



## Horizon 2020 - Research and Innovation Framework Program

### Semantically Connected Semiconductor Supply Chains

Project Acronym:

# SC3

Grant agreement no: 101007312

<b>Deliverable no. and title</b>	<b>D7.7 - Data Management Plan (DMP) First Version</b>	
<b>Work package</b>	7	Dissemination, Exploitation, and Sustainability
<b>Involved Tasks</b>	all	
<b>Lead contractor</b>	Infineon Technologies AG Thomas Gutt, mailto: <a href="mailto:thomas.gutt@infineon.com">thomas.gutt@infineon.com</a>	
<b>Deliverable responsible</b>	Claude YUGMA, mailto: <a href="mailto:yugma@emse.fr">yugma@emse.fr</a>	
<b>Version number</b>	V1.0	
<b>Date</b>	21.04.2021	
<b>Status</b>	Final <sup>1</sup>	
<b>Dissemination level</b>	Public (PU)	

**Copyright: SC3 Project Consortium, 2021**

<sup>1</sup> The Data Management Plan is considered a living document that will be updated whenever needed.

## Authors

Participant no.	Part. short name	Author name	Chapter(s)
	EMSE	Abdelhak KHEMIRI	1, 2, 3

## Document History

Version	Date	Author name	Reason
v0.1	20.03.2020	Abdelhak KHEMIRI	The initial document
v0.2	25.03.2020	Abdelhak KHEMIRI	Added partner contributions
V1.0	21.04.2021	Abdelhak KHEMIRI	Deliverable finalisation and submission

## Publishable Executive Summary

This deliverable refers to the first version of the Data Management Plan for SC3. In this context, this report examines and describe the data that will be produced, collected and/or processed within the project scope and how these data will be handled during and after the project, i.e., the standards that will be used, the ways in which data will be exploited and shared (for verification or reuse), and in which way data will be preserved. This DMP has been prepared by taking into account the template of the [“Template for the Data Management Plan” \[Version 1.0 of 13.10.2016\]](#).

This deliverable is not a fixed document, but it is a living document that will evolve during the lifespan of the project, particularly whenever significant changes arise such as dataset updates or changes. At this stage of the project, some questions are still open for discussion and we will add relevant information to the DMP as soon as it is available.

## Table of contents

1. Introduction .....	6
1.1 SC <sup>3</sup> Project Overview .....	6
2. Data Management Plan .....	7
2.1 Data Summary .....	7
2.2 FAIR Data .....	10
2.3 Allocation of resources .....	14
2.4 Data security and Ethical aspects.....	14
3. Conclusion.....	16
4. Appendix.....	17
4.1 Abbreviations .....	17

**List of tables**

Table 3: Abbreviations ..... 17

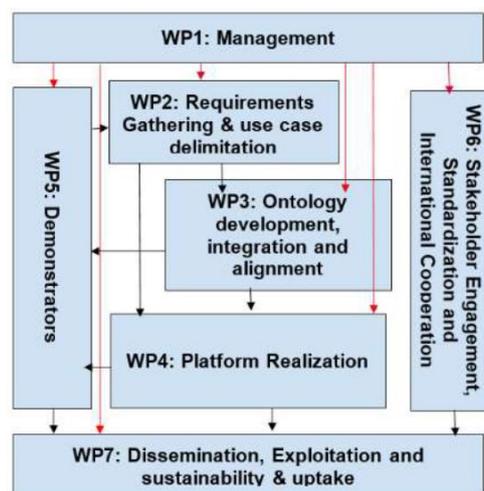
## 1. Introduction

### 1.1 SC<sup>3</sup> Project Overview

SC3 is a Horizon 2020 (H2020) Coordination and Support Action (CSA) project that aims to support seamless communication between various supply chain participants in the semiconductor industry. SC3 relies on allowing industrial and academic stakeholders to work together to ensure interoperability between semiconductor companies and other industrial fields.

The overall project is organized in 7 work packages as described below:

- **WP1** on project management includes coordination, logistics, reporting, communication risk control, and all other necessary actions to ensure that the project is executed properly.
- **WP2** will identify various requirements and needs from different stakeholders in regards to the data documentation in the semiconductor domain as well as others. This work package defines the functionalities for the collaboration platform in question. This WP also identifies relevant existing data models to be re-used.
- **WP3** is about the development of the ontologies and extending the top-level model with new ontologies from semiconductor and other domains, identified in WP2.
- **WP4** implements the platform for collaborative ontology development and maintenance, along with other functionalities.
- **WP5** is to validate the implementations done in WP3 and WP4 via demonstrators,
- **WP6** objective is threefold: (i) to organize, establish and ensure stakeholder engagement, (ii) to support the standardization process, and (iii) to foster international cooperation
- **WP7** is about dissemination & communication, community building, training as well as sustainability & uptake. The figure below illustrates the interactions between the work packages



SC3 aims to produce interoperable data, that is allowing data exchange and reuse between researchers, institutions, organizations, countries, etc. (i.e., adhering to standards for formats, as much as possible compliant with available (open) software applications, and in particular facilitating recombinations with different datasets from different origins). For this purpose, the Industrial reference platform using the semantic web technologies enables seamless data accessibility and exploitability for the public. Using an open innovation approach ensuring a constant upgrade of the Industrial Reference Platform.

## 2. Data Management Plan

### 2.1 Data Summary

#### 2.1.1 Purpose of the data collection/generation and its relation to the objectives of the project

The objective of the project is to overcome challenges related to the complex manufacturing process in the semiconductor domain and related supply chains. The key goals of the SC3 project, as well as those that clarify the function of data collection/generation in the framework of the project, as gathered during the proposal stage, are listed below:

- **OBJ1: Network relevant stakeholders to collect input on existing data documentation for semiconductor domain**  
*Within the SC<sup>3</sup> consortium, the focal domain is semiconductor industry. Such consortium domain knowledge along with our methodological approach towards retrieving, understanding, analysing and categorizing all relevant existing information will enable the implementation of consistent ontologies and taxonomies ensuring their intra-and cross-domain interoperability. The goal of SC<sup>3</sup> is to gather these stakeholders in order to collect, consolidate and confine documentation in a top-level ontology connecting all relevant pillars for semiconductor domain. SC<sup>3</sup> will capitalize on the semiconductor industry inclusive ontologies and taxonomies that will be developed as well as the partners' industrial and research domain experience, and will work systematically towards defining a top-level ontology able to connect to all relevant subdomains.*
- **OBJ2: Examine existing data documentation with respect to the top-level ontology**  
*Digital Reference offers a commonly-agreed top level ontology of relevant manufacturing and supply chain processes for the semiconductor's domain. The goal is to identify the status of Digital Reference as output from H2020/ECSEL/Productive4.0 project. SC<sup>3</sup> will systematically examine Digital Reference with respect to other existing industrial generic data models accepted by other industrial companies, especially semiconductor data models depicted by the major semiconductor manufacturers. The goal of SC<sup>3</sup> is to devise a semantic vocabulary for the domain, thus the project examines different data models relatively to Digital Reference to identify the gaps in the domain vocabulary. Further, we will devise steps to create a complete semantic standard vocabulary for the domain being the updated Digital Reference.*
- **OBJ3: Identify missing semantic data documentation for the domain**  
*Following the definition of the top-level ontology and the identification of existing data models for the domain, SC<sup>3</sup> will proceed in identifying missing semantic data documentation for the domain. A series of co-design workshops will be scheduled, with relevant bodies and wide representation of the industrial sector in order to acquire consensus, receive input and agree on the final missing documentation for the semiconductor domain to extend Digital Reference and reach a full commonly-agreed top level ontology representing the domain. SC<sup>3</sup> will collect input on existing data documentation, identify the existing semantic representation and missing data documentation.*
- **OBJ4: Extend the top-level ontology for the semiconductor domain**  
*After identifying the missing data document, SC<sup>3</sup> will work systematically towards extending a top-level ontology able to connect to all relevant semiconductor models. The project will foster the industry inclusive ontologies and taxonomies that will be developed as well as the partners' industrial and research domain experience, and will work systematically towards defining a top-level ontology able to connect to all relevant semiconductor concepts. Following the gap analysis, SC<sup>3</sup> extracts semantic data from different data structures and link them to the ontology. Existing industrial generic data models serve as an input for the ontology engineering process in order to ensure reuse-ability of existing models. Following the integration of vocabularies into Digital Reference, SC<sup>3</sup> emphasizes how semantic documentation overcomes*

*fragmentation due to lack of consensus in different communities having widely differing requirements for the same domain.*

- **OBJ5: Extend the Digital Reference to other domains**  
*SC<sup>3</sup> will proceed developing ontologies for other relevant industrial domains namely: automotive, industrial domain, communication, pharmaceutical and agricultural. Consensus with a wide number of industries and relevant bodies will be achieved through a series of dissemination events where wide representation of the industrial sector is due in order to acquire consensus, receive input and agree on the final definitions. SC<sup>3</sup> will examine the effect of semiconductor supply chain on other domains by analysing existing ontologies while safeguarding industrial applicability and fitting to Digital Reference. The integration of identified vocabularies into Digital Reference, highlights the impact of semiconductor industry on the widest use of data and, ultimately, reliable end-user products.*
- **OBJ6: Harmonize ontologies into top-level ontology**  
*SC<sup>3</sup> will work towards harmonizing newly created ontologies in order to be able to connect seamlessly to the top-level scheme, Digital Reference. An iterative and recurring process will be deployed in order to adapt ontologies to the new scheme. SC<sup>3</sup> will analyse existing vocabularies, ontologies from other domains and include them into Digital Reference. All modifications will be systematically disseminated to all relevant stakeholders. Digital Reference will be iteratively refined in a stakeholder-in-the-loop approach in order to maximize acceptance, applicability and ability to connect to all relevant sub-domains.*
- **OBJ7: Standardise documentation for the proposed taxonomies and ontologies**  
*SC<sup>3</sup> will liaise with relevant standardisation bodies so that the proposed taxonomies and ontologies are included in a standardised documentation with metadata that will be widely accepted by researchers and the industry. Providing such an approach the proposed top-level and sub-domain ontologies can be easily employed in existing efforts that aim at bringing together European researchers and innovators acting as key enabler for data discovery, access, use and reuse within a broad spectrum of resources for semiconductor domain as well as others.*
- **OBJ8: Establish a communication platform**  
*SC<sup>3</sup> enables collaboration not only among semiconductor manufacturers but beyond the semiconductor supply chain via Industrial Reference platform (i.e. the extended, merged or linked Digital Reference to other accepted or upcoming ontologies), which aims to evolve as an open standard and as a basis for a commercial B2B platform. This framework acts as a key enabler for realizing an agile development, validation, refinement loop of the Digital Reference ontology. Its infrastructure and hosting mechanisms assure accessibility in an open yet suitably controlled and secure manner. A series of co-design workshops will be scheduled, with relevant bodies and wide representation of the industrial sector in order to acquire consensus, receive input and agree on functional specifications.*
- **OBJ9: Sustainability and further development of on the ontologies and data documentation**  
*SC<sup>3</sup> will ensure the maintenance and further development of the ontologies and data documentation after the project duration through working with industry associations and other relevant EU-funded projects. The platform acts as a sustainable platform, which generates to a large extent its own income to ensure that the platform stays alive after the project lifetime.*

Data collection/generation shall augment the ontology development for comprehension. Ontologies are human crafted conceptualization of domain knowledge. Data generation focuses on the creation of ontologies and their documentations. Data collection (e.g., provenance information, who modified what, when, and why) serve the purpose to augment the data with meta-information to install trust and human-validation mechanisms for a successful adoption of the project.

Semantic Web technologies provide machine-readable representations of domain knowledge in the form of ontologies. For a successful realization and adoption, standardized ontologies of high quality are required. This standardization process has to fulfill requirements from various stakeholders and have to reflect a common agreement of the domain. The project aims this achievement through a collaborative ontology development framework. This framework shall enable multiple user groups to refine and create ontologies.

### 2.1.2 Types and formats of data generated/collected

The generated/collected types of data can be divided into two information layers. The first layer addresses the conceptualization (i.e., the ontology) and its documentation. In this layer, the datatypes comprise of .ttl files for ontologies and PDF files for documentation. The second layer addresses the provenance and meta-information (i.e., information about the authority governing the ontology, access rights such as read/write/suggest/accept changes, and related data and user role management information). In this layer, the data will be stored internally within the database and its access will comprise of machine-readable formats such as JSON.

### 2.1.3 Existing data to be re-used

To achieve the objectives of SC<sup>3</sup>, the outcome from previous projects will be used as a starting point. The Digital Reference (DR) was developed during the Productive4.0 project as funded from the Electronic Component Systems for European Leadership Joint Undertaking under grant agreement No 737459. Digital Reference comprises a combination of different ontologies of semiconductor supply chains and supply chains containing semiconductors. The domains have classes and subclasses that are defined by properties as well as descriptions and are connected to other classes in different domains, enabling a consistent representation across all taxonomies. These domains represent all stages of the supply chain with several sub-ontologies that currently represent concepts, hierarchies and organizations, e.g., product ontology, sensor ontology, organization ontology and business process ontology.

Existing data models like the Generic Data Model (GDM) or data sets will be transformed into ontologies and other relevant ontologies will be integrated to further expand the scale of the DR to more areas of semiconductor supply chain processes, consequently resulting in a holistic ontology.

### 2.1.4 Origin of the data

During the project lifetime, SC3 activities will exploit different data from varied sources. A first source of data that will be used relates to already existing and publicly accessible ontologies by using their URI. Another source of data relates to the framework proposed by SC3, and can be further divided into two different sources. The first one corresponds to the uploading of local ontologies using the file upload feature of the framework. The second one refers to the data that are generated by the users through the use of the system. Finally, several datasets coming from the Information System (IS) and Manufacturing Execution System (MES) of actual wafer fab will be used.

In summary, the project will mainly use four different data sources:

- 1) Existing public accessible ontologies using URI,
- 2) Upload of local ontologies using file upload feature of the framework
- 3) User generated data through the use of the system
- 4) Dataset generated by the IS and MES of real wafer fab

### 2.1.5 Expected size of the data

In the current state of the project, the expected size of the data is not yet defined. A first estimate will be provided in the next version of this document and will be refined in future versions. Moreover, ontology size can vary, its documentation representing human-readable explanations can increase in size with complexity of ontologies, modifications and messaging mechanisms increase overtime including all transition states of the ontology.

### 2.1.6 Data utility

A fundamental aspect of Semantic Web is to create and share conceptualizations of domain knowledge. The framework provides a collaborative ontology development approach to be used by academic and industrial stakeholders. Standardized and high-quality ontologies thrive through communication and mutual agreement on conceptualizations. Thus, shaping the standards in the domain will benefit all involved participants.

More specifically, the data collected and generated by the SC3 project is useful for a wide spectrum of stakeholders interested in the semiconductor supply chains or supply chains containing semiconductors. Different groups have been identified in the initial plan for exploitation and dissemination of results (PEDR) and can be divided as follows:

- Industry representatives and associations
- Other relevant EU-funded projects
- Scientific / academic community
- Standardization bodies and community initiatives
- Representatives from the overall (industry) data community / economy
- Society in general

## 2.2 FAIR Data

### 2.2.1 Making data identifiable

#### 2.2.1.1 Outline the discoverability of data (metadata provision)

Within the scope of the project, integration activities will be performed for the realization of the Digital Reference collaborative ontology framework. All the software components that constitute the framework will be put together under a unified online environment. This way the framework will facilitate all the activities involved in the development of ontologies, the support of the ontology governance model, including tools for metadata provision to facilitate the FAIRification of ontologies, curation, issue management, unit and integration testing, workflows and collaborative modelling interfaces, catalogue maintenance, export and federation. Additionally, the necessary UI interfaces will also be implemented, so that the Digital Reference framework to be accessible to all the stakeholders, able to collaborate with each other via this environment.

The framework will also provide indexing and archiving features, allowing for searching of ontologies and related conceptualizations directly in the system. The indexing system provides the name and a human readable description of the ontology, which is publically accessible metadata, allowing for discoverability of data. If the ontology is not yet publically released, users can request access that has to be granted by the governing authority.

### 2.2.1.2 Describe the data identifiability and refer to the standard identification scheme. Do you make use of persistent and unique identifiers such as Digital Object Identifiers?

Ontologies define their own unique identifiers (URI) within the description of their conceptualizations. Typically, most elements are identified by URIs.

Within the project ontologies are going to be uniquely identified using UUID4 tags.

### 2.2.1.3 Description of naming conventions used.

The best practices for naming conventions for ontologies will be used for the creation of classes and properties. Towards this objective, a first document, based on literature research, describing best practices for naming conventions that should be considered when implementing ontologies has been proposed. Also, to guide the different users and enforce the naming convention, the system will provide warnings to the user if naming conventions are not followed.

### 2.2.1.4 Outline the approach towards search keyword

Ontologies reflect a network-like structure with concepts and their corresponding labels. Keyword search will be realized through different dictionaries. The first dictionary will include the machine-readable definitions of concepts and relations based on their URIs. Ontologies typically support multiple languages (i.e., language tags), therefore each language will create own dictionaries. Keyword search will allow the user to specify the individual dictionary or combination of dictionaries.

### 2.2.1.5 Outline the approach for clear versioning

Regarding the versioning approach that will be used, the SC3 project addresses two different aspects.

- 1) Ontologies can provide own metadata about versions. However small changes and the comprehension of what has changed are not reflected by version tags.
- 2) Versioning using version control systems, such as git, will facilitate identifying the changes over time, and also allow users to revert to a specific version in time. Furthermore, differences between two versions are directly available out of the box.

We plan to address versioning using git and additionally allow users to push version tags. The use of Blockchain technologies will provide the history of the ontology and allow installing trust and validation.

### 2.2.1.6 Specify standards for metadata creation (if any). If there are no standards in your discipline describe what metadata will be created and how

Metadata is created to facilitate communication, trust, and validation. Most of the metadata will be created by users. This information will be augmented with timestamps, decisions and comments. Further metadata will contain diffs, facilitating comprehending the changes over time. To increase discoverability, the following metadata is required for all ontologies (also for unpublished ones): Name, Description, and URI. Further classification into domains and sub-domains (e.g., Semiconductor, or Health-Care domains) are optional. Ontologies use owl:imports statements to integrate existing ontologies. If these ontologies have domain specifications in our system we can directly infer these and provide them to the user for validation.

## 2.2.2 Making data readily accessible

### 2.2.2.1 Specify if the data will be accessible and if yes which data will be made openly available? If some data is kept closed provide rationale for doing so.

If ontologies are published using the framework they will be openly available. Additional metadata to install discoverability, trust, and validation will be published alongside. User-related metadata will remain closed to be compliant with GDPR.

### 2.2.2.2 Specify how the data will be made available

The data will be made available within the framework with different granularity of access rights. Publicly available data will be provided as downloadable ontology files (.ttl). Documentations of ontologies will be available via PDFs.

### 2.2.2.3 Specify what methods or software tools are needed to access the data? Is documentation about the software needed to access the data included? Is it possible to include the relevant software (e.g. in open source code)?

To access the data, a web-browser is necessary. The framework will allow to access and interact with the data. Ontologies are typically distributed in textual formats. However, to facilitate development and communication, the framework will provide this features. Most of the developed components will be available as open-source code, allowing for reuse and customization of functionalities to individual use cases.

### 2.2.2.4 Specify where the data and associated metadata, documentation, and code are deposited

Data, documentation, and associated metadata will be deposited in the internal database of the framework. The source code for individual components and documentation will be located on GitLab.

### 2.2.2.5 Specify how access will be provided in case there are any restrictions

Access to ontologies will be governed by users, groups, and organizations themselves. Since conceptualizations developed over time, read/write/suggest access can be granted to users or whole groups during the development process. A governing authority can publish their ontologies within the system, making them publically available for everybody with read-rights. Suggestion rights and editing rights have to be granted by the governing authority to users or groups.

## 2.2.3 Making data interoperable

### 2.2.3.1 Assess the interoperability of your data. Specify what data and metadata vocabularies, standards or methodologies you will follow to facilitate interoperability.

Ontologies are machine-readable representations of domain knowledge. Thus, they already provide interoperability functionalities. To further create the interoperability of data with respect to metadata, necessary metadata will be provide in JSON formats. Furthermore, we plan to create a local Ontology-Lookup-Service (OLS) instance for further interoperability of individual ontologies and related metadata by its Neo4J graph database and RESTful API calls.

### 2.2.3.2 Specify whether you will be using standard vocabulary for all data types present in your data set, to allow inter-disciplinary interoperability? If not, will you provide a mapping to more commonly used ontologies?

Ontologies are developed using RDF, OWL and standard vocabularies such as SKOS, ScorVoc, DataCube, and dcterms. Related metadata will use similarly standard vocabularies that will be stored within the OLS system for further interoperability.

## **2.2.4 Increasing data re-use**

### 2.2.4.1 Specify how the data will be licensed to permit the widest reuse possible

Almost all artefacts developed in SC3 will be made available under CC-4.0 or MIT licenses to maximize exploitation, reuse and market penetration.

### 2.2.4.2 Specify when the data will be made available for re-use. If applicable, specify why and for what period a data embargo is needed

Published ontologies will be made available directly after the publication process. Documentation and provenance metadata will be made available as soon as they are available.

### 2.2.4.3 Specify whether the data produced and/or used in the project is useable by third parties, in particular after the end of the project? If the re-use of some data is restricted, explain why

The central aspect of Semantic Web is to create and share conceptualizations of domain knowledge. Created ontologies if publicly available will also be accessible to third parties. Versioning, history and diff information should also be publicly available. However, we will exclude the details about discussion/comments and user names to preserve small digital foot print and being GDPR compliant.

One of the project's objectives is to ensure the maintenance and further development of the ontologies and data documentation after project duration through working with industry associations and other relevant EU-funded projects. Thus, the produced data will remain usable by third parties after the end of the project.

### 2.2.4.4 Describe data quality assurance processes

From an industrial side there will be several instruments to maintain and curate the data. On the one hand, the data will be updated and extended answering public change requests in yearly release workshops. The entire industrial consortium will decide about the change request and aimed extensions of the Generic Data Model in these workshops. On the other hand, the industrial data basing on the Generic Data Model are structured in different use cases. Therefore, they are targeting specific purposes and only require a specific view of the Generic Data Model. Each use case is maintained by one partner of the industrial contributors. The public data user concept guarantees that the right group of users gets access to the right, use-case based data. In consequence for each use case data sample the public has concrete responsible contacts to get advice concerning the data sample. In this way, also the incoming public change requests can be channeled to the responsible for the industrial use case data. In the frequent version release workshops the industrial consortium decides about the change requests collected by the data use case owners.

Naming conventions and syntax validation for creation of ontologies will be used for data quality assurance. Additional human-in-the-loop validation through the governing authority supported

by message system providing discussions based on information about the suggestion of changes and the decision for accepting/rejecting a change.

#### 2.2.4.5 Specify the length of time for which the data will remain re-usable

We estimate the provision of data to remain re-usable for the duration of ten years.

### 2.3 Allocation of resources

The costs for data curation and preservation will be covered as follows:

- During the project duration, the costs for curation and preservation will be covered directly by the project. In addition, TIB will contribute the use of its services (e.g., long-term preservation) as in-kind contribution to the project.
- After the end of the SC3 project, TIB is committed to continue the preservation and facilitate the curation of the ontology and datasets as part of its role as an information center for science and technology. For that purpose, TIB is currently setting up a sustainable ontology curation and preservation platform, which will also be offered to be used for long term sustainability by similar initiatives in other domains (e.g., the NFDI or Int. Data Spaces). Also, TIB already provides a long-term archiving and preservation service, which is NESTOR certified and will be used for SC3.
- In addition, TIB plans together with the other SC3 consortium partners to identify and target further funding sources for an extended curation and further exploitation of the ontology and datasets after the course of the SC3 project.

### 2.4 Data security and Ethical aspects

The Consortium will ensure respect for the ethical principles and fundamental rights embedded in the regulatory framework of the European Union. It will ensure respect for the Charter of Fundamental Rights of the European Union, the European Convention on Human Rights and the forthcoming European Commission Ethics Guidelines for trustworthy AI. Any data processing activities within the project will be carried out in accordance with the data protection rules as specified in Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data and repealing Directive 95/46/EC (hence: the GDPR) in combination with the e-Privacy Directive 2002/58/EC. The Consortium will pay close attention to the new developments in this area in order to provide the maximum level of privacy and data protection guarantees and create a culture of trust and assurance around data.

Specific attention will be paid towards addressing the ethics of research throughout different stages of the project since they will be handling certain types of personal data. All activities will be carried out ensuring the ethical principles and in compliance with Directive 95/46/EC of the European Parliament and subsequent Regulation 2016/679/EC and Directive 2016/680/EC, about the protection of individuals with regard to the processing of personal data and on the free movement of such data, as well as Directive 2002/58/EC, concerning the processing of personal data and the protection of privacy in the electronic communications sector, as modified by Directive 2009/136/EC.

The consortium guarantees that all personal data collected during the project will be kept secure and unreachable by unauthorized persons. The data will be handled with appropriate

confidentiality, accessibility controls, and technical security. The partners will use state-of-the-art security policies and rules to ensure the integrity and availability of electronic information captured, stored, maintained, and used. In detail, the consortium will adopt the following principles when dealing with personal data:

- Confidentiality and anonymity - confidentiality will be guaranteed whenever possible. This may also be important in dealing with industrial stakeholders. The only exception can be in some cases for the researcher directly interacting with a group of participants (e.g., focus group). Our research team will not make publicly accessible any personal data. Anonymity will be granted through generalization and pseudonymisation. Furthermore, provisions will be taken to avoid the possibility of information linkage.
- Informed consent - the informed consent policy requires that each participant will provide his/her informed consent prior to the start of any activity involving him/her. All people involved in the project's face-to-face research and evaluation activities (e.g., interviews, workshops) will be asked to read and approve an Informed Consent Form explaining how personal data will be collected, managed and stored. For the remaining data collection activities (collection of publicly available patent data and companies' data) informed consent procedure will not be applied.
- Circulation of the information limited to the minimum required to achieve a limited circulation of the information, the database containing in anonymous form the data collected from the users will be distributed to the partners, if needed at all, through protected and encrypted Internet connections; the raw data will only be shared if it is required for the development. The project partners will never pass on or publish the data without first protecting participants' identities. No irrelevant information will be collected; at all times, the gathering of private information will follow the principle of proportionality by which only the information strictly required to achieve the project objectives will be collected.

The project team will store research data on secure repositories protected by password at the head offices of partner organizations. File names will not refer to any personal information. Information that might enable data to be linked to individuals, such as the file linking participants' names to their respective code/pseudonym, will be password protected and encrypted so that access will be restricted to only those with the requisite credentials. The partner organizations will be responsible for ensuring all ethical principles relating to their country and institutional context are adhered to.

### 3. Conclusion

This deliverable has provided the first version of the SC3 project Data Management Plan. It gives an overall picture on the research data management that will be used within the project scope by describing the data that will be produced, collected and/or processed scope and how these data will be handled during and after the project.

## 4. Appendix

### 4.1 Abbreviations

Table 1: Abbreviations

Abbreviation	Meaning
SC3	Semantically Connected Semiconductor Supply Chains