, while the action-cheat condition was perceived as an intentional behaviour of the robot.

Anthropomorphism can also evolve with experience, as suggested by previous research (Fussell et al., 2008; Lemaignan et al., 2014). Fussell et al. (2008) found that participants rated a robot as possessing more traits, moods, and feelings after interacting with it compared to merely imagining one. Thus, interaction with a robot enhances its anthropomorphism. In proposing a formal model of anthropomorphism, Lemaignan et al. (2014) noted that the interaction context also plays a crucial role in shaping anthropomorphism dynamics. While humans can construct an 'anthropomorphic model' of the robot during interaction, the presence of unpredictable behaviours may lead to an increase in anthropomorphism.

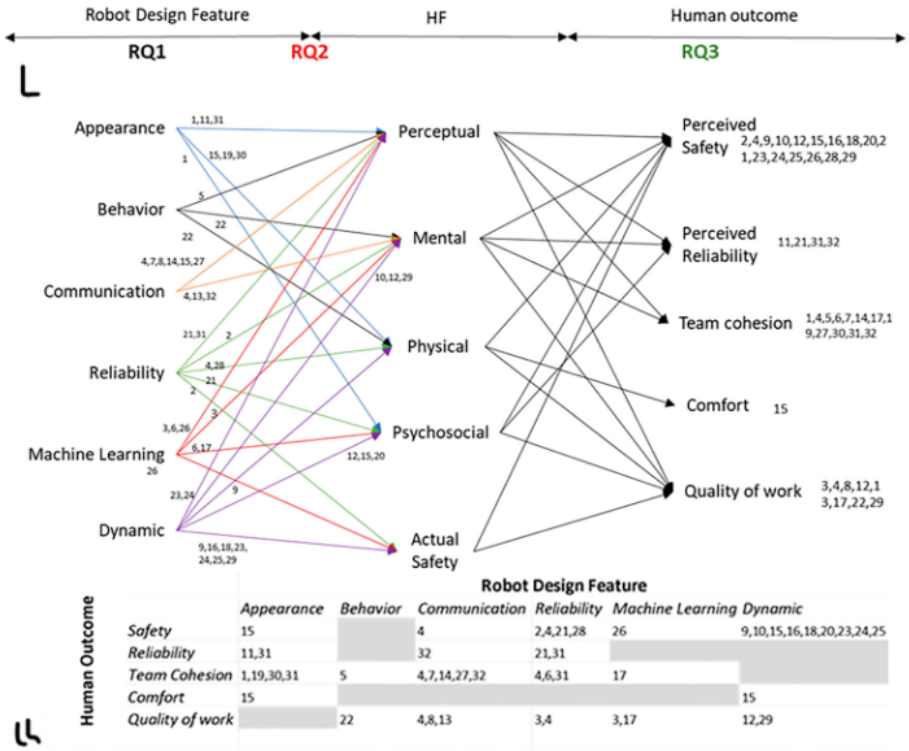


Figure 2 – How the interconnection of different robot design features and the HF capabilities influences humans in various aspects (Panagou et al., 2024).

Funded by the European Union

Table 2. Summary of key findings on the role of trust in HAT.

Categories	*Key Findings*
Nature of Anthropomorphism	• Humans have a natural inclination to anthropomorphize various elements, including animals, planets, and geometric shapes. • Anthropomorphism has been harnessed in marketing products and designing technology, particularly in the case of robots.
Role of Anthropomorphism in HAT	• Anthropomorphic characteristics in robots are suited for contexts involving regular social interactions, while machine-like attributes are preferred for task-oriented roles. • The influence of human-like appearance on the acceptance and collaboration with robots is a crucial aspect of anthropomorphism in MHT. • Verbal and non-verbal communication, including accents, natural voices, gestures, gaze, and joint attention, contribute to making robots more engaging and competent. • The context of interaction significantly influences the dynamics of anthropomorphism. The presence of unpredictable robot behaviors can lead to an increase in anthropomorphism, suggesting that humans construct an 'anthropomorphic model' of the robot during interaction. • Increased anthropomorphism can result in more favourable attitudes and interactions with robots, especially in roles that demand social collaboration.
Factors influencing Anthropomorphism in HAT	• Limited knowledge of non-human agents • Human desire for understanding and predictability • Need for social connection • The context of interaction significantly influences the dynamics of anthropomorphism. The presence of unpredictable robot behaviours can lead to an increase in anthropomorphism, suggesting that humans construct an 'anthropomorphic model' of the robot during interaction.

2.3.3. Agency and Autonomy

Agency is defined as the capacity and authority of an agent to act according to their own discretion and timing and is a crucial attribute of a team member (Lyons & Wynne, 2021; Schlosser, 2019). The concept of agency takes on a multifaceted nature in the context of HAT. It's not merely about executing predefined tasks but about making decisions, exercising judgment, and adapting actions within the broader framework of a collaborative team (Chen & Barnes, 2013). This is where the distinction between automation and autonomy becomes strikingly clear. Agency is the pivotal attribute that separates a machine merely following instructions from one capable of independent decision-making. Lyons et al. (2021) introduce agency as the defining attribute that separates automation from autonomy. Therefore, they argue, agency is one of the essential qualities that computational agents must possess in order to be recognised as teammates by humans. In a survey evaluating perspectives on autonomous weapons conducted by Arkin (2009), a significant majority of the respondents expressed that autonomous robots should possess the capability to decline orders that violate certain rules of engagement, emphasising the importance of agency

Funded by the European Union

in their design. Properly harnessed, agency can lead to more efficient problem-solving, enhanced productivity, and a harmonious human-machine partnership.

Parasuraman ty al. (2000) propose a scale of Level of Automation (LoA), varying on a continuum from the lowest level of fully manual/no automation to the highest level of full automation/ no human involvement.

HIGH	10. The computer decides everything, acts autonomously, ignoring the human.
	9. informs the human only if it, the computer, decides to
	8. informs the human only if asked, or
	7. executes automatically, then necessarily informs the human, and
	6. allows the human a restricted time to veto before automatic execution, or
	5. executes that suggestion if the human approves, or
	4. suggests one alternative
	3. narrows the selection down to a few, or
	2. The computer offers a complete set of decision/ action alternatives, or
LOW	1. The computer offers no assistance: humans must take all decisions and actions.

It is difficult to define on this continuum where the transition is between automation and autonomy, as well as how it evolves over time as today's automation becomes tomorrow's tool (O'Neill et al., 2022). What the literature suggests is that a sufficient level of autonomy is necessary for the agent to be recognised as a separate entity able to complete part of tasks anonymously, a foundational character of HAT (O'Neill et al., 2022).

Table 3. Summary of key findings on the role of agency in HAT.

Categories	*Key Findings*
Nature of Agency	• Involves the ability and authority of an agent to make decisions and adapt actions within a collaborative team • Is a distinguishing attribute that separates automation from autonomy
Role of Agency in HAT	• Shapes cooperation and trust between humans and non-human agents • Increases problem-solving and productivity • Leads to a more efficient HAT
Factors influencing Agency in HAT	• Collaboration levels • Ethical consideration • Level of agency granted to machines

2.3.4. Task Allocation & Interdependency

Task allocation is critical in production endeavours, particularly when designing human-machine teams (Ali et al., 2022). Forming an effective team goes beyond a simple division of functions based on whether humans or machines excel at each of them, balancing efforts and interdependencies should also be considered to foster a functional team. While heterogeneity and having a diverse team are important for its structure, what really matters for team process and functioning is the interdependencies. Having a certain

Funded by the European Union

level of overlap in roles and responsibilities can be also valuable for adaptive teams. This way, team members can support one another or take over tasks in the absence of a colleague (National Academies of Sciences, Engineering, and Medicine et al., 2022).

An essential concern that must be addressed in conversations about human-machine collaboration is understanding the extent and type of interdependence within a team of humans and machines. Human-agent teams can collaborate at different levels, from no coexistence to collaboration. Task allocation methods need to consider the level of collaboration for optimal task allocation and interdependence of such tasks (Ali et al., 2022). Interdependence in Human-autonomy teaming implies (Mosier et al., 1998) that task execution relies on the collaboration between operators and intelligent machines, blending the skills and expertise contributed by each and that operators and intelligent machines hold complementary responsibilities; nevertheless, according to team adaptability, these responsibilities might be rearranged to accommodate changing task requirements.

Engaging in interdependent collaboration implies that both humans and machines will take on roles and share teamwork across tasks that necessitate the participation of both parties. Apart from task allocation, the notion of interdependence also means a collective commitment to achieving tasks. In its most extreme form, shared responsibility extends to shared control; think of scenarios where a human and an AI-enabled agent team up to move an object, or when a passenger switches control of a self-driving car with the car's automated system (Lyons et al., 2021). Ethical considerations also play a crucial role in task allocation, especially in contexts where human lives are at risk (Ali et al., 2022).

Table 4. Summary of key findings on the role of task allocation in HAT.

Categories	*Key Findings*
Nature of Task Allocation	• Involves more than strengths • Interdependencies are crucial • Overlapping roles enhance adaptability • Human-Agent Teams benefit from diverse collaboration
Role of Task Allocation in HAT	• Shapes human-agent collaboration • Influences teamwork and adaptability • Effective task allocation optimizes performance • Interdependence fosters seamless collaboration • Extends to shared control
Factors influencing Task Allocation in HAT	• Ethical concerns and automation appropriateness • Human involvement and trust

2.3.5. Adaptability

The capacity to react adaptively to evolving task requirements and environmental circumstances is essential for successful teamwork, and this applies equally to the realm of Human-autonomy teaming (Mosier et al., 1998). Operators may need to adjust to the degree of support offered by intelligent machines based on their workload. Likewise, the machines should be capable of foreseeing and reacting to new knowledge and

Funded by the European Union

behaviour of operators, deducing their workload and proposing adjustments to their assistance level accordingly (Hagos & Rawat, 2022; Mosier et al., 1998).

These two aspects of adaptability are what Hagos & Rawat (2022) call “human-assisted adaptability”, and “machine-assisted adaptability”. Oppermann(1994) and Miller et al. (2005) also distinguish between these two types of adaptability. They use “adaptable automation” for the cases where the operator modifies the automation level, and assign “adaptive automation” to the cases where the system adjusts the automation level (National Academies of Sciences, Engineering, and Medicine et al., 2022).

Achieving adaptability in human-machine teams requires effective communication within the team and a shared comprehension of their circumstance, tasks, and teamwork (Mosier et al., 1998). The successful implementation of adaptability has been proven to lower workload, enhance operator involvement when manual control is active, boost human-system performance, and improve the retention of skills (National Academies of Sciences, Engineering, and Medicine et al., 2022).

Table 5. Summary of key findings on the role of adaptability in HAT.

Categories	*Key Findings*
Nature of Adaptability	• "Human-assisted adaptability" involves operator-driven adjustments. "Machine-assisted adaptability" involves system-driven adjustments. • Effective communication is crucial for adaptability.
Role of Adaptability in HAT	• Enhances teamwork • Lowers workload • Enhances operator involvement during manual control • Improves human-agent performance and skill retention.
Factors influencing Adaptability in HAT	• Technology advancements shape adaptability • Training and education impact adaptability • Team culture and communication practices influence adaptability • Task complexity affects the need for adaptability

2.3.6. Communication

Effective communication between operators and machines constitutes a fundamental aspect of successful teamwork and overall performance, as they can influence work processes and interdependence of team members (National Academies of Sciences, Engineering, and Medicine et al., 2022; Panagou et al., 2024). Lyons et al. (2021) consider communication a vital factor for team operation because it enables the development of shared situation awareness, shared mental models, and alignment of goals.

Communication can occur in verbal or non-verbal form, utilising a variety of methods such as spoken language, gestures, haptic signals, and screen-based interfaces (National Academies of Sciences, Engineering, and Medicine et al., 2022; Panagou et al., 2024). However, the use of natural language, due to its inherent ambiguities, might not always be the most suitable option for efficient communication in

Funded by the European Union

teamwork. For instance, in scenarios like the military and aviation, different methods of signalling and concise codes are employed (National Academies of Sciences, Engineering, and Medicine et al., 2022).

Because of their limited communication capacities, humans can only handle a finite volume of information. Thus, by sharing only crucial data, humans and machines can efficiently communicate information that enhances their collaborative teamwork (Hagos & Rawat, 2022).

Mosier et al. (1998) argue that communication in human-machine teams needs to be comparable to communication within human teams in terms of content, timing, and efficiency. Therefore, HATs need to ensure optimal communication richness, content relevance, and proper timing in their interactions (Lyons et al., 2021).

Proficient team members adjust their communication to suit the knowledge and information requirements of fellow teammates and the demands of the situation. This capability stems from their aptitude to effectively understand or "read" their teammates (Mosier et al., 2017). Hence, the actions and behaviour of machines should be closely aligned with relevant human expectations to achieve the highest level of effectiveness (Lyons et al., 2021). Machines need to initially replicate how humans process data in order to communicate information in a way that humans can easily grasp (Hagos & Rawat, 2022).

Table 6. Summary of key findings on the role of communication in HAT.

Categories	*Key Findings*
Nature of Communication	• Influencing work processes and interdependence • Occurs through verbal or non-verbal methods, but natural language may not always be the most efficient option, such as spoken language, gestures, haptic signals, and screen-based interfaces • Context may dictate the most suitable communication method.
Role of Communication in HAT	• Vital for shared situation awareness, shared mental models, and goal alignment • Machines need to first familiarise with human data processing to align with human expectations
Factors influencing Communication in HAT	• Communication methods • The amount, context, and timing of information • Explainability best practices

2.3.7. Coordination

Coordination is defined as the organisation and management of members in a team as well as their behaviours and actions in order to accomplish a shared objective (Hagos & Rawat, 2022). Communication in HATs must be coordinated, which essentially means it should be precise and appropriately targeted to the correct team member at the correct time. Successful teamwork involves skilfully arranging the order and timing of interdependent actions (National Academies of Sciences, Engineering, and Medicine et al., 2022). Identifying "the correct team member" and "the correct time" can be nuanced and might only become evident with substantial experience (Demir et al., 2018).

Funded by the European Union

Successful coordination between humans and machines necessitates fulfilling three core prerequisites: establishing common ground, fostering inter predictability, and enabling directability (Hagos & Rawat, 2022). Common ground denotes information that is shared and mutually accepted by all individuals engaged in a conversation. Inter Predictability refers to the capability of team members to reasonably anticipate each other's actions and behaviours. Directability means the capacity of team members to guide, assist, or impact each other's behaviours when there's a sudden change in circumstances or priorities (Hagos & Rawat, 2022).

At its ultimate level of coordination, a team may also achieve implicit coordination, which is characterised as the procedure of harmonising the actions and behaviour of team members by inferring their assumptions and intentions, without relying on explicit behavioural communication (Hagos & Rawat, 2022). Implicit coordination serves to enhance team effectiveness and substantially decrease the need for communication.

Table 7. Summary of key findings on the role of coordination in HAT.

Categories	*Key Findings*
Nature of Coordination	• Involves organizing team members, their behaviours, and actions to achieve a shared objective.
Role of coordination in HAT	• Affects communication, task allocation and adaptability • Enhance team effectiveness
Factors influencing coordination in HAT	• establishing common ground • fostering interpredictability • enabling directability • Experience and training

2.3.8. Shared Mental Model & Situation Awareness

A fundamental requirement for achieving effective team performance is that team members share consistent models of their tasks, teamwork processes, and the operational environment (Mosier et al., 1998). This concept is commonly called a "shared mental model". According to the National Academies of Sciences, Engineering, and Medicine (2022), a shared mental model refers to a uniform comprehension and representation of how systems operate, shared among team members. This encompasses models of technology, equipment, taskwork, teamwork, and even insights into teammates, including their skills, knowledge, attitudes, and preferences.

Research shows that human teams that hold shared mental models are more capable of forming a common understanding of dynamic situations. Possessing such capabilities enables HATs to proactively anticipate team requirements, actions, and forthcoming challenges. Consequently, this improves mutual adaptation and enhances the level of team coordination and synchronisation (Lyons et al., 2021). Moreover, the resemblance and precision of mental models held by team members related to the team's functioning as

Funded by the European Union

well as specific tasks, directly adds to the effectiveness of team processes, which in turn enhances the overall performance of the team (National Academies of Sciences, Engineering, and Medicine et al., 2022).

The presence of shared mental models within teams supports the establishment of shared situation awareness (National Academies of Sciences, Engineering, and Medicine et al., 2022). Situation awareness (SA) can be defined as "the perception of the elements in the environment within a volume of time and space [level 1 SA], the comprehension of their meaning [level 2 SA], and the projection of their status in the near future [level 3 SA]" (Endsley, 1988, p. 98).

It is widely understood that for effective interaction between humans and AI-enabled machines, having a strong situation awareness (SA) is essential. This includes comprehending both the current and anticipated performance, status, and information held by each entity (National Academies of Sciences, Engineering, and Medicine et al., 2022). In HATs, both humans and intelligent machines should develop internal situational awareness (Figure 3) of the world, themselves, and other team members, which will be updated continuously within the framework of more persistent overarching mental models.

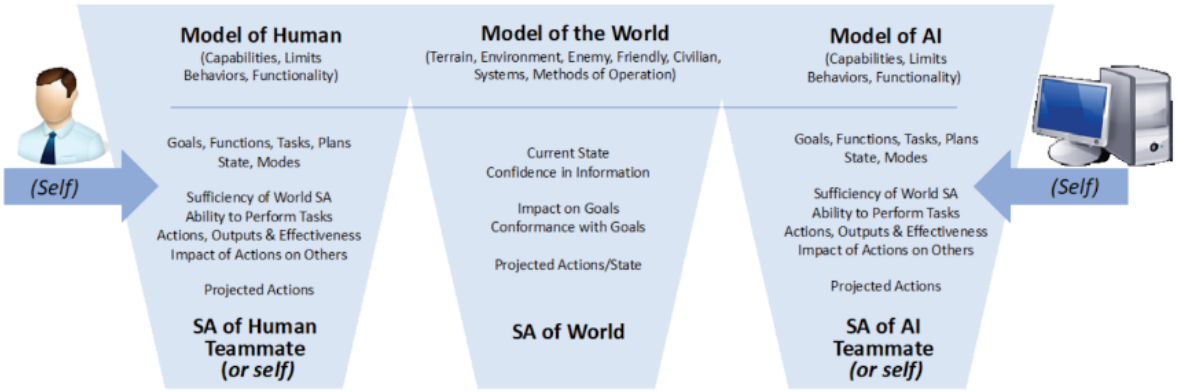


Figure 3 – Mental models and situation awareness of humans and AI-enabled machines (National Academies of Sciences, Engineering, and Medicine et al., 2022).

The mental models that are essential to the development of shared situation awareness between humans and intelligent machines consist of Situation, Task environment, Teammate awareness, and Self-awareness (National Academies of Sciences, Engineering, and Medicine et al., 2022). Situation refers to the mental model of the surrounding world and the environmental factors. Task environment refers to the goals, plans, functional assignments, task allocation, and task status. Teammate awareness refers to the understanding of the knowledge, skills, capability, status, and reliability of the other team member. Self-awareness refers to the understanding of one's own expertise, capabilities, limitations, and status for performing their assigned tasks.

Upholding the self-awareness mental model, people can identify the impact of fatigue, excessive workload, or inadequate training on their performance and decide to shift some tasks to intelligent machines in order to optimise the performance of the team. In a similar way, if AI-enabled machines maintain a mental model of their capabilities and limitations, they can ask humans to intervene when necessary or to assign confidence levels to their outputs (National Academies of Sciences, Engineering, and Medicine et al., 2022).

Funded by the European Union

There are some important concepts related to SA, including “team SA” and “shared SA”. Team SA can be defined as “the degree to which every team member has the SA required for his or her responsibilities” (Endsley, 1995, p. 39). Shared SA is defined as “the degree to which team members possess the same SA on shared SA requirements'' (Endsley & Jones, 2001, p. 48).

In human-machine teams, it is important to tailor the information displayed to the SA needs of each role in order to mitigate the overload. Moreover, to facilitate team coordination, the displays should support team SA by offering visibility into the relevant SA of other team members.

Table 8. Summary of key findings on the role of shared mental models and situation awareness in HAT

Categories	*Key Findings*
Nature of Shared Mental model and Situational Awareness	• A uniform comprehension of how systems operate, including technology, equipment, taskwork, teamwork, and insights into teammates
Role of Shared Mental model and Situational Awareness in HAT	• Enhance mutual adaptation, coordination, synchronisation, and overall team performance in HATs.
Factors influencing Shared Mental model and Situational Awareness in HAT	• Complexity of tasks • Diversity of team members • The need to comprehend current and anticipated performance, status, and information held by each team member

2.3.9. Transparency and Explainability

An effective interaction between humans and intelligent machines requires the machines to be sufficiently transparent (National Academies of Sciences, Engineering, and Medicine et al., 2022). According to Hagos and Rawat (2022) the success of Human-autonomy teaming depends majorly on the transparency of the system and on the user’s trust in the understandability and predictability of it. The concept of transparency encompasses various dimensions, including organisational transparency, process transparency, data transparency, algorithmic (logic) transparency, and decision transparency (National Academies of Sciences, Engineering, and Medicine et al., 2022). The dimension of transparency that enables the human to supervise and collaborate with the intelligent machine is called system transparency and is defined as “the understandability and predictability of the system”(Endsley et al., 2003, p. 146). System transparency consists of two interconnected aspects (Figure 4): Display transparency, which offers real-time visibility into the AI system's operations currently being carried out, contributing to situation awareness (SA); Explainability, which provides retrospective information regarding the system's actions or recommendations, clarifying the underlying logic, processes, factors, or reasoning (National Academies of Sciences, Engineering, and Medicine et al., 2022).

Funded by the European Union

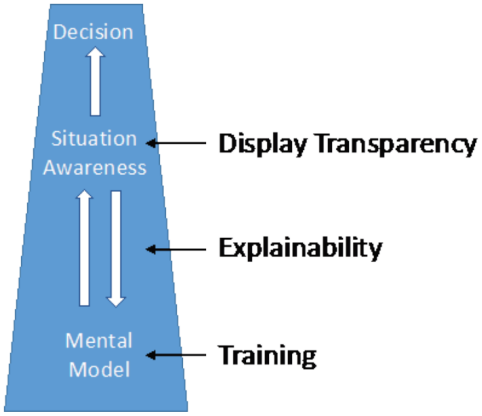


Figure 4 – The influence of transparency and explainability on mental models and situation awareness (National Academies of Sciences, Engineering, and Medicine et al., 2022).

National Academies of Sciences, Engineering, and Medicine (2022) conducted a literature review on the system functions that hold significance in ensuring system transparency, and arranged the results in a table based on the information type and the SA levels to which they correspond (Table 9). In this table, system status (level 1 SA) refers to the current state and real-time operations of the system as well as its purposes, plans, performance, and progress. It also includes information that reveals the limits and biases of the system. The notion of understandability refers to the reasons and logic behind the system's behaviour and recommendations. it also includes the extent of its capabilities and constraints as well as the reliability and confidence of the system outputs. Confidence in the outputs of the intelligent system is an important part of SA which occurs at multiple levels related to human decision-making (Endsley & Jones, 2001): (a) Data certainty; (b) Comprehension certainty; (c) Projection certainty; (d) Decision certainty.

The concept of predictability in the table refers to the planned behaviours and actions and their expected results or outcomes. it comprises confidence in the future projections and the system's ability to work in future situations in a way that is consistent to its usual performance. The last category of functions in the table, team and shared SA, refers to the transparency in the team tasks including task allocation, interdependence of tasks, the effect of ongoing tasks on other team members' status, etc.

According to Lee and See (2004) Providing explanations is influential in shaping different dimensions of trust: (i) Affective trust, which is rooted in emotional reactions, aligning trust with positive feelings and safety; (ii) Analogic trust, which relies on established behaviour patterns, where adherence to domain-specific language and norms can bolster trust, irrespective of content; (iii) Analytic trust, which hinges on comprehending the underlying reasoning behind decisions, underlining the importance of rationale in trust development.

Funded by the European Union

Table 9. System functions that are crucial for system transparency according to the literature (National Academies of Sciences, Engineering, and Medicine et al., 2022).

	Chen et al. (2014a, 2018)	Endsley (2017, 2020b)	Lee and See (2004)	Lyons (2013)	Sarter and Woods (1995)	Wickens et al. (2022)	U.S. Air Force (2015)
Level 1 SA System Status	Factors user is taking into account	State of knowledge				Raw data used by automation	
	Current system state	Key system states and mode transitions			Current system state	What automation is doing	Current system state and modes
	Purpose (goals)		Purpose	Goals			
	Intentions		Process	Intentional (purpose and social intent)			
	Plan of action			Tasks			
	Progress			Progress			
	Performance	Performance effectiveness	Performance	Errors			
				Environmental constraints			
Level 2 SA Understand-ability	Reasoning	Understand-ability of actions (drivers)		Analytical - decision logic	Reasons for current behavior	How automation is doing it	Explanation of reasoning
	Trade-offs						
	Capabilities/ limits	Ability to handle current situations		Capabilities		How automaton might err	
		Confidence in assessments				Degree of uncertainty	Confidence
Level 3 SA Predictability	Planned actions	Predictability of future actions (possible and predicted)			Future behaviors		Projected actions
	Predicted consequences	Ability to handle upcoming situations					
	Predicted outcomes						
	Uncertainty						

Funded by the European Union

Other	History of performance	System reliability	History				
Team and Shared SA	Teammate's roles and responsibilities, goals, projections, and reasoning	Teammate's current goals, priorities, function allocation, plans, and tasks		Team - division of labor			Teammate's current goals, function allocation, plans, and tasks
		Relative capabilities of human and autonomy		Human state - workload, performance, stress			Relative capabilities of human and autonomy
	Contributions to shared tasks	Impact of tasks on others					Impact of tasks on others
		Projected strategies, actions, and plans					Projected strategies, actions, and plans

Table 10. Summary of key findings on the role of Transparency and Explainability in HAT

Categories	*Key Findings*
Nature of Transparency and Explainability	• Transparency and Explainability dimensions include organizational, process, data, algorithmic, and decision processes.
Role of Transparency and Explainability	• Affects the quality of human-agent interaction and trust. • Improve Predictability • Improve Safety
Factors influencing Transparency and Explainability in HAT	• Complexity of algorithm or agent functions • Effective Communication • User training

2.3.10. Bias

Bias in Human-autonomy teaming can manifest from either the human or the machine. Human-induced bias can stem from different sources such as: (1) improper data curation, (2) design of some algorithms, (3) interpretation of the resulting algorithms (Cummings & Li, 2021).

While it's commonly believed that humans can supervise AI systems and spot their errors, performing independent validation, research has demonstrated that this assumption doesn't hold true. The accuracy of an AI system can directly influence human decision-making, resulting in the emergence of bias within the Human-AI team dynamic (National Academies of Sciences, Engineering, and Medicine et al., 2022).

According to the National Academies of Sciences, Engineering, and Medicine (2022) people commonly use the recommendations provided by AI systems as their starting point and then gather additional information to either support or challenge those recommendations. Consequently, AI systems play a direct role in shaping human decision-making, which can heighten the risk of human errors when the system's recommendations are inaccurate, resembling the cognitive bias referred to as confirmation bias. Mosier &

Funded by the European Union

Skitka (1996, p. 205) defined this type of bias as automation bias: “the tendency to use automated cues as a heuristic replacement for vigilant information seeking and processing”. Automation bias can lead to specific errors (Wickens et al., 2015) such as when decision-makers overlook issues because an AI system fails to detect them (referred to as an omission error) or when they unquestioningly follow an incorrect guidance or announcements from an automated decision assistant (referred to as commission errors). The omission error can result in delayed manual intervention in the event of a failure and reduced overall accuracy, while the commission error might lead to more radical loss of accuracy (Wickens et al., 2015).

It has been made clear in the body of research that automation bias adversely affects the performance of human-machine teams. Metzger and Parasuraman's study (2005) revealed that air traffic controllers exhibited superior performance when operating independently compared to when collaborating with an imperfect conflict detection system. Additionally, when AI systems provide incorrect guidance, their human counterparts are significantly more prone to making errors, with error rates ranging from 30% to 60%, compared to situations where they receive no directives from the AI system. Thus, understanding and mitigating bias in Human-autonomy teaming are not only ethical imperatives but also crucial for harnessing the full potential of AI to augment human capabilities. That is why Wickens et al. (2015) advise that automation diagnostic aids should provide information about their level of confidence or certainty in their assessments so that the risk of automation bias can be identified and reduced.

Table 11. Summary of key findings on the role of bias in HAT

Categories	*Key Findings*
Nature of Bias	• Bias can manifest from either the human or the machine • Automation bias can be due to either omission errors or commission errors
Role of Bias in HAT	• Affects the quality of human-agent interaction and trust • Affects the performance of human/machine teams • Heightens the risk of human errors • Can result in delayed manual intervention • Reduces overall accuracy • Impacts safety
Factors influencing Bias in HAT	• Improper data curation • Design of some algorithms • Interpretation of the resulting algorithms • Accuracy of AI system

2.4. Summary and Research Needs

The manufacturing industry has increasingly integrated AI and advanced automation into its processes, the concept of Human-AI teaming has emerged as a critical area of study. This creates the opportunity to push this integration to a point where collaborative partnership between human operators and AI systems can be achieved. In this context, understanding the foundational elements that contribute to effective Human-AI teaming is essential. “*The potential of autonomous agents working with humans opens up an interesting question, which involves both articulating a clear definition of human–autonomy teams (HATs) as well as*

Funded by the European Union

identifying the factors that make these teams successful" (O'Neill et al., 2022). At the core of this transformation lies Trust, a multifaceted concept that goes beyond mere machine reliability. Trust in the industrial context extends to the predictability of outcomes, especially when manufacturing precision is at stake. The trustworthiness of AI systems in predicting defects, optimising processes, and ensuring safety becomes paramount. In this context, the anthropomorphic design of machines can further enhance trust, making interactions more intuitive and fostering a sense of camaraderie on the shop floor. Agency, particularly in automated manufacturing units, defines the spectrum of machine autonomy. Machines equipped with higher agency can independently adapt to changes, ensuring seamless production even in the face of unforeseen challenges. This is particularly crucial in manufacturing, where downtimes can have significant economic implications.

Task interdependence in manufacturing goes beyond the mere division of labour; it's about orchestrating a harmonious ballet of human skills and machine precision. Whether it's a human overseeing quality control after an AI-driven assembly or machines taking over repetitive tasks, understanding this interplay is crucial for efficiency. Adaptability in manufacturing translates to the ability of the production line to recalibrate in real-time. Whether it's a sudden change in raw material quality or an unexpected surge in demand, both humans and machines must pivot swiftly, ensuring that production goals are met without compromising on quality. Effective communication becomes the backbone of HAT in manufacturing. It ensures that both human operators and machines are aligned in their objectives, be it achieving a particular production target or maintaining a specific quality standard. Cooperation and Coordination in a manufacturing setup ensure that there's a seamless flow of operations. Machines need to be predictable to their human counterparts, and humans need to be directable, ensuring that the entire production line functions as a cohesive unit.

Shared mental models in manufacturing ensure that there's a unified understanding of the production process, goals, and potential challenges. This shared understanding is pivotal in ensuring that both humans and machines can anticipate and respond to each other's actions, optimising the production process.

Transparency and Explainability take on added significance in manufacturing. Workers need to understand the logic behind machine actions, especially when anomalies occur. This not only fosters trust but also ensures that any issues can be swiftly addressed, minimising production disruptions. Bias, especially in AI-driven quality control or predictive maintenance, can have tangible repercussions in manufacturing. Ensuring that AI systems are free from biases, be it in detecting defects or predicting machine failures, is crucial to maintain the integrity of the production process. A deep understanding and meticulous implementation of the aforementioned factors will ensure that the manufacturing industry can harness the full potential of HAT, paving the way for a future that is not just productive but also innovative and sustainable.

Funded by the European Union

3. Human-AI Teaming Interaction Model

3.1. Introduction and Context

While this state of the art provides an overview of the elements that can affect HAT, it can be noted how an effective approach in designing teaming is missing (Xu & Gao, 2023). This includes understanding the context in which teaming needs to be introduced, making purposeful and meaningful choices in selecting the appropriate "flavor" of teaming, operationalizing design choices for system implementation, and establishing validation metrics for the process.

This chapter introduces a new model that aims to describe the components that determine HAT in manufacturing: *Agent Characteristics, Operator Characteristics, Mediator Characteristics, and Teaming Characteristics.* These elements not only define the dynamics of Human-AI interaction but also serve as the basis for evaluating the efficacy and viability of such teams in a manufacturing environment. The last part of the chapter also introduces how the model may fit in a design effort, while not being a “how to” guide it informs on how a descriptive model can be used as an active design and development tool.

3.2. Methods

The Human-AI Teaming Interaction Model has been developed with the aim of constructing an actionable structure that can describe HAT, collecting the factors that describe the team and its agents as well as the external factors that can influence its dynamics and ultimately performance. The structure itself, although derived from the literature in its taxonomy and composition, is the result of an internal workshop with members of Deep Blue. The workshop included the use of Speculative Design and Scenario Based Design to start visualising the number of variables that will later be summarised in the schema. The aim was to use transdisciplinary human factors knowledge to give shape to the complexity of human-AI teaming.

The teaming characteristics are the result of the combined influence of the environment, task, team characteristics, operator characteristics, and mediator characteristics.

$$T = f(E, Ta, Tc, Oc, Mc)$$

Where:
T = Teaming characteristics (the resultant of the interaction between all factors)
E = Environment characteristics
Ta = Task characteristics
Tc = Team characteristics
Oc = Operator characteristics
Mc = Mediator characteristics

A workshop and a survey have been carried out with the members of the HARTU project with the aim of operationalizing what design choices respond to the question of what “flavour” of teaming is most appropriate. The project’s partners were involved in this stage to ensure that a system designed using the model can be effectively implemented.

Funded by the European Union

3.2.1 Survey

In the survey partners of the HARTU project were presented with an initial list of Roles, they were asked to comment and expand on the list as well as indicate which roles would best suit the Use Cases they are involved in. The initial list of identified roles was further expanded based on the feedback collected, leading to the delineation of new roles for both AI and human operators.
The complete survey is reported in the annex.

3.2.2 Workshop on strategies of role selections

The survey highlighted the possible human and AI roles for each scenario selected based on the technical limitations and the nature of the task itself. Two were the possible outcomes: desirable and undesirable. The desirable is highlighted in green in the first table, and the undesirables are highlighted in red in the second table (figure 5). The survey was followed by a workshop aimed at fostering a collaborative and multidisciplinary environment and creating a common understanding of the concept of roles.

The aim of the workshop was to collaboratively categorise the different roles. Each of the involved partners could highlight strengths or issues that might affect the implementation of such roles.

The workshop highlighted a strong preference among the participants for a collaborative role, where AI functions as a support rather than a replacement for human operators. Participants agreed that this role best suits tasks requiring both AI's efficiency and human oversight, especially in scenarios that might be affected by unavoidable malfunctions, beyond design basis conditions or including tasks that require supervision.

Another key takeaway was the introduction of an **optimizer role** for both the AI and the Operator. This role allows the operator and AI to proactively identify areas to improve AI performance and avoid recurrent failures. The optimizer role builds a more flexible and resilient partnership, where both AI and human skills are utilized dynamically.

The participants also recognized the importance of role adaptability, the operators should be able to shift between roles, such as passive observer, helper, or supervisor, depending on the AI's performance or the task's complexity. Participants suggested that this adaptability would foster trust by giving operators the control to step in or oversee, when necessary, while also streamlining workflows as tasks evolve.

The workshop emphasized potential risks with certain combinations, such as pairing a faulty AI with a passive or clumsy user, which participants deemed particularly undesirable and imagined that a good step for the design process is to focus on that undesired combination that could increase operational risks due to the user's limited capacity to intervene or correct issues. Fully independent AI was also seen as undesirable in many cases, as it reduces opportunities for human oversight and response, which are critical in real-time operations where unforeseen challenges may arise. Finally, the group emphasized trust-building measures, noting that transparent AI communication is fundamental for effective teamwork. They suggested that the AI should clearly communicate, especially regarding common failure points or adjustments, to reduce the burden on the operator for troubleshooting and maintain a smoother workflow. By increasing transparency, AI can build trust with operators, ensuring they feel confident and informed within the Human-AI team. This proactive approach to communication aims to improve both performance and mutual reliability, contributing to a more effective human-AI partnership.

Funded by the European Union

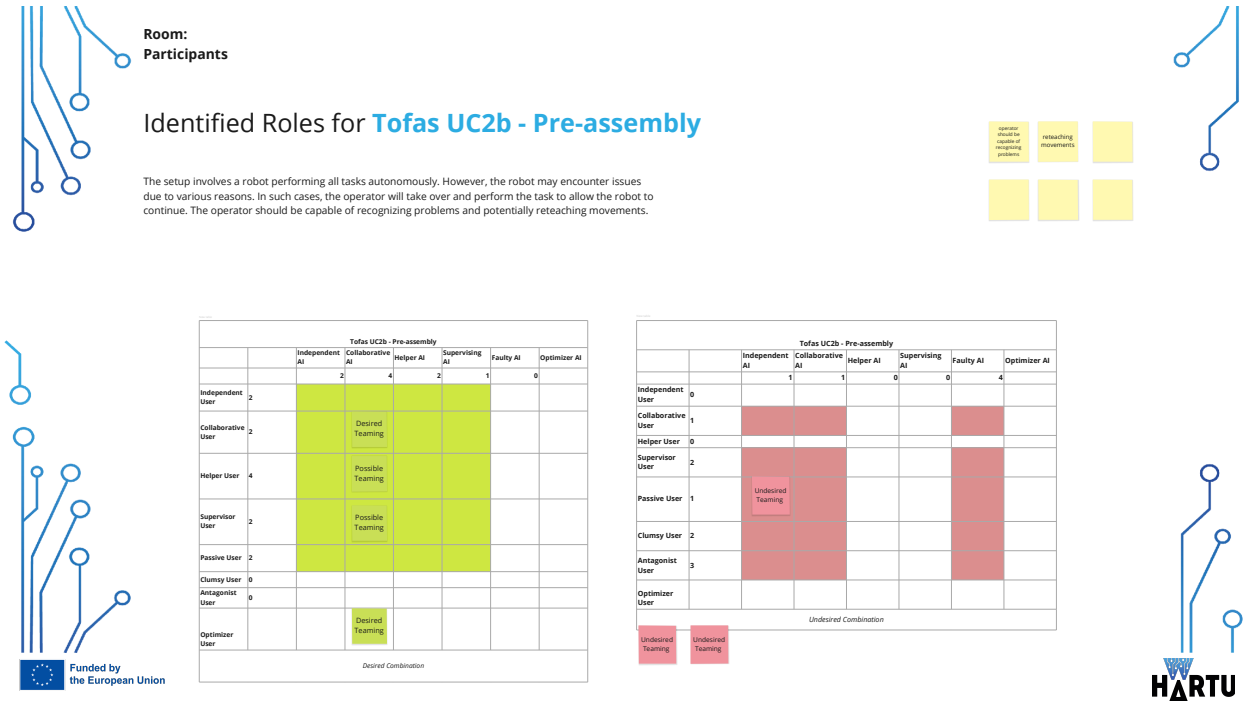


Figure 5 – Sample from the results of the online workshop in Human-AI Teaming, for each UC the desirable and undesirable roles were identified and discussed.

The workshop insights reveal a focus on adapting roles based on task complexity, user expertise, error frequency, and real-time conditions. For straightforward tasks, participants generally supported the AI taking on more autonomous or independent roles, other tasks involving high error rates, promoted problem-solving, or potential safety risks called for supervisory or collaborative roles for humans.

Participants advocated for systems that could shift human and AI roles dynamically in response to emerging issues, such as task complexity or system malfunctions. This adaptive approach also requires clearly defined role boundaries, ensuring both the AI and the human operator understand their specific contributions without overlap or ambiguity. Finally, the workshop underscored the importance of context-aware system design in supporting real-time, situation-dependent shifts in role designation, promoting a robust teamwork dynamic where users are empowered to take on varied roles as needed, which fosters trust, efficiency, and productivity in human-AI interaction.

Funded by the European Union

3.3. The Model

3.3.1. Factor clusters

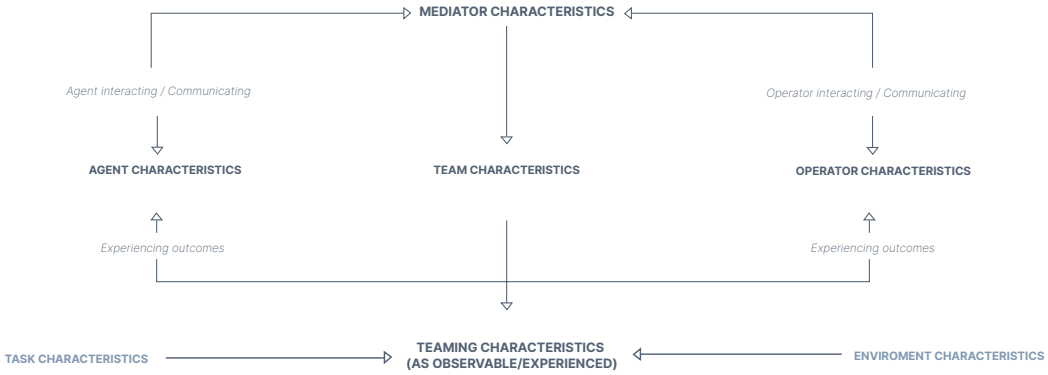


Figure 6 – Simplified HATIM model, representing only the main clusters.

To structure the Human-AI Teaming Interaction Model (HATIM), the main actors and the way they interact are highlighted as the central blocks. Although the Agent and the Operator are the primary protagonists, they do not interact directly. A 'Mediator' layer is introduced between them to represent all the elements that govern their interaction.

Mediator: The Mediator is a complex element as it includes choices regarding the user interface, the modes of interaction (for example voice control, visual feedback etc..) as well as procedures (for example “the AI can respond only if interpellated” or “the user has a veto on AI decisions”) and data processing and visualisation.

From the Agent’s perspective, the mediator layer also includes data feeds and sensors that it can use to perceive the environment, the operator or the state of what they are working on. Although some of these might be part of the technical implementation of an AI system, it is advantageous to maintain this separation to emphasise how the same agent could be implemented with different interfaces or in a different context of procedures and rules, the same way as different operators will interact with the system.

The AI **Agent** itself is described as the first half of the team, how it is developed and set up as well as its emerging characteristics such as its SA, goals and mental models, it interacts with the Operator through the mediator, but this is a two-way channel, depending on the implementation, the Agent may be able to evolve and adapt based on feedback and inputs it receives.

Next, the **Operator** is described, including skills, personality, previous experiences, his assigned role and responsibilities. It is important to note how the operator has many dynamic and evolving characteristics, and elements that will change during the interaction. Concepts such as his mental model of the AI Agent, his SA or his trust in automation will be affected by the interaction within the team, thus it will become apparent how it should be specified, when describing a team, whether the description regards an initial

Funded by the European Union

state (how the operator might be as it joins the team for the first time), or a goal state (how the operator state should evolve to optimise teaming).

Once the Operator and Agent are connected through the Mediator the concept of **Team starts to emerge**. As discussed in the state of the art, there are requirements for a group of entities to become a team. Agency, autonomy, and task interdependence among team members are primary elements. As these foundational elements are established, the emerging characteristics of the team, such as trust, shared mental models, goals and internal dynamics, begin to take shape. While these characteristics can be generally designed and defined in the form of optima combinations, they cannot be directly influenced. **Teaming** describes how the team performs through observable and measurable characteristics of how the team operates. While Team characteristics provide a static description of the team, Teaming is a performative metric influenced by both the team itself and its external context, including the environment and the specific tasks assigned. A team's performance can vary depending on these factors. The environment in which the team operates encompasses not only physical and ergonomic factors but also legal, ethical, organizational, and cultural ones. Consequently, the implementation of a specific team may vary significantly across different cultural and legislative contexts.

In the following section, a more detailed description of the model components is provided. The definition of such components is based on the results of the state of the art, although it has been adapted to the purposes of the manufacturing context.

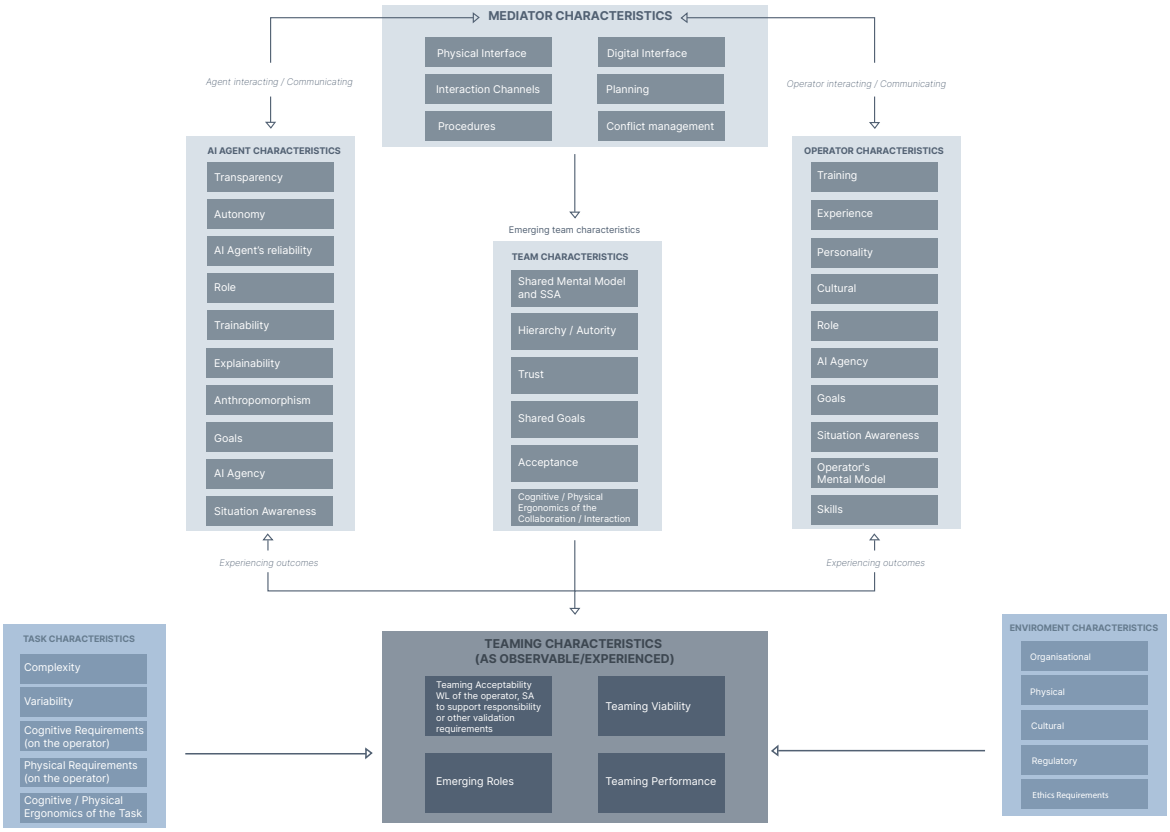


Figure 7 – Complete HATIM model.

Funded by the European Union

3.3.2. Mediator

Mediator characteristics refer to the interface through which the AI and human operator interact. This interface plays an important role in ensuring that both parties can communicate and collaborate effectively. Mediator characteristics include:

Interaction Channels: The methods of communication between the AI and the operator can include visual displays, auditory signals, haptic feedback, or voice commands. An initial step in designing the team might involve defining how members will perceive and interact with each other. This model distinguishes between physical and digital interfaces as subcategories of interaction channels, aligning with traditional design processes. However, it's important to consider non-traditional channels that may not be typically categorized as interfaces, such as observing a robot's operation through a window or an operator's body language for signs of distress or fatigue.

Physical Interface: Tangible elements of the interface through which the operator interacts with the AI: control panels, buttons, and levers. A component of the physical interface could also be part of the robot itself designed to facilitate the assessment of the robot's state.

Digital Interface: Software-based elements that facilitate interaction between the operator and AI: screens, dashboards, and alerts.

Planning and Procedures: Norms, processes, guidelines or other constructs that guide how the AI and operator should coordinate their actions, make decisions and interact. Procedures can have multiple forms, both explicit and implicit. Explicit rules should be well-defined and documented, this includes safety or emergency procedures or quality assurance while implicit rules may just exist as uses and best practices between workers. Procedures in HATs should be implemented in such a way to be shared among the team members, this may be done organically as part of interfaces (timers, interlocks etc..) or baked-in the AI agent’s behaviour.

Conflict Management: As a subset of procedures, conflict management includes all strategies and tools available for resolving disagreements or misunderstandings between the AI and the operator. Whether an operator can override the Agent’s decisions or vice versa should be well-defined and enforceable.

Information modulation allows the AI to adapt communication based on the operator’s workload or to anticipate the operator’s informational needs and "push" relevant data proactively.

3.3.3. Agent

These characteristics include the agent's internal model, goals, and interaction capabilities with both the operator and the environment. These factors significantly influence the overall performance of the team**.**

Transparency: Transparency in autonomous systems refers to how and how much the AI communicates about its operations. At a basic level, transparency might involve only sharing the AI status, while higher transparency includes revealing the reasoning behind its decisions, projecting future outcomes, and communicating uncertainties. As the model focuses on human-AI interactions, the human-centric side of transparency should be prioritised, which emphasises openness to discussing AI-driven decisions. This involves making AI decisions challengeable, falsifiable, and explicable. Transparency is key for building trust between human operators and AI, as it allows operators to understand and predict the AI’s actions, helping them form and refine mental models of the AI based on their experiences. While transparency alone doesn't

Funded by the European Union

guarantee trust, it enables operators to evaluate other critical factors like accuracy and fairness. For instance, even an unbiased AI system could be perceived as biased if it is opaque. In manufacturing, where decisions must often be made rapidly and based on real-time data, transparency allows operators to follow the AI's logic and step in if needed. However, studies show mixed effects of transparency. High transparency can clarify AI reasoning and decision-making, but it may also increase the operator's workload or lead to complacency, reducing vigilance in overseeing or questioning the AI's actions.

Explainability: Explainability is a crucial feature that supports transparency and trust. It refers to the AI's ability to produce human-readable explanations for its decisions and actions, such as identifying the variables that influenced a decision. For explainability to be effective, operators must have the literacy and training needed to interpret the AI's outputs. In manufacturing, explainability is essential for operators to understand the reasoning behind the AI's recommendations or behaviours, which aids in decision-making, troubleshooting, and process improvement.

Autonomy: Autonomy refers to the degree of independence an AI system has in executing tasks. In manufacturing, AI systems can function autonomously to optimise processes, control machinery, or manage logistics. The level of autonomy indicates how much of the task the AI can perform without human intervention, feedback, programming, or monitoring. For a system to be classified as a Human-Autonomy Team (HAT), the autonomous agent must demonstrate at least some level of autonomy. Partial autonomy allows the agent to suggest and carry out actions unless overridden (Levels 5–6 on Parasuraman's scale). Higher autonomy levels (7–10) enable the agent to act independently without human input or prompts.

Anthropomorphism: Anthropomorphism includes all the effort and strategies that can be used to imprint the Agent with human-like characteristics. While not always necessary, anthropomorphism can enhance human-AI interaction by making the AI seem more relatable and easier to understand.

Reliability: The reliability of an AI system refers to its ability to consistently perform tasks accurately over time. Reliability is often measured by the degree of accuracy with which the AI completes its tasks. A decline in reliability tends to produce various negative effects in human-AI teams, such as increased workload, reduced performance and decreased trust. In the design of HAT Reliability can either be a requirement (minimum reliability required for the teaming to be functional) or as a fact, (specification of the machine that needs to be dealt with). Users can cope with limitations in reliability, but this needs to be addressed, though, for example, explainability, teachability or supervision of the user. Transparency can mitigate the effects of lower reliability, as greater openness about the AI's decision-making processes may help maintain trust and understanding. However, transparency itself comes with trade-offs, balancing benefits like improved performance against potential drawbacks like operator complacency.

Goals: The AI's goals must be aligned with the overarching objectives of the manufacturing process. Whether optimising for speed, quality, or cost-efficiency, the AI's goals should be transparent and understood by human operators to ensure cohesive teamwork. Depending on how the AI system is developed it might not be possible to bake in predefined goals, for example, goals might emerge from training data, and should be verified.

Role: In Human-AI teaming, roles define the division of responsibilities and interaction dynamics between humans and AI systems. These roles determine how tasks are managed, decisions are made, and how collaboration is structured. Often Roles are the result of a "leftover allocation" strategy (Roth et al., 2019),

Funded by the European Union

this approach tries to maximise what the AI Agent can take over and subsequently allocates what the AI cannot handle to an operator. The model suggests a different approach, aiming to create functional teams with recognisable members. The roles of the operator and AI can be defined in a process similar to Vladimir Propp's narrative theory (Propp et al., 1968), which categorizes characters based on their functional roles in a story. Just as Propp's character functions can be applied across a wide range of narratives, this approach can be adapted to different use cases of Human-AI collaboration. In this framework, humans and AI systems perform different "narrative" roles, each contributing uniquely to the team's overall performance. For instance, the human might take on a leadership or decision-making role, while the AI functions as a helper, providing insights and automation to assist in achieving the goal. Alternatively, the AI could even take on a more autonomous role, guiding certain processes and reducing the load on the human user.

This role-based approach can be applied across a wide range of industries and tasks. In each of these environments, the nature of the task, the goals of the human, and the capabilities of the AI system will explore the most appropriate distribution of roles. Understanding and identifying these roles enables organizations to create more effective Human-AI teams by tailoring interactions. based on the strengths and limitations of each entity.

By applying this structured role model, derived from Propp's narrative theory, not only these roles can be defined but also analysed in terms of when and where they are most effective.

The identified Roles for the AI are:

Collaborative AI: works alongside human users, sharing responsibilities in tasks and decision-making. It actively engages in a partnership with the human, communicating frequently and adapting to feedback. The AI is open to collaboration and shared responsibility and can take responsibility for a portion of the task. This role is suited to create a balanced collaboration, where the AI provides support by suggesting options, providing insights, and executing tasks based on human input.

Independent AI: has primarily an autonomous role, leading tasks and making decisions without significant human involvement. It is designed to manage workflows, solve problems, and handle complex operations on its own. It can operate with minimal to no human intervention, suitable for situations where human input is limited or not feasible. This type of system might not be designed to accept user input, so it should be decided if and how the AI can allow occasional human oversight or emergency intervention.

Helper AI: supports human users by performing secondary tasks that assist in completing a primary goal. It focuses on simplifying operations, such as handling repetitive tasks, data analysis, or making recommendations based on patterns it detects. Helper AI is particularly useful in reducing the load on the human operator but is not supposed to take responsibility.

Supervising AI: acts as an overseer of the user state and actions, monitoring processes and ensuring compliance with predefined standards, safety regulations, or performance parameters. It provides real-time feedback and alerts, helping to ensure smooth operations and reducing the chances of errors. This AI is not directly involved in task execution but plays a crucial role in maintaining the quality and safety of operations.

Funded by the European Union

Optimizer AI: focuses on identifying inefficiencies in workflows and suggesting improvements. It actively analyses processes, detects areas where performance can be enhanced, and recommends changes that optimize the way tasks are carried out. The Optimizer AI's role is more strategic, supporting continuous improvement and fine-tuning of both human and machine operations.

Faulty AI: Although this may not be an intentional role, it represents scenarios where the AI system experiences errors or malfunctions, leading to incorrect actions or poor decision-making. It might misinterpret data, provide misleading feedback, or fail to detect important issues. When such errors occur, human intervention is required to correct or override the AI's actions to prevent negative outcomes. Understanding the limitations of AI and anticipating potential faults is crucial for effective human-AI collaboration.

Agency: The degree of agency granted to an autonomous agent defines its responsibilities and level of influence. According to previous literature, two key criteria determine whether an autonomous agent is viewed as a true teammate: the level of interdependence with other team members in terms of activities and outcomes, and the agent's degree of independence and proactivity in taking actions. These elements are essential for the agent to function as an active, contributing member of the team.

Trainability: The AI's ability to be trained or adapted to new tasks and environments. Trainability ensures that AI can evolve alongside changes in production processes, technologies, and market demands. When defining trainability it should be also defined when and how training can happen. Learning could be continuous, through interaction, or it could be at defined stages.

Adaptability and Flexibility: Adaptability and flexibility refer to the AI system's capacity to adjust its behaviour and responses based on varying contexts, user needs and states. Examples of Adaptability may include the ability to take on different responsibilities to alleviate the workload, taking over a process if the Operator becomes incapacitated, or simply responding to operations commands and changes in the environment without failing.

Situation Awareness: Situation awareness refers to the AI's ability to perceive and understand its current operating environment. While the concept of S.A is extensively described for human operators it must be adapted to be applicable to an AI agent. The ability of a computer to hold and process data makes concepts such as working memory redundant, rather It should be noted how mere access to information is not necessarily proof of SA in an AI system. Whether an AI has certain information available (for example the readings from a temperature sensor) doesn't necessarily mean that it will act on it in a predictable manner. Explainability helps in understanding what data inputs affect the agent's decision-making process and what instead does not. Therefore, understanding SA is to understand what information affects AI decision making and how this might change in different instances.

Mental Model: The AI's mental model refers to its internal representation of tasks, the environment, the systems it interacts with, and its objectives. A well-developed mental model enables the AI to anticipate changes, predict outcomes, and align its actions with those of the human operator. This human dimension must be adapted to ensure relevance for an autonomous agent. Transparency plays a crucial role in assessing the agent's internal model and helps the operator form a well-constructed understanding of how the agent might respond to different inputs. We can therefore define the **inner mental** model as the agent's schematic representation of the world, encompassing its understanding of various elements and their

Funded by the European Union

interactions. In contrast, the **perceived mental model** reflects the operator's interpretation of this representation. Bridging the gap between these two models is essential for effective collaboration, as it fosters a shared understanding of the agent's capabilities and decision-making processes. This alignment enhances trust and allows for smoother interactions, ultimately improving the effectiveness of human-AI teamwork.

3.3.4. Operator

Human operators bring a set of characteristics that are equally important in the context of Human-AI teaming. These characteristics shape how operators interact with AI systems and determine the effectiveness of the collaboration. Key operator characteristics include:

Training, Experience and Skills: The level of training and experience of the human operator directly impacts their ability to work effectively with AI systems and be a useful member of a team. Skilled operators are better equipped to understand and leverage AI capabilities, interpret its outputs, and respond to its recommendations. Training represents the opportunity for both human and agent team members to learn about, practice, and train with their other team members. Training and experience can refer to either what is technically necessary for the completion of the task (for example the ability to weld), as well as experience in interacting with, servicing or training AI agents.

Human individual differences: factors that may vary from person to person such as personality, culture, and past experiences. Individual differences in personality and cultural background can affect how operators perceive and interact with AI. For instance, operators from cultures that emphasise hierarchical structures may have different expectations regarding decision-making authority compared to those from more egalitarian cultures. Understanding these factors is crucial for designing AI systems that are accepted and trusted by diverse user groups. Some of these characteristics can be documented only through surveys or interviews of potential operators.

Role and Agency: The human operator’s role and level of agency must be clearly defined so that operators can understand their responsibilities, the extent of their decision-making power, and how their actions complement or guide the AI’s functions. The identified roles provide a framework for clarifying these dynamics.

> **Collaborative User**: In this role, the operator shares responsibility with the AI, frequently interacting with the system to provide input, guide processes, and make decisions based on feedback. The level of agency is balanced, allowing the human to take the lead when needed while also relying on AI support for task completion. A “Collaborative” operator is willing to share responsibility with an AI Agent and allocate part of the task outside their control.
>
> **Independent User**: The operator takes full control and leads the task, making all critical decisions and managing the workflow without significant AI intervention. The AI may still assist in a secondary capacity, but the operator retains the primary agency and responsibility for the success of the operation. This type of operator may use AI as a support tool but is in general unwilling to share responsibility with an AI agent.
>
> **Helper User**: Supports the AI by providing additional input or manual adjustments when the AI encounters limitations. The operator steps in to guide, correct, or oversee specific tasks, ensuring that the AI's actions align with the overall objectives. The level of agency may vary depending on the

Funded by the European Union

situation, for example, a helper user might need to take responsibility from the AI in case of an emergency.

Supervisor User: The operator oversees the AI's actions and system performance, intervening only when necessary to correct issues or optimize processes. The human's role is to monitor the bigger picture, stepping in to provide direction or adjustments to ensure that the AI stays on track. In this role, the operator has a high level of situational awareness and agency but exercises it selectively, focusing on strategic oversight rather than continuous involvement.

Optimizer User: The operator actively seeks to improve processes, identify inefficiencies, and propose enhancements for both human and AI operations. This role involves a strategic level of agency, where the user takes on a continuous improvement mindset, analysing workflows, and implementing changes to optimize performance. The operator collaborates with the AI to fine-tune the system, leveraging the strengths of both human insight and machine data analysis.

Passive User: A passive user has limited involvement in a task as well as limited awareness of the task's status. This role can be intentionally implemented as a team member if she is paired with an AI that has a level of agency sufficient to complete the task on its own. In this case, the operator is free to take on different tasks but is still available to support the AI if requested.

A passive user can also be an unintended consequence of incorrect task allocation or overreliance on automation. This role should be used to envision what can happen if the AI agent is left operating unsupervised.

Clumsy User: This is another unintended role, a clumsy operator may unintentionally introduce errors or inefficiencies due to misinterpretation of AI prompts or difficulty with controls. The role involves a low level of effective agency, as the user struggles to complement the AI's functions smoothly. This scenario emphasizes the importance of designing user-friendly interfaces and training to reduce the occurrence of errors and improve human-AI interaction.

Antagonist User: This role is characterized by resistance to the AI's guidance, either through intentional disregard for instructions or negligence. The operator may challenge the AI's decisions or disrupt the workflow, creating friction in the collaboration. The level of agency is high, but it is used in a counterproductive way that undermines the effectiveness of human-AI teaming.

Goals: The operator's goals should align with both the AI's objectives and the broader manufacturing goals. Discrepancies between operator and AI goals can lead to conflicts, inefficiencies, and a breakdown in teamwork.

Situation Awareness: The operator's awareness of the manufacturing environment and the AI's actions is critical. High levels of situation awareness allow operators to anticipate potential issues, make informed decisions, and intervene effectively when necessary. SA requirements should be defined in particular if the operator has supervision responsibilities or is expected to take over in cases of AI failure or emergency.

Mental Model: The operator's mental model represents their understanding of the AI system, the manufacturing process, and the goals to be achieved. A shared mental model between the operator and the AI is essential for effective communication, coordination, and problem-solving. The operator mental model is a dynamic and evolving characteristic, elements that will be affected by the interaction within the team. It

Funded by the European Union

can be den defined both as an initial state, as the operator joins the team for the first time or a goal state, how it should evolve through experience.

3.3.5. Team & Teaming

Team characteristics describe the team, the internal dynamics and roles while **Teaming** characteristics are the outcomes of the interaction between the AI, the operator, and the mediator. These characteristics define the quality and effectiveness of the Human-AI collaboration.

Team Characteristics:

Shared Mental Model (SMM): The degree to which both the AI and the human operator share a common understanding of the task, environment, and goals. A strong shared mental model ensures that both the AI and humans are aligned in their understanding of key aspects of the task, facilitating seamless teamwork and reducing the likelihood of errors. When AI can accurately predict human intentions and when humans can understand the AI's objectives and processes, coordination improves, and task outcomes are enhanced.

Shared Goals: The alignment of objectives between the AI and the human operator to ensure that both the AI and humans are working toward the same outcomes.

Shared Situation Awareness (SSA): The degree to which the AI and the human operator maintain a shared understanding of the current situation, environment, and task status, critical for coordinated action. High SSA minimizes errors by ensuring both parties respond effectively to changes in the environment or task parameters. Poor SSA can lead to miscommunication, errors, and inefficiencies, as one party may be unaware of critical information that the other party has.

Communication: The process of exchanging information between the AI, human operator, and any mediators (e.g., systems or interfaces). Communication can be characterised in terms of **quality, quantity** of information, and **frequency**.

Hierarchy and Authority: The structure of decision-making power within the Human-AI team, including who has the final authority in different situations. Clear hierarchy and authority structures prevent conflicts and ensure that decisions are made efficiently. In some situations, the human may have ultimate decision-making power, while in others, the AI may be granted authority to act autonomously (e.g., in high-speed, high-stakes environments where real-time AI decision-making is necessary).

Trust and Acceptance: The human operator’s confidence in the AI system and their willingness to rely on it. Trust is built through the AI’s consistent performance, transparency in decision-making, and the ability of the system to demonstrate reliability in different scenarios. Without trust, the human operator may ignore or override AI recommendations, potentially reducing team effectiveness. Conversely, too much reliance on AI without sufficient critical oversight can also lead to problems, especially if the AI makes an error.

Cognitive and Physical Ergonomics: The design of Human-AI interaction to optimize both mental and physical effort. Cognitive ergonomics focuses on reducing mental workload, making AI outputs easy to understand, and minimizing information overload. Physical ergonomics ensures that human operators can comfortably interact with AI systems without strain, which is particularly important in physical control environments or when using specialized tools or interfaces.

Funded by the European Union

Adaptability and Flexibility: The ability of the human-AI team to adjust to changing conditions, new tasks, or unexpected challenges. This can be implemented adjusting roles, tactics, or strategies in response to dynamic environments as well as communication structures.

Teaming Characteristics: emerge from the interaction between the AI agent and the operator; therefore they can be used to evaluate if teaming is successful and if it corresponds to what was designed.

Teaming Acceptability: The degree to which the human operator accepts and integrates the AI system as a functional and trusted member of the team.

Surveys and Questionnaires: Use Likert-scale surveys to gauge the operator's trust in the AI system, their comfort in relying on its outputs, and their perception of the AI's usefulness and transparency.

Usage Metrics: monitors how frequently the operator relies on the AI recommendations or overrides them. High override rates might indicate a lack of trust or acceptability.

Decision Adoption Rate: Measure how often the human operator accepts the AI's suggestions versus rejecting or modifying them.

Interviews and Focus Groups: Conduct qualitative assessments through interviews to understand deeper concerns or areas where the operator feels the AI is not fully integrated into the team.

Teaming Viability: Long-term sustainability and effectiveness of the human-AI partnership in the context they need to be deployed.

Longitudinal Performance Tracking: Track the team's performance over time, looking for trends in task completion time, accuracy, error rates, and problem resolution. Improvements or stability over time indicate higher viability.

Adaptability Scores: Evaluate how well the team adjusts to changes in task complexity, workload, or operational environment. This can be assessed by introducing new scenarios and tracking performance.

Stress and Fatigue Monitoring: Measure operator stress levels through wearable devices or self-reports during long-term engagements. Higher stress or fatigue levels over time could indicate low teaming viability.

Operator Satisfaction: Dissatisfaction could signal poor teaming viability. Regular feedback sessions can provide insights into long-term sustainability.

Emerging roles: The new or evolving roles and responsibilities that develop as the human operator and AI work together, to be evaluated against the desired roles to assess if corrective measures are necessary.

Task Allocation Analysis: Measure the distribution of tasks between the human operator and AI. Compare the actual role distribution with the predefined roles and responsibilities.

Role Adaptation Feedback: Gather feedback on how well operators feel their roles have adapted to the AI, whether they feel overburdened, underutilized, or if their responsibilities have shifted unexpectedly.

Error Rates by Role: Monitor whether role evolution leads to changes in error rates. For example, if the human operator is taking on new supervisory roles, how often do manual interventions correct AI errors or failures?

Funded by the European Union

Competency Gaps: Assess whether operators feel they have the necessary skills for their evolving roles through surveys or tests. Training programs can be adjusted based on the gap between the current role and skill set.

Teaming Performance: The overall effectiveness of the human-AI team in achieving its goals and objectives, measured by factors such as task completion speed, accuracy, collaboration quality, and decision-making effectiveness.

Task Completion Time: Measure how quickly tasks are completed with human-AI collaboration compared to benchmarks or past performance without AI assistance.

Task Accuracy and Quality: Evaluate the quality of outputs (e.g., precision, error rates, rework rates) in tasks completed by the human-AI team. High accuracy with fewer errors indicates better teaming performance.

Collaboration Quality Metrics: Assess how smoothly communication and coordination happen between human and AI. This could include the number of miscommunications or misunderstandings or frequency and efficiency of task handoffs between the human and AI.

Decision-Making Effectiveness: Evaluate how well human-AI decisions align with optimal outcomes. Metrics may include decision accuracy (correct or beneficial decisions) and time taken to make decisions.

Resource Utilization: Assess how efficiently human and AI resources (time, energy, cognitive load) are used to achieve objectives. This includes measuring the cognitive workload on human operators during collaboration with AI.

3.3.6. Task

The task can be operationalised as the set of goals, requirements and specifics that describes what the Team is asked to do. The scope of the task in Human-AI Teaming should be sufficiently circumscribed to ensure a stable relationship between team members. This means that the roles and responsibilities of the human and AI agents should remain consistent throughout task execution. Complex tasks with dynamic role and allocation changes can be broken down into a series of subtasks. This process of subdivision continues until a sufficient level of granularity is achieved. If subsequent tasks involve significant changes in roles, it's necessary to analyse how reallocation occurs and whether it introduces critical risks or challenges. For example, suppose an AI, normally autonomous in its operation, encounters an error, it may be required for an operator to quickly take over responsibility and switch between an observer/supervisor role to an active one.

Task Documentation: The task should be documented in a manner that provides sufficient information to inform the design and understanding of team behaviour. Depending on the design and development stage, this documentation can take various forms. Hierarchical Task Analysis or User Journeys, adapted to include the AI's perspective, can provide a clear overview of a task and its evolution. Simply listing task goals and performance metrics (time pressure, throughput, error rate) may be sufficient, particularly in early design stages.

Task Complexity and Variability: The complexity of the tasks the Human-AI team is expected to perform, including the degree of variability. Variability will influence the level of autonomy and adaptability required

Funded by the European Union

from the AI teammate to function and support the teamwork. Likewise, a sub-optimal match between the Agent characteristics and the task complexity will affect the operator's role as "leftover allocation" will force the Operator to take over everything that the AI is unable to handle.

Variability also refers to possible permutations of the task: premises, outcomes and contexts the same task could be performed in. In manufacturing, this may be related to adapting a process to different products, performing the same operation (for example maintenance of machinery) in different contexts or variability in the same process (for example machining the same part in brass or aluminium).

Time pressure: Time pressure and constraints can dramatically affect perceived workload, negatively affecting teaming quality. Low time pressure conditions allow the operator to schedule how much time to dedicate to an operation, whereas high time pressure might require the operator to employ coping tactics to complete a task successfully. Where time pressure is high, an unresponsive or lagging AI solution can impede communication and reduce acceptability.

Interdependence is the extent to which humans and autonomous agents need to engage in intensive exchanges to achieve the team objective, it is a fundamental requirement for teaming. It describes portions of the task where the operator and agent necessarily work together to complete the task. Complete Interdependence may reduce the perception of the agent as a unique team member occupying a distinct role, likewise, a lack of Interdependence results in a complete lack of perception of the agent. Interdependence may be documented as points of contact between the Agent and Operator during the task execution, exchange of information or commands as well as concurrent actions.

Physical and Cognitive Requirements should also be documented as the demands placed on both the AI and the operator in terms of physical labour, cognitive processing, and decision-making will determine the effectiveness of the teaming and the need for ergonomic design.

3.3.7. Environment

The organizational context, physical environment, regulatory requirements, and ethical considerations all play crucial roles in shaping how Human-AI teams function in practice.

Environmental Factors: these include the broader context in which the Human-AI team operates, such as the physical environment (e.g., noise, lighting, workspace layout) and external conditions that can affect performance. A noisy or poorly lit environment may impede both the human operator's ability to interpret information and the AI system's ability to capture accurate inputs. Space constraints may also impact how the team physically interacts with equipment and systems.

Organizational Factors: the structure, policies, and culture of the organization play a significant role in shaping Human-AI teaming. Organizational culture affects openness to innovation, trust in AI systems, and how collaboration between Humans and AI is perceived. An organization with a clear emphasis on technological innovation and training will likely foster more effective human-AI collaboration. Additionally, the organization's policies and workflows influence how tasks are delegated, which roles are automated, and how responsibility is shared between AI and human operators. Change management and the willingness of employees to adapt to AI-driven changes are also crucial in determining the success of the Human-AI team.

Funded by the European Union

Ethics Requirements: Ethical considerations govern how AI systems should be designed, deployed, and interacted with. These include the need for transparency, ensuring that AI systems provide understandable and justifiable decisions, and accountability, ensuring human oversight over AI actions and decisions. Ethical frameworks should prioritize the well-being of human operators, ensuring that the AI system does not inadvertently cause harm, such as by increasing cognitive workload or making decisions without clear human input. Bias mitigation is also essential to ensure fairness in how AI systems process data and make recommendations, particularly in decisions that affect individuals or groups.

Regulatory Context: Human-AI teaming must comply with a wide array of regulatory requirements for the development of AI systems, particularly in sectors like manufacturing, healthcare, or autonomous systems. These regulations include standards for safety, data protection, and liability in case of system failures or accidents. Regulatory bodies may dictate specific constraints on how AI systems are tested, how they handle personal or sensitive data, and who is ultimately responsible for decisions taken by AI. AI deployment may also be subject to certification processes that require rigorous testing and validation to ensure safety and reliability.

Cultural Context: the cultural context shapes attitudes toward Human-AI collaboration, both at the organizational level and within the broader society. Depending on the geographical location and the type of industrial production, trust in automation and AI may vary. Furthermore, scepticism and fear over job displacement could create resistance. For example, in a culture that values hierarchical decision-making, human operators may be more hesitant to rely on AI recommendations or delegate to AI systems. Conversely, in more collaborative cultures, AI might be more readily integrated as an equal team member in decision-making processes. Understanding and adapting to these cultural dynamics is essential for fostering effective and sustainable Human-AI teaming.

Funded by the European Union

3.3.8. Design Layers of the HATIM

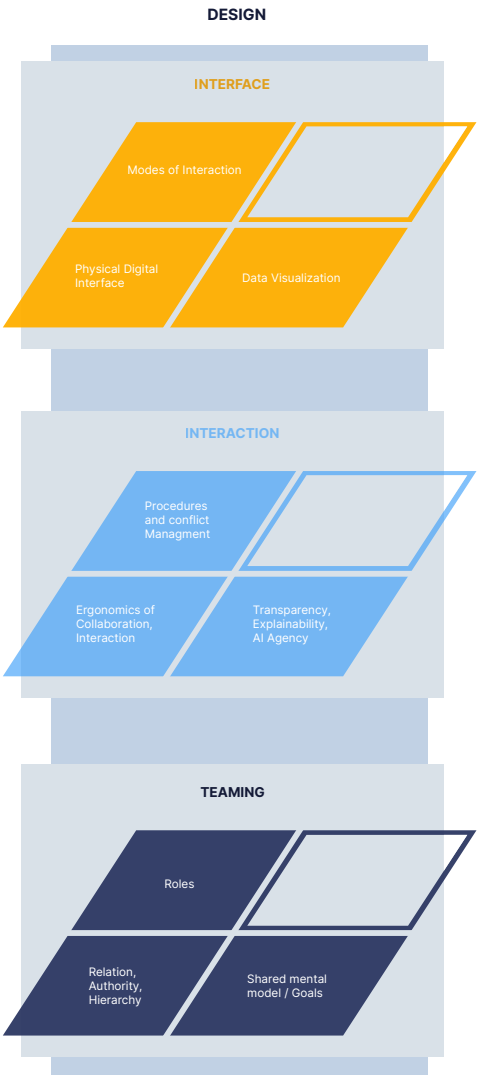


Figure 8 – The three abstraction layers of Human-AI Teaming design: Interface, Interaction and Teaming.

To approach the design of HAT, the model has been divided into three abstraction layers: **Interface, Interaction, and Teaming** (see Figure 8). These layers are not intended to represent layers between the Agent and the Operator but rather they are different parts of the design in relation to each other.

The organization in layers is based on the five levels that summarise user experience design (Strategy, Scope, Structure, Skeleton and Surface) defined by James Garrett (2008), "Experience actually results from a series of decisions... how it behaves and what it allows you to do. These decisions build on each other,

Funded by the European Union

informing and influencing all aspects of the user experience. If we peel back the layers of this experience, we can begin to understand how these decisions are made". The five layers have been rephrased and merged where appropriate to better encapsulate the logic behind the concept of human-AI teaming.

The **Interface** focuses on the design of the User Interface (UI) through which the human operator and AI system communicate. It describes the **modes of interaction** (such as voice commands, text-based inputs, or gestures), the **physical or digital interface** (whether the interaction takes place through touchscreens, control panels, or wearable devices), and the **data visualisation** components (how information such as real-time data, system status, or decision recommendations are presented to the operator).

The **Interaction** layer defines the dynamics of **communication and interaction** between the human operator and the AI system. It includes defining how tasks are coordinated, the procedures for information exchange or action, and strategies for conflict management (how to resolve discrepancies when the AI's suggestions conflict with the human's actions or judgement). Additionally, this level considers the ergonomics of collaboration (ensuring the interaction is comfortable, efficient, and minimises cognitive and physical strain on the human operator). It also covers transparency (the AI's ability to explain and document its reasoning) and **explainability** (how understandable the AI's decisions are to the operator), which are essential for trust and effective teamwork.

Teaming instead defines the structure of the HAT. It defines the roles and responsibilities of each team member (what tasks are automated by the AI versus controlled by the human) and addresses the authority structure (who has the final say in decision-making) and hierarchy (how decisions are escalated or deferred). It also includes expectations on the shared mental models (the common understanding of goals, tasks, and situational awareness between the operator and AI). The design at this level is made to ensure that the human and AI can effectively work together toward shared goals, with clear responsibilities and seamless collaboration.

3.4. Application of the HATIM in a design context

Having established the model and its internal relationships, it becomes crucial to outline **initial guidelines** for its integration into the design and development workflow. Without these guidelines, the model risks remaining a passive/descriptive framework rather than an active design and development tool.

To ensure its practical use, the model should align with the User-Centred Design (UCD) workflow, applying it at various stages of the design process. Importantly, different aspects of the model will be relevant in different stages, influenced by the specific design context and technical constraints. It is therefore the designer's responsibility to understand which parts of the model may be fixed and which are instead open to be defined and specified.

Fixed elements may be predefined or inherited constraints that must be documented as inputs to the design process (e.g., certain hardware requirements, operational environments, or features dictated by the software vendor providing the AI system) while **flexible elements** are open for further definition and specification, allowing designers to tailor the AI-human teaming solution to the specific needs of the task, workflow, or environment.

Funded by the European Union

3.4.1. Example of HATIM application in a UCD design process

The following section illustrates a suggested workflow of application of the HATIM in the development of a HAT. The example represented in the various steps is based on a fictitious scenario where a newly developed AI solution is introduced in an existing production context with the aim of teaming expert operators with an AI agent.

The UCD approach typically involves four key phases in each iteration. First, designers collaborate to understand and document the context in which users will interact with the system. Next, they identify and define the users' specific requirements. Following this, the design phase begins, where the team develops potential solutions. Finally, in the evaluation phase, the design is assessed against the users' context and requirements to determine how effectively it meets their needs.

When dealing with HAT design it is recommend having an extensive initial phase (Figure 9) in which **Context and Requirements** are mapped. This may include surveying the environment in which the Team will operate, understanding which relevant regulations apply and, most importantly, map out which parts of the model are already predefined.

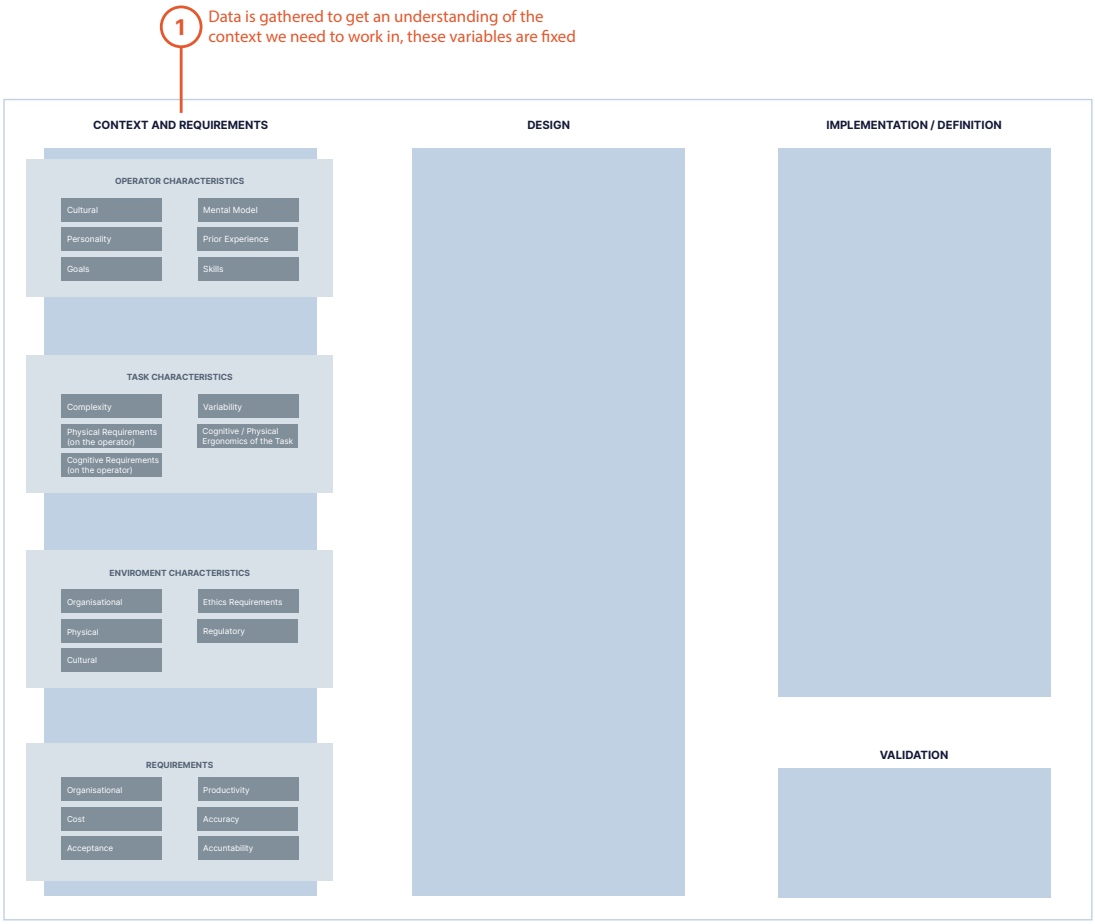


Figure 9 – In the first phase of the application, data is collected from the context and stakeholders to define the basis of our design.

Funded by the European Union

Figures 9 to 12 illustrate the context, which, as previously mentioned, includes pre-existing workforce participating in a HAT. Because of this, some Operator characteristics are predetermined. In contrast, the Agent is here considered as being purposely developed and thus, within reason and technical limitations, it can freely be defined. This is why the Agent characteristics are entirely positioned in the Implementation and Definition section. Note also how part of the Operator characteristics is expected to be open to definition, and thus positioned in the Implementation and Definition section. This is because the design can influence how the operator will act as a member of the team through Training or assigning them roles or responsibilities.

The **Design phase** (Figure 10) refers to the three abstraction layers (Figure 8), where the most desirable Teaming is selected and designed. This phase should define and describe how the Operator and Agent should interact and collaborate as well as how they can communicate. It should also include a definition of the roles. The output of this section should describe these elements but does not necessarily provide specifics on how the relation can be implemented. User Journey Maps, User Task Matrices, User Stories, Proto Personas and likewise Robotic Personas representing the Agent can be used to make this initial definition.

This intermediate step is needed as certain team dynamics could be achieved by manipulating different elements of the teaming model. For example, a design requirement may be *"to ensure high trust for a specific task"* between the Operator and the Agent. As mentioned in the state of the art, trust can be improved ensuring high reliability and accuracy of the Agent. If this is not possible because of technical constraints it is still possible to address the issue through Explainability or the ability of the Agent to respond to feedback.

Funded by the European Union

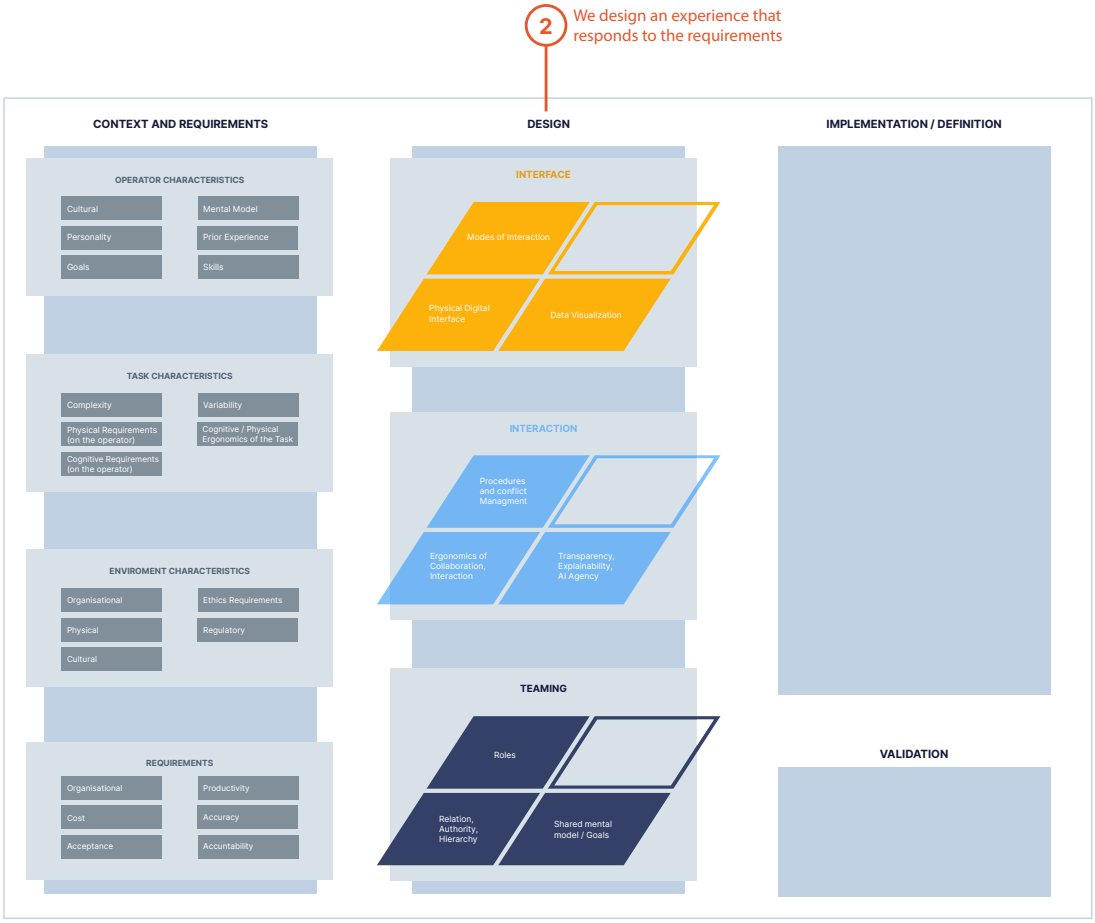


Figure 10 – The second phase of the application consists in designing the teaming to reflect the requirements set in phase one.

Based on the mapping of **Context and Requirements**, it is possible to identify the implementation variables that can be affected (Figure 11). Figures 11 and 12 use colour coding in the Implementation/Definition section to illustrate the process of translating high-level team descriptions into user and functional requirements. Since the literature can only provide suggestions for potentially favourable combinations and expected relationships between variables, designers must formulate specific details to meet design requirements. These specific details should be iteratively validated against the requirements.

Funded by the European Union

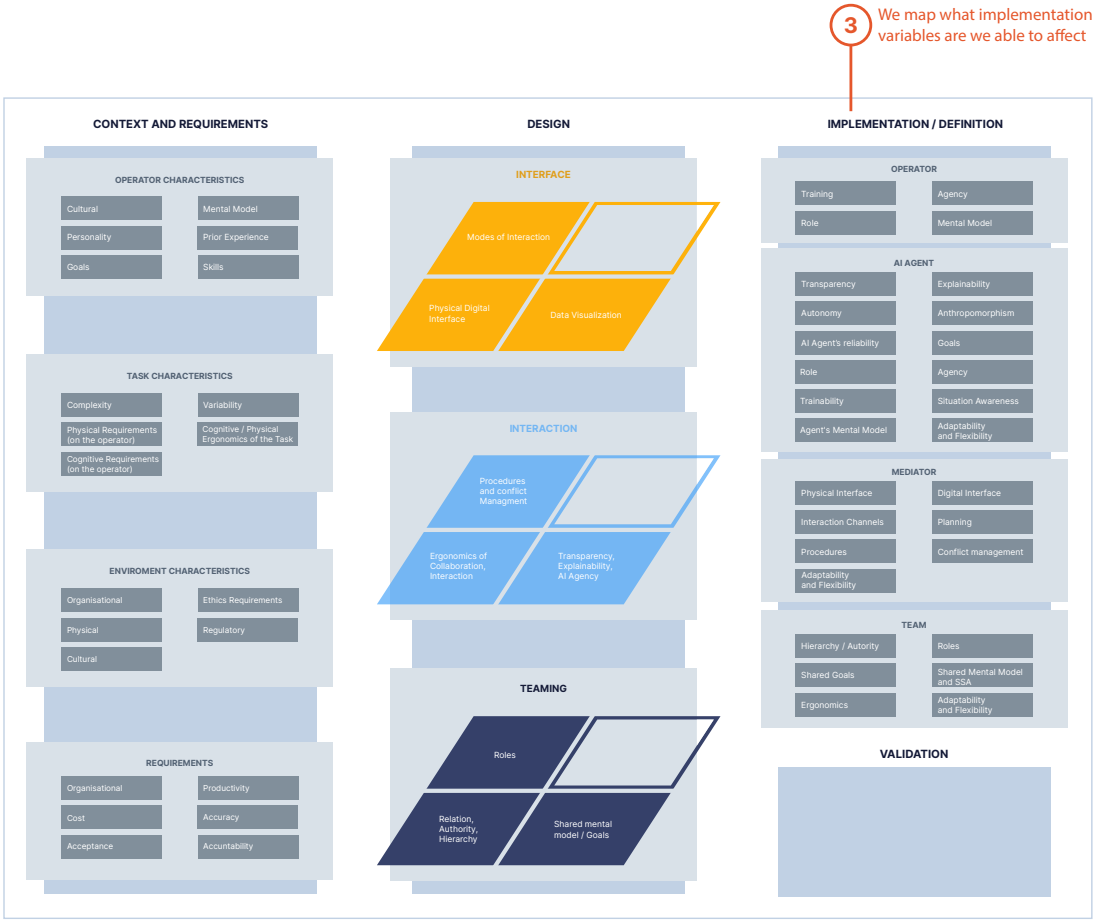


Figure 11 – The third phase of the application. The variables that are left open for definition are mapped on the right (what can be affected trough design), this is based on what variables have been collected in the first phase.

Funded by the European Union

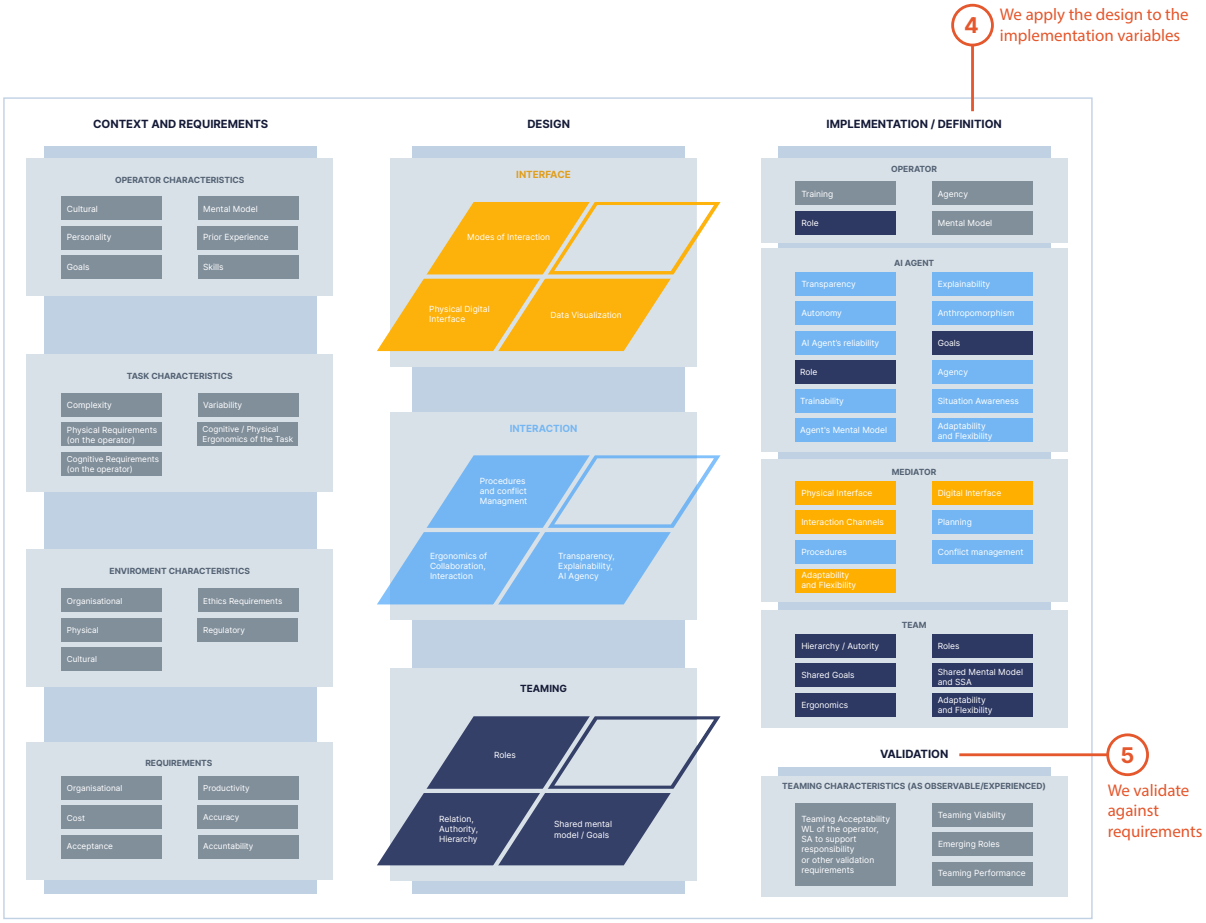


Figure 12 – The last two phases of the application. Here the design is applied to the implementation variables, represented trough the colour coding, and validated against the requirements.

For the **validation phase** (Figure 12) the variables described in Teaming characteristics should be used to evaluate the Teaming. To determine the functionality and effectiveness of the developed team, it's essential to evaluate Team Viability, Performance, and Acceptability. This evaluation should be conducted in conjunction with assessments of traditional Human Factors metrics, such as Operator Workload and Situational Awareness. Roles, hierarchy and collaboration dynamics can also be evaluated, preferably with the support of subject matter experts to evaluate the teaming effectiveness.

A blank canvas to facilitate the application of this process is available in Appendix B1.

Funded by the European Union

4. Application to a Use Case

4.1. Introduction

While the HATIM model development is grounded in established literature, it's not expected to be immediately comprehensive and capable of addressing all potential scenarios. The trial application described in this chapter represents an initial step towards validating and establishing the HATIM as a reliable tool.

The following section focuses on the application of the proposed model to the TOFAS UC1 scenario, where the task implies sorting boxes, plastic bags and spare parts for the logistics sectors. In this use case, the operator initially positions the inbound boxes to be sorted into two categories: those requiring manual handling and those suitable for robotic processing. These boxes are then placed in their designated areas. The robot and AGV then collaborate within this area until all sortable boxes have been processed. Once the robot has completed its task, the operator inserts plastic bags and spare parts that the robot cannot handle. The task, as described, is autonomous and does not necessitate direct human operator intervention. However, several potential 'what-if' scenarios are envisioned. These potential scenarios include a collapsed mosaic, obstacles in the work area, a malfunctioning robot arm causing a crate to fall, or the robot's inability to detect a suitable crate, necessitating operator intervention to reposition crates.

The application of the HATIM model has been carefully developed to address all relevant factors identified in the framework. We have assessed the current state of the UC to determine which factors are already covered and which require further attention.

The representation of the user-task matrix, along with user stories outlining the envisioned tasks is available in the annex.

4.2. Application

Starting from the **context and requirements**, we identified Initial Operator Characteristics, Task Characteristics, Agent Characteristics, Environment Characteristics, and Requirements.

This information is collected at different stages of the project. Some of them, such as the operator characteristics, have been collected during a field visit, as reported in D1.1. Others, such as task characteristics, come from a later stage, mainly reported in D1.2 Initial setup of real-world scenarios.

Funded by the European Union

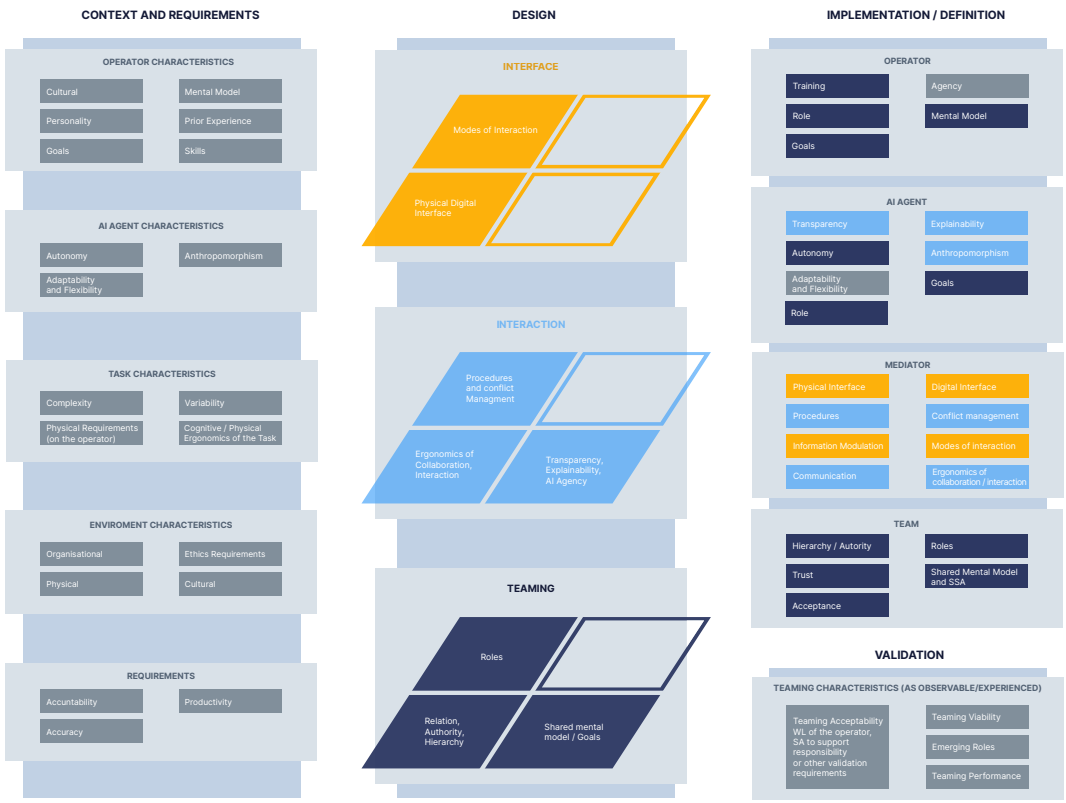


Figure 13 Schematic HAITM application to the current state of the TOFAS UC1 scenario.

4.2.1 Phase 1: Context and Requirements

Operator Characteristics

- **Cultural Characteristics**: The operator's cultural background influences their approach to teamwork and communication, shaping their interactions with both human colleagues and AI systems. In general, operators receive regular training on the new technology being implemented in the factory. The company is always up to date with the latest technology, allowing all employees to provide feedback for improvement. As an example, a new technology had just been implemented and the operators reported difficulties using it with the gloves they were provided, so the management, through the process of continuous improvement, took the feedback and promptly worked on a possible solution.
- **Goals:** The operator aims for efficient sorting and handling of materials, ensuring that the robot and AGV operate smoothly within their designated tasks.
- **Prior Experience:** The operator brings relevant prior experience in logistics such as picking and placing items, addressing and resolving issues with misplaced parts and conduct quality checks on boxes or plastic bags.
- **Personality:** The operator's personality traits, such as being proactive and detail-oriented, contribute to their ability to effectively manage tasks and communicate with the AI.

Funded by the European Union

- **Mental Model:** The operator possesses a clear mental model of the task flow, which helps them anticipate potential issues and adapt their strategies accordingly during the collaboration with the AI.
- **Skills:** The operator has developed specific skills in machine operation and problem-solving, allowing them to intervene effectively when the robot encounters limitations or failures.

Agent Characteristics

As reported in D1.2 "Initial setup of real-world scenarios" the agent characteristics are aimed at autonomous navigation, object handling, and spatial recognition capabilities in an industrial environment. Central to the system is a mobile platform, the Segway RMP omnidirectional mobile base equipped with a KUKA IIWA 14 robotic arm, which combines mobility with precision manipulation. This setup is augmented by a ZED2i camera and several grippers allowing the robot to adaptively handle various package sizes. The system's tool-changing mechanism enables automatic switching between grippers, enhancing flexibility in handling different object types.

An integrated barcode identification system plays a critical role in task execution, ensuring that each item is correctly identified, associated with the appropriate order, and placed accurately within the designated output box. To achieve this, the robot executes a series of detailed actions: it holds the box above a barcode reader, takes images to estimate box positioning, and adjusts its grasp accordingly. Once the correct item is identified, the robot navigates autonomously to the output location, places the object in a specified arrangement, and captures an image for validation. This sequential process, combining navigation, imaging, and adaptive manipulation, demonstrates a robust level of autonomy that can dynamically respond to changes in object orientation, size, and destination, optimizing productivity and accuracy in logistics or assembly scenarios.

- **Autonomy**: The system should be able to navigate autonomously around the work area and most of the boxes. The Segway RMP platform, coupled with the KUKA IIWA 14 robotic arm, should have a high degree of autonomy in performing tasks. According to the Parasuraman Autonomy Scale (2000), this system should operate at around level 7 or 8. At this level, the system should be able to read barcodes, identify objects, plan gripping points and place objects in designated locations without direct human intervention;
- **Anthropomorphism:** While not designed with anthropomorphic features, the robot's role and actions emulate human tasks in handling, sorting and pointing boxes.
- **Adaptability and flexibility:** Equipped with multiple grippers and a ZED2i camera, the system adjusts its handling strategies based on item characteristics, size, and order specifics. The automatic tool-changing feature enables seamless transitions between tasks, allowing the robot to handle varying object types and sizes efficiently. This adaptability makes the system highly versatile across different environments and task requirements, such as logistics and assembly scenarios.

Task Characteristics

- **Complexity:** The task involves a moderate level of complexity, as it requires the operator to categorize inbound boxes, manage the interaction with the robot, and address potential issues such as obstacles or equipment malfunctions. The operator must understand both the manual and automated processes involved in sorting and handling.

Funded by the European Union

- **Variability:** The task exhibits moderate variability. While the sorting process follows a defined procedure, unexpected factors such as box collapse, obstacles, or robot detection failures introduce variability. The operator must adapt to these changes and make real-time decisions to ensure efficient operations.
- **Physical Requirements (for the operator):** The task requires specific movements like bending, lifting, and positioning boxes. The operator must navigate a workspace, which may necessitate agile movements.
- **Cognitive Requirements:** The task demands a moderate level of cognitive engagement. The operator needs to monitor the robot's actions, assess potential risks, and make decisions regarding the positioning of boxes and intervention when necessary. This includes maintaining situational awareness and managing multiple priorities simultaneously.

Environmental Characteristics

- **Organizational:** The organizational environment emphasizes collaboration between human operators and automated systems, fostering a culture of teamwork and efficiency, as delineated in D1.1 Real-world scenarios and metrics for validation definition.
- **Ethics Requirements:** As reported in **Internal Report M18 (D1.4) Ethics and Liability Requirements** from an ethical and legal standpoint, UC1 presents potential issues, particularly regarding the operator's role within the new framework. Key considerations will thus focus on the impact on fundamental rights, especially concerning the worker's autonomy and dignity in their interactions with the robot. Furthermore, potential implications for the configuration of tasks and competencies will be examined.
- **Physical:** The physical environment complies with regulatory standards to ensure the safety and well-being of operators. This includes adhering to guidelines for workspace design, safety equipment, and emergency protocols.
- **Cultural:** The organizational culture promotes a positive attitude towards technology and innovation, encouraging operators to embrace automation as a tool that enhances their work rather than a threat to their roles.

Requirements

The project prioritizes safe, ergonomic human-robot collaboration, emphasizing user acceptance, skill development, and trust in AI-robotics. It includes human-centric design guidance, flexible interfaces, continuous learning for human support, and legal/ethical compliance to ensure seamless adaptation to automated production environments.

- **Productivity:** Productivity requirements include optimising robot speed and precision for tasks like box sorting and barcode verification, minimising transition times between different box types, maintaining continuous operation with minimal downtime, ensuring effective task handover between humans and robot, and enabling rapid adaptability to various scenarios.
- **Accuracy:** As reported in D 1.1 the "Non-Functional Requirements", the proposed target for the successful picking rate is >95% and for the successful part identification rate is >95%;
- **Accountability:** Involves tailored suitability for warehouse operators and performance monitors, enabling situational awareness and system-level oversight. Moderate liability risks include product and operational safety concerns, with a focus on enhancing monitoring and responsive interventions to ensure reliable robot functionality.

Funded by the European Union

4.2.2 Design

After identifying the operator, the task and the environment characteristics and requirements, the next step involved determining the desired type of teaming. To achieve this, we conducted a survey and a workshop (details in Section 3.2) involving the main partners of this UC. The outcomes helped shape the desired teaming structure, encompassing the complexities of human-AI interaction.

Teaming layer

- **Roles:** In Sections 3.3.3 and 3.3.4, it has been identified the possible overall roles that influence and shape work with an AI. Based on the envisioned scenario, the task allocations, and the limitations and opportunities of the technology, the possible roles that have been identified are the following:
  - **Collaborative AI:** The system (robot and AGV) works alongside the operator in sorting the boxes, and handling tasks that the robot can manage while the operator takes over tasks the robot cannot. There's a shared responsibility, with the AI assisting in sorting, and the operator handling leftover tasks or correcting malfunctions, indicating an active partnership and task-sharing dynamic.
  - **Independent AI:** During periods where the AI operates autonomously, handling the box sorting process without needing input or supervision from the operator.
  - **Helper User:** The operator assists the AI and robot by providing manual inputs and making adjustments when the AI encounters limitations.
  - **Passive User:** A passive user delegates all aspects of the task to the AI, due to the current state of technology the total automation of the task is not possible and this this role would lead to issues. Over reliance of the operator should therefore be avoided.
- **Shared Mental Model:** Once the possible roles were identified, we defined the type of possible shared mental model and the authority and hierarchy for the proposed scenario. This information is not directly the result of the workshop, but rather the result of the various iterations carried out throughout the design process. SMM ensures that the operator and AI (robot/AGV) understand their distinct roles, task completion expectations, and potential failure points.
  - **For the AI:** Accurately identify and sort boxes, Handle most of the sorting process autonomously, minimizing the operator's involvement.
  - **For the User:** The operator should have a clear understanding of what the AI can and cannot do. AI's Perspective: The AI should be programmed to recognize its own limitations and understand when to alert the operator for assistance or fallback strategies.
- **Shared Goals:** Shared goals involve both the AI and the human operator aiming to sort the inbound boxes. The AI autonomously handles boxes it can process, while the operator ensures completion by managing boxes the robot cannot handle, resolving issues, and maintaining overall task flow, ensuring smooth operations.
- **Hierarchy and Authority:** The AI (robot and AGV) handles sorting tasks autonomously until a problem arises. The operator retains final decision-making authority, especially in error recovery situations, such as moving misplaced boxes or addressing robot malfunctions.

Funded by the European Union

Interaction layer

On the interaction layer, at this stage of the process, we have defined three main points covered by the design, which are the ergonomics of collaboration, the procedures and conflict management, and we delineated the transparency and explainability to ensure a correct SA.

- **Ergonomics of Collaboration / Interaction:**
  - **Cognitive Ergonomics:** is crucial for reducing mental strain, ensuring that AI outputs are intuitive, easily understood, and free from information overload. For instance, the information provided by the AI could be structured in a way that prioritizes only the most critical data for the operator, with alerts and recommendations displayed clearly and simply to facilitate quick decision-making during the possible issues that are delineated during the execution of the tasks.
  - **Physical ergonomics:** focuses on creating user interfaces and control devices that the operator can comfortably and safely interact with over prolonged periods. In environments requiring physical control, like operating machinery, this could involve ergonomic design in positioning screens, control panels, and input devices to reduce strain. The scenario envisioned does not require prolonged interaction with any interface. However, we can say that during the "what-if" scenarios the operator may need to interact with the interface. The physical ergonomics of the task itself is profoundly changed, as the operator only finishes filling the boxes. Even though the parts and plastic bags he needs to interact with are heavier and slippery, they are fewer in number. Due to the increased height of the endpoint of the task, there is an improvement from a physical ergonomics point of view. The reduced number of parts to be moved could even encourage the operator to use the forklift to adjust the height of the gripping point. Increasing the start point and allowing the endpoint pose to be set to a reduced delta difference reduces the physical strain on the operator, as suggested in ISO 11228:1. In addition, the reduced delta difference between the start point and end point creates another countereffect: the reduced bending or hard bending to finalise the task. For the above reason, the fact that the operator completes the task rather than starting it is an efficient improvement that has a positive effect on the operator's well-being.

- **Procedures and conflict management:**
  - **Regular Procedures:** Routine tasks such as positioning the inbound boxes, operating after the robot and AGV have completed the task, and introducing plastic bags and spare parts after the robot completes its task.
  - **Emergency Procedures:** Steps to be taken in case of urgent situations such as a malfunction of the robot arm, obstacles in the area or safety concerns. All these possible scenarios were defined during a series of co-creation workshops, with all the partners involved.
  - **Error/Recovery Procedures:** Specific actions to take when the robot fails to detect boxes or a handled box fall. This would include manually repositioning the boxes and restarting or assisting the robot.
  - **Planning:** Procedures for setting up the working area, ensuring all resources and tools are in place before starting the task.

Funded by the European Union

- **Transparency and Explainability:** The operator must comprehend the AI's decision-making process in real time, especially during critical tasks like error recovery or when dealing with a robot's malfunctions. If only part of the AI's logic is visible or accessible (e.g., diagnostics are available but not in real-time), there might be gaps in understanding during fast-paced or complex decision-making. However, the level of explainability and transparency of the system is not fully defined yet. The process to ensure the right level of transparency and explainability should be considered in later stages of implementation according to the UCD process and testing.
  Ultimately, on the interface layer, we defined what the interface should provide to facilitate interaction and valorise teaming. The interface is the most tangible layer and the one that is directly used by the operator.

Interface Layer

- **Physical, Digital Interface:**
  - **Digital Interface:** Visual displays could be used to monitor the robot and AGV status, provide feedback and alert the operator. The interface might show areas where human intervention is needed, such as moving boxes.
  - **Physical Interface:** includes tangible components enabling direct interaction between the operator and the system. The ergonomics of emergency controls is critical to facilitate intervention during failures like robot malfunctions or box collapses. Additionally, strategically placed visual or auditory alarms can promptly alert operators to address safety risks effectively.

- **Modes of Interaction:**
  - **Sequential Collaboration**: The operator and robot work in stages, with the operator first positioning the boxes for robotic handling, followed by autonomous sorting by the robot and AGV, and the operator concluding by adding unhandled items.
  - **Supervisory Monitoring**: The operator oversees the sorting process, remaining alert for "what-if" scenarios such as obstacle removal, robotic errors, or unexpected conditions that require manual intervention.
  - **Reactive Assistance**: To ensure workflow efficiency, the operator must step into repositioning boxes or address malfunctions.

4.2.3 Implementation / Definition

In this scenario, it can be assumed that the mapping of context and requirements allows for the mapping of implementation variables that can be influenced. The process of translating a high-level description of the teaming into user and functional requirements is a task for the researcher. At this stage, we believe that it cannot be operationalised, as the literature can only suggest favourable combinations or expected relationships between variables. It is therefore the responsibility of the designers to formulate specifics that may satisfy the design requirements, and then to iterate and validate against the requirements.

Funded by the European Union

Agent

- **Adaptability and Flexibility:** The use case could benefit from dynamically adjustable roles where the AI takes a lead role in executing well-defined, repetitive tasks, while the human operator steps in more actively when complex decisions, troubleshooting, or exceptions arise. This approach would keep the human engaged in critical thinking tasks, while the AI focuses on predictable operations, reducing mental fatigue and boosting overall efficiency.
- **Autonomy:** The AI operates autonomously for routine tasks, without requiring human intervention. However, selective overrides allow human input when needed, creating a balance between autonomous function and real-time control.
- **Anthropomorphism:** Anthropomorphic cues must be carefully calibrated to balance user engagement with realistic expectations of AI capabilities. Overly humanized AI could mislead operators about its decision-making depth, while subtle anthropomorphic features can foster intuitive interactions Develop a non-verbal communication system that mimics familiar human expressions, such as visual LED cues or sound notifications, to indicate task status or need for human intervention. A flashing light for “help needed” or a specific sound for “task complete” can enhance human understanding and response time.
- **Role:** The roles in the TOFAS UC1 scenario were initially defined during a collaborative workshop, providing a foundational structure for human-robot teaming. However, there is flexibility to further refine these roles or reconfigure them in implementation to better fit specific task demands. This adaptability allows the team to valorise certain roles based on real-time requirements, optimizing the balance between autonomous robot functions and human intervention. As the project progresses, some roles may be reshaped to enhance task efficiency, improve user experience, and address evolving operational needs. This iterative approach ensures that the final role assignments align closely with both user and functional requirements.
- **Explainability:** This could be implemented into the system by having the robot that provides visual or textual feedback on its actions, helping the operator understand the AI’s rationale for item placement or failure conditions. For example, if the robot fails to sort an item due to a positioning issue, it could display a simple message on a connected interface explaining why the grasp attempt failed.
- **Transparency:** To enhance transparency in this use case, a dashboard displaying the robot’s task progress, system status, and any flagged anomalies could be accessible to the human operator. For example, if the robot encounters an obstacle or malfunctions, the dashboard could immediately highlight the issue and suggest solutions.
- **Goals:** The agent’s implementation goals focus on autonomous task execution, adaptive role responsiveness for complex scenarios, and clear, explainable feedback to operators. Iterative adjustments will align agent actions with operational requirements and enhance human-agent collaboration.

Operator

- **Training:** To prepare operators for effective collaboration with AI, future training should emphasize the development of core competencies in interpreting AI actions, understanding system feedback, and adapting to AI-driven processes. This training should incorporate flexible learning approaches, such as simulations and real-time scenario exercises, allowing operators to engage with various interaction styles and response protocols. By fostering these skills, training can improve overall

Funded by the European Union

situational awareness (SA), enabling operators to recognize shifts in AI behaviour and adapt as necessary. In this way, the training ensures that operators are equipped to make well-informed decisions, enhancing both safety and efficiency in human-AI teamwork.

- **Agency:** Empowering the operator with decision-making authority at key points in the workflow is essential to establish trust and engagement with the AI. Providing options for manual override and intervention, especially in instances of unexpected AI behaviour (e.g., sorting errors or task prioritization issues).
- **Role:** The operator's role, as defined in preliminary workshops, is largely supervisory with occasional intervention in error correction and task initiation. However, this role can be dynamically re-evaluated in implementation, with the potential to expand the operator's involvement in situational adjustments, problem-solving, or quality assurance tasks as the system evolves. This flexibility allows the role to adapt as operational needs shift, ensuring that the human presence remains essential and valuable within the AI-assisted environment.
- **Mental Model:** Implementation strategies should focus on developing intuitive interfaces and clear visual feedback that reflect the AI's current status, task progress, and intended actions. For example, using visual indicators, color-coded alerts or progress bars, can help the operator easily understand the system's state and anticipate its actions.
- **Goals:** The operator's implementation goals focus on ensuring effective collaboration with the AI through targeted training, empowering decision-making authority, and fostering adaptability in role definition.

Mediator

- **Communication/ Information modulation:** Must prioritize quality by delivering clear, actionable insights without overwhelming the operator. The quantity of information should be adjusted to avoid cognitive overload while providing all essential data, allowing the operator to maintain situational awareness and respond as needed.
- **Physical interface:** At this stage, evaluating the implementation of the physical interface is challenging, as the design has already incorporated preventive measures to minimize ergonomic issues and interaction barriers. However, it is possible that unforeseen challenges may arise during implementation. Our role will be to closely monitor these interactions to identify any emerging issues and ensure the physical interface remains ergonomically sound, prioritizing user comfort and accessibility in all operator interactions with the system.
- **Digital Interface:** This should deliver essential data through a clear, user-friendly dashboard that emphasizes readability and intuitive navigation. To ensure accessibility, the interface should use culturally sensitive language, employing familiar or easy-to-learn terms to facilitate seamless understanding and reduce learning curves. This approach helps bridge diverse operator backgrounds, fostering a more inclusive and effective interaction with the AI system.
- **Modes of interaction:** Interaction will primarily occur through the digital interface, providing a structured and accessible communication between the operator and AI. However, there remains the opportunity to uncover and address latent or "hidden" needs that may only become apparent during the implementation phase, allowing for adjustments that enhance usability and adaptability in real-world settings.
- **Procedure/Conflict management:** Procedures for managing conflicts are defined in this phase of the project, including predefined protocols for error handling, decision-making, and conflict

Funded by the European Union

resolution to ensure smooth human-AI collaboration. These procedures will need to be evaluated and refined later, based on data gathered during the implementation phase.

- **Ergonomics of Collaboration/interaction:** The ergonomics of collaboration will be later evaluated to ensure that the physical and digital interfaces support the operator's comfort and efficiency. The goal is to minimize physical strain and cognitive overload, enhancing the overall user experience in interacting with the system.

Team

- **Trust and Acceptance:** Are foundational to the human-AI relationship, where the operator's confidence in the AI's capabilities directly affects team dynamics. Trust is nurtured by the AI's consistent, reliable performance and transparency, especially when explaining decision-making or alerting the operator to potential issues. When the AI reliably demonstrates its effectiveness, the human operator is more likely to accept its guidance and integrate its suggestions, strengthening collaborative efforts.
- **Roles:** Are initially defined, but they remain flexible and subject to change as the system is implemented and refined. The operator's role may evolve from a supervisory position to a more active participant in problem-solving and decision-making, depending on the operational needs and the AI's capabilities.
- **Shared Mental Model and SSA:** To develop a shared mental model, the system should provide real-time feedback, interactive training, and transparent explanations of AI decisions. Training should focus on familiarizing operators with AI behaviour, decision-making processes, and error handling.
- **Authority and Hierarchy:** The hierarchy is designed to be flexible, allowing the operator to step in and take control when necessary, while the AI autonomously handles predefined tasks. This structure will be iteratively refined, with continuous feedback from both the system and the operator, ensuring a balance of control that optimizes efficiency, autonomy, and operator trust over time.

4.2.4 Validation

Parameters for the validation: Given the general aim of the innovation project, it is possible to define a series of variables that can define the quality and acceptability of the solution regarding the teaming, specifically as observable or experienced factors are important to highlight the **Teaming Acceptability, WL of the operator, SA to support responsibility, Teaming Viability, Emerging Roles, Teaming Performance.** For the highlighted factors we envisioned possible validation metrics on the proposed solution that will have a significant impact on the task and improve the working condition of human operator, so:

- Effective reduction of operator's ergonomics overload due to bending and hard bending for pick and place and repetitive actions, to consider ISO 11228:1
- Effective reduction of man working hours dedicated to the task (maximization of time available for other tasks).
- Effective use of machine hours (minimization of idle time).
- Assessment of the operator's cognitive overload and ability to understand the different situations that can arise in the different scenarios, and evaluation of SA in error/failure scenarios.

Funded by the European Union

- Evaluation of the operator's acceptability and trust in the AI, willingness to trust the robot to operate independently and willingness to support the robot when required.
- Evaluate how well the operator and AI roles are defined, understood, and executed in real-time scenarios. This includes assessing the clarity of task delegation, the adaptability of both parties in dynamic situations, and the operator's ability to comfortably switch between supervisory and active roles as needed. The operator's trust in the AI's autonomy, willingness to allow the robot to operate independently, and readiness to intervene, when necessary, should also be considered. Key metrics will include the operator's satisfaction with the defined roles, the perceived efficiency of the division of labour, and the system's ability to adjust roles based on evolving tasks or environmental conditions.

4.2.5 Conclusions

The results of the test application of the HATIM model and its application methodology support its utility in organising and evaluating the information regarding HATs and their context of the application. The tool allows mapping the available information, fixed points and design opportunities. Used early in the early stages of a process it will elicit key questions that will support it throughout its development.
By defining adaptive roles for humans and AI through narrative-driven approaches, the method promotes a collaborative environment where humans and intelligent agents can dynamically share tasks, enhancing overall productivity and responsiveness. This narrative-based role delineation encourages seamless integration, facilitating not only technical adaptability but also user acceptance and trust.
The method helps to map out available information and generate relevant questions but does not provide recommendations or safeguards, it is still up to the design team to test and validate. This stage of the project does not allow for validation of the method as it is not possible to directly observe the teaming and compare it to what was designed.
Future work will consist in applying the model to other UCs and collecting feedback on its use and utility, revising the variables and methods when necessary.

Funded by the European Union

A.Bibliography

Admoni, H., & Scassellati, B. (2017). Social Eye Gaze in Human-Robot Interaction: A Review. *Journal of Human-Robot Interaction*, *6*(1), 25. https://doi.org/10.5898/JHRI.6.1.Admoni

Akash, K., Reid, T., & Jain, N. (2019). Improving Human-Machine Collaboration Through Transparency-based Feedback – Part II: Control Design and Synthesis. *IFAC-PapersOnLine*, *51*(34), 322–328. https://doi.org/10.1016/j.ifacol.2019.01.026

Ali, A., Tilbury, D. M., & Robert Jr, L. P. (2022). *Considerations for Task Allocation in Human-Robot Teams* (No. arXiv:2210.03259). arXiv. http://arxiv.org/abs/2210.03259

Arkin, R. C. (2009). *Governing lethal behavior in autonomous robots*. CRC Press.

Breazeal, C., Kidd, C. D., Thomaz, A. L., Hoffman, G., & Berlin, M. (2005). Effects of nonverbal communication on efficiency and robustness in human-robot teamwork. *2005 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 708–713. https://doi.org/10.1109/IROS.2005.1545011

Burghardt, G. M. (2017). Anthropomorphism. In J. Vonk & T. Shackelford (Eds.), *Encyclopedia of Animal Cognition and Behavior* (pp. 1–4). Springer International Publishing. https://doi.org/10.1007/978-3-319-47829-6_1042-1

Campion, M. A., Medsker, G. J., & Higgs, A. C. (1993). Relations Between Work Group Characteristics and Effectiveness: Implications for Designing Effective Work Groups. *Personnel Psychology*, *46*(4), 823–847. https://doi.org/10.1111/j.1744-6570.1993.tb01571.x

Chen, J. Y., & Barnes, M. J. (2013). *Human-Agent Teaming for Multi-Robot Control: A Literature Review:* Defense Technical Information Center. https://doi.org/10.21236/ADA583900

Chiou, E. K., & Lee, J. D. (2023). Trusting Automation: Designing for Responsivity and Resilience. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *65*(1), 137–165. https://doi.org/10.1177/00187208211009995

Cummings, M. L., & Li, S. (2021). Subjectivity in the Creation of Machine Learning Models. *Journal of Data and Information Quality*, *13*(2), 1–19. https://doi.org/10.1145/3418034

Demir, M., Cooke, N. J., & Amazeen, P. G. (2018). A conceptual model of team dynamical behaviors and performance in human-autonomy teaming. *Cognitive Systems Research*, *52*, 497–507. https://doi.org/10.1016/j.cogsys.2018.07.029

Duffy, B. R. (2003). Anthropomorphism and the social robot. *Robotics and Autonomous Systems*, *42*(3–4), 177–190. https://doi.org/10.1016/S0921-8890(02)00374-3

Funded by the European Union

Dzindolet, M. T., Peterson, S. A., Pomranky, R. A., Pierce, L. G., & Beck, H. P. (2003). The role of trust in automation reliance. *International Journal of Human-Computer Studies*, *58*(6), 697–718. https://doi.org/10.1016/S1071-5819(03)00038-7

Eddy, T. J., Gallup, G. G., & Povinelli, D. J. (1993). Attribution of Cognitive States to Animals: Anthropomorphism in Comparative Perspective. *Journal of Social Issues*, *49*(1), 87–101. https://doi.org/10.1111/j.1540-4560.1993.tb00910.x

Endsley, M. R. (1988). Design and Evaluation for Situation Awareness Enhancement. *Proceedings of the Human Factors Society Annual Meeting*, *32*(2), 97–101. https://doi.org/10.1177/154193128803200221

Endsley, M. R. (1995). Toward a Theory of Situation Awareness in Dynamic Systems. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *37*(1), 32–64. https://doi.org/10.1518/001872095779049543

Endsley, M. R., Bolté, B., & Jones, D. G. (2003). *Designing for situation awareness: An approach to user-centered design*. Taylor & Francis.

Endsley, M. R., & Jones, W. M. (2001). *A Model of Inter and Intra-Team Situation Awareness: Implications for Design, Training and Measurement*.

Epley, N., Waytz, A., & Cacioppo, J. T. (2007). On seeing human: A three-factor theory of anthropomorphism. *Psychological Review*, *114*(4), 864–886. https://doi.org/10.1037/0033-295X.114.4.864

Eyssel, F., Kuchenbrandt, D., & Bobinger, S. (2012). *'If You Sound Like Me, You Must Be More Human': On the Interplay of Robot and User Features on Human- Robot Acceptance and Anthropomorphism*.

Ezer, N., Bruni, S., Cai, Y., Hepenstal, S. J., Miller, C. A., & Schmorrow, D. D. (2019). Trust Engineering for Human-AI Teams. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *63*(1), 322–326. https://doi.org/10.1177/1071181319631264

Forster, Y., Naujoks, F., & Neukum, A. (2016). Your Turn or My Turn?: Design of a Human-Machine Interface for Conditional Automation. *Proceedings of the 8th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, 253–260. https://doi.org/10.1145/3003715.3005463

Fussell, S. R., Kiesler, S., Setlock, L. D., & Yew, V. (2008). How people anthropomorphize robots. *Proceedings of the 3rd ACM/IEEE International Conference on Human Robot Interaction*, 145–152. https://doi.org/10.1145/1349822.1349842

Funded by the European Union

Garrett, J. J. (2008). *The elements of user experience: User-centered design for the Web* (Repr.). New Riders [u.a.].

Goetz, J., Kiesler, S., & Powers, A. (2003). Matching robot appearance and behavior to tasks to improve human-robot cooperation. *The 12th IEEE International Workshop on Robot and Human Interactive Communication, 2003. Proceedings. ROMAN 2003.*, 55–60. https://doi.org/10.1109/ROMAN.2003.1251796

Gustavsson, L., Augustsson, S., & Hult, H. V. (2022). *Trigger Points of Fear and Distrust in Human-Robot Interaction: The Case of Cooperative Manufacturing*. *13*.

Guthrie, S. E. (1995). *Faces in the clouds: A new theory of religion*. Oxford Univ. Pr.

Hagos, D. H., & Rawat, D. B. (2022). Recent Advances in Artificial Intelligence and Tactical Autonomy: Current Status, Challenges, and Perspectives. *Sensors*, *22*(24), 9916. https://doi.org/10.3390/s22249916

Hancock, P. A., Billings, D. R., Schaefer, K. E., Chen, J. Y. C., De Visser, E. J., & Parasuraman, R. (2011). A Meta-Analysis of Factors Affecting Trust in Human-Robot Interaction. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *53*(5), 517–527. https://doi.org/10.1177/0018720811417254

Hoff, K. A., & Bashir, M. (2015). Trust in Automation: Integrating Empirical Evidence on Factors That Influence Trust. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *57*(3), 407–434. https://doi.org/10.1177/0018720814547570

Huang, C.-M., & Thomaz, A. L. (2010). *Joint Attention in Human-Robot Interaction*.

Kahn, P. H., Ishiguro, H., Friedman, B., Kanda, T., Freier, N. G., Severson, R. L., & Miller, J. (2007). What is a Human?: Toward psychological benchmarks in the field of human–robot interaction. *Interaction Studies. Social Behaviour and Communication in Biological and Artificial Systems*, *8*(3), 363–390. https://doi.org/10.1075/is.8.3.04kah

Kiesler, S., Powers, A., Fussell, S. R., & Torrey, C. (2008). Anthropomorphic Interactions with a Robot and Robot–like Agent. *Social Cognition*, *26*(2), 169–181. https://doi.org/10.1521/soco.2008.26.2.169

Lee, J. D., & See, K. A. (2004). Trust in Automation: Designing for Appropriate Reliance. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *46*(1), 50–80. https://doi.org/10.1518/hfes.46.1.50_30392

Lemaignan, S., Fink, J., & Dillenbourg, P. (2014). The dynamics of anthropomorphism in robotics. *Proceedings of the 2014 ACM/IEEE International Conference on Human-Robot Interaction*, 226–227. https://doi.org/10.1145/2559636.2559814

Funded by the European Union

Lyons, J. B., Sycara, K., Lewis, M., & Capiola, A. (2021). Human–Autonomy Teaming: Definitions, Debates, and Directions. *Frontiers in Psychology*, *12*, 589585. https://doi.org/10.3389/fpsyg.2021.589585

Lyons, J. B., & Wynne, K. T. (2021). Human-machine teaming: Evaluating dimensions using narratives. *Human-Intelligent Systems Integration*, *3*(2), 129–137. https://doi.org/10.1007/s42454-020-00019-7

Malle, B. F., & Ullman, D. (2021). A multidimensional conception and measure of human-robot trust. In *Trust in Human-Robot Interaction* (pp. 3–25). Elsevier. https://doi.org/10.1016/B978-0-12-819472-0.00001-0

Markowitz, J. A. (2017). *Speech and Language for Acceptance of Social Robots: An Overview*. *2*.

McDermott, P., Dominguez, C., Ryan, M., Trahan, I., Nelson, A., & Kasdaglis, N. (2018). Human-Machine Teaming Systems Engineering Guide. *MP180941*, *17*.

Metzger, U., & Parasuraman, R. (2005). Automation in Future Air Traffic Management: Effects of Decision Aid Reliability on Controller Performance and Mental Workload. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *47*(1), 35–49. https://doi.org/10.1518/0018720053653802

Miller, C., Funk, H., Wu, P., Goldman, R., Meisner, J., & Chapman, M. (2005). *The Playbook (Trademark) Approach to Adaptive Automation:* Defense Technical Information Center. https://doi.org/10.21236/ADA444096

Milstein, T. (2011). Nature Identification: The Power of Pointing and Naming. *Environmental Communication*, *5*(1), 3–24. https://doi.org/10.1080/17524032.2010.535836

Mirnig, N., Stollnberger, G., Miksch, M., Stadler, S., Giuliani, M., & Tscheligi, M. (2017). To Err Is Robot: How Humans Assess and Act toward an Erroneous Social Robot. *Frontiers in Robotics and AI*, *4*, 21. https://doi.org/10.3389/frobt.2017.00021

Mosier, K. L., & Skitka, L. J. (1996). Human decision makers and automated decision aids: Made for each other? In *Automation and human performance: Theory and applications.* (pp. 201–220). Lawrence Erlbaum Associates, Inc.

Mosier, K. L., Skitka, L. J., Heers, S., & Burdick, M. (1998). Automation Bias: Decision Making and Performance in High-Tech Cockpits. *The International Journal of Aviation Psychology*, *8*(1), 47–63. https://doi.org/10.1207/s15327108ijap0801_3

National Academies of Sciences, Engineering, and Medicine, Committee on Human-System Integration Research Topics for the 711th Human Performance Wing of the Air Force Research Laboratory, Board on Human-Systems Integration, & Division of Behavioral and Social Sciences and Education. (2022). *Human-AI Teaming: State-of-the-Art and Research Needs* (p. 26355). National Academies Press. https://doi.org/10.17226/26355

Funded by the European Union

O'Neill, T., McNeese, N., Barron, A., & Schelble, B. (2022). Human–Autonomy Teaming: A Review and Analysis of the Empirical Literature. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *64*(5), 904–938. https://doi.org/10.1177/0018720820960865

Oppermann, R. (Ed.). (1994). *Adaptive user support: Ergonomic design of manually and automatically adaptable software*. Erlbaum.

Panagou, S., Neumann, W. P., & Fruggiero, F. (2024). A scoping review of human robot interaction research towards Industry 5.0 human-centric workplaces. *International Journal of Production Research*, *62*(3), 974–990. https://doi.org/10.1080/00207543.2023.2172473

Parasuraman, R., & Riley, V. (1997). Humans and Automation: Use, Misuse, Disuse, Abuse. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *39*(2), 230–253. https://doi.org/10.1518/001872097778543886

Parasuraman, R., Sheridan, T. B., & Wickens, C. D. (2000). A model for types and levels of human interaction with automation. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, *30*(3), 286–297. https://doi.org/10.1109/3468.844354

Pinto, A., Sousa, S., Simões, A., & Santos, J. (2022). A Trust Scale for Human-Robot Interaction: Translation, Adaptation, and Validation of a Human Computer Trust Scale. *Human Behavior and Emerging Technologies*, *2022*, 1–12. https://doi.org/10.1155/2022/6437441

Portal, S., Abratt, R., & Bendixen, M. (2018). Building a human brand: Brand anthropomorphism unravelled. *Business Horizons*, *61*(3), 367–374. https://doi.org/10.1016/j.bushor.2018.01.003

Propp, V. J., Wagner, L. A., & Dundes, A. (1968). *Morphology of the folktale* (2. ed., rev.22. pbk. print). Univ. of Texas Press.

Ragni, M., Rudenko, A., Kuhnert, B., & Arras, K. O. (2016). Errare humanum est: Erroneous robots in human-robot interaction. *2016 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, 501–506. https://doi.org/10.1109/ROMAN.2016.7745164

Riek, L. D., Rabinowitch, T.-C., Chakrabarti, B., & Robinson, P. (2009). Empathizing with robots: Fellow feeling along the anthropomorphic spectrum. *2009 3rd International Conference on Affective Computing and Intelligent Interaction and Workshops*, 1–6. https://doi.org/10.1109/ACII.2009.5349423

Robinette, P., Li, W., Allen, R., Howard, A. M., & Wagner, A. R. (2016). Overtrust of robots in emergency evacuation scenarios. *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, 101–108. https://doi.org/10.1109/HRI.2016.7451740

Funded by the European Union

Roth, E. M., Sushereba, C., Militello, L. G., Diiulio, J., & Ernst, K. (2019). Function Allocation Considerations in the Era of Human Autonomy Teaming. *Journal of Cognitive Engineering and Decision Making*, *13*(4), 199–220. https://doi.org/10.1177/1555343419878038

Salas, E., Dickinson, T. L., Converse, S. A., & Tannenbaum, S. I. (1992). Toward an understanding of team performance and training. In *Teams: Their training and performance.* (pp. 3–29). Ablex Publishing.

Schlosser, M. (2019). Agency. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Winter 2019). Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/win2019/entries/agency/

Schmidbauer, C., Hader, B., & Schlund, S. (2021). Evaluation of a Digital Worker Assistance System to enable Adaptive Task Sharing between Humans and Cobots in Manufacturing. *Procedia CIRP*, *104*, 38–43. https://doi.org/10.1016/j.procir.2021.11.007

Short, E., Hart, J., Vu, M., & Scassellati, B. (2010). No fair!!: An interaction with a cheating robot. *Proceeding of the 5th ACM/IEEE International Conference on Human-Robot Interaction - HRI '10*, 219. https://doi.org/10.1145/1734454.1734546

Tamagawa, R., Watson, C. I., Kuo, I. H., MacDonald, B. A., & Broadbent, E. (2011). The Effects of Synthesized Voice Accents on User Perceptions of Robots. *International Journal of Social Robotics*, *3*(3), 253–262. https://doi.org/10.1007/s12369-011-0100-4

Tuškej, U., & Podnar, K. (2018). Consumers' identification with corporate brands: Brand prestige, anthropomorphism and engagement in social media. *Journal of Product & Brand Management*, *27*(1), 3–17. https://doi.org/10.1108/JPBM-05-2016-1199

Wickens, C. D., Clegg, B. A., Vieane, A. Z., & Sebok, A. L. (2015). Complacency and Automation Bias in the Use of Imperfect Automation. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *57*(5), 728–739. https://doi.org/10.1177/0018720815581940

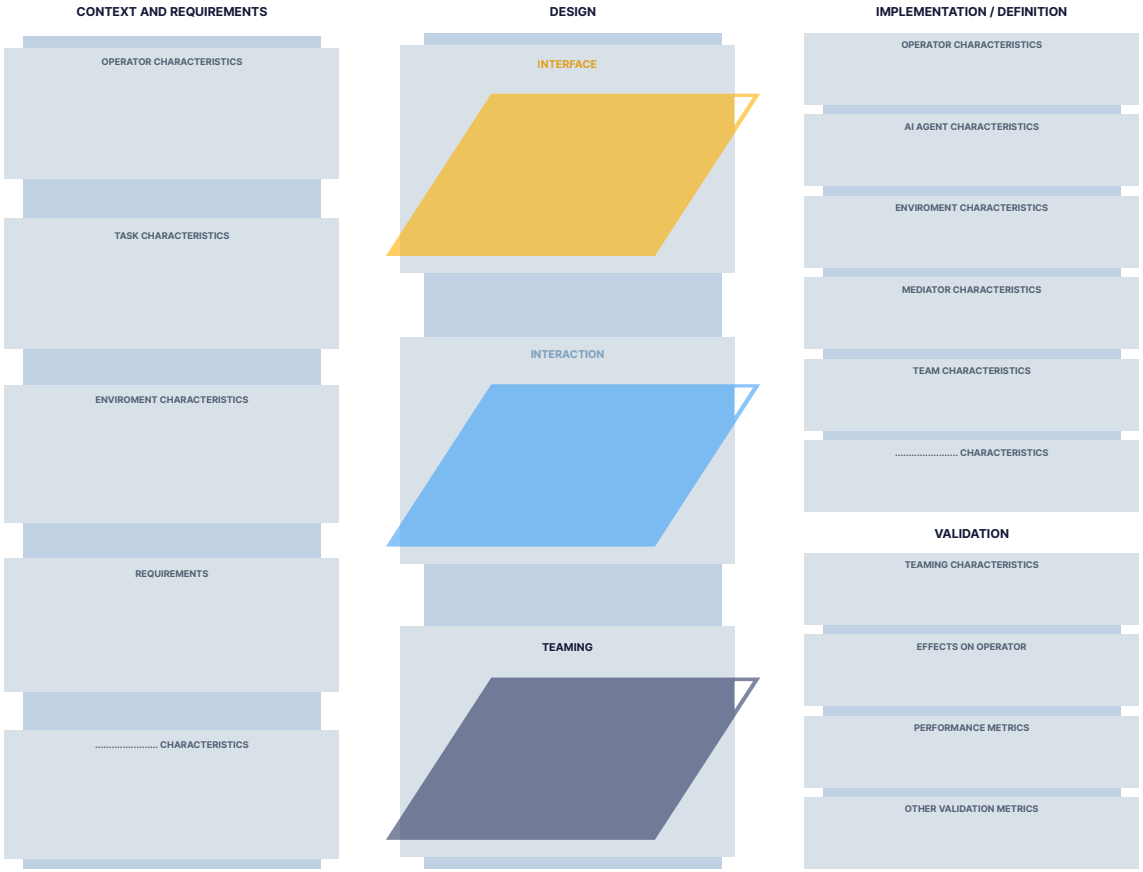
Xu, W., & Gao, Z. (2023). *Applying HCAI in developing effective human-AI teaming: A perspective from human-AI joint cognitive systems* (Version 5). arXiv. https://doi.org/10.48550/ARXIV.2307.03913

Zanatto, D. (2019). *When do we cooperate with robots?* University of Plymouth. https://doi.org/10.24382/1015

Funded by the European Union

B. Appendix

B.1. Blank Application Canvas



Funded by the European Union

B.2. Survey

Roles in the Human-AI Teaming Interaction Model

This survey aims to identify common definitions of the different roles that humans and AI can take when interacting in a team.
You will be presented with a set of role definitions for both humans and AI and asked for insights, with a particular focus on the various use cases.

Please read this section carefully, as it outlines key role definitions for both humans and AI when interacting in a team, which will be used as the basis for the survey.

* Indicates required question

Funded by the European Union

B.2. Survey

Teaming between humans and AI agents refers to the collaboration between humans and autonomous agents working interdependently toward a shared goal. For Teaming to function effectively, both humans and autonomous agents must be defined as distinct team members with specific roles. Each member, whether human or autonomous, contributes to the task with different degrees of autonomy, where the autonomous agent is not merely following instructions but also making decisions, adapting to changes, and exercising judgment.

Autonomy is central to the concept of teaming. **Agents must possess agency**, which allows them to act independently and fulfill parts of tasks that would otherwise require a human. This shift from automation to autonomous agents is crucial to the design of a true teammate, rather than just a tool. The level of autonomy can vary, but for effective teaming, the agent must possess sufficient decision-making capabilities to interact meaningfully with human counterparts and contribute to achieving shared outcomes.

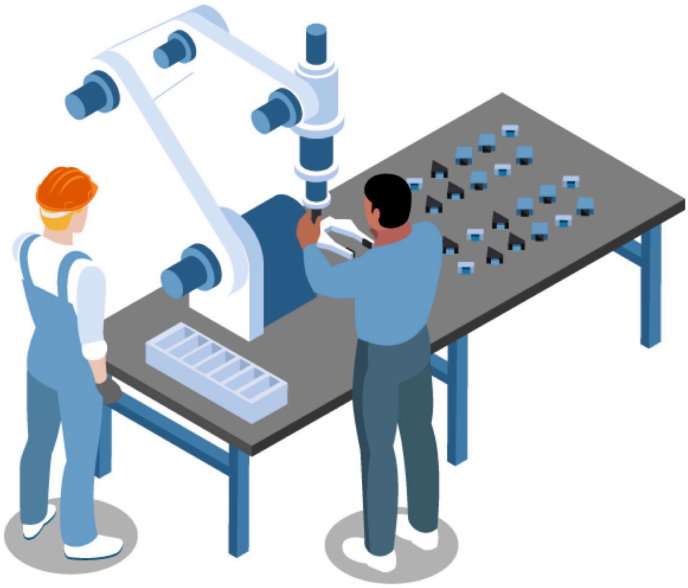
The definition of the different roles played by humans and autonomous agents is foundational to the construction of a human-AI teaming interaction model. Therefore, a set of roles is defined below, for you to assess and revise.

Human roles include:
Collaborative: Humans work with AI in shared decision-making.
Independent: Humans take full control, guiding the process, AI tends to be excluded or under-utilised.
Helper: Supporting AI with oversight and input.
Supervisor: Monitoring AI without direct intervention.
Clumsy: Human actions unintentionally hinder the process.
Passive: Minimal engagement or involvement in the task.
Antagonist: Resisting or causing conflicts in the process.

AI roles include:
Collaborative AI: Works alongside humans to assist in decisions.
Independent AI: Takes primary responsibility in task execution, human tends to be excluded from the process.
Helper AI: Supports by completing routine tasks or providing insights.
Supervising AI: Observes and advises on actions.
Faulty AI: Causes errors or requires human intervention due to malfunction or incorrect decisions.

Funded by the European Union

B.2. Survey



AI Agent Roles

In this section, the role of the AI Agents is considered.

Funded by the European Union

B.2. Survey

1. **Collaborative AI** works in harmony with human users, actively sharing tasks and decision-making. It relies on frequent communication with users, adjusting to feedback and providing support for a balanced partnership in completing tasks. *

For each use case, please state whether this role is desirable, undesirable, not applicable, or if you cannot answer.

Mark only one oval per row.

	Desirable	Undesirable	Not Applicable for the UC	Cannot answer
Tofas UC1 - Spare parts delivery preparation	◯	◯	◯	◯
Tofas UC2a - Kitting	◯	◯	◯	◯
Tofas UC2b - Pre-assembly	◯	◯	◯	◯
PCL UC3 - Handling for mass customization	◯	◯	◯	◯
TCA UC4 – Packaging operation in food sector	◯	◯	◯	◯
INFAR UC5 – Fixtureless assembly in hand tool	◯	◯	◯	◯
ULMA UC6 – Pallet to pallet order preparation	◯	◯	◯	◯
ULMA UC7 – Box to box order preparation				

Funded by the European Union

B.2. Survey

ULMA UC7 – Box to box order preparation				

2. If you like, please justify your choices.

Funded by the European Union

B.2. Survey

3. The **Independent AI** takes primary responsibility for completing tasks, making decisions, and leading the workflow. As it evolves into Independent AI, it autonomously handles complex operations, human tends to be excluded from the process. *

For each use case, please state whether this role is desirable, undesirable, not applicable, or if you cannot answer.

Mark only one oval per row.

	Desirable	Undesirable	Not Applicable for the UC	Cannot answer
Tofas UC1 - Spare parts delivery preparation	◯	◯	◯	◯
Tofas UC2a - Kitting	◯	◯	◯	◯
Tofas UC2b - Pre-assembly	◯	◯	◯	◯
PCL UC3 - Handling for mass customization	◯	◯	◯	◯
TCA UC4 – Packaging operation in food sector	◯	◯	◯	◯
INFAR UC5 – Fixtureless assembly in hand tool	◯	◯	◯	◯
ULMA UC6 – Pallet to pallet order preparation	◯	◯	◯	◯
ULMA UC7 – Box to box order preparation				

Funded by the European Union

B.2. Survey

ULMA UC7 – Box to box order preparation				

4. If you like, please justify your choices.

B.2. Survey

5. **Helper AI** provides assistance by performing secondary tasks, simplifying user operations. It takes care of repetitive or time-consuming actions, offers suggestions, or analyzes data to make the user's job easier, enhancing overall productivity. *

For each use case, please state whether this role is desirable, undesirable, not applicable, or if you cannot answer.

Mark only one oval per row.

	Desirable	Undesirable	Not Applicable for the UC	Cannot answer
Tofas UC1 - Spare parts delivery preparation	◯	◯	◯	◯
Tofas UC2a - Kitting	◯	◯	◯	◯
Tofas UC2b - Pre-assembly	◯	◯	◯	◯
PCL UC3 - Handling for mass customization	◯	◯	◯	◯
TCA UC4 – Packaging operation in food sector	◯	◯	◯	◯
INFAR UC5 – Fixtureless assembly in hand tool	◯	◯	◯	◯
ULMA UC6 – Pallet to pallet order preparation	◯	◯	◯	◯
ULMA UC7 – Box to box order preparation				

Funded by the European Union

B.2. Survey

ULMA UC7 – Box to box order preparation				

6. If you like, please justify your choices.

Funded by the European Union

B.2. Survey

7. **Supervising AI** monitors processes, ensuring compliance with predefined parameters and safety standards. It provides real-time feedback and alerts, ensuring that tasks are performed correctly, while overseeing the performance of both human operators and other systems. *

For each use case, please state whether this role is desirable, undesirable, not applicable, or if you cannot answer.

Mark only one oval per row.

	Desirable	Undesirable	Not Applicable for the UC	Cannot Answer
Tofas UC1 - Spare parts delivery preparation	◯	◯	◯	◯
Tofas UC2a - Kitting	◯	◯	◯	◯
Tofas UC2b - Pre-assembly	◯	◯	◯	◯
PCL UC3 - Handling for mass customization	◯	◯	◯	◯
TCA UC4 – Packaging operation in food sector	◯	◯	◯	◯
INFAR UC5 – Fixtureless assembly in hand tool	◯	◯	◯	◯
ULMA UC6 – Pallet to pallet order preparation	◯	◯	◯	◯
ULMA UC7 – Box to box order	◯	◯	◯	◯

Funded by the European Union

B.2. Survey

preparation

preparation

8. If you like, please justify your choices.

Funded by the European Union

B.2. Survey

9. **Faulty AI** experiences malfunctions or provides incorrect feedback, causing potential errors or disruptions. It may misinterpret data or fail to alert users to critical issues, requiring human intervention to correct or override its actions to avoid negative outcomes. *

For each use case, please state whether this role is desirable, undesirable, not applicable, or if you cannot answer.

Mark only one oval per row.

	Undesirable	Not Applicable for the UC	Cannot Answer
Tofas UC1 - Spare parts delivery preparation	◯	◯	◯
Tofas UC2a - Kitting	◯	◯	◯
Tofas UC2b - Pre-assembly	◯	◯	◯
PCL UC3 - Handling for mass customization	◯	◯	◯
TCA UC4 – Packaging operation in food sector	◯	◯	◯
INFAR UC5 – Fixtureless assembly in hand tool	◯	◯	◯
ULMA UC6 – Pallet to pallet order preparation	◯	◯	◯
ULMA UC7 – Box to box order	◯	◯	◯

Funded by the European Union

B.2. Survey

preparation

preparation

10. If you like, please justify your choices.

Human Roles

In this section, the role of the human is considered

Funded by the European Union

B.2. Survey

11. The **Collaborative User** works alongside the AI, sharing responsibility. They frequently interact with the system, offering input, guiding processes, and making decisions based on feedback from the AI. *

For each use case, please state whether this role is desirable, undesirable, not applicable, or if you cannot answer.

Mark only one oval per row.

	Desirable	Undesirable	Not Applicable for the UC	Cannot Answer
Tofas UC1 - Spare parts delivery preparation	◯	◯	◯	◯
Tofas UC2a - Kitting	◯	◯	◯	◯
Tofas UC2b - Pre-assembly	◯	◯	◯	◯
PCL UC3 - Handling for mass customization	◯	◯	◯	◯
TCA UC4 – Packaging operation in food sector	◯	◯	◯	◯
INFAR UC5 – Fixtureless assembly in hand tool	◯	◯	◯	◯
ULMA UC6 – Pallet to pallet order preparation	◯	◯	◯	◯
ULMA UC7 – Box to box order preparation				

Funded by the European Union

B.2. Survey

ULMA UC7 – Box to box order preparation				

12. If you like, please justify your choices.

Funded by the European Union

B.2. Survey

13. The **Independent User** takes full control, actively leading the task. They are decision-makers and problem-solvers, responsible for managing the workflow and making key adjustments to ensure success. *

For each use case, please state whether this role is desirable, undesirable, not applicable, or if you cannot answer.

Mark only one oval per row.

	Desirable	Undesirable	Not Applicable for the UC	Cannot Answer
Tofas UC1 - Spare parts delivery preparation	◯	◯	◯	◯
Tofas UC2a - Kitting	◯	◯	◯	◯
Tofas UC2b - Pre-assembly	◯	◯	◯	◯
PCL UC3 - Handling for mass customization	◯	◯	◯	◯
TCA UC4 – Packaging operation in food sector	◯	◯	◯	◯
INFAR UC5 – Fixtureless assembly in hand tool	◯	◯	◯	◯
ULMA UC6 – Pallet to pallet order preparation	◯	◯	◯	◯
ULMA UC7 – Box to box order preparation				

Funded by the European Union

B.2. Survey

ULMA UC7 – Box to box order preparation				

14. If you like, please justify your choices.

Funded by the European Union

B.2. Survey

15. The **Helper User** supports the system by providing additional inputs or manual adjustments. They intervene when the AI faces limitations, offering guidance, oversight, or small corrections to ensure smooth operations and successful outcomes. *

For each use case, please state whether this role is desirable, undesirable, not applicable, or if you cannot answer.

Mark only one oval per row.

	Desirable	Undesirable	Not Applicable for the UC	Cannot Answer
Tofas UC1 - Spare parts delivery preparation	◯	◯	◯	◯
Tofas UC2a - Kitting	◯	◯	◯	◯
Tofas UC2b - Pre-assembly	◯	◯	◯	◯
PCL UC3 - Handling for mass customization	◯	◯	◯	◯
TCA UC4 – Packaging operation in food sector	◯	◯	◯	◯
INFAR UC5 – Fixtureless assembly in hand tool	◯	◯	◯	◯
ULMA UC6 – Pallet to pallet order preparation	◯	◯	◯	◯
ULMA UC7 – Box to box order	◯	◯	◯	◯

Funded by the European Union

B.2. Survey

preparation

preparation

16. If you like, please justify your choices.

Funded by the European Union

B.2. Survey

17. The **Supervisor User** monitors the system's performance, stepping in only to oversee and ensure everything runs smoothly. They focus on big-picture operations, intervening when necessary to correct or optimize. *

For each use case, please state whether this role is desirable, undesirable, not applicable, or if you cannot answer.

Mark only one oval per row.

	Desirable	Undesirable	Not Applicable for the UC	Cannot Answer
Tofas UC1 - Spare parts delivery preparation	◯	◯	◯	◯
Tofas UC2a - Kitting	◯	◯	◯	◯
Tofas UC2b - Pre-assembly	◯	◯	◯	◯
PCL UC3 - Handling for mass customization	◯	◯	◯	◯
TCA UC4 – Packaging operation in food sector	◯	◯	◯	◯
INFAR UC5 – Fixtureless assembly in hand tool	◯	◯	◯	◯
ULMA UC6 – Pallet to pallet order preparation	◯	◯	◯	◯
ULMA UC7 – Box to box order preparation				

Funded by the European Union

B.2. Survey

ULMA UC7 – Box to box order preparation				

18. If you like, please justify your choices.

Funded by the European Union

B.2. Survey

19. For the **Passive User**, the role focuses on minimal interaction and reliance on the system. The user waits for prompts or instructions from the AI, rarely taking initiative. They depend on automated systems to manage tasks and intervene only when necessary. This user trusts the system to maintain control, contributing to a smooth but passive workflow. *

For each use case, please state whether this role is desirable, undesirable, not applicable, or if you cannot answer.

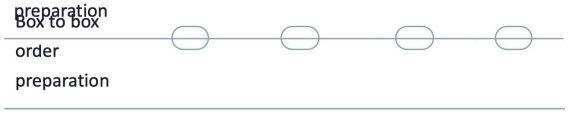
Mark only one oval per row.

	Desirable	Undesirable	Not Applicable for the UC	Cannot answer
Tofas UC1 - Spare parts delivery preparation	◯	◯	◯	◯
Tofas UC2a - Kitting	◯	◯	◯	◯
Tofas UC2b - Pre-assembly	◯	◯	◯	◯
PCL UC3 - Handling for mass customization	◯	◯	◯	◯
TCA UC4 – Packaging operation in food sector	◯	◯	◯	◯
INFAR UC5 – Fixtureless assembly in hand tool	◯	◯	◯	◯
ULMA UC6 – Pallet to pallet order preparation	◯	◯	◯	◯
ULMA UC7 – Box to box				

Funded by the European Union

B.2. Survey

order
ULMA UC7 – preparation
Box to box



20. If you like, please justify your choices.

Funded by the European Union

B.2. Survey

21. The **Clumsy User** inadvertently introduces mistakes into the process. They may misinterpret system prompts or struggle with controls, leading to inefficiencies or errors that need correction. *

For each use case, please state whether this role is desirable, undesirable, not applicable, or if you cannot answer.

Mark only one oval per row.

	Undesirable	Not Applicable for the UC	Cannot Answer
Tofas UC1 - Spare parts delivery preparation	◯	◯	◯
Tofas UC2a - Kitting	◯	◯	◯
Tofas UC2b - Pre-assembly	◯	◯	◯
PCL UC3 - Handling for mass customization	◯	◯	◯
TCA UC4 – Packaging operation in food sector	◯	◯	◯
INFAR UC5 – Fixtureless assembly in hand tool	◯	◯	◯
ULMA UC6 – Pallet to pallet order preparation	◯	◯	◯
ULMA UC7 – Box to box order preparation			

Funded by the European Union

B.2. Survey

ULMA UC7 – Box to box order preparation			

22. If you like, please justify your choices.

Funded by the European Union

B.2. Survey

23. The **Antagonist User** resists the system, either intentionally or through negligence. They may bypass instructions, ignore feedback, or disrupt the workflow, creating friction in human-AI collaboration. *

For each use case, please state whether this role is desirable, undesirable, not applicable, or if you cannot answer.

Mark only one oval per row.

	Undesirable	Not Applicable for the UC	Cannot Answer
Tofas UC1 - Spare parts delivery preparation	◯	◯	◯
Tofas UC2a - Kitting	◯	◯	◯
Tofas UC2b - Pre-assembly	◯	◯	◯
PCL UC3 - Handling for mass customization	◯	◯	◯
TCA UC4 – Packaging operation in food sector	◯	◯	◯
INFAR UC5 – Fixtureless assembly in hand tool	◯	◯	◯
ULMA UC6 – Pallet to pallet order preparation	◯	◯	◯
ULMA UC7 – Box to box order preparation			

Funded by the European Union

B.2. Survey

ULMA UC7 – Box to box order preparation			

24. If you like, please justify your choices.

In this section, please provide feedback on the delineated roles, indicating your thoughts.

Funded by the European Union

B.2. Survey

25. To summarise, the identified roles for the Autonomous Agent are: *

Collaborative AI: Works in harmony with human users, actively sharing tasks and decision-making. It relies on frequent communication, adjusting to feedback and providing support for a balanced partnership in completing tasks.

Independent AI: Takes primary responsibility for completing tasks, making decisions, and leading the workflow. As it evolves, it autonomously handles complex operations, often excluding humans from the process.

Helper AI: Provides assistance by performing secondary tasks, simplifying user operations. It handles repetitive or time-consuming actions, offers suggestions, or analyses data to enhance overall productivity.

Supervising AI: Monitors processes to ensure compliance with predefined parameters and safety standards. It provides real-time feedback and alerts, overseeing the performance of both human operators and other systems.

Faulty AI: Experiences malfunctions or provides incorrect feedback, leading to potential errors or disruptions. It may misinterpret data or fail to alert users to critical issues, requiring human intervention to correct or override its actions to avoid negative outcomes.

In your opinion, are there any additional roles for the Autonomous Agent that have not been identified? If so, please describe them and explain how they could contribute to human-AI teaming.

Funded by the European Union

B.2. Survey

26. To summarise, the identified roles for the Autonomous Agent are: *

Collaborative User: Works alongside the AI, sharing responsibility for tasks. They frequently interact with the system, providing input, guiding processes, and making decisions based on AI feedback. Please select the appropriate option for each use case.

Independent User: Takes full control of the task, actively leading it. They are decision-makers and problem-solvers, responsible for managing the workflow and making key adjustments to ensure success. Please select the appropriate option for each use case.

Helper User: Supports the AI by providing additional inputs or manual adjustments. They step in when the AI faces limitations, offering guidance, oversight, or minor corrections to ensure smooth operations and successful outcomes. Please select the appropriate option for each use case.

Supervisor User: Monitors the system's performance, intervening only to ensure everything runs smoothly. They focus on big-picture operations, stepping in when necessary to correct or optimize. Please select the appropriate option for each use case.

Passive User: Has minimal interaction with the AI, relying heavily on the system for task management. They wait for prompts or instructions from the AI, rarely taking initiative and depending on automation to maintain control. Their contribution creates a smooth but passive workflow. Please select the appropriate option for each use case.

Clumsy User: Inadvertently introduces errors or inefficiencies into the process. They may misinterpret system prompts or struggle with controls, requiring corrections to maintain smooth operations. Please select the appropriate option for each use case.

Antagonist User: Resists the system, either intentionally or through negligence. They may ignore instructions, bypass feedback, or disrupt the workflow, creating friction in the human-AI collaboration. Please select the appropriate option for each use case.

In your opinion, are there any additional roles for the User that have not been identified? If so, please describe them and explain how they could contribute to human-AI teaming.

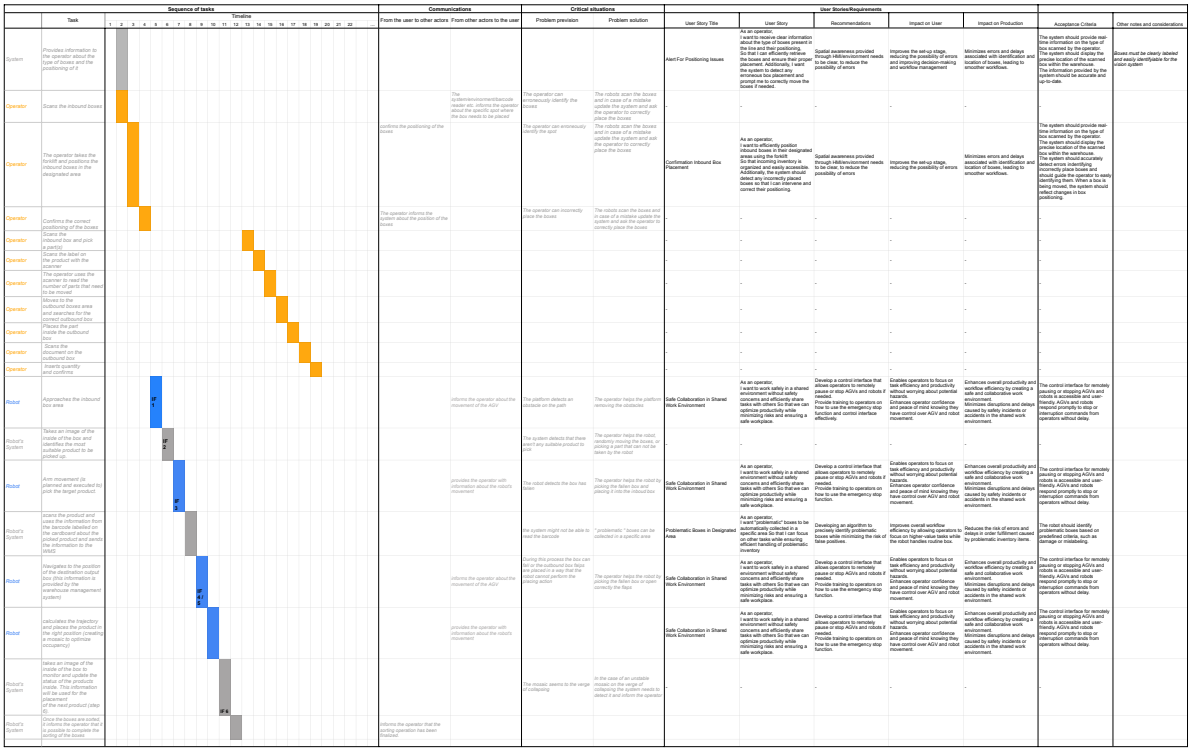
Funded by the European Union

B.2. Survey

Thank you for your valuable insights, we will discuss these results at the WP1 meeting on 16 October.

Funded by the European Union

B.3. TOFAS User Task Matrix



Sequence of tasks			Communications		Critical situations		User Stories/Requirements						
	Task	Timeline	From the user to other actors	From other actors to the user	Problem prevision	Problem solution	User Story Title	User Story	Recommendations	Impact on User	Impact on Production	Acceptance Criteria	Other notes and considerations
System	Provides information to the operator about the type of boxes and the positioning of it						Alert For Positioning Issues	As an operator, I want to receive clear information about the type of boxes present in the area and their positioning, So that I can efficiently retrieve the boxes and ensure their proper placement. Additionally, I want the system to detect any misplacement or misalignment and prompt me to correctly move the boxes if needed.	Spatial awareness provided through HMI/environment needs to be clear, to reduce the possibility of errors	Improves the set-up stage, reducing the possibility of errors and improving decision-making and workflow management	Minimizes errors and delays associated with identification and location of boxes, leading to smoother workflows	The system should provide real-time information on the type of box scanned by the operator. The system should display the precise location of the scanned box within the warehouse. The information provided by the system should be accurate and up-to-date.	Boxes must be clearly labeled and easily identifiable for the vision system
Operator	Scans the inbound boxes			The system/environment/barcode reader etc. informs the operator about the specific spot where the box needs to be placed	The operator can erroneously identify the boxes	The robotic scan the boxes and in case of a mistake update the system and ask the operator to correctly place the boxes	-	-	-	-	-	-	
Operator	The operator takes the forklift and positions the inbound boxes in the designated area		confirms the positioning of the boxes		The operator can erroneously identify the spot	The robotic scan the boxes and in case of a mistake update the system and ask the operator to correctly place the boxes	Confirmation Inbound Box Placement	As an operator, I want to efficiently position inbound boxes in their designated areas using the forklift So that incoming inventory is organized and easily accessible. Additionally, I want the system to detect any incorrectly placed boxes so that I can intervene and correct their positioning.	Spatial awareness provided through HMI/environment needs to be clear, to reduce the possibility of errors	Improves the set-up stage, reducing the possibility of errors	Minimizes errors and delays associated with identification and location of boxes, leading to smoother workflows.	The system should provide real-time information on the type of box scanned by the operator. The system should display the precise location of the scanned box within the warehouse. The system should accurately detect errors indentifying incorrectly place boxes and provide clear information for easily identifying them. When a box is being moved, the system should reflect changes in box positioning.	
Operator	Confirms the correct positioning of the boxes		The operator informs the system about the position of the boxes		The operator can incorrectly place the boxes	The robotic scan the boxes and in case of a mistake update the system and ask the operator to correctly place the boxes	-	-	-	-	-	-	
Operator	Scans the inbound box and pick a part(s)						-	-	-	-	-	-	
Operator	Scans the label on the product with the scanner						-	-	-	-	-	-	
Operator	The operator uses the scanner to read the labels on the parts that need to be moved						-	-	-	-	-	-	
Operator	Moves to the outbound boxes area and searches for the correct outbound box						-	-	-	-	-	-	
Operator	Places the part inside the outbound box						-	-	-	-	-	-	
Operator	Scans the document on the outbound box						-	-	-	-	-	-	
Operator	Inserts quantity and confirms						-	-	-	-	-	-	
Robot	Approaches the inbound box area			informs the operator about the movement of the AGV	The platform detects an obstacle on the path	The operator helps the platform removing the obstacles	Safe Collaboration in Shared Work Environment	As an operator, I want to work safely in a shared environment without safety concerns and efficiently share tasks with others So that we can optimize productivity while minimizing risks and ensuring a safe workplace.	Develop a control interface that allows operators to remotely pause or stop AGVs and robots if needed. Provide training to operators on how to use the emergency stop function and control interface effectively.	Enables operators to focus on task efficiency and productivity without worrying about potential hazards. Enhances operator confidence and peace of mind knowing they have control over AGV and robot movement.	Enhances overall productivity and workflow efficiency by creating a safe and collaborative work environment. Minimizes disruptions and delays caused by safety incidents or accidents in the shared work environment.	The control interface for remotely pausing or stopping AGVs and robots is accessible and user-friendly. AGVs and robots respond promptly to stop or interruption commands from operators without delay.	
Robot's System	Takes an image of the inside of the box and identifies the most suitable product to be picked up				The system detects that there aren't any suitable product to pick	The operator helps the robot randomly moving the boxes, or picking a part that can not be taken by the robot	-	-	-	-	-	-	
Robot	Arm movement (is planned and executed in) pick the target product			provides the operator with information about the robot's movement	The robot detects the box has been	The operator helps the robot by picking the fallen box and placing it into the inbound box	Safe Collaboration in Shared Work Environment	As an operator, I want to work safely in a shared environment without safety concerns and efficiently share tasks with others So that we can optimize productivity while minimizing risks and ensuring a safe workplace.	Develop a control interface that allows operators to remotely pause or stop AGVs and robots if needed. Provide training to operators on how to use the emergency stop function.	Enables operators to focus on task efficiency and productivity without worrying about potential hazards. Enhances operator confidence and peace of mind knowing they have control over AGV and robot movement.	Enhances overall productivity and workflow efficiency by creating a safe and collaborative work environment. Minimizes disruptions and delays caused by safety incidents or accidents in the shared work environment.	The control interface for remotely pausing or stopping AGVs and robots is accessible and user-friendly. AGVs and robots respond promptly to stop or interruption commands from operators without delay.	
Robot's System	scans the product and uses the information from the barcode labelled on the cardboard about the picked product and saves the information to the WMS				the system might not be able to read the barcode	"problematic " boxes can be collected in a specific area	Problematic Boxes in Designated Area	As an operator, I want "problematic" boxes to be automatically collected in a specific area So that I can focus on other tasks while ensuring efficient handling of problematic inventory	Developing an algorithm to precisely identify problematic boxes while minimizing the risk of false positives.	Improves overall workflow efficiency by allowing operators to focus on higher-value tasks while the robot handles routine box	Reduces the risk of errors and delays in order fulfillment caused by problematic inventory items.	The robot should identify problematic boxes based on predefined criteria, such as damage or mislabeling.	
Robot	Navigates to the position of the destination output box (this information is provided by the warehouse management system)			informs the operator about the movement of the AGV	During this process the box can fall off the outbound box traps are placed in a way that the robot cannot perform the placing action	The operator helps the robot by picking the fallen box or opens correctly the flaps	Safe Collaboration in Shared Work Environment	As an operator, I want to work safely in a shared environment without safety concerns and efficiently share tasks with others So that we can optimize productivity while minimizing risks and ensuring a safe workplace.	Develop a control interface that allows operators to remotely pause or stop AGVs and robots if needed. Provide training to operators on how to use the emergency stop function.	Enables operators to focus on task efficiency and productivity without worrying about potential hazards. Enhances operator confidence and peace of mind knowing they have control over AGV and robot movement.	Enhances overall productivity and workflow efficiency by creating a safe and collaborative work environment. Minimizes disruptions and delays caused by safety incidents or accidents in the shared work environment.	The control interface for remotely pausing or stopping AGVs and robots is accessible and user-friendly. AGVs and robots respond promptly to stop or interruption commands from operators without delay.	
Robot	calculates the trajectory and places the product in the right position (creating a mosaic to optimize occupancy)			provides the operator with information about the robot's movement			Safe Collaboration in Shared Work Environment	As an operator, I want to work safely in a shared environment without safety concerns and efficiently share tasks with others So that we can optimize productivity while minimizing risks and ensuring a safe workplace.	Develop a control interface that allows operators to remotely pause or stop AGVs and robots if needed. Provide training to operators on how to use the emergency stop function.	Enables operators to focus on task efficiency and productivity without worrying about potential hazards. Enhances operator confidence and peace of mind knowing they have control over AGV and robot movement.	Enhances overall productivity and workflow efficiency by creating a safe and collaborative work environment. Minimizes disruptions and delays caused by safety incidents or accidents in the shared work environment.	The control interface for remotely pausing or stopping AGVs and robots is accessible and user-friendly. AGVs and robots respond promptly to stop or interruption commands from operators without delay.	
Robot's System	takes an image of the inside of the box to monitor and update the status of the products inside. This information will be used for the placement of the next product (step 4)				The mosaic seems to the verge of collapsing	In the case of an unstable mosaic on the verge of collapsing the system needs to detect it and inform the operator	-	-	-	-	-	-	
Robot's System	Once the boxes are sorted it informs the operator that it is possible to complete the sorting of the boxes		informs the operator that the sorting operation has been finished										

Funded by the European Union