Accelerated deployment of integrated CCUS chains based on solvent capture technology

Deliverable D7.2 – Data Management Plan

WORK PACKAGE N°7 – Project and IPR Management

TASK 7.4: Data Management

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or CINEA. Neither the European Union nor the granting authority can be held responsible for them.

Deliverable Report

Start date of project:	01/01/2023
Duration of project:	42 months
Deliverable nº & name:	D7.2 – Data Management Plan
Version	**1.0**
Work Package n°	**7**
Due date of D:	M6
Participant responsible:	**SINTEF**
Main authors:	Hanne Kvamsdal, Inna Kim, Asmira Delic all SINTEF
Website:	https://aurora-heu.eu/

Nature of the Deliverable		
R	Document, report (excluding the periodic and final reports)	X
DEM	Demonstrator, pilot, prototype, plan designs	
DEC	Websites, patent filing, press & media actions, videos, etc.	
OTHER	Software, technical diagram, etc.	

Dissemination Level		
PU	Public, fully open, e.g. web	X
CO	Confidential, only for members of the consortium (including the Commission Services)	

Quality procedure			
Date	Version	Reviewers	Comments
	V1.0	Consortium members	First draft
26.06.2023	V1.0	Karl Anders Hoff (SINTEF)	QAM SINTEF

ACKNOWLEDGMENTS

This report is one of the 25 deliverables of the project "AURORA" which has received funding from the European Union's HORIZON EUROPE research and innovation programme under Grant Agreement n° 101096521.

PROJECT SUMMARY

Rapid up-scaling and deployment of more cost-efficient and sustainable carbon capture solutions are needed to reduce the emissions of CO_2-intensive industries. Solvent-based carbon capture is an important technology that can be readily adopted to many emission sources. Such technology can achieve high capture rates and deliver CO_2 at high purity with relatively low energy demand. In AURORA, the open and non-proprietary CESAR1 solvent technology will be optimised and qualified for commercial deployment. The technology will be demonstrated at TRL7-8 for three CO_2 intensive industries: refining, cement, and materials recycling, for which there are few other options to achieve climate neutrality. The partners will demonstrate negligible environmental impact (emissions being a potential issue for solvent technology), capture rates at 98%, and capture costs reduced by at least 47% compared to a benchmark process with the MEA solvent.

This will be achieved due to the following innovations: 1) Holistic optimisation of solvent composition, process design, emission monitoring and control, and solvent management, 2) Validated models for use in commercial process simulators 3) enhanced waste heat integration with carbon capture for reduced external heat demand and operational costs 4) Improved and integrated advanced control system for reduced OPEX and optimised performances.

These innovations will be integrated in four optimised capture processes, and various aspects will be demonstrated in pilots of various size and complexity. The partners will ensure transferability of results to other CO_2 intensive industries thanks to the large variations in CO_2 source and developed clusters addressed in the project and a strong stakeholder participation. The project will also do full CCUS chain assessments for its end-users. It is noteworthy that the end-users are situated in two different regions of Europe, offering different conditions for the implementation of CCUS value chains.

More information on the project can be found at https://aurora-heu.eu/

DOCUMENT OBJECTIVE AND EXECUTIVE SUMMARY

The purpose of this deliverable is to make an initial plan for data management in the AURORA project. This deliverable is the first version of the Data management Plan (DMP) and will be updated every 6 month during the project.

It considers the data that will be generated in the project and suggests how to make the data findable, accessible, interoperable, and reusable in accordance with the concept of FAIR data management. The document follows the European Commission's Guidelines on Data Management in Horizon 2020 Online Manual (https://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-data-management/data-management_en.htm) and provides information about the collected data, such as its purpose, utility, accessibility and re-usability.

LIST OF PARTNERS

N°	Participant organisation name	Acronym	Country
1	SINTEF AS	SINTEF	NO
2	Norwegian University of Science and Technology	NTNU	NO
AP	University of Cambridge	UCAM	UK
3	Technology Centre Mongstad	TCM	NO
4	Sapienza Università di Roma	UNIROMA1	IT
5	Total Energies OneTech	TOTAL	FR
6	Aker Carbon Capture AS	ACC	NO
7	MOTOR OIL	MOH	GR
8	Heracles Group	HERACLES	GR
9	UMICORE	UMICORE	BE
10	Cybernetica	CYB	NO
11	Euroquality SAS	EQY	FR

TABLE OF CONTENTS

Acknowledgments 2
Project Summary 2
Document Objective And Executive Summary 2
List Of Partners 3
Table Of Contents 4
List Of Figures 5
List Of Tables 5
1. Introduction 6
2. Data Summary 6
2.1. The purpose of the data collection/generation 6
2.2. Relation to the objectives of the project 7
2.3. Specification of types and formats of the generated and collected data 7
2.4. Reuse of existing data 8
2.5. Origin of the data 8
2.6. SharePoint and metadata provision 8
2.7. Expected size of the data 10
2.8. Data utility: to whom will it be useful 10
3. Fair Data MANAGEMENT 10
3.1. Making data findable, including provisions for metadata 10
3.1.1. Depository of the data and associated metadata, documentation and Discoverability of data (metadata provision) 10
3.1.2. Identifiability of data and standard identification mechanism 11
3.1.3. Assignment of search keywords 12
3.2. Making data openly accessible 12
3.2.1. Data to be made openly avAIlable 12
3.2.2. Methods or software tools needed to access the data 12
3.2.3. ACCESS in case of any Restrictions 13
3.3. Making data interoperable 13
3.3.1. Interoperability of your data 13
3.3.2. Data formats 13
3.3.3. Use of standard vocabulary for all data types present in the data setS to allow inter-disciplinary interoperability 13

3.4. Increase data reuse (through clarifying licenses) 13
3.4.1. How the data will be licenced to permit the widest reuse possible 13
3.4.2. Data availability for reuse 13
3.4.3. Use of the data produced and/or used in the project by third parties 14
3.4.4. Data quality assurance processes 14
3.4.5. The length of time for which the data will remain reusable 14
4. Allocation of Resources 14
4.1. Cost estimates for making project data FAIR 14
4.1.1. Responsibilities for data management in AURORA project 14
4.1.2. Costs and potential value of long-term preservation 14
5. Data Security 15
5.1. Active Project - Data security as specified for SINTEF SharePoint 15
5.2. Repository 15
6. Ethical aspects 15
6.1. Contact information 16
6.2. Pictures and videos for communication purposes 16
7. Other issues 17

LIST OF FIGURES

Figure 2.1 Illustration of the process for management of results in AURORA 9

LIST OF TABLES

Table 2.1 Instructions for uploading datasets to SharePoint. 9

1. INTRODUCTION

The research data management in AURORA will be performed in accordance with the guidelines on FAIR data management using the Horizon 2020 FAIR Data Management Plan (DMP) Template. The dissemination of project data is subject to a dissemination and approval procedure as specified in the Consortium Agreement.

The data management plan (DMP) will describe the life cycle for data to be collected, processed and generated by the project; it will ensure that research data are retrievable, accessible, interoperable and reusable. The DMP will be established in collaboration with all partners so that it is consistent with the consortium's exploitation and IP requirements.

The WP7 leader (SINTEF), through the WP7 activities, will be responsible for data management and development and updating the DMP over the course of the project whenever significant changes arise, such as when:

- there is new data available
- there are changes in the consortium policies (e.g. new innovation potential, decision to file for a patent)
- there are changes in consortium composition and external factors (e.g. new consortium members joining or old members leaving).

For commercial reasons, some of the project data may be kept confidential within the consortium and not available for verification and reuse (in particular for solvent management technology where patenting is ongoing). Project data used for scientific articles will be made publicly available through the project depository (Zenodo or other), general project website or by other means.

2. DATA SUMMARY

2.1. The purpose of the data collection/generation

The AURORA project will demonstrate and qualify an advanced and innovative solvent-based CO2 capture technology for large scale cost-efficient integration in at three CO_2-intensive industries: refining, cement, and materials recycling as sectors. As these sectors represent a large range in flue gas composition and conditions, this will be exploited to assess the effect on solvent performance. The results will be utilised for the validation of process models to be established for assessment of full-scale integrated CO_2 capture plants for all three sectors. Transport and storage solutions, as well as CO_2 utilisation options, will be developed for end-user (in AURORA) business cases, and the societal impact of the proposed solutions will be assessed.

For the lab and pilot scale testing and for the assessment of the full Carbon Capture, Utilization, and Storage (CCUS) chains, information/data are needed from the project end-user, suppliers of the technologies and materials, as well as from the developers of the technologies. For the assessment of societal impact, information/data on public and stakeholder opinion needs to be collected.

2.2. Relation to the objectives of the project

The overall objective of AURORA is to develop advanced CO_2 capture technologies to the point of commercial deployment. By optimising and qualifying the CESAR1 solvent and associated CO_2 capture process, AURORA will offer a highly competitive solution compared to the conventional MEA-based benchmark. Full CCUS chain assessments will be carried out by taking a holistic approach that considers large-scale application in different industrial clusters, which will in turn complete conceptual studies applied to the four CO_2-intensive end-users of AURORA.

The data collection/generation is necessary for the execution of the project tasks and for the fulfilment of the sub-objectives (SO) set in the project:

SO1: Qualify and validate the most optimal composition for the CESAR1 solvent

SO2: Close knowledge gaps related to degradation, emission, and emission control

SO3: Demonstrate in three pilots to validate flexible, controllable, and cost-efficient operation on relevant industrial sources

SO4: Provide validated thermodynamic, absorption kinetic and property models for commercial tools

SO5: Establish CESAR1 as the new benchmark at the same level as MEA

SO6: Ensure the sustainability of the demonstrated capture technology and of the integration in CCUS chains

SO7: Establish a plan for further actions towards final integration into several industrial processes and clusters

2.3. Specification of types and formats of the generated and collected data

A large variety of data will be generated and collected in the project. The AURORA project covers the full CCUS value chain. Different types and formats of data will be generated/collected along the whole value chain, such as:

- Documents/Reports/Publications: .PDF/A, txt, doc/docx
- Spreadsheets: .xls/.xslx
- Databases: .cvs
- Audio files: .mp3, .wav, .wma, .ra
- Pictures: jpg, png
- Video: avi, flv, mov, mp4, wmv

Data from experimental and modelling work will be generated in the technical work packages (WP1-WP3) in the lab and pilot scale experiments and from the field tests.

Qualitative and quantitative data will also be generated through interviews, questionnaires, and workshops for social science-related activities (WP5). This data will be anonymised and processed in the project and is dealt with under the EPQ report.

2.4. Reuse of existing data

Technology developments in AURORA are based on the innovations developed in earlier national, ACT and EU projects, and data from those projects will be reused in AURORA where necessary.

2.5. Origin of the data

For the demonstration of the capture technology at pilot scale (three different scales) and for the techno-economic assessment of the full-scale capture plant integration for the project end-users, minimum necessary information will be collected, including:

- Flue gases' composition and volume
- Available utilities
- Layout of the facilities

This information will be kept confidential within the project consortium and used as an input to the process models and techno-economic assessment of CO_2 capture integrated with the host facilities/industrial cluster, for establishing requirements for the flue gas pre-treatment and waste handling.

For the assessment of the full CCUS (Carbon Capture, Utilization and Storage) chain, open-access information will be collected on possible CO_2 storage sites (both intermediate storage and permanent) and CO_2 use and transport opportunities in relevant regions/countries.

Data to be generated in the project include:

- Experimental data on solvent properties, e.g. phase equilibrium, mass transfer, thermo-physical properties, etc.
- Experimental data on process performance, including performance of the emission and solvent management technologies, both testing in the laboratories and the operational environment in pilot scale.
- Data/results for testing of an advanced control system.
- Data/results from process simulations and techno-economic assessment of full-scale CO2 capture cost.
- Data/results from the assessment of CO_2 utilisation options.
- Data/results from the assessment of CO_2 transport and storage options.
- Anonymised data collected through interviews/questionnaires addressing societal aspects

Most of the data is planned to be published in open-access journals and presented to a broad audience in workshops and conferences. Some of the data might be kept confidential due to possible patenting.

2.6.SharePoint and metadata provision

Except for some data as indicated in Section 6, all datasets in the AURORA project will be stored on a SINTEF SharePoint project site. This will be the project's online collaboration area during the 42 months the project is active. All partners will be responsible for uploading the datasets they have collected/generated during the project. Each dataset will be uploaded to a dedicated research data folder in the SharePoint site. All datasets will use standard SharePoint version

control, and access control is available to enable limited access to certain types of data. These metadata will be provided for each data set:

- File name
- Date
- Version
- File type
- Description
- WP number
- Responsible person
- Dissemination level

Management, protection, and dissemination of results in AURORA is illustrated in Figure 2.1.

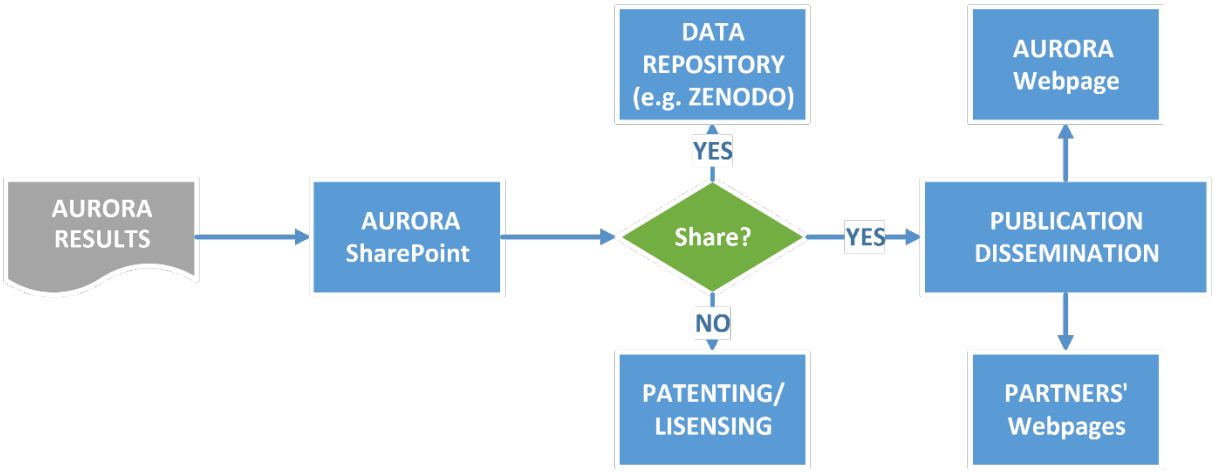

Figure 2.1 Illustration of the process for management of results in AURORA.

Table 2.1 provides instructions to project participants on how to upload datasets to SharePoint.

Table 2.1 Instructions for uploading datasets to SharePoint.

Upload instructions – AURORA SharePoint Site
WP leads are responsible for data repository on SharePoint. - Please upload all data sets to the appropriate WP folder in the AURORA SharePoint site: - Use this naming convention (for details, see 2.1.2): o DeliverableNumber Descriptive text_versionNumber

- Publication Description text_versionNumber

- Be sure to use the same file name when uploading later versions

2.7. Expected size of the data

This is presently not clear.

2.8. Data utility: to whom will it be useful

The data generated/collected in the AURORA project will be useful for different groups of the CCUS stakeholders:

- CO_2 emitting industries and power generation plants – end users – will use data for techno-economic assessment of an integrated capture plant at their facilities.
- CCUS technology providers, equipment and chemicals vendors – suppliers (lead users) – will use data for the design of the equipment/process, requirements of the equipment, chemicals.
- Universities and R&D organisations active in the area of CCUS technologies development – education and research providers – will use the data in education and further research work.
- Municipalities, NGO, etc. – decision makers – will use data to support decisions to be made regarding deployment of CCUS.

3. FAIR DATA MANAGEMENT

AURORA will manage data in accordance with the principles of FAIR data management[1] (Findable, Accessible, Interoperable and Reusable data). The project aims to maximise access to, and reuse of research data generated by the project. At the same time, there are datasets, or parts of datasets, generated in this project that cannot be shared in order to protect the privacy of voluntary participants, to protect business-sensitive information and/or to protect confidential/classified information.

3.1. Making data findable, including provisions for metadata

3.1.1. DEPOSITORY OF THE DATA AND ASSOCIATED METADATA, DOCUMENTATION AND DISCOVERABILITY OF DATA (METADATA PROVISION)

There are currently hundreds of repositories available for hosting scientific data. Some of the partners in AURORA use their own repositories for archiving the data (e.g. SINTEF open[2], NTNU open[3], Apollo (the University of Cambridge repository as indicated in Section 6) etc.), some countries provide repositories for

[1] http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf

[2] https://sintef.brage.unit.no/sintef-xmlui/

[3] https://ntnuopen.ntnu.no/ntnu-xmlui/

research data (e.g. NSD[4] – Norwegian Centre for Research Data, DRI[5] – Digital Repository of Ireland, UK Data Service[6], DANS[7] - Data Archiving and Network Services in the Netherlands, etc.), there are also open globally-scoped repositories, e.g. UDdat[8], Zenodo[9], Mendeley Data[10], etc.

It will be discussed in the project and decided if a specific repository will be used in the project and what data and when will be archived in a separate workshop during a project meeting in M12. The DMP will then be updated.

The Zenodo repository will be suggested as the main tool to make our research data findable in accordance with the H2020 Open Access Mandate. All the public datasets and deliverables, as well as scientific publications, can be uploaded to this community. In addition, we will link all our uploads to the European Commission Funded Research (OpenAIRE) community for maximum findability.

Metadata associated with each published data set in Zenodo contains the following:

- Digital Object Identifiers
- Version numbers
- Bibliographic information
- Keywords
- Abstract/description
- Associated project and community
- Associated publications and reports
- Grant information
- Access and licensing info
- Language

3.1.2. IDENTIFIABILITY OF DATA AND STANDARD IDENTIFICATION MECHANISM

Where appropriate, Digital Object Identifiers (DOIs) will be applied to archived datasets, which enables citation of the information both in project reports and journal publications. This is a prerequisite of leading science journals. Assigning a DOI to datasets:

- allows data to be cited in the same manner as a scientific journal article
- allows credit to be assigned to the dataset creators
- recognises the value of the data and the effort that has gone into its creation
- ensures the discoverability, permanence, and stability of the dataset

[4] https://nsd.no/nsd/english/index.html

[5] https://www.dri.ie/

[6] https://reshare.ukdataservice.ac.uk/

[7] https://dans.knaw.nl/en/

[8] https://www.eudat.eu/

[9] https://zenodo.org/

[10] https://data.mendeley.com/

It is not appropriate or necessary to have a DOI for all data, and data can be archived without a DOI. However, a DOI gives assurance to future users that the dataset is:

- stable
- complete
- permanent
- of good technical quality

3.1.3. ASSIGNMENT OF SEARCH KEYWORDS

Relevant keywords will be assigned to the public datasets. Dataset-specific keywords must be descriptive of the content of the dataset. The project has defined a set of general keywords that can be used when descriptive of the public datasets, scientific publications and public deliverables. These are as follows:

- Keyword 1: Carbon capture and storage (CCS)
- Keyword 2: Absorption
- Keyword 3: Solvent
- Keyword 4: Non-Linear Model Predictive Control (NMPC)
- Keyword 5: CCUS value chains

3.2. Making data openly accessible

3.2.1. DATA TO BE MADE OPENLY AVAILABLE

Most of the (processed) data from the project will be openly available. The experimental data for the solvent will be made available, which in turn is needed to establish process models able to describe the CESAR1 system in all conditions relevant for full-scale processes.

Data detailing end-user-specific information will be available only for project partners.

Some of the data on, e.g., solvent management technologies, emission measurements (ACEMS), control system (CENIT) and heat integration options, will be confidential to protect IP, although data from demonstration of these technologies may be open.

Raw data from this project will not be open as the data have to be processed taking into account used methodologies and experimental facilities (mainly in-house developed) to avoid misinterpretation. The processed data will be made available where it underpins publications and validates research findings.

3.2.2. METHODS OR SOFTWARE TOOLS NEEDED TO ACCESS THE DATA

Microsoft Office documents, pdf, and text files will be available. These types of documents can be opened using different software, both requiring a license and free (e.g. Adobe Reader, Google Docs, etc.).

3.2.3. ACCESS IN CASE OF ANY RESTRICTIONS

As for the repository, this will be discussed and decided in a workshop in M12. The DMP will then be updated.

3.3. Making data interoperable

3.3.1. INTEROPERABILITY OF YOUR DATA

Data should be interoperable, and standard open formats will be used to allow data reuse. The data will be reusable without the need for communication with the data creator.

This will be discussed and decided in a workshop in M12. The DMP will then be updated.

3.3.2. DATA FORMATS

Data formats will conform to widely accepted standards and be readable by tools that are freely available now and are likely to remain so.

Audio/video files will be stored in MP3 or WAV format, digital images will be stored as JPEG or PNG. Microsoft Word or LaTeX files will be used for text-based documents. Microsoft Excel (.xxls or .csv) or text files (.txt) will be used for data tables. For sharing, the documents will be converted to PDF format. These file formats have been chosen because they are accepted standards and are in widespread use.

Raw data during the tests are recorded using data logging software like LabView or software provided by different equipment providers. The data are recorded in broadly used standards, e.g., TXT or CSV files. As mentioned previously in the document, these data will not be open access.

3.3.3. USE OF STANDARD VOCABULARY FOR ALL DATA TYPES PRESENT IN THE DATA SETS TO ALLOW INTER-DISCIPLINARY INTEROPERABILITY

This will be discussed and decided in a workshop on M12. The DMP will then be updated.

3.4. Increase data reuse (through clarifying licenses)

3.4.1. HOW THE DATA WILL BE LICENCED TO PERMIT THE WIDEST REUSE POSSIBLE

This will be discussed and decided in a workshop on M12. The DMP will be updated.

3.4.2. DATA AVAILABILITY FOR REUSE

Data embargoes may be required whilst agreement is reached on ownership and Intellectual Property protection or until journal publications are released.

Otherwise, the data will be available for reuse as soon as they are published.

3.4.3. USE OF THE DATA PRODUCED AND/OR USED IN THE PROJECT BY THIRD PARTIES

There may be some data that will be deemed to be Intellectual Property of the creators or be commercially sensitive. In these cases, the release of limited or processed data will be explored.

There are no International partners/third parties in the project.

3.4.4. DATA QUALITY ASSURANCE PROCESSES

All partners are responsible for the quality of the data generated by the partner, following quality assurance procedures at the partner organisations.

Typically, the data sets are provided with an overview of uncertainties in the data.

In addition, all deliverables are subjected to the quality assurance specified in the Project Management Plan as described in D7.1.

3.4.5. THE LENGTH OF TIME FOR WHICH THE DATA WILL REMAIN REUSABLE

The data are reusable as long as they are searchable and available for download.

4. ALLOCATION OF RESOURCES

4.1. Cost estimates for making project data FAIR

No cost is allocated for data repositories, as many of them are free of charge.

Cost for gold open-access publications in high-impact scientific journals is included in the partners' budget where such publications are planned. Manuscripts may also be published in other open-access journals.

4.1.1. RESPONSIBILITIES FOR DATA MANAGEMENT IN AURORA PROJECT

Currently, the Project Coordinator, supported by the Executive Board (WP-leaders), is responsible for data management in AURORA. A Data Management officer may be appointed later in the project.

4.1.2. COSTS AND POTENTIAL VALUE OF LONG-TERM PRESERVATION

No cost is allocated for data repositories, as many of them are free of charge.

It is believed that the value of the data from the AURORA project will remain high for a long term after the end of the project, taking into account that full-scale CCUS deployment is currently taking off in Europe and other parts of the world. The CCUS technologies development will be continued, and cost reduction will be addressed in the coming decades.

5. DATA SECURITY

Publication/Dissemination of the project results is covered by the Consortium Agreement. Only data that have been approved for publication by all partners of the consortium will be made openly available.

5.1. Active Project - Data security as specified for SINTEF SharePoint

SINTEF SharePoint is the online collaboration platform used for the AURORA project. A dedicated project site has been established on this platform, accessible only by the partner representatives in the consortium. Furthermore, a dedicated folder for research datasets is set up, allowing for stricter access control than on the main project site.

The AURORA SharePoint site has the following security settings:

- Access level: Restricted to persons (project members only). Further access restrictions on specific folders are enabled.
- Encryption with SSL/TLS protects data transfer between partners and the SINTEF SharePoint site.
- Threat management, security monitoring, and file-/data integrity prevent and/or register possible manipulation of data.

Documents and elements in the SINTEF SharePoint sites are stored in Microsoft's cloud solutions, based in Ireland and the Netherlands. There will be no use of data centres in the US or outside EU/EEA and associated countries.

Nightly back-ups are handled by SINTEF's IT operations contractor. As a baseline, all project data will be stored for ten (10) years according to SINTEF's ICT policy unless otherwise agreed in contracts and data processing agreements.

5.2. Repository

If and when a specific repository is chosen for the project, The DMP will be updated with the data security settings for that specific repository.

6. ETHICAL ASPECTS

The proposed work in AURORA will fully comply with the regulations set out in Regulation (EU) 2016/679, the General Data Protection Regulation (GDPR). In addition, AURORA complies with the principles of the European Charter for Researchers, the European Code of Conduct for Research Integrity, including ethical standards and guidelines, regardless country in which research is carried out.

Nothing in this project shall be deemed to require a party to breach any mandatory statutory law under which the party is operating. This includes any national or European regulations, rules and norms regarding ethics in conducting research.

Currently, no ethical or legal issues that can have an impact on data generation and sharing have been identified. Ethical and legal aspects related to research data generated by the project will be considered as the work proceeds.

With regard to the public and stakeholder data collected, Cambridge will maintain the highest standards with regard to data management. All Cambridge datasets will follow the University of Cambridge policy on data protection: https://www.data.cam.ac.uk/university-policy

Data will be archived on the network drive of Judge Business School, University of Cambridge and will be accompanied by **metadata** based on Dublin Core, (https://www.dublincore.org/specifications/dublin-core/dces/) which is a 'basic, domain-agnostic standard which can be easily understood and implemented, and as such is one of the best known and most widely used metadata standards'. All Cambridge-led publications and data sets will be submitted to **Apollo** (the University of Cambridge repository), which is publicly accessible and searchable (https://www.repository.cam.ac.uk/).

Data from quantitative studies (i.e. representative public surveys) will be collected by UCAM as numeric data. There will also be room in the questionnaires for people to write in their own words, which will be saved as text in the same datasets as the numeric data. Personal data such as income, education, gender, and age will be collected via the questionnaires. The survey data collected by the market research firm that collects the data will do so in compliance with the ICC/ESOMAR International Code on Market, Opinion and Social Research and Data Analytics: https://esomar.org/uploads/attachments/ckqtawvjq00uukdtrhst5sk9u-iccesomar-international-code-english.pdf as well as any appropriate national codes of conduct for market research (e.g., ADM in Germany or MRS in the UK). The data policy of the University of Cambridge will be followed to secure such data. For every empirical study, we will apply for an approval from the Ethics committee of Judge Business School at the University of Cambridge. The committee assesses case by case whether additional security measures are needed.

The data will be accompanied by the respective questionnaires, so that it is evident what questions are represented by the variables in the dataset and units of measurement. In addition, syntax files will be prepared that include the functions of data screening and, in the case of publication, data analysis.

6.1. Contact information

Use of e-mail addresses in AURORA SharePoint Site:

An e-mail address is, by definition, personal information and is covered by GDPR. The e-mail addresses of project participants are stored on the AURORA SharePoint Site. Only the project participants invited will have access. The e-mail address is a prerequisite to accessing the project's working area. By accepting the SharePoint invitation, the participants consent to use and store their e-mail addresses for the purpose of online collaboration in the project.

The e-mail addresses will be deleted when access to the project area is no longer needed.

SINTEF has signed GDPR data processing agreements with both Microsoft and the IT operations contractor handling the SharePoint platform.

6.2. Pictures and videos for communication purposes

AURORA will collect pictures and videos for use in communication activities (website, newsletter, social media). Such data will only be collected with prior consent from the people involved and only used for as long as consent is given. Pictures and videos can contain personal data if an individual is the focus of the image or video. Examples include: 1) pictures/videos of individuals stored together with personal details (e.g., identity cards); 2) pictures/videos of individuals posted on the project website along with biographical

details; 3) individual images published in a newsletter. Examples of pictures and videos that are unlikely to contain personal data are: 1) pictures/video where people are incidentally included in an image or are not the focus (e.g. at a big conference/workshop); 2) images of people who are no longer alive (the GDPR only applies to living people, see section 2.2).

When collecting pictures and videos, AURORA will follow established guidance and best practice on collecting and processing such data to ensure that we adhere to the legal requirements (e.g., guidance established by the University of Reading, UK). Under no circumstances will pictures containing personal information be publicly shared without the subject's explicit consent.

The AURORA project partners are obliged by European and national law (GDPR) to protect personal data.

The coordinator of the AURORA project, SINTEF, follows ethical guidelines in its work, and all work conducted by SINTEF is subject to the SINTEF Ethics Council and the appointed Ethics Representative. SINTEF will also ensure that all participants in the AURORA project follow the ethical guidelines of SINTEF. Important aspects with respect to this are:

- The ethical guidelines are based on the vision of using science and technology to create a better society and are reviewed continuously to ensure they stay up to date with developments in society and the challenges of today. They generally fall into these categories: research ethics, business ethics, and ethics in interpersonal relationships.
- SINTEF is a member of the UN Global Compact and Transparency International, and SINTEF's ethics are guided by the principles highlighted by these organisations, as well as based on the regulations of the national ethics committees, the principles promoted by the European Group on Ethics in Science and New Technologies (EGE), and on international conventions such as the Vancouver Convention. When dilemmas of research ethics require an assessment beyond the scope of our guidelines, our Ethics Council and Ethics Representative, we refer to statements from the EGE.
- All SINTEF's employees are expected to act in accordance with the ethical guidelines and principles. As coordinator of the AURORA project, SINTEF will ensure that any ethical issues, which may arise, will be handled appropriately and in a transparent and fair manner.

7. OTHER ISSUES

This will be discussed and decided in a workshop in M12. The DMP will then be updated.