Explainable Manufacturing Artificial Intelligence

WP9: Project Management and Coordination

D9.2: Ethics and Data Management Plan

Deliverable Leader: Deep Blue

Due Date: April the 31st 2021

Dissemination Level: PU

Version: F 1.0

Short Abstract

This document outlines the main elements of the data management policy with regard to all research datasets that will be generated during the pilots' execution and the efficient management of publications will be agreed and followed by the Consortium, in accordance with the H2020 guidelines regarding Open Access to Scientific Publications and Research Data.

Further Information: www.ai4manufacturing.eu

Document Log

Contributors	Marika Rullo (DBL), Linda Napoletano (DBL), Andrea Capaccioli (DBL)
Internal Reviewer 1	Michele Sesana (TXT)
Internal Reviewer 2	Fenareti Lampathaki (SUITE5)
Type	ORDP
Delivery Date	M6

History

Versions	Description
D0.1	Set up the ToC
D0.2	Input from the pilots
D0.3	Completion of the sections of the document
R0.1	Revision of internal reviewers: Michele Sesana (TXT), Fenareti Lampathaki (SUITE5)
F1.0	Final

Executive Summary

This document describes the main elements of the data management policy with regard to all the **research datasets that will be generated during the pilots' execution.**

The report presents the first version of the Data Management Plan for the XMANAI Project- The DMP is produced to report on the main elements of the data management policy that the consortium will use. The XMANAI Data Management Plan is designed to monitor and report on industrial data such as:

- types of industrial data are to be handled during & after the project;
- the conditions under which data are made available in the demonstrators (for use only in the XMANAI project context); the applicable methodology & standards;
- the responsibilities of the XMANAI DPO (Data Protection Officer) and the liabilities of the XMANAI platform; the maintenance and sustainability of the data.

The efficient management of publications will be agreed and followed by the Consortium, in accordance with the H2020 guidelines regarding Open Access to Scientific Publications and Research Data.

This DMP follows the guidelines provided by the European Commission on FAIR-findable, accessible, interoperable and re-usable- Data Management in Horizon 2020.
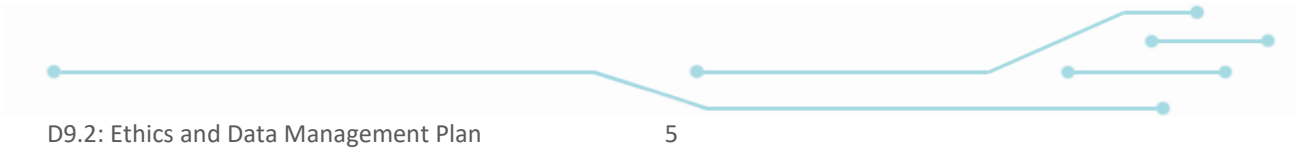
Table of Contents

1 Introduction

1.1 XMANAI Project Overview

What is artificial intelligence (AI) and how does it work? For many people, these questions are not easy to answer: this is due to the fact that many machine learning and deep learning algorithms cannot be examined after their execution. The EU-funded XMANAI project will focus on explainable AI, a concept that contradicts the idea of the 'black box' in machine learning, where even the designers cannot explain why the AI reaches at a specific decision. XMANAI will carve out a 'human-centric', trustful approach that will be tested in real-life manufacturing cases. The aim is to transform the manufacturing value chain with 'glass box' models that are explainable to a 'human in the loop' and produce value-based explanations.

1.2 Deliverable Purpose and Scope

This deliverable will serve as the basis for the main guidelines that partners will have to respect towards ethics and privacy protection and will be constantly updated during the course of the project. The DMP is built in accordance to the "Guidelines on FAIR Data Management in Horizon 2020" (1), which lists the following points that are to be addressed by a DMP that promotes FAIR data:

- the handling of research data during and after the end of the project,
- what data will be collected, processed and/or generated,
- which methodology and standards will be applied,
- whether data will be shared/made open access, and
- how data will be curated and preserved (including after the end of the project).

The DMP is a living document, which will evolve during the lifespan of the project. Particularly whenever significant changes arise, such as dataset updates or changes in consortium policies.

This document is the first version of the DMP, delivered in M6 of the project. It provides a framework for data documentation. Although this report already covers a broad range of aspects related to the XMANAI data management, the upcoming versions will get into more detail on particular issues such as data interoperability and practical data management procedures implemented by the XMANAI project consortium.

1.3 Nomination of Data Protection Officer

Data Protection Officer (DPO) for the project will be responsible for the handling and management of personal data in accordance with the existing provisions of GDPR and other relevant EU and national legislations. The responsibilities of the DPO will be in line with Article 39 of the GDPR and will include, at a minimum, the following:

1. Advise the data controller and processors within the XMANAI project on the processing of personal data and data privacy issues, training of researchers and assignment of responsibilities.
2. Provide support on the performance and tracking of Impact Assessments
3. Assist in risk assessment of personal data processing.
4. Cooperate with any national or European supervisory authority and act as contact point for the project with such authorities.
5. Act as contact point for the purposes of consent

Table 1 shows the assigned DPO manager, chosen by the project partners.

Partner name company	Name	Email
DBL	Linda Napoletano	linda.napoletano@dblue.it
Linda Napoletano (F) holds a Ph.D. in Human Computer Interaction. She has been working since 2002 on EU co-funded projects aiming at designing and validating humans' integration into technological innovation processes. She has worked for Deep Blue since 2008 as a human factors, validation, dissemination and training expert working in EU, EUROCONTROL and ENAV projects. Linda is Associate Researcher in Human Computer Interaction and Service Design at the University of Sassari and Siena. Since 2009, Linda worked as Dissemination Manager in two large cooperative EU projects ALICIA and ACROSS (IP-FP7) and the CSA Mobility4EU (H2020). With 15+ years of experience with EU projects, Linda acts as project coordinator on behalf of Deep Blue. By coordinating the team of graphic designers and communication experts, she supported the definition of the dissemination strategy and plan, and its implementation throughout the project life cycle. She supported the constitution and moderation of the external advisory boards and she ensured their role within the projects in close collaboration with the coordinators and the technical leaders. Linda contributed to many data collection trials, including a number of trials in which automations and innovative technologies and devices were evaluated.			

1.4 Deliverable Methodology

The Data Management Plan of XMANAI is realized within the Work Package 1. The XMANAI project data management plan follows the principle of Open Access according to the Horizon 2020 guideline.

The others main principles for the XMANAI project DPM are the following:

- This Data Management Plan (DMP) has been prepared by taking into account the template of the "Guidelines on Data Management in Horizon 2020"
- The DMP is an official project Deliverable (D1.9) due in Month 6 (April 2021), but it will be updated throughout the project. The first initial version will evolve depending on significant changes arising and periodic reviews at relevant reporting stages of the project.
- The consortium complies with the requirements of Regulation (EU) 2016/679 and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation).

1.5 Document Structure

This document is divided into six sections:

- **Section 1** provides a short introduction to the project and to the purpose of this document, including its structure.
- **Section 2** introduces main aspects relevant to the project data, procedure of data collection, exposing FAIR data principles in the management of XMANAI data (i.e., aspects associated with making data findable, openly accessible, interoperable and re-usable).
- **Section 3** describes the security aspects (i.e., protection of personal data).
- **Section 4** discusses key ethical aspects associated with data collection with human participants and refers European legislation related to ethics in research
- **Section 6** lists key references supporting this document and its preparation.

- **Section 7** contains acronyms and abbreviations
- **Section 8** includes the template of the consent form

2 XMANAI data management policy

1

2.1 Description of the data used and produced in the XMANAI project

XMANAI aims at placing the indisputable power of Explainable AI at the service of manufacturing and human progress, carving out a "human-centric", trustful approach that is respectful of European values and principles, and adopting the mentality that "our AI is only as good as we are". XMANAI, demonstrated in 4 real-life manufacturing cases, will help the manufacturing value chain to shift towards the amplifying AI era by coupling (hybrid and graph) AI "glass box" models that are explainable to a "human-in-the-loop" and produce value-based explanations, with complex AI assets (data and models) management-sharing-security technologies to multiply the latent data value in a trusted manner, and targeted manufacturing apps to solve concrete manufacturing problems with high impact.

This section describes the types of industrial data are to be handled during & after the project; the conditions under which data are made available in the demonstrators (for use only in the XMANAI project context); the applicable methodology & standards; the responsibilities of the XMANAI DPO (Data Protection Officer) and the liabilities of the XMANAI platform; the maintenance and sustainability of the data, etc.

Two levels of data should be identified in the process:

Industrial pilot data including private data coming from end-users assets which will be used to train AI models and they will be the source of the analysis of the trained models to provide predictions and any other analysis needed in the project. These data are private data covered by industrial secret and so they will be not shared with the scientific community. They will be used as exploitation assets in the business development of XMANAI, please refers to XMANAI exploitation deliverables (D7.1) at chapter 2.3.7 "Data Assets" and chapter 2.3.9 "manufacturing data model".

Human-XAI collaboration data focusing on the interaction happening between the XAI as decision support system and the human user. The scope of this data collection is of high importance for the scientific community focusing on the project focal point: transform the AI black box in a XAI glass box providing a human-centric trustful approach. The type of data managed includes:

- data coming from the decision support provided by the explainable AI including the suggestions to be executed
- user behavior in approaching XAI including AI trust (how many times the human agree with the suggestion provided or the most confident suggestion) the ability to reach explanation (human actions in accessing explainability functions), the overall trust in AI (including also qualitative analysis)
- the process of AI evolution thanks to explainability usage including the algorithm/system improvements led by XAI analysis by the user

These data will be collected, anonymized and released to the scientific community.

2.2 Data summary

2.2.1 Data collection process

XMANAI will both collect existing data from demonstrators and will create new data within the project. XMANAI will collect and produce data for design and evaluation of explainable AI, research data and metadata, manuscripts and dissemination material. A complete list of all data to be collected and created is shown in Table 1. The additional information about each item will be explained in the following sections.

2.2.2 Data collection description

In order to provide an overview of the different "**Human-XAI collaboration data**" data sets that are currently and will be produced in the XMANAI project, the following table shows the data type, the origin of the data, the related WP number and the format, in which the data will be presumably stored.

Table 1. data collection description

Ref. nr.	Responsible Partner	Data type	Origin	WP	Format
1	FORD	Qualitative analysis (online questionnaire, interview)	FORD end-users	WP6 (T6.4)	.xls and .docx
2	WHIRLPOOL	Qualitative analysis (online questionnaire, interview)	WHIRLPOOL end-users	WP6 (T6.5)	.xls and .docx
3	CNH	Qualitative analysis (online questionnaire, interview)	CNH end-users	WP6 (T6.6)	.xls and .docx
4	UNIMETRIK	Qualitative analysis (online questionnaire, interview)	UNIMETRIK end-users	WP6 (T6.7)	.xls and .docx
5	SUITE5	Quantitative Human-XAI collaboration data (User Interactions path, DSS suggestions, user final choice, etc.)	XMANAI platform	WP5 (technical platform), WP6 user interactions	Sql DB
6	KBIZ	dissemination materials	presentation, flyers, video	WP8	.pdf, .mov

Table 2 describes the data set and the purpose of the data collection of data generation in relation with the objectives of the project. Additionally, it shows the data utility for clarifying to whom the data might be useful.

Ref. nr.	Responsible Partner	Data type	Description and Purpose	Utility
1,2,3,4	INNOVALIA	Qualitative analysis (online questionnaire, interview)	The qualitative analysis provide a feedback from humans about their experience in collaborating with AI and their relation/feeling in having the explainability	Scientists can use the data to understand more about the usefulness of XAI in creating trustiness and full adoption and scale-up for the AI
5	SUITE5	Quantitative Human-XAI collaboration data (User Interactions path, DSS suggestions, user final choice, etc.)	end-to-end technical log of the human interaction with XAI. From the prediction presentation to the user, and human usage of explanation functionalities to the user whether or not to trust AI and follow the suggestion or to execute a new training of the algorithm.	Scientists can use the data to understand more about the usefulness of XAI in creating trustiness and full adoption and scale-up for the AI. Moreover the different HMI interfaces can be evaluated and improved accordingly
6	KBIZ	dissemination materials (presentation, flyers, video-interviews, etc.)	The material includes interpretation and summarisation of data from human-XAI interaction complementing the rough data	Scientists can use the data to understand more about the usefulness of XAI in creating trustiness and full adoption and scale-up for the AI

2.2.3 Data storage and back-up

Data will be stored and managed complying with the jurisdictions of the countries where the participants reside. XMANAI will activate best practice in the technology employed by the project to collect, store and process data that is in any way attributable. In relation to data retention, the contracted researchers will destroy all data containing private data as soon as the study/research task will be completed.

It is the responsibility of the project partner who collects the data to ensure that the data is regularly backed-up and stored securely for the lifetime of the project. Storage indicates the medium and location of the backups.

Example of location for storage are Network drives, local drives, remote or cloud storage and external storage devices (e.g. usb drives)

Backup indicates the location and frequency of the backups. The data in this study will be stored and backed up as described in **Table 4.3.**

2.2.4 Making data findable, including provisions for metadata

XMANAI will use the Zenodo repository as the main tool to make our research data findable in accordance with the H2020 Open Access Mandate.

A H2020 XMANAI community has been established on the Zenodo website, and the project will upload all our public datasets and deliverables as well as scientific publications to this community. In addition, we will link all our uploads to the European Commission Funded Research (OpenAIRE) community for maximum findability. All uploads will be enriched with standard Zenodo metadata, including Grant Number and Project Acronym. Zenodo provides version control and assigns DOIs to all uploaded elements.

As described in the project Grant Agreement (Article 29.2), metadata associated with each published data set in Zenodo will by default be as follows:

1.   Digital Object Identifiers
2.   Version numbers
3.   Bibliographic information
4.   Keywords
5.   Abstract/description
6.   Associated project and community
7.   Associated publications and reports
8.   Grant information
9.   Access and licensing info
10.   Language

2.2.5 Making data openly accessible

Due to privacy and security concerns and the commercial interests of the industrial partners the research data will not be made openly accessible as primary data but in a processed form. It is planned to make the anonimised "**Human-XAI collaboration data"** openly accessible on the data repository Zenodo, with the definition *human-AI collaboration*.

Furthermore, reports working with the qualitative data (e.g. interview, workshops) will also be accessible. The consortium will also aim at open access when publishing papers and articles.

Concrete dissemination, communication and stakeholder engagement actions will be released to build awareness around XMANAI (both as a project and as a "product"). In workshops, webinars and demonstrations during relevant events and industry fairs, the XMANAI results will be validated by external stakeholders, providing further insights on the market readiness of the XMANAI exploitable assets. Any publications from XMANAI should be available as open access (as far as practicable possible, including within the limitations of the project's budget). In alignment with the EC Guidelines on Open Access to Scientific Publications and Research Data in Horizon 2020, XMANAI will also follow a combination of Gold and Green Open Access strategy to its scientific publications, with a potential embargo period for peer-reviewed publications that will be agreed during the first months of project execution. Gold open access means the article is available as open access by the scientific publisher. Some journals require an author-processing fee for publishing open access. Green open access or self-archiving means that the published article or the final peer-reviewed manuscript is archived by the researcher in an online repository, in most cases after its publication.

2.2.6 Making data interoperable

Partners will observe OpenAIRE guidelines for online interoperability, including OpenAIRE Guidelines for Literature Repositories, OpenAIRE Guidelines for Data Archives, OpenAIRE Guidelines for CRIS Managers based on CERIF-XML. These guidelines can be found at: https://guidelines.openaire.eu/en/latest/.

As already introduced DMP is a living document, which will evolve during the lifespan of the project; for example by M18, once the pilots will be set up, the document will include specific information on data and metadata vocabularies, standards or methodology to facilitate interoperability and whether the project uses standard vocabulary for all data types present to allow interdisciplinary interoperability.

2.2.7 Allocation of Resources

The consortium opted for the free-of-charge Zenodo repository for making the dataset Human-XAI collaboration data accessible in order to ensure a safely storage of data and a long-term preservation and curation.

The handling of the Zenodo repository on behalf of XMANAI as well as all data management issues related to the project fall in the responsibility of the coordinator.

Each partner is responsible for the data they produce. Any fee incurred for Open Access through scientific publication of the data will be the responsibility of the dissemination manager budget.

3 Data security

3.1 Data management during and after the project execution

Data will be stored and managed complying with the jurisdictions of the countries where the participants reside. XMANAI will embed legal best practice in the technology employed by the project to collect, store and process data that is in any way attributable. Wherever possible, data on user experiences collected by demonstrators will be anonymized at source, such that nowhere in the data a direct connection to the individual can be drawn. For all members of the Consortium, these are routine procedures for privacy and data protection.

In relation to data retention, the contracted researchers will destroy all data containing private data as soon as the study/research task will be completed. In any case all personal data will be destroyed automatically in a maximum of two years after the end of the project if the data owners don't provide explicit consent to keep them and only anonymous or non-identifiable data will be retained after the completion of the final report.

Where applicable, for highly sensitive industrial pilot data, staff involved in the project will sign a 'Declaration of confidentiality'. In the unlikely event that staff hired in this project have given away private information from anyone working on the project or complaints about such a behaviour will be handled by the Consortium and the appropriate disciplinary action will be taken through them, including, if necessary, the withdrawal of the staff from the project due to professional misconduct and following the rules of the Consortium Agreement.

Validation activities envisaging the involvement of users will perform specific training/awareness activity to these users before the start of the activities. This training will both explain the objectives and organization of the trial or validation activity, as well as clearly detail which information will be acquired, using which means and how the acquired information will be managed and processed.

Industrial pilot data and Human-XAI collaboration data will be stored in repositories with the highest security standards. The Consortium is currently assessing different options:

- Servers made available by the technical/project coordinator,
- Servers onsite in the four different use-cases,
- Cloud services provided by AWS (Amazon Web Services) or Google Cloud depending on which of the providers is able to ensure data storage and back-up within the EU

The policies to react to cyber-attacks will be set-up by DBL according to the risk assessment which will be performed before the deployment.

In addition to software and hardware security standards, data will be protected by applying strict data protection procedures.

3.2 Protection of personal data

Any personal data will be processed legally and fairly: collection of data will be adequate, relevant and not excessive in relation to the purposes of the project; data that identifies individuals (personal data) will not be kept any longer than necessary: once the project has finished, data will be completely anonymised if possible, meaning irreversibly preventing identification of the data subject. Any personal data will be destroyed two years past the termination of the project. The Consortium will comply with European (i.e. GDPR) and national legislation relevant to the countries where data collection is taking place.

Data collected for research purposes will be subjected to current European regulations on matters of data handling and privacy (GDPR, Regulation (EU) 2016/679). The research outcomes will always be reported without contravening the right to privacy and data protection.

- **Secure Access Policy**: Data will be encrypted, and password protected. Only members of the team directly working with the data ("need to know") will have authorisation to access the data. Each Data Controller will be responsible for de-identification of the data or establishing a procedure to be followed by other partners in charge of personal data. This data might be transferred for further processing to other project members.
- **Secure Storage**: Location and Hardware. All personal data will be stored on digital hard disks on computers that are not connected to WAN Internet. Removable storage will include large capacity hard drives that will be kept in locked cabinets.

- **Monitoring of Data Transfer:** The data will not be transferred outside the XMANAI Consortium without prior authorisation.

In any case all personal data will be destroyed automatically in a maximum of two years after the end of the project if the data owners don't provide explicit consent to keep them and only anonymous or non-identifiable data will be retained after the completion of the final report. Staff involved in the project will sign a 'Declaration of confidentiality'. In the unlikely event that staff hired in this project have given away private information from anyone working on the project or complaints about such a behaviour will be handled by the Consortium and the appropriate disciplinary action will be taken through them, including, if necessary, the withdrawal of the staff from the project due to professional misconduct.

Validation activities envisaging the involvement of users will perform specific training/awareness activity to these users before the start of the activities. This training will both explain the objectives and organization of the trial or validation activity, as well as clearly detail which information will be acquired, using which means and how the acquired information will be managed and processed. The data subject's rights, as set forth by GDPR, will be guarantee to him/her, including:

- the right of access to own personal data;
- the rights of erasure, blocking or rectification of the data, which do not comply with the provisions of the Directive, are incomplete or inaccurate;
- the right to be informed of all relevant details relating to the data processing and the rights granted to self;
- the right to a judicial remedy for any breach of the above-mentioned rights.

4 Ethical aspects

4.1 Ethical and Legal Framework

The XMANAI Consortium is fully aware of the ethical implications of the proposed research and respects the ethical rules and standards of HORIZON 2020, and those reflected in the Charter of Fundamental Rights of the European Union. Ethical, social and data protection considerations are crucial to this project and will be given all due attention to them. This section provides an ethics self-assessment performed by the partners, which guarantees that the ethical issues that may possibly arise are well thought and dealt with before the start of the project.

XMANAI will involve human participants and might deal with personal data during the activities of researching the collaboration between humans and AI models and the mapping of the human knowledge to the domain and the decisions taken by business experts in manufacturing systems, towards creating more trusted, explainable and robust AI models and machine learning algorithms. However, it will be ethically driven and privacy-aware by design: the collected data belong to the business, who will remain the owner of them and will be able to freely and consciously choose what to share and with whom and what to remove. During the implementation phase, the consortium will guarantee the proper handling of ethical issues and the adherence to national, European and international law, regulations and directives and more specifically:

- Regulation (EU) 2016/679 (General Data Protection Regulation)
- The Universal Declaration of Human Rights and the Convention 108 for the Protection of Individuals with Regard to Automatic Processing of Personal Data
- The European Convention for the Protection of Human Rights and Fundamental Freedoms.
- Directive on Privacy and Electronic Communications (2002/58/EC) and the upcoming "ePrivacy Regulation", aiming at repealing it for reinforcing trust and security in the Digital Single Market, as well as at ensuring alignment with GDPR new rules;
- Article 19 "Ethical principles" of Regulation No. 1291/2013 /EC of the European Parliament and of the Council which states the fundamental principles of the H2020 Ethics in research;
- Directive 96/9/EC of the European Parliament and the Council of 11 March 1996 on the legal protection of databases;
- The Charter of Fundamental Rights of the European Union (2000);
- Directive EU/2016/1148 "NIS Directive", and the upcoming Regulation on Privacy and Electronic Communications

In order to conform to national and international ethical guidelines, every participant involved in the project will receive an introduction and continued support from the ethical board. The partners involved in the pilots will receive this information before involvement in any XMANAI activities dealing with data related to piloting or data they will use for research. This is the responsibility of each partner organisation.

4.2 Ethics research principles and guidelines

The research activities during XMANAI involve the participation of humans and will collect human-AI interactions; to the extent that the project will interact with individuals in the demonstrators and gather, process and share data on their work environment. The purpose of this research fully respects human dignity and ethical values are fully to the side of the individual who will be involved on a voluntary basis: the project will contribute to improve the general quality of wok of business experts and will serve them with value added information and services. The free and aware choice of the data subjects is assured by the provision to them of the risk assessment and privacy exposure report regarding their shared data. In this way, both in the validation phase and in the post-project uptake of XMANAI technologies, it will be concretely ensured that the individual's consent is freely given, specific, informed and unambiguous, as requested by GDPR art. 7 and Recitals 32. Likewise, situations where individuals feel pressurized to share are avoided, while safeguarding human dignity at the maximum extent.

The data processing foreseen by XMANAI does not involve the collection of any sensitive personal data, tracking and observation of participants. In such a context, the results of the research have zero risk and burden to the participants and will benefit the individual or the group represented by the participant. Actually, the individuals are recognized as the starting point and key players of XMANAI value chain and their significant role in the business decision making is coupled with the attribution of data's (and its derivatives) value to them. They will gain benefits as employees (e.g. through insights provided by better and more explainable AI models). Furthermore, they will even profit from XMANAI as member of a group of an educated workforce, which will take advantage from the improved data-driven efficient, effective and value-added services, thanks to the availability of additional tons of personal data, complementing the ones they already manage.

The XMANAI Ethical Policy relies on the following actions:

- Ethical Issues will be addressed at a first level within the consortium, for potential ethics issue arising from the activities to be performed during the project
- Organization of open discussion and consultation with stakeholders on the privacy and ethics issues arising from project research. These participatory events will also address ethical and legal concerns and expectations;
- All the subjects involved in the demonstration activities will be strictly volunteers. The voluntary participants will be recruited from the pilot organizations following recruitment procedures and inclusion/exclusion criteria previously set, including also details on the sampling and recruitment process for each of the demonstrators. The XMANAI (scientific and technological) objectives and project advancement, as well as use cases' refinement and pilots' planning and operation, will drive the identification of voluntary subjects. The demonstrators will serve as a basis for identifying the number of volunteers (i.e., the sample size), as well as relevant statistical parameters of the sample. All necessary steps will be taken to eliminate bias within the selection process in order to avoid discrimination based on physical and cognitive aptitude and political, social, religious and cultural, gender orientation. Under no circumstances will vulnerable subjects be selected as XMANAI voluntary participants; this includes persons under the age of 18 and any other person unable to give the informed consent.
- Informed consent procedures for the participation of humans will be defined and implemented (see Annex. XMANAI will maintain transparency, i.e. it will never intentionally

deceive, mislead or withhold information from participants as to the purpose and nature of the investigation. All participants in the demonstrators will be informed of all aspects of the activities that might reasonably be expected to influence willingness to participate. The foundation of ethical research is the principle of informed consent. The voluntary participants, who agree to take part in a specific exercise, after the described information session and delivery of an information sheet, will be asked to fill and sign a consent form. The informed consent form provided to, and signed by, the volunteers involved in the use cases will explain all details of the use case or validation activity according to specific National and EU legislation. The project researchers will also discuss before and after each experiment with participants to maintain on-going consent. Informed consent procedures for data collection, storage and processing will be also defined and implemented. Each potential voluntary participant will be adequately informed, before being involved in XMANAI, on the project and on the ethics and regulatory framework. They will be provided with an information sheet describing the project, its objectives and status, an explanation on the particular research activity related to the pilot and the test, the information to be collected and how that information will be used. Then they will be invited to sign a consent form. The Consent Form will clearly state that the project will preserve user's right to privacy and anonymity, and that the data arising from his/her participation will be only used for scientific research and system's validation purposes. This form will detail the objectives, all elements (e.g., purpose, duration, procedures), and all the steps of the trial or activity, as well as data subject's rights. Participants will have the right to withdraw from the trial at any time, and to withdraw retrospectively any consent given and to demand that their personal data be destroyed. The Consent Form will be tailored specifically to each different test/technology and demonstrator's features. In all cases volunteers can withdraw at any time during any workshop and test session.

- If applicable, the ethics approvals for the involvement of humans or any other relevant ethics authorisation/approval in each demonstrator sites will be obtained and kept of file.

As regards data collection, processing, storing and sharing, the XMANAI Ethical Policy will outline adequate procedures and safeguards ensuring efficient data and privacy protection, based on the following:

- After the end of the project, all collected data that can relate to individuals will be put on hold and users will be noticed to extend their presence from the platform or get a copy of their data and remove it completely from the platform.
- All the test subjects will be informed and given the opportunity to provide their consent to any monitoring and data acquisition process.
- No personal or sensitive data will be centrally stored without consent. In addition, data will be scrambled where possible and abstracted in a way that will not affect the final project outcome.
- No data collected will be sold or used for any purposes other than the current project.
- A data minimisation policy will be adopted at all levels of the project. This will ensure that no data which is not strictly necessary to the completion of the current study will be collected.
- Anonymisation of the users' data right at their source for sensitive information used within the XMANAI demonstrators, by removing the direct identifiers (e.g., name, address etc.),
- The data to be stored in the platform will be held securely using state of the art encryption methods and the other techniques described in the project (see section 7.1).

- Any shadow (ancillary) personal data obtained during the course of the research will be immediately cancelled. However, the plan is to minimize this kind ancillary data as much as possible.
- Any incidental findings will be kept strictly confidential and erased from files under request from the enrolled subject.

In order to carry this process forward, the consortium will work from the very early stages of the project together with the demonstrator users towards understanding the complete procedures needed to obtain consent for researching such data. Apart from the already research information, all other possible information sources will be analysed and included in the process of identifying:

- What data and information are relevant to the XMANAI concept that will benefit the whole platform differentiating "**Industrial pilot data**" and "**Human-XAI collaboration data**"
- Whether consent to access this kind of data is necessary (e.g. open data publicly available vs. personal data protected by various laws).

All of the above will be discussed internally within the consortium and will be checked against national laws and GDPR and other EU legislations towards understanding the differences in each country and identifying the process to be followed for each pilot. A continuous interaction will be established and maintained between the Coordinator and the WP Leaders and the Ethics Manager in order to ensure the efficient management of ethical, privacy and data protection issues. This will allow to regularly seek input from all the WP Leaders, prepare and regularly update the Ethical Policy, as well as provide guidance and recommend workflows that comply with current European and national legislation and ethical standards, and are well accepted by society. This:

- Will raise awareness by leaders of all relevant work packages of salient aspects of the project's impact on society, to couple scientific excellence with social awareness and responsibility and to align innovative outcomes to the values, needs and expectations of European society, and
- Will ensure that leaders of all relevant work packages include in their work flows the processes needed to make their project solutions and activities compliant with existing EU and national legislations and H2020 ethical guidelines.
- Any complaints or adverse incidents will be referred to the Coordinator and the Ethics Manager, for independent review and appropriate disciplinary action if necessary.

4.3 XMANAI Pilots – Ethical research standards

4.3.1 Informed consent

Informed consent represents a strategy to preserve confidentiality and ensure anonymization that are strongly related to the privacy and data protection rights (European Commission, 2010). Informed consent implies a consent of subjects to voluntary participate in the research and to release any personal data. In this sense, prior to take part to the research activity, participants should have obtained all the information needed to make an informed decision to participate.

All participants taking part in research activities within the XMANAI project, must give explicit consent to participate. For this purpose, they will receive an information letter and a consent form (see Annex X) containing information on:

- What data will be collected
- Means of collection (interview, observation, etc.)
- The purpose of the project and the expected results
- Information on anonymization of data and confidentiality
- Information on secure data storage, retention period and procedures for de-identification of any personal data
- Information on access rights and authorised personnel
- The rights of participants (i.e. participation is voluntary and can be withdrawn at any time without explanations and repercussions)
- Contact information for requests

Moreover, a consent to use data in aggregated form to publication purpose will be acquired.

Participants should provide the explicit consent to participate in written form. The **informed consent form (Annex 1)** has to include the information sheet and a certificate of consent.

The research activities during XMANAI involve the participation of humans and personal data collection and handling, to the extent that the project will interact with individuals in the demonstrators and gather, process and share data on their work environment. The purpose of this research fully respects human dignity and ethical values are fully to the side of the individual who will be involved on a voluntary basis: the project will contribute to improve the general quality of work of business experts and will serve them with value added information and services. The free and aware choice of the data subjects is assured by the provision to them of the risk assessment and privacy exposure report regarding their shared data. In this way, both in the validation phase and in the post-project uptake of XMANAI technologies, it will be concretely ensured that the individual's consent is freely given, specific, informed and unambiguous, as requested by GDPR art. 7 and Recitals 32. Likewise, situations where individuals feel pressurized to share are avoided, while safeguarding human dignity at the maximum extent.

4.3.2 Gender Issues

The consortium partners are committed to active promotion of gender equality within XMANAI and fully support the European Union policy on equal opportunities between women and men, as well as between people with social/cultural backgrounds. Participation of women in XMANAI is continually encouraged and further supported by the fact in the modern times, the AI industry is gender agnostic. From a technological perspective, though, lack of high-qualified professionals from both genders is widely acknowledged, thus it is essential to attract, promote and retain well-trained and highly

qualified women in the XMANAI-related technologies. Social, protective and emotional skills that are especially attributed to women should be proactively supported during the project and contribute to a higher acceptance of technological solutions by both genders. Hence, XMANAI takes the diverse realities of both sexes and different social/cultural factors into account during the development of innovative and usable technologies and solutions for the manufacturing industries, through the following specific steps: (i) developing and implementing equal opportunities policy; (ii) implementing selection procedures transparent and free of gender bias; (iii) setting targets to achieve a gender balance less/equal skewed compared to the present representation of female/male personnel participating in decision-making. In XMANAI, the gender aspect will be considered through different requirements and context uses, during the planning, design and evaluation activities, to ensure that the acquired results are valid and applicable for both genders:

- The composition of the project partners' members will be constantly monitored by the coordinator to ensure a social equality basis, engaging a balanced number of women and men. The participation and equality of women in management positions (WP leadership and / or partners' team management) is promoted while the Technical Coordinator is female.

- The composition of the team involved in the AI models design and evaluation, as well as in the piloting activities, is based on the current gender distribution of the population (at maximum degree, as far as possible) and shall be characterized by diversity in terms of in age, education and socio-cultural backgrounds.

- The execution of dissertation research related to XMANAI will be promoted to the female PhD students of the universities involved in the project.
  Finally, the XMANAI gender balancing approach shall follow the Responsible Research and Innovation (RRI) approach, which is defined as an inclusive approach to Research & Innovation (R&I) aiming at better aligning both the process and outcomes of R&I with the values, needs, and expectations of society.

Also, the findings of the project will not affect women and men differently. The outputs (e.g. papers, reports) are expected to do not contain any assumptions or biases with regard to sex and/or gender.

4.3.3 Copyright and Intellectual Property Rights (IPR)

This project is being carried out in collaboration with an industrial partner. The intellectual property rights are set out in the consortium agreement signed by all partners at the beginning of the project.

5 Conclusions

This document describes the man principles and guidelines for the Data Management for the XMANAI project. The DPM is a living document and it will be updated throughout the project lifetime. The document describes the two levels of data identified in the process:

Industrial pilot data including private data coming from end-users assets which will be used to train AI models and they will be the source of the analysis of the trained models to provide predictions and any other analysis needed in the project. These data are private data covered by industrial secret and so they will be not shared with the scientific community. They will be used as exploitation assets in the business development of XMANAI.

Human-XAI collaboration data focusing on the interaction happening between the XAI as decision support system and the human user. The scope of this data collection is of high importance for the scientific community focusing on the project focal point: transform the AI black box in a XAI glass box providing a human-centric trustful approach. These datasets are the focus of this document involving human-AI interaction observation and being the target of open access publication.

Within the updated Data Management Plan the consortium will update information on online research data repository, which and how the data will be made available to others within the Pilot on Open Research Data in order to comply with the FAIR principles.

6 References

Article 19. Ethical Principles., Regulation (EU) No 1291/2013 of the European Parliament and of the Council of 11 December 2013 establishing Horizon 2020 - the Framework Programme for Research and Innovation (2014-2020) and repealing Decision No 1982/2006/EC

Directive 96/9/EC of the European Parliament and of the Council of 11 March 1996 on the legal protection of databases OJ L 77, 27.3.1996, p. 20–28

Directive 2002/58/EC of the European Parliament and of the Council of 12 July 2002 concerning the processing of personal data and the protection of privacy in the electronic communications sector (Directive on privacy and electronic communications)

Dvořák, Jan, Bollini, Andrea, Rémy, Laurent, & Schirrwagen, Jochen. (2018). OpenAIRE Guidelines for CRIS Managers 1.1 (Version 1.1.1). Zenodo.

EU General Data Protection Regulation (GDPR): Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), OJ 2016 L 119/1.

European Commission. H2020 Programme. Guidelines on FAIR Data Management in Horizon 2020. Version 3.0, 26 July 2016

European Commission. Horizon 2020 Programme. Guidelines to the Rules on Open Access to Scientific Publications and Open Access to Research Data in Horizon 2020. Version 3.2, 21 March 2017

European Union, Charter of Fundamental Rights of the European Union, 26 October 2012, 2012/C 326/02, available at: https://www.refworld.org/docid/3ae6b3b70.html [accessed 24 May 2021]

Regulation (EU)No 1290/2013of the European Parliament and of the Council of 11 December 2013 laying down the rules for participation and dissemination in "Horizon 2020 - the Framework Programme for Research and Innovation (2014-2020)" and repealing Regulation (EC) No 1906/2006, OJ L347/81

European Commission, Directorate-General for Research & Innovation (2016). Guidelines on FAIR Data Management in Horizon 2020.

7 List of Acronyms/Abbreviations

Acronym/Abbreviation	Description
AI	Artificial Intelligence
DPO	Data Protection Officer
DOI	Digital Object Identifier
FAIR data	Findable, Accessible, Interoperable, Re-usable data
GDPR	General Data Protection Regulation EU regulation (regulation (EU) 2016/679) on the protection of natural persons with regard to the processing of personal data and on the free movement of such data.
ORDP	Open research data pilot
WP	Work package
XAI	eXplainable Artificial Intelligence

8 Annexes

8.1 Annex I: Informed consent

Information Sheet

Name/description of the Organization/Consortium (To be completed with a general description) XMANAI Consortium Contact Person (s):

This research is organized by [research leader] and within the [state the Institution]

Please take the time to read the information carefully.

Your participation in the project is completely **voluntary**. You are free **to withdraw** from the project, without giving a reason for your withdrawal and without any consequences to your future treatment by the researcher.

This research aims to (**Purpose of the research and of data collection**). This [**method of collection: eg. Questionnaire, interview, observation**] will [describe the request] about [describe the topic of investigation]. It will take approximately [**time for completing**] to complete this study.

Potential risks from partecipating: [if any].

Incentives for partecipating : [if any].

Pre-condition to participate: You must be currently [describe pre-conditions to participate]

What happens to your data?

Your answers will be kept completely confidential and you will never be named in any research. If you agree to participate, all the data that you have provided will remain anonymous, will not be disclosed to other parties and will only be presented in an aggregated form.

Can you withdraw from this research?

Yes, you can withdraw at any point before, during and after the study. If you wish to withdraw please contact one of the following:

- [Contact name]

 Contact for further information

If you require any further information or have any queries about this study please contact the lead researcher:

If you have any serious concerns about the ethical conduct of this study, please inform the [information about the DPO] in writing, providing a detailed account of your concern.

Consent sheet

I confirm that I have read and understand the information sheet for this study.

I understand that my participation is voluntary and that I am free to withdraw at any time, without

giving any reason.

I understand that only the research team from the [**Institution**] who are working on the project will have access to my answers.

I understand that any data or information used in any publications which arise from this study will be anonymous.

I understand that all data will be stored securely.

I understand that [topic of the study] will be investigated within this [**method of collection**]