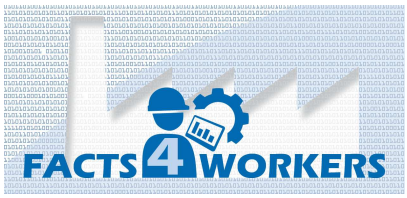

Data Management Plan

Document Type	Deliverable
Document Number	D8.3
Primary Author(s)	Stelios Damalas and Martin Wifling | ViF
Document Version / Status	1.0 | Initial Version
Distribution Level	PU (public)

Project Acronym	FACTS4WORKERS
Project Title	Worker-Centric Workplaces in Smart Factories
Project Website	www.facts4workers.eu
Project Coordinator	Martin Wifling | ViF | martin.wifling@v2c2.at
Grant Agreement Number	636778

This project has received funding from the European Union's **Horizon 2020** research and innovation programme under Grant Agreement n° 636778

CONTRIBUTORS

Name	Organization	Name	Organization
Stelios Damalas	ViF		
Martin Wifling	ViF		
Alexander Stocker	ViF		

REVIEWERS

Name	Organization	Date
Alexander Richter	UZH	2015-06-23

DOCUMENT HISTORY

Revision	Date	Author / Organization	Description
First draft	2015-05-04	Stelios Damalas / ViF	

TABLE OF CONTENTS

1 FACTS4WORKERS: PROJECT DESCRIPTION 4

2 REQUIREMENTS FOR A DATA MANAGEMENT PLAN 5

2.1 Data and Data Management in Horizon 2020 5
2.2 Elements of a Data Management Plan 6
2.3 Data Management and Dissemination Management 7

3 DATA MANAGEMENT IN FACTS4WORKERS 8

3.1 Participation in the Open Data Research Pilot 8
3.2 Data Management Strategy and Template 8
3.3 FACTS4WOERKERS Research Data 9

3.3.1 Dataset 1 – Interviews_Workers_THK 9
3.3.2 Dataset 2 – Interviews_Workers_SCA 10

4 REFERENCES 11

1 FACTS4WORKERS: PROJECT DESCRIPTION

The project FACTS4WORKERS will develop and demonstrate workplace solutions that support the inclusion of increasing elements of knowledge work on the factory floor. These solutions will empower workers on the shop floor with smart factory ICT infrastructure. Advancement will be gained through integrating several building blocks from a flexible smart factory infrastructure, focusing on workers' needs, expectations and requirements, and being supported by organisational measures and change management.

The solutions will be developed according to the following four industrial challenges which are generalizable to manufacturing in general: personalised augmented operator (IC1), worked-centric rich-media knowledge sharing/management (IC2), self-learning manufacturing workplaces (IC3) and in-situ mobile learning in the production (IC4).

Moreover, FACT4WORKER's objectives in terms of measureable indicators are:

- To increase problem-solving and innovation skills of workers;
- To increase cognitive job satisfaction of workers participating in the pilots;
- To increase average worker productivity by 10% for workers participating in pilots;
- To achieve TRL 5-7 on a number of worker-centric solutions through which workers become the smart element in smart factories.

The smart factory demonstrator will be run within the automotive supply chain. The consortium is composed by 15 partners from 7 different EU member states including tier-1, -2 and -3 suppliers, large production enterprises and individual factories that are representative of many medium-sized manufacturing operations, manufacturing SMEs, universities and excellent research institutes.

2 REQUIREMENTS FOR A DATA MANAGEMENT PLAN

2.1 Data and Data Management in Horizon 2020

According to the EC all project proposals submitted to "Research and Innovation actions" and "Innovation actions" have to include a **section on research data management** which is evaluated under the criterion 'Impact'. Projects participating in the **pilot action on open access to research data** have to develop **a data management plan (DMP)** to specify what data will be open.[1]
The DMPs are introduced in the Horizon 2020 Work Programme for 2014-15[2]:

> *"… Horizon 2020 ... use of Data Management Plans (DMPs) detailing what data the project will generate, whether and how it will be exploited or made accessible for verification and re-use, and how it will be curated and preserved. The use of a Data Management Plan is required for projects participating in the Open Research Data Pilot. Other projects are invited to submit a Data Management Plan if relevant for their planned research."*

The **purpose of a DMP** is to provide an analysis of the main elements of the data management policy that will be used by the applicants with regard to all the datasets that will be generated by the project.

Considering privacy and data protection issues **scientific research data** should be easily **discoverable**, **accessible**, **assessable and intelligible**, **useable** beyond the original purpose for which it was collected and **interoperable** to specific quality standards.

What is research data?	*Research data refers to information, in particular facts or numbers, collected to be examined and considered as a basis for reasoning, discussion or calculation.*
What is open research data?	*Openly accessible research data can typically be accessed, mined, exploited, reproduced and disseminated, free of charge for the user.*

Table 1: Research Data and Open Data as defined by the EC[3]

[1] Guidelines on Data Management in Horizon 2020, EC (2013),
http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf

[2] Horizon 2020 Work Programme 2014-2015, EC (2014),
http://ec.europa.eu/research/participants/portal/doc/call/h2020/common/1617601-part_1_introduction_v2.0_en.pdf

[3] Guidelines on Open Access to Scientific Publications and Research Data in Horizon 2020, EC (2013),
https://www.fosteropenscience.eu/project/images/presentations/H2020-open-data-pilot.pdf

2.2 Elements of a Data Management Plan

According to the EC application guideline documents, the DMP to develop within Horizon 2020 projects has to address the aspects below on a dataset by dataset basis and to reflect the current status within the consortium about the data that will be produced:

Data set reference and name	*Identifier for the data set to be produced.*
Data set description	*Description of the data that will be generated or collected, its origin (in case it is collected), nature and scale and to whom it could be useful, and whether it underpins a scientific publication. Information on the existence (or not) of similar data and the possibilities for integration and reuse.*
Standards and metadata	*Reference to existing suitable standards of the discipline. If these do not exist, an outline on how and what metadata will be created.*
Data sharing	*Description of how data will be shared, including access procedures, embargo periods (if any), outlines of technical mechanisms for dissemination and necessary software and other tools for enabling re-use, and definition of whether access will be widely open or restricted to specific groups. Identification of the repository where data will be stored, if already existing and identified, indicating in particular the type of repository (institutional, standard repository for the discipline, etc.).* *In case the dataset cannot be shared, the reasons for this should be mentioned (e.g. ethical, rules of personal data, intellectual property, commercial, privacy-related, security-related).*
Archiving and preservation	*Description of the procedures that will be put in place for long-term preservation of the data. Indication of how long the data should be preserved, what is its approximated end volume, what the associated costs are and how these are planned to be covered.*

Table 2: Minimum Elements of a DMP[4]

Project piloting in the **Open Research Data** activity, have to consider the following aspects[5]:

*Regarding the **digital research data** generated in the action ('data'), the beneficiaries must deposit in a **research data repository** and **take measures** to make it possible for **third parties** to **access, mine, exploit, reproduce and disseminate** — **free of charge** for any user — the following:*

a) the data, including associated metadata, needed to validate the results presented in scientific publications as soon as possible;

[4] Guidelines on Data Management in Horizon 2020, EC (2013), http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf

[5] AGA - Annotated Model Grant Agreement, EC (2015), http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/amga/h2020-amga_en.pdf

b) other data, including associated metadata, as specified and within the deadlines laid down in the data management plan.

Research data repositories are online archives for research data. They can be subject based/thematic, institutional or centralised. Useful listings of research data repositories include the Registry of Research Data Repositories (*www.re3data.org*) and Databib (*http://databib.org*). It is expected that the Open Access Infrastructure for Research in Europe – OpenAIRE (*https://www.openaire.eu*) will become an entry point for linking publications to underlying research data.

Projects have to take measures to **enable** for **third parties** to **access**, mine, exploit, reproduce and disseminate research data, thereby attach Creative Commons Licence (CC-BY or CC0 tool) to the data deposited.[6] More information on Creative Commons Licenses can be found on the respective Web site *creativecommons.org*[7].

A DMP is a living document outlining how research data will be handled during a research project and evolves until the project completion. It gains more precision and substance during the project lifecycle.

2.3 Data Management and Dissemination Management

Both, the **Data Management Plan**, and **the Dissemination Plan**, are key deliverables for the strategy of H2020 projects towards openness of scientific results, i.e. for access to scientific publication and research data in the wider context of dissemination and exploitation.

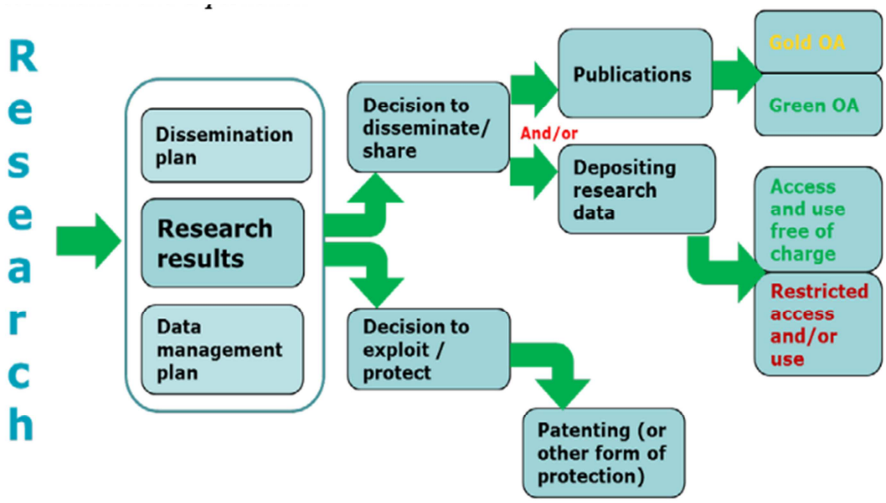

Figure 1: Open Access Strategy in H2020 projects[8]

[6]Guidelines on Open Access to Scientific Publications and Research Data in Horizon 2020, EC (2013), http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-pilot-guide_en.pdf

[7] http://creativecommons.org/about/cc0

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

3 DATA MANAGEMENT IN FACTS4WORKERS

3.1 Participation in the Open Data Research Pilot

As described in the project proposal, FACTS4WORKERS participates in the open data research pilot.

"Apart from publications and patents, project results will mainly be in the form of practical, implicit knowledge gained by partners—and it is this that the partners are most interested in exploiting. Most project data generated will be from the evaluation of our pilots. This data relates to worker job satisfaction, skills and productivity—information that must be considered both private (from the worker's perspective) and company confidential (from the partner's perspective). Therefore, in general we only make anonymised, aggregate data available—under non-disclosure agreement—to parties with a legitimate research interest. **We will create a Data Management Plan (D8.3) for the Open Data Pilot that will address the details of data collection, storage and access.** A special mention must be made of Database Rights on the knowledge data that will be the content of the Social KMS we will prototype and demonstrate. Knowledge data (e.g. on steel processing in the case of THK) must be considered "company confidential" and will not be published—it will be used/exploited internally by the party contributing the subject matter knowledge. Most internal project working documents will be marked "confidential"—these might contain commercially sensitive information e.g. on production line performance or quality issues. These will not be made available to the public and will be subject to confidentiality clauses in the Consortium Agreement."[9]

"Our project takes part in the Open Research Data Pilot and develops a data management plan (M6)**."**[10]

3.2 Data Management Strategy and Template

The project will generate research data by e.g. surveying, interviewing, observing workers, collecting sensor data from in production processes, etc.

FACTS4WORKERS proposes the following strategy regarding openness of research data management.

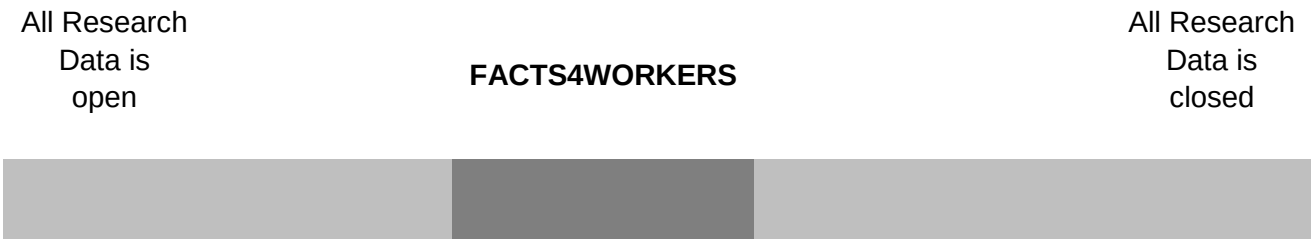

Figure 2: Openness of research data in FACTS4WOERKERS

[9] FACTS4WORERS proposal, pg. 39

[10] FACTS4WORKERS proposal, pg. 62

The project consortium suggests the following **template for data management**:

Project name	What is the project name?
ID	What is the project id?
Project coordinator	Who is coordinating this project?
Project description	A short description of the project
Funder	From whom is this project funded and in which research programme?
Principal researcher	Who is the researcher responsible for this data set?
Data set reference and name	What data are collected and how are they named?
Data set description	Description of the data, the origin, nature and scale and for what purpose they were generated. Are there any similar data?
Standards and metadata	Are there any suitable standards and what metadata will be crated?
Data sharing	Describe if and how the data will be shared (access, procedure, embargo periods, technical mechanisms, necessary software and tools, repositories). In case the data cannot be shared describe the reason.
Archiving and preservation	Where and how long will the data be preserved? What is the approximated end volume, what are the associated costs and how will they be covered?

3.3 FACTS4WOERKERS Research Data

3.3.1 Dataset 1 – Interviews_Workers

Project name	FACTS4WORKERS
ID	636778
Project coordinator	Martin Wifling (ViF)
Project description	Worker-centric workplaces for smart production
Funder	European Commission / Horizon 2020
Principal researcher	**Dr. Alexander Richter (UZH)**

Data set reference and name	41 interviews with workers at 6 industrial partners. Photos, Videos, Data from focus groups.
Data set description	The data set consists of the voice recorded interviews with the workers at the partners participating. This data set will be used to gain insights about how the worker on concrete use cases for smart factories information system support.
Standards and metadata	Anonymized data will be stored using descriptive metadata.
Data sharing	No. The interviews contain a lot of personal data and utterances and will therefore not even be shared with respective collaborating industrial partner. The videos and photos contain confidential information of the industrial partners and will therefore not be shared neither.
Archiving and preservation	At the University of Zurich

3.3.2 Dataset 2

to be added when relevant

4 REFERENCES

[1] Guidelines on Data Management in Horizon 2020, Version 1.0, European Commission, 11. December 2013, http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf

[2] Guidelines on Open Access to Scientific Publications and Research Data in Horizon 2020, Version 1.0, European Commission, 11. December 2013, http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-pilot-guide_en.pdf

[3] AGA - Annotated Model Grant Agreement, Version 2.0, European Commission, 30 March 2015, http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/amga/h2020-amga_en.pdf

[4] Horizon 2020 Work Programme 2014-2015, European Commission Decision C (2014)4995 of 22 July 2014, http://ec.europa.eu/research/participants/portal/doc/call/h2020/common/1617601-part_1_introduction_v2.0_en.pdf