



## Deliverable 1.1

### Data Management Plan

DISSEMINATION LEVEL		
<b>PU</b>	Public	<b>X</b>
<b>CO</b>	Confidential, only for members of the consortium (including the Commission Services)	



COVER AND CONTROL PAGE OF DOCUMENT	
Project Acronym:	CPSELabs
Project Full Name:	CPS Engineering Labs - expediting and accelerating the realization of cyber-physical systems
Grant Agreement No.:	644400
Programme	ICT-1: Cyber-Physical Systems
Instrument:	Innovation Action
Start date of project:	01.02.2015
Duration:	36 months
Deliverable No.:	D1.1
Document name:	Data Management Plan
Work Package	WP1
Associated Task	Task 1.3
Nature <sup>1</sup>	R
Dissemination Level <sup>2</sup>	PU
Version:	1.0
Actual Submission Date:	2015-07-31
Contractual Submission Date	2015-07-31
Editor: Institution: E-mail:	Name: Meike Reimann Name (short name): SEZ email: reimann@steinbeis-europa.de

The CPSELabs project is co-funded by the European Community's Horizon 2020 Programme under grant agreement n° 644400.

The author is solely responsible for its content, it does not represent the opinion of the European Community and the Community is not responsible for any use that might be made of data appearing therein.

<sup>1</sup> R=Report, DEC= Websites, patents filling, etc., O=Other

<sup>2</sup> PU=Public, CO=Confidential, only for members of the consortium (including the Commission Services)

## Change Control

### Document History

Version	Date	Change History	Author(s)	Organization(s)
0.1	2015-06-08	Document drafted	Meike Reimann	SEZ
0.2	2015-07-06	Document updated, partner inputs included	Meike Reimann, all Partners	SEZ, all partners
0.4	2015-07-21	Document updated	Meike Reimann	SEZ
1.0	2015-07-30	Document finalised	Holger Pfeifer	FOR

### Distribution List

Date	Issue	Group
2015-06-15	Collection of inputs	Project consortium
2015-07-12	Revision v0.2	Project consortium
2015-07-24	Revision v0.4	Project consortium
2015-07-30	Finalisation	Project coordinator
2015-07-31	Submission	EC

## Table of Contents

<b>Executive Summary .....</b>	<b>5</b>
<b>1 Introduction and Context .....</b>	<b>6</b>
<b>2 Project Goals and Implications on Data Sharing.....</b>	<b>7</b>
<b>3 Data Generation and Management in the Scope of CPSELabs .....</b>	<b>8</b>
<b>3.1 Data and Information generated and shared by the Consortium.....</b>	<b>9</b>
<b>3.2 Data generated in conjunction with Third Parties Experiments.....</b>	<b>13</b>
<b>4 Conclusion .....</b>	<b>19</b>

## Executive Summary

The purpose of this Data Management Plan (DMP) is to provide an analysis of the main data foreseen to be generated in the course of the project and to describe the data management policy that will be applied by CPSE Labs. The project consortium fully supports the endeavour to improve access to scientific information and research data and will make information and data generated within the project available on a voluntary basis, whenever possible.

CPSE Labs pursues the goal to contribute to establishing an open eco-system, and the project plan has been conceived to broadly disseminate the project findings and to contribute to the generation of broader knowledge in the field. Therefore, the vast majority of the project deliverables are public, containing information and data that can be used or re-used by various target groups.

A variety of data and information will also be generated in CPSE Labs experiments, which involve 'third parties' through 'cascading funding'. CPSE Labs perceives it as its' role to accompany the third parties in aspects of data management supporting open access of the generated research results, along with publications, so they can be easily discovered, identified and re-used, whenever possible.

# 1 Introduction and Context

As stated in the *'Guidelines on Open Access to Scientific Publications and Research Data in Horizon 2020'*, fuller and wider access to scientific publications and data helps to:

- build on previous research results (improved quality of results)
- foster collaboration and avoid duplication of effort (greater efficiency)
- accelerate innovation (faster to market = faster growth)
- involve citizens and society (improved transparency of the scientific process)

The CPSELabs consortium fully supports the endeavour to improve access to scientific information and research data in order to enhance the benefits of public investment. Especially, if the information and data have been derived with the help of public funding, CPSELabs agrees, that this should benefit European companies and citizens to the full.

The Open Research Data Pilot, aims to improve and maximize access to and re-use of research data generated by projects. As defined in the guidelines, openly accessible research data can typically be accessed, mined, exploited, reproduced and disseminated free of charge for the user.

The CPSELabs project participates in the 'Open Research Data Pilot' and will make its research data available on a voluntary basis, whenever possible.

The role of this Data Management Plan (DMP; D1.6), is to drive the policy towards providing open access to the data generated in the scope of the CPSELabs project, along with publications and other project results, so they can be easily discovered, identified and re-used.

## 2 Project Goals and Implications on Data Sharing

A variety of data and information will be generated in CPSE Labs, ranging from interview outcomes, guidelines and best practices to software artefacts and (raw) sensor data.

Whereas a part of that data will be generated by the consortium itself, and will be made freely available via the website and public deliverables, much of the data that could be categorized as 'digital research data' will be generated in conjunction with 'third parties', participating via cascading funding. (A part of the CPSE Labs project funding is used to involve project external 'third parties' through open calls in 'experiments'.)

A major goal of the project is to build an open eco-system supporting the whole stakeholder community (from CPS developers, integrators, and suppliers to users), to enhance technology transfer by providing existing open platforms and tools for application experiments and to enable stakeholders to benefit from use and re-use of experiences, data and information.

A second major aim of the project is to efficiently involve SMEs and mid-caps (as third parties), to help them in the development and commercialization efforts of CPS enabled/related technologies and products (through open call experiments) and with this increase European competitiveness.

While in the first case, opening information and data is well in-line with the project goal, in the second case, sharing of data and information might jeopardize the endeavour of exploiting and commercializing the results or products developed in CPSE Labs experiments with third parties.

In the course of the project, the CPSE Labs consortium will have to carefully consider and agree with the third parties on a case-by-case basis if, how and to what extent data can be shared. Especially in the following cases, above others, the collected/generated data will not be shared:

- if the results can be expected to be commercially or industrially exploited by the project partners or third parties (or if sharing would contradict intellectual property rights and commercial exploitation in any way);
- if sharing would jeopardize efficient the involvement of SMEs;
- if incompatible with the need for confidentiality in connection with privacy-related or security-related issues;
- if incompatible with existing rules concerning the protection of personal data ;
- if incompatible with existing rules concerning ethical issues. (In the project's ethical report, the consortium has clearly defined how sensitive / personal data will be treated.)

### 3 Data Generation and Management in the Scope of CPSE Labs

The CPSE Labs partners expect that most likely it will be mostly software artefacts that are being produced (mainly within the third parties experiments), rather than true research data. The latter includes simulated data that can be used to test physical systems based on a simulation platform that will be made available on an open source basis. Moreover, performance analysis data, based on several configurations of the platform or systems based on the Design Centre platform may provide a basis for decision making for other third parties. More explicitly, the generated data within the third parties experiments might include:

- open-source software, either as standalone tools, or as libraries/plugins extending other existing (not necessarily open source) tool sets;
- experimental artefacts like use case descriptions, exemplary analysis or design models, exemplary analysis results
- descriptions of domains, co-models, descriptions of metrics used, tool extension data
- experimental sensor data sets (anonymized, if required) to repeat executed experiments
- reports on executed experiments, public deliverables synthesizing the experiments; and scientific publications
- reports on best practices using the Design Centre platforms

Next to this, the Design Centres will yield other data, such as results of interviews carried out with stakeholders in the context of eco-system analysis, needs in professional training, measurement data on performance of the Design Centres (in terms of KPI) and summary data related to literature surveys (coded sources, categorizations, etc.). Moreover, results will include information like contributions to the contacts database and information and data on the Open Call content, process, outcome and response.

More explicitly, the information captured and data generated by the project partners with the help of interviews, surveys and other investigations with project internal and external participants will derive, among others:

- overview of existing innovation practices and opportunities in innovation eco-systems
- inventory of existing professional training, good practices and effects
- investigation of needs for professional training as perceived by relevant stakeholders
- overview of CPS areas of relevance for stimulating innovation by means of Market Places (MP's) including an overview of existing MP's and best practices of MP's
- stakeholder needs and considerations for the MP pilots within CPSE Labs
- information on open call process, (including FAQs), outcome, statistics, feedback

A very important factor for sharing information and data within the CPSE Labs project is, as described above, by what means and with which aim this data is generated. This can be sub-divided into two categories:

- 1) Data collected / generated by the project consortium, with the aim of broad dissemination
- 2) Data generated by / with the third parties in the scope of open call experiments

As these categories differ substantially, they will be described in two separate sub-chapters.



### 3.1 Data and Information generated and shared by the Consortium

As CPSELabs pursues the goal to contribute to establishing an open eco-system supporting the whole community of stakeholders, the vast majority of the projects deliverables are public, containing information and data that can be used or re-used by different target groups. Besides being publicized via the CPSELabs website, the documents will be spread via the CPSELabs partners networks. The following table gives an overview on a selection of the CPSELabs (planned) public deliverables, conceived to broadly disseminate the projects findings and contribute to the generation of broader knowledge in the field. The table below lists the documents per work package (WP), deliverable number (Del) and contains the publicizing month (M); M1 corresponds to February 2015. A brief information is given on the type of data and the target group. More detailed information in terms of information / data content, can be obtained from the deliverables themselves.

WP Del	Document name, Publicizing month	Data / type of information  Target group for (re-) use
<b>WP1 Project Management</b>		
D1.6	'Collaboration plan with other Smart Anything Everywhere projects' (M2)	The public deliverable provides a shared vision of the current SAE coordinators/teams on collaboration within and future evolvement of the SAE initiative.  (Targeted to SAE stakeholders, EC, policy makers)
<b>WP2 Communication and Outreach</b>		
D2.1 D2.2 D2.4	'Web portal' (M2) 'Communication Plan' (M3, 12, 18, 24, 30, 36) 'Public Materials' (M3)	The data/information provides detailed information on the CPSELabs Design Centres and their 'open tools and platforms', the CPSELabs Vision as well as practical information and guidance for applicants of the open call process (e.g. FAQs).  (Targeted to experiment proposers, stakeholders of the CPS eco-system, broad public)
<b>WP3 Open Call Process for Experiments</b>		
D3.1 D3.3	'Open Call Process Documents' (M3) 'Call Texts' (M3, 9, 15)	Information on the open calls content and the process: The data/information provides detailed information and guidance on structuring and handling of the open call process. Next to giving guidance to proposers and evaluators, these documents provide 're-usable' information on the call process and templates for future open-call projects.  (Targeted to experiment proposers, evaluators, EC, other projects with cascade funding)
D3.2	'Information events and coaching activities' (M16)	The data/information includes experiences and best practices from 'Information events and coaching activities', which can be valuable in terms of 'lessons learnt' for future endeavours.  (Targeted to experiment proposers, other projects with cascade funding)

WP4 Design Centres		
D4.1	'Centre handbook' (M4)	<p>The data includes information on centre management and exchange of best practices among Design Centres, promoting synergies among them and their regional eco-systems by:</p> <ul style="list-style-type: none"> <li>- establish a learning network among the Design Centres to exchange best practices in creating innovation eco-systems;</li> <li>- carry out cross-centre opportunity scouting in which the research, industrial and business profiles of centres and their regional eco-systems are examined to identify innovation and other collaboration opportunities</li> </ul> <p>It also includes templates and guidelines for basic processes (Targeted to Design Centres, regional eco-systems, educational institutions, policy makers...)</p>
D4.2	'Report on best practices and professional training' (M12, M24, M36)	<p>Information will include results of:</p> <p>Analysis of best practices and professional training within partner eco-systems and exchange of best practices within the regional eco-systems of each centre by</p> <ul style="list-style-type: none"> <li>- establishing regional learning networks</li> <li>- identifying industrial needs for professional training of particular relevance for CPSELabs</li> <li>- matching these needs with existing competences and courses;</li> <li>- implementing selected training</li> </ul> <p>(Targeted to Design Centres, regional eco-systems, educational institutions, stakeholders of CPS eco-system, policy makers...)</p>
D4.3	Innovation management including 'Annual report on innovation management activities' (M12, M24, M36).	<p>Information will include results of the innovation management activities of CPSELabs</p> <ul style="list-style-type: none"> <li>- participating in reviews of experiments and marketplace efforts, including categorization, TRL assessments, and mapping and analysis of collaborative innovation activities using social network analysis,</li> <li>- identifying business opportunities and improvements in practices for CPS innovation management.</li> <li>- interview studies of firms having central roles in the innovation eco-systems based on cyber-physical systems in order to identify existing best practices for managing networked and open innovation in this field</li> <li>- preparing an action plan for commercialization / standardization</li> </ul> <p>(Targeted to Design Centres, stakeholders of CPS eco-system, regional eco-systems, EC, policy makers...)</p>

D4.4	'Strategic Innovation Agenda for CPS' (M8, 14)	<p>The Strategic innovation agenda for CPSE Labs contains information on setting out the overall direction for experiments and other eco-system promoting interactions, and provides plans for the open calls for experiments. The CPS-SIA will also consider existing agendas as far as relevant, including for example the Artemis strategic research agenda and the EIT ICT Labs strategic innovation agenda.</p> <p>(Targeted to Design Centres, stakeholders of CPS eco-system, EC, policy makers)</p>
D4.5	'Market Place Report' (D4.5)	<p>The report will contain information about the creation of marketplaces for selected CPS technology platforms, such as middleware platforms for CPS and tool integration platforms. Information on suitable models for a marketplace (e.g. in terms of IP rights, open source, governance, codex, best practices) will be presented. A first marketplace pilot will address the sharing of software assets and best practices to promote interoperability for CPS engineering environments. An early survey will identify the willingness of research and industrial organizations to contribute to and take-up assets from the marketplace.</p> <p>(Targeted to stakeholders of CPS eco-system, regional eco-systems, EC, policy makers...)</p>
D4.6	'Design Centres final report' (M36)	<p>The report will include the final evaluation and impact assessment. Additional information on identified "take-aways" and further evolution of innovation eco-systems in general, and for CPSE Labs in particular; an overall evaluation of the goals, methodology and achievements of CPSE Labs will be included.</p> <p>(Targeted to stakeholders of CPS eco-system, EC, policy makers, other projects with cascade funding ...)</p>
<b>WP5 Dissemination and Exploitation</b>		
D5.1 D5.2	'Dissemination and Exploitation Plan' (M3) 'Annual Report on Dissemination Activities' (M12, M24, M36)	<p>WP5 will make projects outcomes public and will build an eco-system for sharing information and exploiting the knowledge generated during the projects lifetime.</p> <p>Next to publishing direct results the information will be related to: relevant conferences and workshop outcomes, influencing research programs, influencing standards bodies, influencing educational institutions, raising awareness and setting up communities, open access; new or improved products and services, incubation of business ideas, creation of start-up and spin-offs.</p> <p>(Targeted to stakeholders of CPS eco-system, standardization bodies, educational institutions, policy makers...)</p>

WP6 Execution of Experiments		
	Public outcomes from experiments	Experiments will produce a publishable summary of their work and results (not including any confidential information). Moreover, a mean of 1 scientific or market-oriented publication per experiment is expected. Additionally, research data might be provided in an open database (to be decided on a case-by-case basis).  (Targeted to stakeholders of CPS eco-system, academia, industry, EC, policy makers, broad public)
D6.3	'Final Experiments Report' (M36)	This report will contain the main publishable outcomes of the experiments, including an assessment of outcomes and extraction of exploitable results.  (Targeted to stakeholders of CPS eco-system, academia, industry, EC, policy makers, broad public)

Table 1: Overview of data generated and shared by the consortium

Regarding additional (peer-reviewed) publications, which are foreseen to be academic or market-related, the CPSE Labs general policy is to require open access for all publications. Self-archiving ("green" open access) is expected. Partners will be required to ensure before submission that publications will be eligible for archiving on institutional repositories of at least one of the co-authors. It is recognized that, in a very few exceptional cases, "gold" open access may be required. Data used in publications will be made available, either on the web portal or by application to the CPSE Labs Service Centre.

Besides the public deliverables and other publications, CPSE Labs will create an interactive open marketplace: The CPSE Labs launches a marketplace for sharing software assets related to integrated CPS engineering tools and environments. CPSE Labs is aiming at maximizing input with this marketplace by establishing an enlarged forum of developers, integrators, and users from global powerhouses as well as SMEs and mid-caps.

Moreover, CPSE Labs aims at contributing to standardization: Relationships with standardization body and open platform groups are planned to make the results available and acceptable for a wider audience. This includes presentation and visit to specific groups such 'The Open Group Open Platform 3.0' (as <http://www.opengroup.org/>) which is cross-domain, or Autosar, which is dedicated to automotive standard. The Open Group has committed to support (non-funded) the CPSE Labs in identifying standardization opportunities and to also participate in open call evaluation and in reviews of experiments for identifying standardization opportunities.

### 3.2 Data generated in conjunction with Third Parties Experiments

Third party experiments are carried out in close collaboration with the partners of one of the CPSELabs Design Centres in South Germany (fortiss), North Germany (Offis), France (ONERA and LAAS-CNRS), Sweden (KTH), the UK (Newcastle Univ.) and Spain (Univ. Politécnica de Madrid and Indra Sistemas). The Design Centres offer expertise and training in developing cyber-physical systems, as well as development environments, tool chains, architectural frameworks, and technology platforms that form the basis for the experiments, including:

- 4DIAC framework for distributed industrial automation and control
- FMI-based virtual co-simulation
- eMIR - open source test platform for maritime systems ([www.emaritime.de](http://www.emaritime.de))
- Model-based safety assessment techniques (AltaRica, Hazop UML)
- GenoM and Mauve-OROCOS frameworks for robotics systems programming
- Open Services for Life-Cycle-Collaboration (OSLC) open standard
- Overture family of VDM-based technologies (Overture, Crescendo, Symphony)
- SOFIA2 interoperability platform for smart spaces

At the time of the deliverable (M6, July 2015), the first round of calls has been closed, but the process of experiment selection, invitation and confirmation has not yet been completely concluded. The collection/generation and sharing of data heavily depends on the experiments performed, and the third parties involved. Considering this, a detailed analysis of the data foreseen to be collected and possibly shared can only be performed at a later time point.

Nevertheless, the Design Centres, based on their calls and platforms available, have made some assumptions on the data that could be generated and the handling of it. The results of a first survey and discussions amongst the consortium is shown in the following table.

Centre/ Partner	Type of data expected to be generated in conjunction with third parties Plan for sharing the data
Design Centre Germany South	
FOR	<p>The data generated will most likely be software artefacts, rather than true research data. “Data” in the stricter sense might come from simulation runs that will be executed within the virtual co-simulation experiments, which will enable to improve the key technologies of the Design Centre. In cases where third parties generate the data (or participate therein) consent to share the data will be required.</p> <p>Developments on some of the core technologies provided by the Design Centre will be provided as open-source software artefacts, or in the form of publishable research reports. The example below illustrates how this will be mapped to the data management plan.</p> <p><b>Data set reference and name</b></p> <p style="padding-left: 40px;">4DIAC: Framework for Distributed Industrial Automation and Control</p> <p><b>Data set description</b></p> <p style="padding-left: 40px;">4DIAC presents an open source software solution implementing IEC 61499. It consists of a run-time and a GUI part. The run-time is called FORTE and is deployed to the individual controllers as a basic execution framework allowing the execution of the applications in real-time on top. The GUI represents an IDE</p>

	<p>realized in Eclipse. It supports the developers by creating their applications and deploying them to the controllers running an instance of FORTE.</p> <p>As 4DIAC is licensed under EPL (Eclipse Public License) all extensions and adaptations to it have to be provided under EPL again. This ensures that all improvements performed within the project are offered to all the other users of 4DIAC, which allows them to benefit from the modifications as well.</p> <p><b>Standards and metadata</b></p> <p>The 4DIAC GUI is implemented using Java in Eclipse and consists of a set of individual Eclipse plug-ins distributed as RCP (Rich Client Platform) and source code.</p> <p>FORTE itself is implemented using C/C++ and is currently ported to a set of different platforms, e.g. Raspberry Pi, Beaglebone Black, Wago PLCs, etc.</p> <p><b>Data sharing</b></p> <p>All the code of 4DIAC is publicly available for download under <a href="http://www.fordiac.org">http://www.fordiac.org</a>. Within the near future the code base of 4DIAC will be ported into the Eclipse repository, where it is even better visible to the public.</p> <p><b>Archiving and preservation (including storage and backup)</b></p> <p>Currently the individual code versions are handled using Hg. Later on the individual development states will be supported using a Git repository.</p>
Design Centre Germany North	
OFF	<p><b>Type of Data</b></p> <p>OFF will focus on architecture development. Work will be based on eMIR (open source test platform for maritime systems (<a href="http://www.emaritime.de">www.emaritime.de</a>)).</p> <p>The following types of data sets may be provided: Simulated data that can be used to test physical systems based on the simulation platform, which will be made available on an open source basis. Examples of simulated data can be simulated traffic data or engine performance simulation. Performance analysis data based on several configurations of the platform may provide a basis for decision making for other third parties.</p> <p><b>Data sharing</b></p> <p>Experiments done by OFF without restrictions of third parties will be published corresponding to DFG guidelines (Deutsche Forschungsgemeinschaft, German Research Foundation) for scientific best practices, most likely in conference papers or journal articles. However, OFF will also archive and publish digital data, if available. In cases where third parties generate the data (or participate therein) consent to share the data will be required.</p> <p>The collection/generation and sharing of data heavily depends on the experiments performed, data collected and third parties involved. As the first experiments at the Design Centre North Germany are only foreseen in the second round of calls, further details can only be elaborated at a later stage.</p>
Design Centre France	

ONR	<p><b>Type of Data</b></p> <p><u>Data foreseen to be generated by ONR include:</u></p> <p>Numerical models of concept of operation of robots (AltaRica CONOPS model). The concept of operation can be devised by the labs or in answer to an actual business case of external companies.</p> <p>Numerical models of software and hardware architecture of robots (AltaRica system model and MAUVE system architecture). The robots are owned either by the CPSE-Labs or by external companies.</p> <p>Software, which implements the robot function for ONERA robots / external companies robot with a focus on the implementation of: safety functions, decision making functions, real-time execution management.</p> <p>Update of the design tools that have been used to build or analyse the numerical models or the embedded software: safety assessment tools owned by ONERA (e.g. DAL-culator, EPOCH), decision making libraries, MAUVE to OROCOS translator</p> <p>Publishable materials are foreseen to include:</p> <ul style="list-style-type: none"> <li>- AltaRica Libraries</li> <li>- MAUVE Libraries</li> <li>- Decision making libraries</li> <li>- Update of ONERA design tools</li> <li>- Simplified version of models / software developed for ONERA or other company use cases</li> </ul> <p><u>Data foreseen to be generated jointly by third parties and ONR include:</u></p> <p>Simplified models of robots of the external the companies + specification of the software embedded in ONERA robot to mimic company use case</p> <p><u>Data foreseen to be generated by third parties include:</u></p> <p>Specialization of the publishable results for the robots owned by the company</p> <ul style="list-style-type: none"> <li>- Detailed AltaRica / Mauve models</li> <li>- Adaptation of the decision making algorithms</li> </ul> <p>OROCOS modules derived from the detailed Mauve models</p> <p><b>Data sharing</b></p> <p>The data are interesting for different focused communities of end users and are considered to be put in, e.g., <a href="http://www.orocos.org/">http://www.orocos.org/</a> or <a href="http://altarica.labri.fr/wp/">http://altarica.labri.fr/wp/</a> and general platforms that exist to deliver open source software (e.g. <a href="https://www.polarsys.org/">https://www.polarsys.org/</a>).</p> <p>Consent of the third parties participating in the data generation will be required.</p>
LAAS-CNRS	<p><b>Type of Data</b></p> <p>Most of the experimental artefacts will be produced together with the third parties.</p> <p>The experiments are expected to generate (i) open-source software, either as standalone tools, or as libraries/plugins extending other existing (not necessarily open source) toolsets; (ii) experimental artefacts like use case descriptions, exemplary analysis or design models, exemplary analysis results; (iii) Public deliverables synthesizing the</p>

	<p>experiments; (iv) scientific publications. Who generates what depends on the experiments.</p> <p>The experimental artefacts produced together with third parties include:</p> <ul style="list-style-type: none"> <li>- GenoM: GenoM3 templates, exemplary verification results</li> <li>- HAZOP UML : exemplary models and safety analysis results</li> <li>- SMOF : exemplary SMOF models and monitoring strategies generated from the models</li> <li>- MORSE: exemplary test experiments and robustness evaluation results</li> </ul> <p>LAAS-specific data during the project will be:</p> <ul style="list-style-type: none"> <li>- GenoM: updated open-source distribution, updated tutorial, scientific papers;</li> <li>- HAZOP UML: scientific papers and tutorial;</li> <li>- SMOF: updated version and open source distribution by the end of the project, scientific papers, tutorial;</li> <li>- MORSE: distribution of exemplary test components, scientific papers and tutorial for MORSE-based testing;</li> </ul> <p><b>Data sharing</b></p> <p>The definition call topics for experiments by the French CPSE-labs Design Centre will include explicit concerns for delivering publicly available material, e.g. by focusing on extending an open source framework, called GenoM<sup>3</sup>, or by requiring that applied techniques shall be illustrated on artefacts derived from use cases that are representative enough but do not raise IP or confidentiality issues, and that results must be summarized in a publishable experiment description document.</p> <p>Industrial experiment partners may be less used to sharing data. CPSELabs perceives it as its role to accompany them in this opening process. CPSELabs will help them in scoping their use cases and demonstration artefacts, in order to extract information that is sufficient to exemplify the concepts and problems, while not disclosing too much about their systems and know how.</p> <p>Not only the code of the tools/libraries but also tutorials can be made available. For the most mature tools, there also are mailing lists gathering a community of users. A repository of courseware material would be useful to the community.</p>
Design Centre Sweden	
KTH	<p><b>Type of Data</b></p> <p>The collaboration with third parties will probably generate software artefacts, rather than research data.</p> <p>Moreover, interview transcripts and measurement data (KPI-related) will be the most relevant research data generated. Possibly, there will also be summary data related to literature surveys (coded sources, categorizations, etc.)</p> <p><b>Data sharing</b></p> <p>KTH will push for open source software artefacts related to CPS marketplaces.</p> <p>Consent of the third parties participating in the data generation will be required.</p>

<sup>3</sup> <https://git.openrobots.org/projects/genom3>



Design Centre UK	
UNEW	<p><b>Type of Data</b></p> <p>Descriptions of domains, co-models, descriptions of metrics used, tool extension data.</p> <p><b>Data sharing</b></p> <p>Consent of the third parties participating in the data generation will be required.</p>
Design Centre Spain	
IND	<p><b>Type of Data</b></p> <p>Data foreseen to be collected/generated in conjunction with third parties include:</p> <ul style="list-style-type: none"> <li>- Environmental data: Data automatically taken from sensors</li> <li>- Personal data: Data taken from sensor that can identify an individual. Information privacy must be considered for these data.</li> <li>- Generated data: Data not taken from sensors, but inferred from the previous kinds of data using traditional or non-traditional processing application</li> <li>- Test data: as above but generated under lab conditions</li> </ul> <p>Indra's activities are foreseen to generate logs of automatically generated data (such as 1 reading of temperature per every thermometer and minute during the period); information that can be inferred from these (the results of CEP engine using the previous sensors as inputs) and the commands sent by human agents answering this (to provide non-automatic answers or to perform forensic analysis).</p> <p>The data in and by themselves are probably not suitable to be published as such. The results will nevertheless most likely provide opportunities to manually generate material that will be interesting - such as documentation for new functionalities, video tutorials, etc. The details of third-party (including SME's and public institutions) activities about data generation will depend on the specifics of their experiment. We do think their functional interests will lead to more visible and friendlier data.</p> <p>Besides this, IND considers the data generated as potential input for pattern-inference analysis if possible. Moreover, it could provide input to identify potential shortages or improvement points for our technology.</p> <p><b>Data sharing</b></p> <p>Not all the data will be offered to the public <i>as is</i>. Some data must be protected due to legislation and/or to ethical concerns. This includes, but is not restricted to, personal data. Moreover, consent of the third parties generating the data will be required.</p> <p>The collection/generation and sharing of data heavily depends on the experiments performed, data collected and third parties involved. As the first experiments at the Design Centre Spain are only foreseen in the second round of calls, further details can only be elaborated at a later stage.</p>
UPM	<p><b>Type of Data</b></p> <p>The following data is foreseen to be generated in conjunction with the third party experiments</p> <ul style="list-style-type: none"> <li>- Data from IED</li> <li>- Data from humans-CPS interactions</li> <li>- Data from Social networks in the context of CPSs</li> </ul>

	<ul style="list-style-type: none"> <li>- Data from assessing the work performed by 3rd parties</li> <li>- Data from applying changes to the work performed by 3rd parties suggested by conclusions from assessment</li> <li>- Code that could be shared</li> <li>- Models</li> </ul> <p><b>Data sharing</b></p> <p>The implications of releasing data will be checked on a case by case basis. To publish the “raw data”, will, in many cases, not be possible. Moreover, consent of the third parties generating the data will be required. While third parties might object to publish raw data, their consent to publish the research results conclusions, with processed data, might still be possible.</p> <p>The collection/generation and sharing of data heavily depends on the experiments performed, data collected and third parties involved. As the first experiments at the Design Centre Spain are only foreseen in the second round of calls, further details can only be elaborated at a later stage.</p>
--	--

Table 2: Overview of data generated in conjunction with third party experiments

Data sets will be provided by the Design Centres, whenever possible. Data generated together with or by third parties will only be shared upon their consent. The Design Centres will provide the data for the execution open experiments during the course of the project. The Design Centres will also publish the data on open access data platforms to ensure availability of data also after the end of the project. In order to manage these data, partner hosted repositories, as well as external repositories will be used to ensure maximum visibility, serve as backups, and ensure availability well after the end of the project. The CPSE Labs market place will also be considered for providing a repository for some of the data or to have links to specific forges in the market place. As the data need to be easily updated by their producers and to avoid fragmentation of open data platforms some data are considered to be put in the general market place or focused places (e.g. <http://www.orocos.org/> or <http://altarica.labri.fr/wp/>). Moreover, general platforms exist to deliver open source software (e.g. <https://www.polarsys.org/>). The usage of well-known platforms like OpenAIRE (or even opendata.eu in future) would be advantageous in some cases.

With respect to software developed during the course of the project, whenever possible (i.e. not violating IPR) it will also be provided under open-source license to allow for their re-use, adaptation and further enhancement to match possibly different application contexts and serve as a baseline for future business and research endeavours. The project will consider using the Open Access Infrastructure for Research in Europe (OpenAIRE)<sup>4</sup> as well as exploiting the expected support to be provided on research data management for projects funded under Horizon 2020.

<sup>4</sup> <https://www.openaire.eu/>

## 4 Conclusion

The CPSE Labs consortium fully supports the endeavour to improve access to scientific information and research data in order to enhance the benefits of public investment. To fully exploit possibilities of data sharing, the project participates in the 'Open Research Data Pilot' and will make its research data available on a voluntary basis, whenever possible.

A variety of data and information will be generated in CPSE Labs, whereof a part will be generated by the consortium itself, and will be made freely available e.g. via the website and public deliverables, while another part, will be generated in conjunction with 'third parties', participating via cascading funding in so called 'experiments'. In the course of the project, the CPSE Labs consortium will have to carefully consider and agree with the third parties on a case to case basis if, how and to what extent data can be shared. CPSE Labs perceives it as its role to accompany the mainly industrial third parties in this opening process. CPSE Labs will help them in scoping their use cases and demonstration artefacts, in order to extract information that is sufficient to exemplify the concepts and problems, while not disclosing too much about their systems and know how.

As the project is just on the way of concluding the first round of 'open call' selection and invitation process, only assumptions on the data generated through the 'third party experiments' and possible ways of sharing this data could be provided within this deliverable. Future version of this document may provide more refined policies to manage and share such data when the scope and contents of experiments can be assessed more clearly. In addition, the deliverable also aims at giving a brief overview on other data and information elaborated by the project consortium that could be useful for specific stakeholders or other projects pursuing similar aims in future.