Industrial Symbiosis^2 Hubs 4 Circularity

Deliverable title:	D1.2 Data Management Plan
Document type:	R — Document, report
Dissemination level:	PU - Public
Version:	1.0. Full
Date:	30 June 2024

Funded by the European Union

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 HADEA. Neither the European Union nor the granting authority can be held responsible for them.

Document Information

Table 1: Document Information

Project Number:	101138473
Project Acronym:	IS2H4C
Project Title:	Sustainable Circular Economy Transition: From Industrial Symbiosis to Hubs for Circularity (IS2H4C)
Deliverable Title:	Data Management Plan (D1.2)
Due Date of Deliverable:	30 June 2024
Work Package:	1
Author(s):	Maren Bödding (UT)
Reviewer(s):	Devrim Yazan (UT)
Document Type:	R — Document, report
Dissemination Level:	PU - Public
Total Number of Pages:	22

List of Beneficiaries

1. UNIVERSITEIT TWENTE (UT),
2. MONTANUNIVERSITAET LEOBEN (MUL),
3. TECHNISCHE UNIVERSITAT DORTMUND (TUDO),
4. FUNDACION TECNALIA RESEARCH & INNOVATION (TEC),
5. INSTITUTO DE CIENCIA E INOVACAO EM ENGENHARIA MECANICA E ENGENHARIA INDUSTRIAL (INEGI),
6. FRAUNHOFER GESELLSCHAFT ZUR FORDERUNG DER ANGEWANDTEN FORSCHUNG EV (FHG),
7. FUNDACION CIRCE CENTRO DE INVESTIGACION DE RECURSOS Y CONSUMOS ENERGETICOS (CIRCE),
8. EREVNITIKO PANEPISTIMIAKO INSTITOUTO SYSTIMATON EPIKOINONION KAI YPOLOGISTON (ICCS),
9. INSTITUT FUR NACHHALTIGE TECHNOLOGIEN (AEE),
10. FUNDACION ZARAGOZA LOGISTICS CENTER (ZLC),
11. ENERGY EFFICIENCY IN INDUSTRIAL PROCESSES ASBL (EEIP),
12. ISTANBUL MADEN VE METALLER IHRACATCI BIRLIKLERI (IMMIB),
13. COOPERATIE WAARDEMAKERS IN WATERSTOF U.A. (H2HT),
14. CLUSTER DE ENERGIA (ACE),
15. CONSORCIO DE AGUAS BILBAO BIZKAIA (CABB),
16. COOPERATIE DUURZAAM AADORP U.A. (ALMELO),
17. KPMG & ASSOCIADOS SOCIEDADE DE REVISORES OFICIAS DE CONTAS SA (KPMG),
18. TURKIYE PETROL RAFINERILERI ANONIM SIRKETI (TUPRAS),
19. ARCELIK A.S. (ARCELIK),
20. SMURFIT KAPPA NERVION SA (SKN),
21. NORTEGAS ENERGIA GRUPO SOCIEDAD LIMITADA (NORTEGAS)
22. SIDENOR ACEROS ESPECIALES SL (SIDENOR),
23. PETRONOR INNOVACION SOCIEDAD LIMITADA (PETRONOR),
24. CALERA DE ALZO SL (CALERA),
25. LOINTEK INGENIERIA Y TECNICAS DE MONTAJES SL (LTK),
26. SPOUTED BED SOLUTIONS SOCIEDAD LIMITADA (SBS),
27. INFRASERV GMBH & CO. HOCHST KG (IH),
28. CROWDHELIX LIMITED (CHX),
29. IRIS TECHNOLOGY SOLUTIONS, SOCIEDAD LIMITADA (IRIS),
30. WORKDECK DIGITAL WORKPLACE SL (WDK),
31. SOLAR ENERGY CONVERSION POWER CORPORATION (SOLENCO),

Funded by the European Union

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 HADEA. Neither the European Union nor the granting authority can be held responsible for them.

Abstract

The deliverable (D) 1.2 Data Management Plan (DMP), within work package (WP) 1 is meant to describe the type of data and datasets produced within the innovation action IS2H4C.

The DMP will determine how the data is collected and used and how it is planned to be processed and stored within the project duration and beyond. Different strategies concerning the collection, sound processing and protection of sensitive, confidential and personal data according to the General Data Protection Regulation (GDPR) will be outlined within the following paragraphs. Furthermore, the technical and organizational measures as well as the security measures planned to be implemented by the IS2H4C project partners, to guarantee the rights and freedoms of the data subjects/ participants and prevent unauthorized access, are given.

While the basic procedures and measures concerning the correct data treatment are outlined in the current report, the DMP remains a living document that will be updated continuously during the IS2H4C.

Funded by the European Union

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 HADEA. Neither the European Union nor the granting authority can be held responsible for them.

Table of Content

Document Information 2
List of Beneficiaries 2
Abstract 3
Table of Content 4
List of Tables 6
List of Abbreviations 7
1. Introduction 8
1.1. Major project objective 8
1.2 Purpose of this document 8
1.3 Related documents 8
2. Data Summary 9
3. Principles of FAIR Data Management 11
3.1 FAIR Data Management 11
3.2 Data Processing according to GDPR 11
4. Data Management within IS2H4C 12
4.1 Making data findable, including provisions for metadata 12
4.2 Making data accessible 12
4.3 Making data interoperable 13
4.4 Increase of re-usability 13
5. Data processing within IS2H4C according to GDPR 14
5.1 Lawfulness, fairness and transparency 14
5.2 Purpose limitation 14
5.3 Data minimization 14
5.4 Accuracy 14
5.5 Storage limitation 14
5.6 Integrity and confidentiality 14
5.7 Accountability 15
6. Sound Data Management within the lifecycle – General Guidelines 16
6.1 Data collection 16
6.1.1 (Sensitive) Personal data 16
6.1.2 Purpose limitation & Data Minimization 16
6.I.3 Informed consent 16

Funded by the European Union

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 HADEA. Neither the European Union nor the granting authority can be held responsible for them.

6.2 Data processing **16**
6.2.1 Accuracy 16
6.2.2 Findability & Inoperability 16
6.2.3 Data Encryption 17
6.2.4 Anonymisation & Pseudonymisation 17
6.2.5 Team access and data sharing within the consortium 17
6.3 Data storage **17**
6.4 Data access (beyond the IS2H4C consortium) **17**
6.5 Data re-use (beyond the IS2H4C consortium) **17**
7. Data protection and security ***18***
7.1 Data protection and security in general **18**
7.2 Potential risks and mitigation measures **18**
8. Ethical data management ***19***
Bibliography ***20***
Appendix I : ***21***

Funded by the European Union

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 HADEA. Neither the European Union nor the granting authority can be held responsible for them.

List of Tables

Table 1: Document Information 2

Table 2: Potential data risks and mitigation measures foreseen 18

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 HADEA. Neither the European Union nor the granting authority can be held responsible for them.

List of Abbreviations

CC BY	Creative Attribution International Public License
CSA	Coordination Support Action
D	Deliverable
DCMI	Dublin Core Metadata Initiative
DMP	Data Management Plan
DoA	Description of Action
DOI	Digital Object Identifier
DPO	Data Protection Officer
EC	European Commission
ECo	Ethical Committee
FAIR	Findable, Accessible, Interoperable, Reusable
GA	Grant Agreement
GDPR	General Data Protection Regulation
H4C	Hubs for Circularity
IPR	Intellectual Property Rights
IS	Industrial Symbiosis
ISM	IT Security Manager
ISO	Information Security Officer
JRC	Joint Research Centre
NDA	Non-Disclosure Agreement
OA	Open Access
ORE	Open Research EU
OS	Open Science
PO	Project Officer
WP	Work Package

Funded by the European Union

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 HADEA. Neither the European Union nor the granting authority can be held responsible for them.

1. Introduction

1.1. Major project objective

As also outlined in the Grant Agreement (GA) IS2H4C prioritizes resource efficiency, use of renewable energy, and waste prevention, and promotes reuse and recycling of unavoidable solid, liquid, and gas waste streams via deployment of innovative sustainable technologies and via systemic integration of existing and new infrastructure in order to develop Hubs for Circularity (H4C).

In order to achieve the overall project objectives, the partners within IS2H4C will among other activities:

- Develop several solutions for the development of H4C in diverse industrial areas of process industry in the Netherlands, Germany, Spain, and Turkey.
- Deploy most innovative sustainable technologies and perform infrastructure integration in four the hubs.
- Scale up industrial areas to H4C
- Implement a digital collaboration platform to manage the resource, infrastructure, and information-sharing within H4C

1.2 Purpose of this document

The deliverable (D) 1.2 Data Management Plan (DMP), within work package (WP) 1 is meant to describe the type of data and datasets produced within the innovation action IS2H4C. In addition, the DMP will determine how the data is collected and used as well as how the gathered data is planned to be processed and stored within the project duration and beyond. Different strategies concerning the collection, sound processing and protection of sensitive, confidential and personal data according to the General Data Protection Regulation (GDPR) will be outlined within the following paragraphs. Furthermore, the technical and organizational measures as well as the security measures planned to be implemented by the IS2H4C project partners, in order to guarantee the rights and freedoms of the data subjects/ participants and prevent unauthorized access, are given. Moreover, the potential data security risks are evaluated, risk mitigation measures are depicted, and ethical aspects are considered. It has to be pointed out that while the basic procedures and measures concerning the correct data treatment are outlined within the current report, the DMP remains to be a living document that will be updated continuously during the course of the IS2H4C project, since at this stage of the project the exact implementation of the FAIR principles and the various data protection measures are still subject to further discussion. The DMP will be updated before the next periodic evaluation expected for month 14 or as soon as e.g. new data is generated or changes in consortium policies are made.

1.3 Related documents

The current report is based on the terms and conditions given in the Grant Agreement number: 101138473 and its Annexes. The specifications stated in the associate Consortium Agreement (CA) are applicable. The current deliverable is closely related to the deliverable D1.1 Management and quality plan within WP1. Sound data management and processing of all datasets collected and produced during the IS2H4C project remains to be of utmost importance. The Guidelines regarding Open Science (OS) under Horizon Europe as well as European law and national law will serve as the framework for the activities performed within IS2H4C.

Funded by the European Union

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 HADEA. Neither the European Union nor the granting authority can be held responsible for them.

2. Data Summary

Within the different WPs of the project multiple data sets will be produced, collected and processed in order to obtain the overall project goal of paving the way towards the development of Hubs for Circularity (H4C) based on the circularity and resilience requirements of existing and future industrial zones and their surrounding ecosystems. It has to be underlined that the DMP will be updated on regular basis, as soon as the consortium gains further insights regarding the collection of datasets, the data management as well as the allocation of resources during the further course of the project. The partners within IS2H4C are expected to collect, produce and process various datasets in order to achieve the following objectives within the various WPs (according to the GA):

WP1 Management: The purpose of data collection WP1 is to enable to smooth progression of all the processes, so that excellent output is generated and that the overall objectives of the IS2H4C project can be achieved.

WP2 Technology development and technical aspects: Within WP2 data sets will be collected and processed in order to develop, analyse, and optimise key technologies to maximise industrial synergies and contribute to the circular economy by focusing on energy conversion, waste recovery and utilisation, and CCUS options. Simulation models of technologies will enable the forecasting of system behaviour and optimisation of the circularity pathways on a technical level.

WP3 Societal engagement and governance: Within the scope of this WP the collection and processing of different data sets containing personal and sensitive data is expected to enable the identification relevant stakeholders and their connection to H4Cs. These include industrial companies within the hub, citizens living in the surrounding of the hub, policy makers & the general public. WP3 considers these stakeholders and identifies non-technical (regulatory, ethical, societal) challenges and barriers that arise in connection with the establishment of hubs.

WP4 Modelling and optimisation for H4C development: Within WP4 data sets containing personal and sensitive data will be collected and processed in order to develop methodology to identify, coordinate, and evaluate IS collaborations. Actors and companies are characterized based on their resource and geolocation properties, and an intelligent matchmaking system is developed to recommend IS opportunities. H4Cs are optimized by ensuring that material resources flow efficiently via logistics and supply chain analytics, and that energy resources are utilized & exchanged in the hub and with the energy networks efficiently. It also analyses how H4Cs can successfully evolve from IS networks to full circular hubs and studies their impact on decarbonisation. Methodology will include state-of-the-art techniques from AI, operations research, and agent-based simulation.

WP5 Business and financial models: Within WP5 the collection and processing of data sets will contribute significantly to designing and developing an impact valuation model, using the “True Value Methodology”, to monetise and assess the potential of impact valuation as a project screening tool in IS2H4C investment vs. other existing screening methodologies. Economic, environmental, and social impact indicators (WP3 and WP4) will be applied to the pilots being demonstrated.

WP6 Integration and demonstration in hubs: Within WP6 data sets are collect and processed in order to advancing the industrial zones into H4C via applying the technologies in the operational

Funded by the European Union

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 HADEA. Neither the European Union nor the granting authority can be held responsible for them.

environment under the principals of circularity and IS and integrating the existing and new infrastructure.

WP7 Digital Collaboration Platform for H4C: The main objective of this WP is to collect and process data in order to develop a user-friendly, scalable, and digital collaboration platform for H4C. The platform consists of a modular suite of tools, features, and data analytics to facilitate IS identification, coordination, and evaluation. Beyond a physical platform, it aims to derive a theory of IS-oriented digital platform development from the implementation evidence collected across the hubs.

WP8 Dissemination, communication & exploitation: Within WP8 data sets concerning effective communication of project findings, results, and outcomes to targeted audiences and promotion of the project to a broader audience are generated. The objective of this WP is to design and implement strategies to engage stakeholders and pursue collaborations with other pertinent initiatives. The majority of data (except from personal and confidential data) will be published via the website of the project.

Due to the fact that the project is aimed prioritizes resource efficiency, use of renewable energy, and waste prevention, and promotes reuse and recycling of unavoidable solid, liquid, and gas waste streams via deployment of innovative sustainable technologies and via systemic integration, most of the data collected and managed will originate from the various parties operating in the hubs as well as the stakeholders of the IS networks as well as the society. Vice versa these actors within the various countries will also be the parties to whom the data collected and processed will be most useful, enabling them to collaborate and to expand their businesses to further developing H4C. IS2H4C will generate data and research outputs from technological, societal, and financial disciplines. Technological data will be associated with the types of technologies developed/integrated, main metrics related to their technical development/integration, and their performance/operation capabilities. The societal and financial data will be collected based on the engagement and participation in the digital platform, adoption of the use of new technologies or practice of industrial symbiosis in any form with the hubs or in their surroundings, and business and financial institutions (e.g., collaboration rate to the H4C)

General data formats that are most likely be used during the course of the project are:

- Adobe reader (.pdf)
- Microsoft Office formats (Microsoft 365 packages) as well as Google Drive formats (spreadsheet; docs; forms etc.)
- Graphical formats (.eps, .jpg, .png etc.)
- Video formats (MP4)

Furthermore, data sets will be gathered through questionnaires and forms to be filled in WP2, WP3, WP4, WP5, WP6, and WP7. A suitable database format is discussed under discussion, whether it will be rational, Non-SQL or semantic. At the current stage it is difficult to make an estimation concerning the expected size of the data generated since many aspects, including the willingness of the market to adopt the platform, are still unclear. What can be stated is that the data generated in within the majority of WPs is not expected to exceed 4GB.

Funded by the European Union

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 HADEA. Neither the European Union nor the granting authority can be held responsible for them.

3. Principles of FAIR Data Management

3.1 FAIR Data Management

Within the context of sound data management, the term “FAIR” describes the fact that data is expected to be: *Findable, Accessible, Interoperable and Reusable*. A description of the characteristics that need to be fulfilled will be given in the following sections. (Wilkinson, et al., 2016)

Findable: “a globally unique and persistent identifier” should be assigned to any kind of (meta)data, to be findable. In addition, data has to be described “with rich metadata” and the metadata should “clearly and explicitly include the identifier of the data described.” Finally, (meta)data are to be “registered or indexed in a searchable resource.”

Accessible: Accessible (meta)data are “retrievable by their identifier using a standardized communications protocol.” In order to guarantee this, the protocol should be “open, free, universally implementable” and enable “an authentication and authorization procedure, where necessary.” Furthermore, it has to be ensured that the “metadata are accessible, even when the data are no longer available.”

Interoperable: In order to be interoperable (meta)data should “make use of a formal, accessible, shared, and broadly applicable language for knowledge representation.” Moreover, (meta)data should “use vocabularies that follow FAIR principles” and should “include qualified references to other (meta)data.”

Reusable: Reusable “(meta)data are richly described with a plurality of accurate and relevant attributes.” Furthermore, they are “released with a clear and accessible data usage license, are associated with detailed provenance and meet domain-relevant community standards.”

3.2 Data Processing according to GDPR

According to the General Data Protection Regulation 2016/679, the basic principles of personal data processing are:

Lawfulness, fairness and transparency: personal data shall be processed lawfully, fairly and in a transparent manner in relation to the data subject.

Purpose limitation: personal data shall be collected for specified, explicit and legitimate purposes and not further processed in a manner that is incompatible with those purposes; further processing for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes shall not be considered to be incompatible with the initial purposes.

Data minimization: personal data shall be adequate, relevant and limited to what is necessary in relation to the purposes for which they are processed.

Accuracy: personal data shall be accurate and, where necessary, kept up to date; every reasonable step must be taken to ensure that personal data that are inaccurate, having regard to the purposes for which they are processed, are erased or rectified without delay.

Funded by the European Union

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 HADEA. Neither the European Union nor the granting authority can be held responsible for them.

Storage limitation: personal data shall be 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; personal data may be stored for longer periods insofar as the personal data will be processed solely for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes, subject to implementation of the appropriate technical and organizational measures required in order to safeguard the rights and freedoms of the data subject.

Integrity and confidentiality: personal data shall be processed in a manner that ensures appropriate security of the personal data, including protection against unauthorized or unlawful processing and against accidental loss, destruction or damage, using appropriate technical or organizational measures.

Accountability: the controller shall be responsible for, and be able to demonstrate compliance with, the principles.

4. Data Management within IS2H4C

The partners within IS2H4C are aimed to implement all the basic principles concerning FAIR data management and GDPR stated above upon collecting and processing the various data sets needed for the different WPs. Therefore, these principles remain to be an important element of this DMP. However, it has to be pointed out that not all aspects regarding the implementation of FAIR data management and the allocation of resources are clear yet, at this very moment in time. Accordingly, the DMP will be updated on regular basis, as soon as the IS2H4C project partners gained further insights regarding the implementation of the principles. In addition, a new version of the DMP will be provided to the project officer (PO) timely before the different reviews given in the GA. Information specialist of the different institutions will be consulted throughout the duration of the project, in order to ensure FAIR Data management, as generally described in section three of this document. The experts at the University of Twente (UT) offer support and advice regarding Data management, Open Science and Data Archiving. In addition, the ICT department of the UT will provide a project server, which incorporates the latest technology according to GDPR for the storage of the data collected and produced during the course of the IS2H4C project, so that the protection of the various datasets can be safeguarded at any time.

4.1 Making data findable, including provisions for metadata

A unique Digital Object Identifier (DOI) will be allocated to each data set collected or produced. In addition, the WP leaders will define applicable structural meta data for each data set, dependent on the origin and scientific context of the data set and in accordance with a predetermined suitable metadata standard. A general naming convention and document visioning standards for the IS2H4C project have already been defined in D1.1. Keywords will be added according to data set content and common terminology within the field and registered with UT Research Information to remain findable. The data concerning project progress and documentation as well as the deliverables defined in the GA will be made available to the European Commission via the EC Funding & Tenders portal.

4.2 Making data accessible

Following the ambition to be as open as possible and as closed as necessary, the IS2H4C consortium is planning to ensure full and immediate Open Access to peer reviewed scientific publications related to their research outputs, according to the GA. The knowledge, outputs, tools, and instruments gathered from this deployment will be free and available OA for the public (considering IP restrictions) following the “as open as possible and as closed as necessary” principle at the latest upon and will be deposited in a trusted repository (ZENODO), under the Creative Attribution

Funded by the European Union

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 HADEA. Neither the European Union nor the granting authority can be held responsible for them.

International Public License (CC BY). Likewise, public deliverables will be made available as soon as approved by the EC. Parallelly, the consortium will be encouraged to publish on the Open Research EU (ORE) platform. An exemption to this will be made for the data sets gathered throughout the different WPs, that are containing personal and sensitive data as these are protected by privacy policies (GDPR) and Intellectual property rights (IPR) or Non-Disclosure Agreements (NDA). These data sets might be partly or completely restricted from being openly accessible. Furthermore, other documentation marked as confidential (e.g. certain deliverables) will only be accessible to the EC and the consortium members. For those data sets that are not in one way or another restricted from open access the consortium is currently considering different options certified by the CoreTrustSeal. Regarding the UT, it can be stated that the preferred trusted repository for archiving research data remain to be 4TU.ResearchData, Zenodo or GitHub. In addition, the data sets might also be stored in trusted repository of consortium partners throughout the EU. A suitable access level, either Open Access, Open Access for restricted users, or completely restricted will be determined by the WP leaders, together with the Data Protection Officers (DPO) team, for each individual data set.

4.3 Making data interoperable

At the moment the metadata standard of the Dublin Core Metadata Initiative (DCMI) appears to be most applicable, being 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.” However, the final decision in this regard has not been made yet, as additional metadata standards may be added in future. The interoperability of the accessible data must be pointed out that the consortium strives to align the meta data standards used with the CSA, where possible, to be able to also supply the data to the Joint Research Centre (JRC) in the future.

4.4 Increase of re-usability

The IS2H4C consortium is aimed to make as many data sets reusable for third parties as possible. As also stated in the DoA, to facilitate the validation of scientific publications, IS2H4C is aimed to enable the disclosure and reusability of the data, acknowledging the convenience of enabling the review, scrutiny, and verification from external scientific actors. Creative Common Licensing will be used to protect the ownership of datasets. Despite the data openness, an embargo period might be applied to the data if used in Green OA journals. The datasets located in ZENODO will have unlimited access, although for some data restrictions to the re-use policy may apply. This can be the case for green OA embargo periods or Creative Commons licensing restrictions (attribution, prohibition of use for commercial purposes, to use the same license as the original). In addition, it must be pointed out that some degree of the data collected is expected to be protected by GDPR, IPR or NDAs. For the open access data, bad record clearance will be executed, and the data sets receive a clear naming convention and metadata associated with their content and appropriate metadata to be easily reusable for third parties. The duration of dataset storage will be variable, highly dependent on the type of data stored.

Funded by the European Union

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 HADEA. Neither the European Union nor the granting authority can be held responsible for them.

5. Data processing within IS2H4C according to GDPR

Regarding the collection, processing and storage of personal data, we foresee that personal data will be collected and processed throughout the different WPs.

5.1 Lawfulness, fairness and transparency

Being aimed to be as transparent as possible, all individuals taking part in the IS2H4C project will be asked to give their informed consent to process their necessary personal data collected for the given project purpose. The informed consent requests will also provide the individual with a guide, containing all relevant information concerning the objective and purpose of the respective data collection, details of data processing, data storage, safe data sharing and the type of data dissemination, the project policies as well as the individual's rights to request all information that have been stored, request an update or the deletion of their personal data at any time. The consortium will provide copies of Information Sheets and Informed Consent forms (and any other relevant document) to the Executive Agency. Furthermore, it is planned to anonymize personal data directly, wherever possible. In cases, the personal data collected cannot be anonymized for project purposes additional consent for open research purposes will be requested from the individual, where necessary.

5.2 Purpose limitation

The researchers and partners within IS2H4C will not collect any personal data that is beyond the scope of the WP. Only personal data necessary to achieve the central objective of the project will be collected and processed by the different partners within the different WPs.

5.3 Data minimization

Despite the fact that in the current project stage, it is not yet clear what methodology will be used to collect the personal data required, it can be stated that only personal data relevant to reach the central objective of the project will be collected and processed within the different WPs during the project.

5.4 Accuracy

All personal data collected during the project will be checked for accuracy regularly and updated where needed.

5.5 Storage limitation

In cases where the personal data collected will no longer be necessary for the purpose for which it has originally been collected, it will be deleted directly. Until this moment the personal data collected will be anonymized, where possible. The anonymised data will be stored in an open repository after the end of the project. In case personal data cannot be anonymised it will be pseudonymised. According to the data policy of the University of Twente (UT), all data sets will be archived for 10 years.

5.6 Integrity and confidentiality

The personal (sensitive) data collected during the IS2H4C project will be stored on specific IS2H4C project server that is ISO 27001 certified. In any case, personal datasets will be stored encrypted to safeguard their full data protection also in the case of data loss or theft. Access will only be granted by the project management to authorized persons. Authorized access can be retracted directly, if necessary. The datasets containing personal information are solely transferred among the IS2H4C project partners in encrypted form. The key to the encryption will be provided to authorized persons and can be changed immediately, to withdraw access if necessary. All members of the consortium who will receive access to the datasets containing personal data will have to sign a confidentiality agreement beforehand.

Funded by the European Union

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 HADEA. Neither the European Union nor the granting authority can be held responsible for them.

Copies of the datasets are only allowed with project management authorization. Furthermore, any copies will have to be stored encrypted on a password protected device. Personal data is only allowed to be shared within the consortium, not beyond. Any copies made will have to be deleted as soon as possible.

5.7 Accountability

While the project management will be responsible to safeguard the compliance with the data management guidelines and procedures set for the project, the WP leaders as well as the researchers collecting and processing the datasets bear the responsibility to handle the datasets prudently following the guidelines set in this document.

Funded by the European Union

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 HADEA. Neither the European Union nor the granting authority can be held responsible for them.

6. Sound Data Management within the lifecycle – General Guidelines

This section is meant to provide the IS2H4C consortium members with a quick reference guide, for prudent data collection and processing, in order to perform sound data management according to the FAIR principles and GDPR guidelines throughout the life cycle of the project.

6.1 Data collection

6.1.1 (Sensitive) Personal data

Before starting to collect data sets, please determine if the specific data set aimed to be collected is expected to include any (sensitive) personal data or information that otherwise need to remain confidential; if so, please follow the GDPR. Please note: Even if the current data set is not expected to comprise any personal data (such as e.g. contact details) but can, combined with other data sets, lead to a situation where data can be traced back to a specific individual, the current data set is classified as comprising personal data.

6.1.2 Purpose limitation & Data Minimization

Please ensure that only personal data necessary and relevant to achieve the given objective of the project (such as e.g. contact details) is collected and processed within the different WPs.

6.I.3 Informed consent

All individuals who are taking part in a research under the IS2H4C project will have to provide their informed consent to process their necessary personal data collected for the given project purpose. Being aimed to be as transparent as possible the informed consent requests should –next to the consent form that will have to be signed by the individual – also provide the individual with a guide, containing all relevant information concerning:

- the objective and purpose of the respective data collection,
- details of data processing,
- data storage,
- safe data sharing and the type of data dissemination,
- the project policies,
- individual's rights to request all information that have been stored, request an update or the deletion of their personal data at any time.

Please provide copies of information sheets and informed consent forms and any other relevant document to the coordinator.

6.2 Data processing

6.2.1 Accuracy

Please ensure that all personal data collected is checked for accuracy regularly and is updated where necessary.

6.2.2 Findability & Inoperability

Please allocate a unique Digital Object Identifier (DOI) to every dataset collected or produced. Applicable structural meta data is to be defined for each dataset by the WP leaders, dependent on the origin and scientific context of the dataset and in accordance with a predetermined suitable metadata standard. The metadata should at least contain the date, time and occasion of data collection, the name

Funded by the European Union

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 HADEA. Neither the European Union nor the granting authority can be held responsible for them.

of the institute and the author. Please use the following general naming convention and visioning guideline as also outlined in D1.1:

- Number of data set
- Title (briefly describing the content of the dataset)
- Initials of the author/ researcher
- Version of document (0.1.0 for drafts; 0.1.1 for reviewed drafts; 1.0 for final version)

Keywords are to be added according to the content of the dataset also taking into account the common terminology within the field.

6.2.3 Data Encryption

Please only store personal data sets in encrypted in order to safeguard their full data protection also in the case they are lost or stolen. The transfer of data sets containing personal information can only be be performed within the consortium and in encrypted form. The key to the encryption will be provided to authorized persons and can be changed immediately, in order to withdraw access if necessary. Copies of the data sets are only allowed to be made with the authorization of the and will have to be stored encrypted on and password protected device. Any copies made will have to be deleted as soon as possible.

6.2.4 Anonymisation & Pseudonymisation

All personal data collected will have to be anonymised directly, where possible. The anonymised data will be stored in an open repository after the end of the project. In case personal data cannot be anonymised please make sure that it is pseudonymised.

6.2.5 Team access and data sharing within the consortium

Access will be granted to authorized persons by the project management. The data sets containing personal information can only be transferred within the consortium in encrypted form. The key to the encryption will be provided to authorized persons and can be changed immediately, in order to withdraw access if necessary. All members of the consortium who will receive access to the data sets containing personal data will have to sign a confidentiality agreement. Copies of the data sets are only allowed to be made with the authorization of the project management. Any copies will have to be stored encrypted on and password protected device and will have to be deleted as soon as possible. It is only allowed to share the personal data within the consortium.

6.3 Data storage

The personal (sensitive) data collected during the IS2H4C project will be stored on specific IS2H4C project server that is ISO 27001 certified. As soon as the personal data collected will no longer be necessary for the purpose for which it has originally been collected it has to be deleted. According to the data policy of the University of Twente (UT) all data sets will be archived for a period of 10 years.

6.4 Data access (beyond the IS2H4C consortium)

Following the commission's ambition of being "as open as possible and as closed as necessary" (European Commission, 2019) the IS2H4C consortium is planning to make as much data as possible openly accessible. A suitable access level and trusted repository will be determined by the WP leaders, together with the DPO team, for each individual data set.

6.5 Data re-use (beyond the IS2H4C consortium)

The IS2H4C consortium is aimed to make as many datasets as possible reusable for third parties. For the open access data (free from GDPR, IPR or NDAs) bad record clearance will have to be executed

Funded by the European Union

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 HADEA. Neither the European Union nor the granting authority can be held responsible for them.

and the data sets will have to receive a clear naming convention and metadata associated with their content as well as appropriate metadata in order to be easily reusable for third parties.

7. Data protection and security

7.1 Data protection and security in general

The UT as coordinator of the project has set up a multidisciplinary team of Data Protection Officers (DPO) guiding compliance with privacy legislation, in order to safeguard personal (sensitive) data. This team consists of an Information Security Officer (ISO), an IT Security Manager (ISM) and a Legal Advisor (LA). The contact details of these people will be made available to all data subjects involved in the research. Furthermore, the data will be stored on a UT server of which regular backups are made. As already explained in previous sections of this document several technical as well as practical measures are taken by the IS2H4C consortium to ensure sound data protection and security, both during the project and after the end of the project.

7.2 Potential risks and mitigation measures

Table 2: Potential data risks and mitigation measures foreseen

Potential risks	Mitigation measures
Unauthorised data sharing (copying)	Only authorized persons get data access Access can be withdrawn by project management immediately if necessary Any copies are deleted as soon as possible
Misuse of data	Only authorized persons get data access Access can be withdrawn by project management immediately if necessary Any copies made are deleted immediately
Loss of data	Data is stored on an ISO certified server Data is stored encrypted Regular backups of the data are made Data is stored anonymised and pseudonymised
Data Theft (leakage)	Only authorized persons get data access Data is stored encrypted, and password protected Regular backups are made Data is stored anonymised and pseudonymised

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 HADEA. Neither the European Union nor the granting authority can be held responsible for them.

8. Ethical data management

As also outlined in the GA, the IS2H4C project partners are aimed ensure that all ethics issues related to the project activities are addressed in compliance with ethical principles, the applicable international and national law, and the provisions set out in the Grant Agreement. Regarding personal data it can be stated that IS2H4C will adhere to legislative requirements of the General Data Protection Regulation. In case any activity of project requires the collection and processing of personal data during the project lifetime, the beneficiaries collecting, processing and hosting the data will fulfil the specific laws (i.e. General Data Protection Regulation, GDPR, EU2016/679) This will also be safeguarded in close cooperation with the Data Protection Officer (DPO). The beneficiaries confirm that they are aware of the following rules for collection and processing of personal data:

- The host institution must confirm that it has appointed a DPO and the contact details of the DPO are made available to all data subjects involved in the research. For host institutions not required to appoint a DPO under the GDPR a detailed data protection policy for the project must be kept on file.

- A description of the technical and organisational measures that will be implemented to safeguard the rights and freedoms of the data subjects/research participants must be submitted as a deliverable.

Detailed information on the informed consent procedures regarding data processing must be kept on file. In case any activity of project requires to be performed in a non-EU country, also these activities will be performed according to highest ethical standards, and applicable international, EU and national law, also taking all relevant ethical aspects into account. In addition, the EC's Artificial Intelligence Act will be applied. This act outlines new rules & actions aiming to turn EU into the global hub for trustworthy AI, whereby by setting the standards, EU can pave the way to ethical technology worldwide & remains competitive along the way. Following the Ethics Guidelines for Trustworthy AI trustworthy:

i) lawful - respecting all applicable laws and regulations,
ii) ethical - respecting ethical principles and values and
iii) robust - from a technical perspective while considering its social environment.

The IS2H4C consortium has appointed an independent external Ethics Advisor (EA) who has proven competence and expertise, providing advice on the ethical aspects appropriate to the complexity of the ethical aspects, in the project context. The recommendations of the Ethics Advisor will be followed up closely and all activities performed in relation to ethics will be summarized.

Funded by the European Union

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 HADEA. Neither the European Union nor the granting authority can be held responsible for them.

Bibliography

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 HADEA. Neither the European Union nor the granting authority can be held responsible for them.

Appendix I :

Data requirements from hubs:

Alignment of data collection from the hubs

General lists of data to be collected from all hubs

Work package	Objective	Boundaries	Description of data required	Type of data collection foreseen (site visit or online)
WP2	to describe and detail hub Germany (block scheme)	Infraserv Mgt. -Responsibility	list of all Infraserv-Processes and Energy Supply Systems	online
WP2	to describe and detail hub Turkey	Tüpraş <--> Arcelik	list of all branches, main processes and energy supply systems	online
WP2	to describe and detail hub Basque	all involved companies <--> urban waste	industrial sites including main processes and products and energy supply systems	online
WP2	to describe and detail hub Netherlands	H2-Hub <--> ?Biogas? <-->HW treatment <--> Almelo <--> H2 fee	companies including main processes and products and energy supply systems	online
WP2/WP5/WP6	to identify energy demand of processes (flows between blocks)	for all processes listed	annual average data on energy demand for all processes (gas, electr., heat recovery, coal, RDF, etc.)	online
WP2/WP5/WP6	to identify mass flows of processes (flows between blocks)	for all processes listed	annual average data on mass input and output for all processes (raw material, products, waste)	online
WP2/WP5	to identify and describe stakeholders of the German hub	hub representatives and other hub-related	list of relevant stakeholder groups, their interest in the project, their	online - taking stock of recent stakeholder analyses in the hubs
WP3	to identify and describe stakeholders of the Dutch hub	project partners with good access	attributes to the project, their influence, their relationships to	online - questionnaire to hub-related organisations (hub-representatives, partners with access to stakeholders)
WP3/WP5	to identify and describe stakeholders of the Spanish hub	other stakeholder groups	other stakeholder groups (conflict potential)	online - interviews with hub-related organisations
WP3	to identify and describe stakeholders of the Turkish hub			online/on site - interviews with selected stakeholders
WP3	to analyse and monitor job creation potential in all hubs	HR, regional actors interested in job creation	mapping of key occupations, job profiles and skills, qualitative and quantitative data for employment baseline	online/on site - interviews with selected stakeholders
WP3 - T3.1.2	to collect stories (upstream) of products from hubs	can also be used as input for T6.2	(upstream) story of products to be included in (to be developed) labels	online - interviews with hub-related organisations (e.g., Turkish hub: Tupras (oil refinery) and Arcelik (fridges); Basque hub: steel slags from Sidenor (steel slags) and Calcinors industry's CO2 by SBS
WP3 - T3.1.2	to identify if labels increase(d) trust and acceptance	from key stakeholders of hubs	find ways how to communicate to stakeholders about sustainability/circularity actions taken in the upstream processes	online/on site - interviews with selected stakeholders
WP3 - T3.1.3	to conduct a social network analysis (SNA)	hub representatives and stakeholders	node data (size, location,...), edge data (relationships, SC links, collaboration...), network (relation: direct, indirect), centrality,..	online - questionnaire/interviews to hub-related organisations (hub-representatives, partners with access to stakeholders)
WP3 - T3.3/WP5	to develop metrics and indicators for measuring non-technologic	hub representatives and stakeholders	metrics and indicators to measure non-technological regional impact	online - checking back on and refining metrics and indicators
WP4/WP5/WP6	to characterize and model a hub	for all hubs	List of companies in the hub (in and out consortium), name and type of industry	online
WP4	to characterize and model a hub	for all hubs	Hub-level infrastructure (road network, pipelines, energy grid connection) and their corresponding capacity	online
WP4/WP5/WP6	to characterize and model a hub	for all hubs	List of hub-level production/storage/processing/transport units (e.g., wind/solar generators, storage systems, electrolyzers, transport vehicles), with characteristics (e.g., amount, capacity, operational costs and emissions)	online
WP4/WP5/WP6	to characterize and model a hub	for all companies in the hub	Basic company data: type of industry, whether part of consortium, location coordinates. It is useful to have this data for companies outside the consortium to identify matchmaking opportunities	online
WP4/WP5	to characterize production, storage, transport units in a hub	for all companies in the hub	List of company-level production/storage/transport units (e.g., wind/solar generators, storage systems, electrolyzers, transport vehicles), with characteristics (e.g., amount, capacity, operational costs and emissions)	online
WP4/WP5/WP6	to identify resources exchanged in a hub	for all companies in the hub	List of resources that are used as input and are generated as output (energy, gas, material, water, waste)	online
WP4/WP5	to identify resource flows within the hub and external	for all companies in the hub	Suppliers and customers of input/output resources (e.g., other companies in hub, external suppliers, markets). Note that suppliers/customers names can be anonymized	online
WP4/WP5	to identify technologies used in industrial symbiosis	for all companies in the hub	List of technologies involved in industrial symbiosis (e.g., for processing or treating resources), with their characteristics (e.g., capacity, processing rates, operating costs, possible ramping rates and emissions etc.)	online
WP4/WP5	to describe key processes of sourcing, manufacturing, etc.	for all companies in the hub	List of supply chain processes involved (*sourcing* [purchasing, receiving, storing, transferring], *manufacturing* [receiving, transforming, quality inspection, packing, transferring], *distribution* [storing, picking, packing, shipping], *waste management*)	online
WP4	to describe key processes of sourcing, manufacturing, etc.	for all companies in the hub	For each supply chain process, operational data and technologies involved from the list above	online
WP4	data to be included in the digital product passport	to be decided	to be defined	online, possibly followed by onsite at later stage
WP4/WP5	to identify resource characteristics	for all resources exchanged in hub	Classification of resource (input/output, energy/gas/material/water, waste with code from European waste catalogue)	online
WP4/WP5	to identify resource characteristics	for all resources exchanged in hub	Details of resource (e.g. water quality, material characteristics) / products (BOM, quality, price, costs, % waste,. etc.)	online
WP4/WP6	to identify resource processing	for all resources exchanged in hub	Treatment / processing requirements / synergies and industrial symbioses where the resource is or can be used	online
WP4	to develop models to optimize resource/energy flows	for all resources exchanged in hub	series). Granularity depends on resource, e.g., energy hourly/daily production/consumption, daily/weekly waste	online
WP4/WP6	to develop models to optimize resource/energy flows	for all resources exchanged in hub	Annual demand/supply for resource (ideally minimum 1 year, e.g., 2023, if possible several years).	online
WP4/WP5/WP6	to account for environmental impact in the models	for all resources exchanged in hub	Costs and emissions information associate with the resource purchasing/selling/utilization/transport/treatment/disposal	online
WP6	to account for environmental metrics	for all hubs	Total CO2 emissions; Breakdown of emissions by source (e.g., industrial processes, energy production); Historical data on CO2 emissions; Total energy consumption; Energy mix (percentage from renewable sources vs. fossil fuels); Energy efficiency metrics of current processes; Air and water pollution levels; Waste generation and management practices.	online
WP6	to identify industrial plant requirements	for all hubs	Specifications for new technologies to be implemented; Compatibility requirements with existing infrastructure; Scalability potential of new technologies; Training and skills required for the workforce; Changes in operational procedures and maintenance; Safety and regulatory compliance requirements.	online
WP6	to identify infrastructure needs	for all hubs	Upgrades needed for energy grids and distribution networks; Storage facilities for hydrogen and other relevant materials; Adjustments needed in supply chain management; Transportation and delivery infrastructure for raw materials and finished products.	online
WP7	to develop the platform	for all hubs	Are there already synergies or circularities running on the platforms?	online
WP7	to develop the platform	for all hubs	How many resources are expected to be in the hubs in a few years? expected growth	online
WP7	to develop the platform	for all hubs	Will the creation of synergies/circularities be in the new platform to be developed, or can they be created in the hubs as well as in the new platform? Will they be mutually exclusive?	online
WP7	to develop the platform	for all hubs	We will need access to the current hub's websites.	online
WP7	to develop the platform	for all hubs	Will we have access to webmaster technicians regarding maintenance of these websites?	online
WP7	to develop the platform	for all hubs	Current technologies used in these websites	online
WP7	to develop the platform	for all hubs	What kind of data access will be provided?	online
WP7	to develop the platform	for all hubs	Will the new platform allow creation of new companies, or we will just need to import current companies from already existing websites? How often these websites are updated with new companies?	online
WP7	to develop the platform	for all hubs	How often these websites are updated with new companies?	online
WP7	to develop the platform	for all hubs	Data from each company: Basic data: type of industry, whether part of consortium, location coordinates (exemplies). Can this data be collected through an API?	online
WP5		for all hubs	Operational costs (taxes, shareholders dividends, interest on loans, wages)	
WP5		for all hubs	Human resourses working in each industry sector of each hub (gender, age, education level, etc)	
WP5		for all hubs	Which benefits are offered to the workers (health ensurance, etc)	
WP5		for all hubs	List and quantity of resources that are used as input and are generated as output (energy, gas, material, water, waste)	
WP5		for all hubs	List of processes of the industry units (waste treatment, etc)	

Funded by the European Union

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 HADEA. Neither the European Union nor the granting authority can be held responsible for them.

Alignment of data collection from the hubs

General lists of data to be collected from all hubs

Work package	Objective	Boundaries
WP2	to describe and detail hub Germany (block scheme)	Infraserv Mgt.-Responsibility
WP2	to describe and detail hub Turkey	Tüpras <--> Arcelik
WP2	to describe and detail hub Basque	all involved companies <--> urban waste
WP2	to describe and detail hub Netherlands	H2-Hub <--> ?Biogas? <-->WW treatment <--> Almelo <--> H2 feed
WP2/WP5/WP6	to identify energy demand of processes (flows between blocks)	for all processes listed
WP2/WP5/WP6	to identify mass flows of processes (flows between blocks)	for all processes listed
WP2/WP5	to identify and describe stakeholders of the German hub	hub representatives and other hub-related
WP3	to identify and describe stakeholders of the Dutch hub	project partners with good access
WP3/WP5	to identify and describe stakeholders of the Spanish hub	other stakeholder groups
WP3	to identify and describe stakeholders of the Turkish hub	
WP3	to analyse and monitor job creation potential in all hubs	HR, regional actors interested in job creation
WP3 - T3.1.2	to collect stories (upstream) of products from hubs	can also be used as input for T6.2
WP3 - T3.1.2	to identify if labels increase(d) trust and acceptance	from key stakeholders of hubs
WP3 - T3.1.3	to conduct a social network analysis (SNA)	hub representatives and stakeholders
WP3 - T3.3/WP5	to develop metrics and indicators for measuring non-technologic	hub representatives and stakeholders
WP4/WP5/WP6	to characterize and model a hub	for all hubs
WP4	to characterize and model a hub	for all hubs
WP4/WP5/WP6	to characterize and model a hub	for all hubs
WP4/WP5/WP6	to characterize and model a hub	for all companies in the hub
WP4/WP5	to characterize production, storage, transport units in a hub	for all companies in the hub
WP4/WP5/WP6	to identify resources exchanged in a hub	for all companies in the hub
WP4/WP5	to identify resource flows within the hub and external	for all companies in the hub
WP4/WP5	to identify technologies used in industrial symbiosis	for all companies in the hub
WP4/WP5	to describe key processes of sourcing, manufacturing, etc.	for all companies in the hub
WP4	to describe key processes of sourcing, manufacturing, etc.	for all companies in the hub
WP4	data to be included in the digital product passport	to be decided
WP4/WP5	to identify resource characteristics	for all resources exchanged in hub
WP4/WP5	to identify resource characteristics	for all resources exchanged in hub
WP4/WP6	to identify resource processing	for all resources exchanged in hub
WP4	to develop models to optimize resource/energy flows	for all resources exchanged in hub
WP4/WP6	to develop models to optimize resource/energy flows	for all resources exchanged in hub
WP4/WP5/WP6	to account for environmental impact in the models	for all resources exchanged in hub
WP6	to account for environmental metrics	for all hubs
WP6	to identify industrial plant requirements	for all hubs
WP6	to identify infrastructure needs	for all hubs
WP7	to develop the platform	for all hubs
WP7	to develop the platform	for all hubs
WP7	to develop the platform	for all hubs
WP7	to develop the platform	for all hubs
WP7	to develop the platform	for all hubs
WP7	to develop the platform	for all hubs
WP7	to develop the platform	for all hubs
WP7	to develop the platform	for all hubs
WP7	to develop the platform	for all hubs
WP7	to develop the platform	for all hubs
WP5		for all hubs
WP5		for all hubs
WP5		for all hubs
WP5		for all hubs
WP5		for all hubs

Funded by the European Union

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 HADEA. Neither the European Union nor the granting authority can be held responsible for them.