



D 1.4

Social, Ethical, Legal, Privacy issues identification and monitoring

Authors: Elisa Spiller, Nikolas Giampaolo (DBL)

Information Table

Deliverable Number	1.4
Deliverable Title	Social, Ethical, Legal, Privacy issues identification and monitoring
Version	1.0
Status	Final
Responsible Partner	DBL
Contributors	ECTL; FerroNATS; CHPR; LiU; TAVS; Bordeaux INP; DFKI; ENG; LFV; ENAC; TUI; Suite5; EMBRT; CERTH; LLA
Contractual Date of Delivery	August 31st, 2025
Actual Date of Delivery	August 26th, 2025
Dissemination Level	Public

Document History

Version	Date	Status	Author	Description
0.1	05/06/2025	Outline	Nikolas Giampaolo (DBL)	Structure of the document
0.2	29/07/2025	First draft	Elisa Spiller, Nikolas Giampaolo (DBL)	Draft for review
0.3	07/08/2025	Second draft	Vanessa Arrigoni (DBL)	Review with minor comments
1.0	26/08/2025	Final version	Elisa Spiller, Nikolas Giampaolo (DBL)	Final version for submission

List of Acronyms

Acronym	Definition
AI	Artificial Intelligence
CA	Consortium Agreement
EC	European Commission
EU	European Union
GDPR	General Data Protection Regulation
GA	Grant Agreement
HF	Human Factors
HLEG	High-Level Expert Group (on AI)
KPIs	Key Performance Areas
XAI	Explainable AI
WP	Work Package

Executive Summary

This deliverable describes the measures taken to ensure that project activities are respectful of human rights, particularly the right to privacy and data protection, and do not generate ethically unwanted individual, collective or social effects. The deliverable describes privacy, ethical and other legal concerns, and proposes mitigation measures to address them.

The report is divided into four sections including a short introduction, a description of the main points regarding ethical issues arising from AI, an overview of the ethical issues stemming from the HAIKU research project and the way the consortium plans to mitigate them, involving the processing of personal data and a set of measures that could be implemented to reduce the risk of misuse of the research results. The Appendix contains the AI Ethics issue checklist proposed by the EC (EU GRANTS, 2021).

This report has been a living document over the project and has been iteratively updated when new information regarding data collection activities and details regarding specific aspects of Social, Ethical, Legal, Privacy issues become available. This final version, released at M36.

Table of Contents

Information Table.....	2
Document History.....	2
List of Acronyms.....	3
Executive Summary.....	4
Table of Contents.....	5
1. Introduction.....	6
1.1. Deliverable structure.....	6
2. Ethical Considerations.....	7
2.1. HAIKU Ethical Issues.....	7
2.2. Research activities concerning the use of AI.....	7
2.3. Ethical principles for the volunteer participants to HAIKU activities.....	8
3. Personal Data Protection.....	10
3.1. Ethical and legal framework.....	10
3.2. Principles and general rules.....	10
3.3. Measures to Prevent Potential Misuse of Research Results.....	12
4. AI Ethical Issues.....	13
4.1. Ethical and legal framework.....	13
4.2. Definition.....	13
4.3. Principles and General Rules.....	14
5. HAIKU Ethical Issues and Mitigation Measures.....	17
5.1 Human- AI relationship.....	18
5.2 Explainable AI.....	20
5.3 Safety and Security of AI.....	21
5.4 Accountability.....	21
5.5 AI Biases.....	22
5.6 Measures to Prevent Potential Ethical Risks.....	23
6. Conclusions.....	24
7. References.....	25

1. Introduction

This deliverable describes the measures taken to ensure that project activities are respectful of human rights, particularly the right to privacy and data protection, and do not generate ethically unwanted personal or social effects. The deliverable describes privacy, ethical, and other legal concerns, and proposes mitigation measures to address them.

1.1. Deliverable structure

This document is divided into eight sections:

- Section 1 contains a short introduction and document structure.
- Section 2 addresses main points regarding ethical issues arising in the HAIKU project.
- Section 3 provides an overview of the ethical and regulatory framework concerning the protection of personal data.
- Section 4 provides an overview of the ethical and regulatory framework concerning trustworthy AI.
- Section 5 lists a set of measures and recommendations that could be implemented to reduce the ethical and legal risks related to HAIKU research activity.
- Section 6 presents the measures to prevent potential misuse of research results.
- Section 7 provides the conclusions stemming from the ethical assessment .

2. Ethical Considerations

2.1. HAIKU Ethical Issues

According to the EC Ethical Self-Assessment Checklist, the ethical issues that arise in the HAIKU project may concern:

- Participation of humans in ‘experiments’ and data collection activities
- Collection of personal data of front-end operators and staff of the end-user organisations through field observation, interviews - as primary data
- Personal data that come from secondary data—surveys, company video recordings, system logs, internal documents, etc.—provided by end user organisations.
- HAIKU involves a non-EU partner from Brazil and stakeholders providing data from the UK. Non-EU countries are required to comply with the legal and ethical requirements for participation in EU research and H2020 guidelines
- As in every project that collects and analyses personal data and relevant business data, the potential misuse of the research results are possible
- Development, deployment and/or use of Artificial Intelligence-based systems able to interact, replace or influence human decision-making processes

In light of the above, the following paragraphs will specifically approach each of these issues. Particular attention will be devoted to the new issues related to the use of AI-based solutions. In this regard, consistently with the approach suggested by the EC and HLEG AI, this document will introduce the AI Ethical Assessment Methodology adopted for the specific purposes of HAIKU.

2.2. Research activities concerning the use of AI

Ethical and legal issues are always a delicate matter for a study that involves users and especially when collecting personal data and involving the use of AI such as the proposed study. When it comes to technology and data sharing in general, where personal data are collected, there exist obvious public concerns: the difficulty of respecting privacy and confidentiality when third parties have a strong interest in getting access to electronically recorded and stored personal health data, or the difficulty in ensuring the security of shared personal monitoring data. Moreover, the use of AI raises additional concerns about human-machine interactions mainly related to the respect of human dignity and autonomy, non-discrimination and liability apportionment in human-AI teaming.

In light of the above, this document primarily relies on the following references and guidelines:

- Ethical Guidelines for Good Research Practice
- Code of Human Research Ethics
- UNESCO Recommendations on the Ethics of Artificial Intelligence
- HLEG AI Ethics Guidelines For Trustworthy AI
- HLEG Assessment List for Trustworthy AI (ALTAI) for self-assessment

- European Commission Note for Identifying serious and complex ethics issues in EU-funded research, published on 5 July 2021
- European Commission Note on Ethics and Data Protection, published on 5 July 2021
- European Commission Note on Ethics By Design and Ethics of Use Approaches for Artificial Intelligence, published on 25 November 2021

The Consortium will conform to the rules and legislation here listed.

The members of the HAIKU project consortium followed the national legislation and the non-EU countries were asked to comply with the legal and ethical requirements for participation in EU research and Horizon Europe guidelines.

2.3. Ethical principles for the volunteer participants to HAIKU activities

This deliverable describes the ethical considerations and procedures for user studies and data processing within the HAIKU project.

We obtained necessary ethics approvals and free and fully informed consent from research participants. Focus groups and structured interviews were strictly limited to volunteers, mainly employees of partner organizations.

The HAIKU Consortium processes personal data on a lawful basis, as provided in Articles 6 and 9 of the GDPR. Notably, the processing of personal data is based on **consent** (for common Personal data) and **explicit consent** (for special categories of Personal data).

The procedure for volunteers strictly adhered to the conditions below:

- Every volunteer has the right to remain anonymous;
- Every volunteer has his or her data protected as stated in the applicable data protection legislation;
- All data released by the eventual volunteers participating in the project results evaluation brings a duty of confidentiality;
- Every volunteer's free and informed consent is required also while doing interviews, ethnographic observation, non-invasive experimentation, neurophysiological data collection (if used), and accessing personal data records. The purpose of informed consent is to empower individuals to make a voluntary informed decision about whether or not to participate in the research based on knowledge of its goals, procedures and outcomes and therefore no volunteers that are minors and with reduced autonomy or vulnerability will be involved;
- Every volunteer also has the right to know who will benefit from his or her participation in the experiment and, if they so wish, to receive information of any potential commercial exploitation of the research that involves their participation;
- Participants are informed about the purpose of the research, the vehicles for the dissemination of the results and people, organizations and stakeholders involved in the research, and also about data treatment and handling;
- Volunteers willing to be informed on the results of the research are timely updated

- with the outcomes of the study;
- Volunteers are also informed of the appropriate insurance cover that is in place, if applicable;
 - It is clearly explained to the volunteers that they can withdraw from the project at any time;
 - The data is collected in a way that the researcher will not impose any of their own bias on the data itself;
 - Fair involvement, equal opportunities and equal treatment among people with a different socio-cultural background (e.g. gender, nationality, religion, age, etc.) are guaranteed.

HAIKU obtained written consent from all research participants, who are informed about the research and about the consequences of their participation. They received information about the methods, the timeframe and the environment in which their data will be studied. The individuals have a contact point (a responsible authority within their organization) in case they need to receive additional information regarding their participation. The participants were informed about the opt-out procedure.

The information sheets, accompanying the informed consent form, are written in a language and terms fully understandable to the participant. The sheets describe the aims, methods and implications of the research, the nature of the participation and any benefits, risks or discomfort that might be involved. The information sheet explicitly states that participation is voluntary and that anyone has the right to refuse to participate and to withdraw their participation, samples or data at any time - without any consequences. Only when the research participant has fully understood the information sheet and has given consent without any pressure having been put on them, then consent is considered valid. For more information about the data collection and the forms utilized in the different experiments please refer to D1.5 "Data Management Plan 2".

3. Personal Data Protection

3.1. Ethical and legal framework

HAIKU research activities are aligned with the principles of dignity, freedom, equality, solidarity, citizens' rights and justice, as stated in the Charter of Fundamental Rights of the European Union. When performing experiments and collecting data, the HAIKU project used particular care to ensure privacy and data protection rights, considering them as fundamental rights to be protected.

In particular, the HAIKU Consortium has ensured ethical and legal compliance with current regulations:

- the EU Charter of Fundamental Rights (art. 7 and 8);
- the European Convention for the Protection of Human Rights and Fundamental Freedoms (art. 8);
- the General Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons concerning the processing of personal data and on the free movement of such data (GDPR);
- the Directive on Privacy and Electronic Communications (2002/58//EC);
- the Council of Europe Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data No.108/1981;
- the Article 29 WP Guidelines on Consent under Regulation 2016/679 (wp259rev.01);
- the EDBO Guidelines 5/2020 on Consent under Regulation 2016/679 (v. 1.1 adopted on 4 May 2020);
- the Article 29 WP Guidelines on Transparency under Regulation 2016/679 (wp260rev.01);
- the Article 29 WP Opinion 05/2014 on Anonymization techniques (wp216);
- the Article 29 WP Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679 (wp251rev.01).

3.2. Principles and general rules

For the purposes of this deliverable, from now on the definitions adopted are the following. These latter refer to those contained in Articles 4 and 9 of the GDPR.

- **Personal data** is intended as any information relating to an identified or identifiable natural person (data subject); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person.
- **Processing** includes any operation or set of operations which is performed on personal data or on sets of personal data, whether or not by automated means, such as collection, recording, organisation, structuring, storage, adaptation or alteration,

retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, restriction, erasure or destruction.

- **Controller** is intended as the natural or legal person, public authority, agency or another body which, alone or jointly with others, determines the purposes and means of the processing of personal data; where the purposes and means of such processing are determined by Union or Member State law, the controller or the specific criteria for its nomination may be provided for by Union or Member State law. In the HAIKU project, every partner is a controller and responsible for the data that it collects and processes. All the partners are committed to the respect of the principles and rules listed in the present document and in the following versions of it.
- **Processor** is intended as a natural or legal person, public authority, agency or other body which processes personal data on behalf of the controller. On the other hand, **recipient** may be any natural or legal person, public authority, agency or another body, to which the personal data are disclosed, whether a third party or not.

Generally, in light of what provided by the GDPR as well as by FAIR practices [4], personal data has been process according the principles outlined by Article 5 of the GDPR, namely:

- **Lawfulness, fairness and transparency** in relation to the data subject;
- **Purpose limitation:** Personal data were collected for specified, explicit and legitimate purposes and not further processed in a manner that is incompatible with those purposes;
- **Data minimization:** Personal data were adequate, relevant and limited to what is necessary in relation to the purposes for which they are processed;
- **Accuracy:** Personal data were accurate and, where necessary, kept up to date; data were not to be modified or falsified.
- **Storage limitation:** Personal data were kept in a form which permits identification of data subjects for no longer than is necessary for the purposes for which the personal data are processed;
- **Integrity and confidentiality:** Personal data were processed in a manner that ensures appropriate security of the personal data, including protection against unauthorised or unlawful processing and against accidental loss, destruction or damage, using appropriate technical or organisational measures;
- **Data protection by design and by default:** The controller implemented appropriate technical and organisational measures which are designed to implement data-protection principles in an effective manner and to integrate the necessary safeguards into the processing in order to protect the rights of data subjects. The controller implemented appropriate technical and organisational measures for ensuring that, by default, only personal data which are necessary for each specific purpose of the processing are processed;
- **Accountability:** The controller is responsible for and is able to demonstrate compliance with the above-mentioned principles.

The legal basis for the processing of personal data is **consent of the data subject**. Consent is considered valid if freely given, specific and informed. The data subjects are free to

withdraw their consent according to Article 6 GDPR. HAIKU research did NOT include children, adults unable to give informed consent, nor vulnerable individuals/groups.

Within the HAIKU research activities, data subjects are able to exercise their rights on the processing of personal data concerning them according to the rules and conditions provided by the GDPR and the related internal policies.

The HAIKU Consortium is committed to ensuring the effective protection of Personal data of all participants to its activities. More detailed information, including those related to the security measures, will be available in D1.5 "Data Management Plan 2".

3.3. Measures to Prevent Potential Misuse of Research Results

The HAIKU consortium tried to avoid the misuse of the research results. In particular, the HAIKU partners favoured a correct and rightful use of the results of the research, as well as of the knowledge of the techniques and theories applied, they avoided any negative repercussion against the public reputation of each partner and of any other involved entity or individual and improper exploitation of research results for obtaining commercial and business advantages, personal profits and any other kind of unlawful application of the research results.

Main mitigation measures to avoid the potential misuse of research results, both from partners of the Consortium and from external malicious actors included:

- Monitor the access to data and the researcher's actions on data.
- Define and implement policy access based on least privilege.
- Protect data from outsiders.

Multiple security mechanisms and technologies ensured protection from malevolent/criminal/terrorist abuse. For more information please refer to section 4 of D1.5 "Data Management Plan 2".

In addition to that, strict procedures were defined and applied for publication of research results, any publication of final or intermediate results must comply with the following statements:

1. Everyone who participated in the collection of data has been acknowledged in the acknowledgements section of the publication, either individually or collectively. Co-authorship of publications has been determined and agreed based on standard academic conventions.
2. Sensitive information has been presented presented to competent, responsible and scientifically sound audiences;
3. Political and socio-economic concerns must be carefully considered in presenting the results as any opinion expressed could affect the public opinion feelings about safety in aviation and maritime transport.

4. AI Ethical Issues

4.1. Ethical and legal framework

As anticipated, the HAIKU research activities are guided by a strong commitment to compliance and the protection of fundamental rights, particularly with regard to privacy and data protection.

In particular, the HAIKU Consortium ensures ethical and legal compliance with current regulations and guidelines:

- the EU Charter of Fundamental Rights (spec. art. 1, 7, 8 and 21);
- the European Convention for the Protection of Human Rights and Fundamental Freedoms (art. 8 and 14);
- Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act)
- Directive (EU) 2024/2853 of the European Parliament and of the Council of 23 October 2024 on liability for defective products and repealing Council Directive 85/374/EEC
- HLEG AI Ethics Guidelines For Trustworthy AI
- HLEG Assessment List for Trustworthy AI (ALTAI) for self-assessment
- EASA Concept Paper: guidance for Level 1 & 2 machine learning applications Issue 02, especially the Ethics-based assessment and the Annex 5 — Full list of questions from the ALTAI adapted to aviation
- Proposal for a Directive of the European Parliament and of the Council on adapting non-contractual civil liability rules to artificial intelligence (AI Liability Directive) (COM(2022) 496 final) (withdrawn in February 2025)

4.2. Definition

For the purposes of this deliverable, from now on the definition adopted is the following.

Artificial intelligence (AI) generally refers to “to systems designed by humans that, given a complex goal, act in the physical or digital world by perceiving their environment, interpreting the collected structured or unstructured data, reasoning on the knowledge derived from this data and deciding the best action(s) to take (according to pre-defined parameters) to achieve the given goal. AI systems can also be designed to learn to adapt their behaviour by analysing how the environment is affected by their previous actions”.

However, “as a scientific discipline, AI includes several approaches and techniques, such as machine learning (of which deep learning and reinforcement learning are specific examples), machine reasoning (which includes planning, scheduling, knowledge representation and reasoning, search, and optimization), and robotics (which includes

control, perception, sensors and actuators, as well as the integration of all other techniques into cyber-physical systems)" [10].

In light of what required by HAIKU research activities, the general definition adopted from the HLEG on AI will be tailored to the different necessities of each use case taking into account that HAIKU integrates different AI applications (e.g. Machine learning, digital assistant, etc).

4.3. Principles and General Rules

Generally, in light of what provided by the HLEG AI Ethics Guidelines For Trustworthy AI [10], within the HAIKU research project AI relies on the following principles:

1. **Respect for Human Agency:** To allow individuals to make their own choices and act in accordance with them, respect is required. Three additional specific principles that define fundamental human rights include respect for human agency: dignity, freedom, and autonomy;
2. **Privacy and Data governance:** Privacy and data security are fundamental human rights that must be upheld at all times;
3. **Fairness:** Equal rights and opportunities should be afforded to everyone, and unfair advantages and disadvantages should be avoided;
4. **Individual, Social and Environmental Well-being:** AI systems should contribute to, and not harm, individual, social and environmental well-being;
5. **Transparency:** Stakeholders should be able to comprehend AI programs' objectives, inputs, and operations;
6. **Accountability and Oversight:** humans should be able to understand, supervise and control the design and operation of AI-based systems.

In light of the above, the development, deployment and use of AI is aligned with the following requirements:

1. **Human agency and oversight:** AI systems must have human oversight, allowing users to override automated decisions. Developers should include technical measures like stop buttons to ensure human control.
2. **Technical robustness and safety:** Trustworthy AI algorithms must be secure, dependable, and robust, handling errors and inconsistencies. Vulnerabilities, hacking, and cyberattacks should be addressed through testing and safety assessment procedures.
3. **Privacy and data protection:** All EU AI stakeholders must comply with GDPR, prioritizing data privacy and security. Citizens should control their data, which AI developers should protect through anonymization and encryption.
4. **Transparency:** Transparency is crucial for impartial AI. This includes traceable data sets and procedures, clear identification of AI systems, and explainability of AI and human decisions.
5. **Diversity, non-discrimination, and fairness:** AI design must avoid unfair bias, ensuring algorithms are not biased (e.g., due to insufficient data). Stakeholders should be consulted, and AI systems should accommodate all human abilities, ensuring accessibility for people with disabilities.

6. **Societal and environmental wellbeing:** AI systems should promote positive social change, sustainability, and environmental responsibility (e.g., energy efficiency). Their impacts on physical and mental health, democracy, and elections should be monitored.
7. **Accountability:** Mechanisms for accountability are essential. AI systems violating fundamental rights require internal and external audits. Conflicts between ethical requirements necessitate continuous evaluation of trade-offs. Accessible redress mechanisms should be implemented.

In particular, considering the ethical principles and the fundamental rights involved in the development, deployment and use of the AI-powered solutions within HAIKU research, the Consortium opts for a proactive approach. This choice is consistent with the three-pronged approach suggested by the EC and HLEG AI, and relying on the so-called Ethics by Design [5].

Ethics by Design can be used to make sure that the ethical requirements of an AI system or technique are properly addressed during its development. The purpose of this strategy is to make it as clear as possible what is required, to describe the specific tasks that must be carried out, and to assist developers in considering ethical considerations when creating an AI system [2].

In some AI projects, the relevant ethical issues may be discovered during the development phase, while in others, they may be discovered after the system has been deployed. By addressing ethical issues during the development phase rather than attempting to fix them later in the process, Ethics by Design aims to prevent ethical issues from occurring in the first place. By proactively utilizing the principles as system requirements, this is accomplished. A five-layer model describes Ethics by Design [2].



Figure 1. The 5-layered model of Ethics by Design. Taken from Dainow & Brey, 2021.

The aim of Ethics by Design is to make people think about and address potential ethics concerns, while they are developing a system [2].

The main ethical requirements for AI and robotics systems above can be summarised as:

- The dignity, freedom, and autonomy of human beings must not be compromised by AI systems.
- The right to personal data protection and privacy must not be violated by AI systems. They must use data that is necessary, representative, and free of bias.
- AI systems must be developed without discrimination and with an inclusive, fair agenda.
- It is necessary to take precautions to ensure that AI systems do not inflict harm on individuals, societies, or the environment, rely on harmful technologies, persuade others to take harmful actions, or lend themselves to function creeps.
- AI systems ought to be as open as possible to their users and stakeholders.
- To address noncompliance and ensure compliance with these principles, human oversight and accountability are required.

5. HAIKU Ethical Issues and Mitigation Measures

Considering the specific research activities proposed by the HAIKU project, the use of AI is aimed at the development of Intelligent Assistants. In particular, the different solutions presented in the use cases involve technologies with different levels of automation and autonomy. Their analytical, decisional and operational scope of action may significantly vary according to the specific needs of the contexts of use.

The general concerns emerging from EC Ethical Self-Assessment Checklist about the development, deployment and/or use of Artificial Intelligence-based systems able to interact, replace or influence human decision-making processes here should be furtherly detailed. More specifically these are the main AI ethical issues related to the HAIKU project:

- Human-AI relationship
- Explainability
- Safety
- Accountability
- Biases

The HAIKU project dedicated a specific Work Package (WP7) to developing an integrated, risk-based assessment framework for AI in aviation. This framework aimed to ensure that AI-enabled Intelligent Assistants were evaluated not only for technical reliability but also for compliance with ethical principles, legal obligations, and fundamental rights - especially regarding safety, security, accountability, and human oversight.

The methodology was structured around four interdependent Key Performance Areas (KPA's):

- Safety, to ensure systems behaved predictably and failed safely;
- Security, to protect systems against cyber threats and data misuse;
- Human Performance, to evaluate impacts on cognitive workload, trust, and usability;
- Liability, to anticipate accountability gaps and ensure legal defensibility.

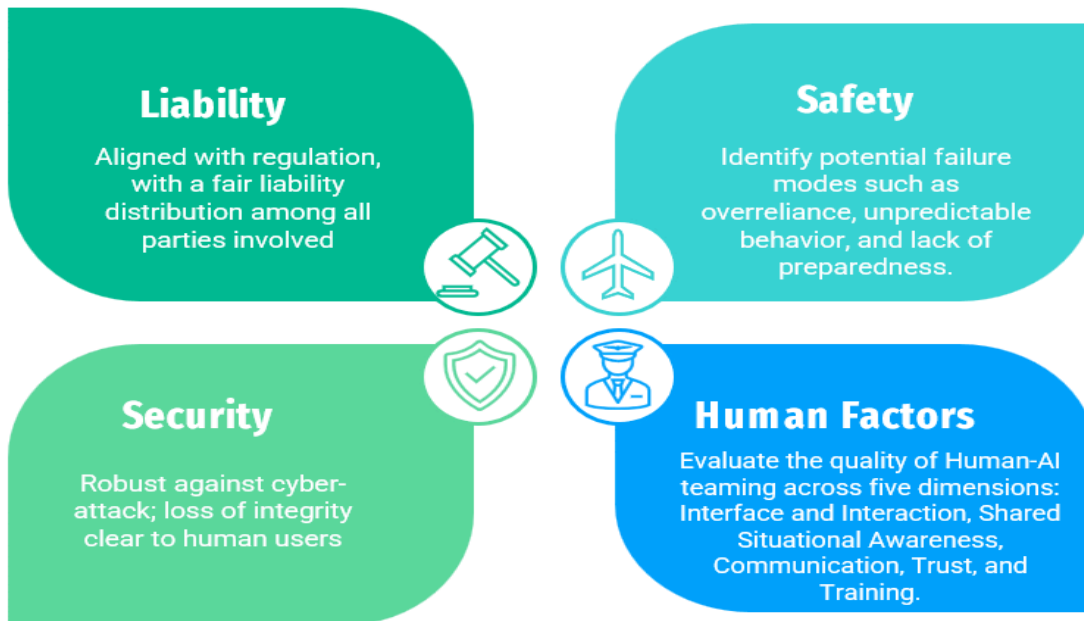


Figure 2. KPAs Dimensions

The assessment process unfolded through four methodological stages, each of which contributed to the early identification and mitigation of safety, security, human performance and legal risks:

1. Operational Sequence Diagrams (OSDs) were first developed to model realistic human-AI interactions [2]. These diagrams captured the sequence of operational steps, allowing analysts to identify where cognitive overload, ambiguity, or coordination breakdowns could arise.
2. Based on these scenarios, a tailored HAZOP (Hazard and Operability) analysis was conducted [12]. This step systematically explored deviations from expected behaviour, including unintended AI actions, loss of human oversight, or failures in communication. Hazards were examined through guidewords and structured team discussions, involving both domain and technical experts.
3. The third step involved a cross-cutting risk analysis across the four KPAs, where each hazard was evaluated in terms of its implications for safety, human factors, security resilience, and legal accountability. For security, the SESAR SecRAM methodology [15] was applied to assess risks to data Confidentiality, Integrity, and Availability (CIA). In parallel, a Legal Case methodology was used to identify potential product, organisational, or personal liability arising from system design or operational misuse.
4. Finally, mitigation strategies were formulated and prioritised. These included human-centred interface improvements, training enhancements, explainability requirements, cybersecurity safeguards, and measures for trust calibration. Particular attention was paid to ensuring that mitigation actions evolved in line with system maturity (TRL progression) and remained compatible with regulatory frameworks such as the EU AI Act and EASA AI Roadmap [5].

Crucially, this assessment approach was iterative and adaptable. It allowed emerging risks to be identified and addressed early in the development process, thus avoiding reliance on post hoc compliance checks. It also recognised that AI safety and security are not merely technical concerns but involve socio-technical interactions where ethical, legal, and human factors intersect. By embedding these considerations into each step of the methodology, the HAIKU framework supported the development of AI assistants that are not only operationally robust but also legally sound, ethically acceptable, and socially trustworthy.

The identified ethical issues will be specifically addressed in the following pages, explaining the mitigation measures and intended requirements the Consortium aims at implementing in HAIKU research activities. Eventually, the document reports the list of measures adopted to mitigate emerging ethical risks over the project.

5.1 Human- AI relationship

The scope of HAIKU is to develop a human-AI relationship for different AI aviation applications that preserves and enhances **Human Oversight**. In particular, HAIKU aims to answer the following research question: How to support the human with AI, without the human feeling under-valued or losing motivation and 'investment' in safe operations, or becoming 'out-of-the-loop', losing situation awareness and the ability to take back control when needed?

The answer relates to deciding where the AI can best help without detrimental impact on human performance. HAIKU is particularly focused on this aspect, which is also in line with leading European thinking on the way we would all like AI to evolve, as stated by Humane AI (<https://www.humane-ai.eu/>) [11]. The core of the humane AI vision is as follows: "The development of robust, trustworthy AI systems capable of understanding humans, adapting to complex real-world environments, and appropriately interacting in complex social settings. The overall vision is to facilitate AI systems that enhance human capabilities and empower individuals and society as a whole while respecting human autonomy and self-determination."

The HAIKU project is implemented via six aviation Use Cases, capturing a range of aviation segments, time horizons (from tactical to strategic), task complexity, and human impact. Each Use Case requires a tailored AI concept - with specific purposes, expected AI benefits for operations and for humans, underlying values and design principles. The Use Cases proceed as parallel strands but adopt a harmonized approach in terms of methods, validation exercises, and expected outcomes. This ensures cross-fertilisation among the Use Cases, whenever lessons learnt, and findings allow. The rationale for this approach is that AI-powered assistants can support humans in a variety of tasks. They can support the execution of simple tasks - be them for perception, decision-making or execution, using the established cognitive framework - or they can partner with humans to monitor and manage overall system status. In some Use Cases, AI will simply speed up human performance or substitute for the humans, while in other instances the use of AI will transform the current tasks and deeply change the nature of work. Effective human-AI partnerships should be designed differently in all of these different cases, depending on the task complexity and on the expected benefits of "work as done with AI."

The deployment and use of AI systems may implicate fundamental rights and their underlying values. In order to ensure an adequate Human-AI interaction, the consortium abides to the following principles, following from the HLEG on AI[10]:

- **Respect for human dignity.** Human dignity encompasses the idea that every human being possesses an “intrinsic worth”, which should never be diminished, compromised or repressed by others – nor by new technologies like AI systems. In this context, HAIKU makes sure that all people are treated with respect due to them as moral subjects. The AI system is hence designed to augment, complement and empower human cognitive, social and cultural skills. HAIKU aims to do this by following human-centric design principles securing human oversight over work processes in AI systems.
- **Freedom of the individual.** Human beings should remain free to make life decisions for themselves. In an AI context, freedom of the individual for instance requires mitigation of (in)direct illegitimate coercion, threats to mental autonomy and mental health, unjustified surveillance, deception and unfair manipulation. HAIKU commits to enabling individuals to wield control over their lives, including (among other rights) protection of the freedom to conduct a business, freedom of expression, the right to private life and privacy, and freedom of assembly and association.
- **Respect for democracy, justice and the rule of law.** AI systems should serve to maintain and foster democratic processes and respect the plurality of values and life choices of individuals. In this respect, HAIKU ensures that AI systems do not undermine democratic processes, human deliberation and mandatory laws and regulation.
- **Equality, non-discrimination and solidarity - including the rights of persons at risk of exclusion.** Equal respect for the moral worth and dignity of all human beings must be ensured. HAIKU commits to the fact that the system’s operations do not generate unfairly biased outputs. In order to mitigate this ethical issue, the consortium respects potentially vulnerable persons and groups, such as workers, women, persons with disabilities, ethnic minorities, children, consumers or others at risk of exclusion.
- **Prevention of harm** AI systems should neither cause nor exacerbate harm or otherwise adversely affect human beings. This entails the protection of human dignity as well as mental and physical integrity. HAIKU commits to the fact that the AI systems and the environments in which they operate must be safe and secure, technically robust, and not open to malicious use.

5.2 Explainable AI

HAIKU aims to maintain the human-in-the-loop in AI learning for each of the aviation applications. Human-in-the-loop learning can lead to more explainable AI (XAI), as the user will have more visibility on the underlying mechanisms, ideally steering a continuous learning process.

In order to ensure an adequate Human-AI interaction, the consortium abides to the following principles:

- **Explicability.** Explicability is crucial for building and maintaining users’ trust in AI systems. HAIKU ensures that AI processes are transparent, the purpose of AI systems openly communicated, and decisions – to the extent possible – explainable

to those directly and indirectly affected. Without such information, a decision cannot be duly contested.

- **Traceability.** The HAIKU consortium ensures that the data sets and the processes that yield the AI system's decision, including those of data gathering and data labelling as well as the algorithms used, is documented to the best possible standard to allow for traceability and an increase in transparency. This enables identification of the reasons why an AI-decision was erroneous which, in turn, could help prevent future mistakes. Traceability facilitates auditability as well as explainability.
- **Explainability.** It concerns the ability to explain both the technical processes of an AI system and the related human decisions (e.g. application areas of a system). The HAIKU consortium ensures that the decisions made by an AI system can be understood and traced by human beings.
- **Communication.** AI systems should not represent themselves as humans to users; humans have the right to be informed that they are interacting with an AI system. The HAIKU consortium makes sure that AI systems are identifiable as such. In addition, the HAIKU consortium provides the option to decide against this interaction in favour of human interaction. Beyond this, the AI system's capabilities and limitations are communicated to end-users in a manner appropriate to the use case at hand.

5.3 Safety and Security of AI

AI is a game-changer, and it cannot be assumed that existing safety, security and Human Factors methods will be sufficient to give assurance that a new system is ready for use in live aviation operations.

The HAIKU consortium abides to the following principles:

- **Technical robustness.** It requires that AI systems be developed with a preventative approach to risks and in a manner such that they reliably behave as intended while minimising unintentional and unexpected harm, and preventing unacceptable harm.
- **Resilience to attack and security.** AI systems, like all software systems, should be protected against vulnerabilities that can allow them to be exploited by adversaries, e.g. hacking.
- **Fallback plan and general safety.** AI systems should have safeguards that enable a fallback plan in case of problems. This can mean that AI systems switch from a statistical to rule-based procedure, or that they ask for a human operator before continuing their action.
- **Accuracy.** Accuracy pertains to an AI system's ability to make correct judgements, for example to correctly classify information into the proper categories, or its ability to make correct predictions, recommendations, or decisions based on data or models.
- **Reliability and Reproducibility.** It is critical that the results of AI systems are reproducible, as well as reliable. A reliable AI system is one that works properly with a range of inputs and in a range of situations.

5.4 Accountability

To ensure the accountability of the AI systems developed in the project, the HAIKU consortium abides to the following principles:

- **Auditability.** Auditability entails the enablement of the assessment of algorithms, data and design processes. Evaluation by internal and external auditors, and the availability of such evaluation reports, can contribute to the trustworthiness of the technology.
- **Minimisation and reporting of negative impacts.** Identifying, assessing, documenting and minimising the potential negative impacts of AI systems is crucial for those (in)directly affected.

The consortium committed to addressing and mitigating potential liability risks from the outset of the design process, employing DBL's "liability by design" approach. This was implemented through the Legal Case methodology [18], an associated tool designed to facilitate the integration of automated technologies into complex organizations.

The Legal Case methodology aimed to proactively identify and manage liability concerns that arose from the interaction between humans and automated tools throughout the design, development, and deployment phases. It provided a structured framework for identifying, analyzing, and mitigating liability attribution issues, particularly when introducing new operational concepts and tools in intricate environments. While initially developed for Air Traffic Management, its application was readily transferable to other domains.

The Legal Case process was comprised of four key steps:

- **Understood Context and Concept:** This involved gathering information on the operational concept, its deployment context, and relevant legal and regulatory frameworks. It also included identifying the system's automation level, its impact on roles, tasks, and responsibilities, and defining use cases pertinent to the subsequent legal analysis.
- **Identified Liability Issues:** This step focused on pinpointing potential liabilities and assessing associated liability risks.
- **Addressed Liability Allocation:** This involved analyzing the acceptability of liability risks for all stakeholders, proposing mitigation strategies to improve liability allocation, and providing corresponding design recommendations.
- **Collected Findings and Systemic Analysis:** The final step involved documenting study results, highlighting identified liability issues and methods for managing legal risks, and formulating further recommendations. The Legal Case process was further illustrated below, demonstrating how Human Factors and Safety analysis informed the overall assessment.

5.5 AI Biases

To ensure that the AI systems developed in the project do not present unwanted and unjustified biases, the HAIKU consortium abides to the following principles:

- **Avoidance of unfair bias.** Data sets used by AI systems (both for training and operation) may suffer from the inclusion of inadvertent historic bias, incompleteness and bad governance models. The continuation of such biases could lead to unintended (in)direct prejudice and discrimination against certain groups or people,

potentially exacerbating prejudice and marginalisation. The HAIKU consortium aims to remove identifiable and discriminatory bias in the collection phase where possible. Bi Gender aspects are taken into account when developing HAIKU Use Cases, Intelligent Assistants, and validation activities. Gender has been shown to be relevant in the domain of digital assistance and AI [7]. Gender inequalities might indeed be reinforced due to the application of AI [5]. For instance, Bansal, Kockelman, and Singh showed that men are more willing to accept Intelligent Assistants in their vehicles than women [1]. Furthermore, the consortium puts in place oversight processes to analyse and address the system's purpose, constraints, requirements and decisions in a clear and transparent manner.

- **Accessibility and universal design.** Systems should be user-centric and designed in a way that allows all people to use AI products or services, regardless of their age, gender, abilities or characteristics. The HAIKU consortium considers Universal Design principles addressing the widest possible range of users, following relevant accessibility standards. This enables equitable access and active participation of all people in existing and emerging computer-mediated human activities and with regard to assistive technologies.
- **Stakeholder Participation.** In order to develop AI systems that are trustworthy, it is advisable to consult stakeholders who may directly or indirectly be affected by the system throughout its life cycle. In this regard, the HAIKU consortium solicits regular feedback and sets up longer term mechanisms for stakeholder participation throughout the whole process of implementing AI systems.

The HAIKU project aims to deliver solutions that are not gender-specific, however gender differences are considered in enrolling experimental subjects and controlling for potential gender variations with the appropriate sample size. The project takes concrete actions to meet all the three objectives on gender equality:

- Fostering gender balance in Horizon Europe research teams;
- Ensuring gender balance in decision-making, to reach the Commission's target of 40% of the under-represented sex in panels;
- Take into account gender/sex analysis in R&I.

5.6 Measures to Prevent Potential Ethical Risks

Specifically, regarding the misuses of AI results, the HAIKU consortium took the following actions:

- An ethical advisors committee was established.
- External ethical advisors were asked to provide an ethical assessment of each use case (UC) at the end of M12.
- Training and seminars on the ethical issues arising from the ethical assessment were provided to all the UC leaders and technical providers to assist them with the monitoring of the ethical issues.
- In case unforeseen ethical issues emerged over the research activities, the Consortium asked the Ethical advisors committee for an expert opinion.

- DBL ultimately carried on the monitoring of the ethical issues at M22 and M34 and synthesized the results in D1.4.
- The mapping of these issues was done on the basis of the ALTAI methodology defined by the HLEG on AI.
- The ALTAI definition was used throughout the development of the project to assess potential emerging issues.

6. Conclusions

This deliverable describes the measures taken to ensure that project activities were respectful of human rights, particularly the right to privacy and data protection, and did not generate ethically unwanted personal or social effects.

The report presented an overview of the main ethical issues arising from AI, an overview of the ethical issues stemming from the HAIKU research project and the way the consortium planned to mitigate them, involving the processing of personal data, and a set of measures that could be implemented to reduce the risk of misuse of the research results. Furthermore, the deliverable assessed the possible issues of the AI systems relating to human autonomy, explainability, safety and security, accountability, and biases.

7. References

- [1] Bansal, P., Kockelman, K. M., & Singh, A. (2016). Assessing public opinions of and interest in new vehicle technologies: An Austin perspective. *Transportation Research Part C: Emerging Technologies*, 67, 1-14.
- [2] Brooks, F. P. (1960). Operational sequence diagrams. *IRE Transactions on Human Factors in Electronics*, HFE-1(1), 33–34.
- [3] Dainow, B., & Brey, P. (2021). *Ethics By Design and Ethics of Use Approaches for Artificial Intelligence*. Brussels: European Commission.
- [4] EU Grants (2021). How to complete your ethics self-assessment. *Resource document*. European Commission. https://ec.europa.eu/info/funding-tenders/opportunities/docs/2021-2027/common/guidance/how-to-complete-your-ethics-self-assessment_en.pdf. Accessed, 14.
- [5] European Aviation Safety Agency. (2023). *EASA Artificial Intelligence Roadmap 2.0*. Cologne: EASA.
- [6] European Commission (2017). H2020 Programme–Guidelines on FAIR Data Management in Horizon 2020.
- [7] European Commission. (2021, July 5). *Note on ethics and data protection*. Publications Office of the European Union.
- [8] European Commission (2018). High-level expert group on artificial intelligence. In *Ethics Guidelines for Trustworthy AI*; European Commission: Brussels, Belgium. Available online: <https://op.europa.eu/en/publication-detail/-/publication/d3988569-0434-11ea-8c1f-01aa75ed71a1> (accessed on 11 December 2022).
- [9] Friedman, B., & Hendry, D. G. (2019). *Value sensitive design: Shaping technology with moral imagination*. Mit Press.
- [10] Humane AI. (2020). *Humane AI: Toward AI systems that augment and empower humans by understanding us, our society and the world around us*. Retrieved from <https://www.humane-ai.eu/>
- [11] Leavy, S., O'Sullivan, B., & Siapera, E. (2020). Data, power and bias in artificial intelligence. arXiv preprint arXiv:2008.07341.
- [12] Kletz, T. A. (1983). *HAZOP and HAZAN: Identifying and assessing process industry hazards*. Rugby, UK: Institution of Chemical Engineers.
- [13] Madiaga, T. (2019). *EU guidelines on ethics in artificial intelligence: Context and implementation*.
- [14] Nass, C., Moon, Y., & Green, N. (1997). Are machines gender neutral? Gender-stereotypic responses to computers with voices. *Journal of applied social psychology*, 27(10), 864-876.
- [15] SESAR Joint Undertaking. (2016). *SecRAM – Security Risk Assessment Methodology*. Brussels: SESAR JU.

- [16] Stahl, B. C. (2021). Ethical issues of ai. In Artificial Intelligence for a Better Future (pp. 35-53). Springer, Cham.
- [17] Timan, T., & Mann, Z. (2021). Data Protection in the Era of Artificial Intelligence: Trends, Existing Solutions and Recommendations for Privacy-Preserving Technologies. In The Elements of Big Data Value (pp. 153-175). Springer, Cham.
- [18] Contissa G. et al., Liability and automation: Issues and challenges for socio-technical systems, Journal of Aerospace Operations, vol. 2, no. 1-2, pp. 79-98, 2013