



SATISFACTORY

Project Acronym: **SatisFactory**
Project Full Title: **A collaborative and augmented-enabled ecosystem for increasing satisfaction and working experience in smart factory environments**
Grant Agreement: **636302**
Project Duration: **36 months (01/01/2015 - 31/12/2017)**

DELIVERABLE D6.2

Data Management Plan

Deliverable Status: **Final**
File Name: **SatisFactory-D6.2-DMP-v1.pdf**
Due Date: **June 2015 (M6)**
Submission Date: **June 2015 (M6)**
Task Leader: **SIGMA**

Dissemination level	
Public	X
Confidential, only for members of the Consortium (including the Commission Services)	



This project has received funding from the European Union's Horizon 2020 Research and innovation programme under Grant Agreement n°636302

The SatisFactory project consortium is composed of:		
CERTH ¹	Centre for Research and Technology Hellas	Greece
SIGMA	Sigma Orionis SA	France
FRAUNHOFER	Fraunhofer-Gesellschaft zur Foerderung der Angewandten Forschung E.V	Germany
COMAU	Comau SPA	Italy
EPFL	Ecole Polytechnique Fédérale de Lausanne	Switzerland
ISMB	Istituto Superiore Mario Boella sulle tecnologie dell'informazione e delle telecomunicazioni	Italy
ABE	Atlantis Engineering AE	Greece
REGOLA	Regola srl	Italy
SUNLIGHT	Systems Sunlight Industrial & Commercial Company of Defensive, Energy, Electronic and Telecommunication Systems S.A.	Greece
GlassUP	GlassUp srl	Italy

Disclaimer

This document reflects only the author's views and the European Union is not liable for any use that may be made of the information contained therein.

¹ Project Coordinator

AUTHORS LIST

Leading Author				
First Name	Surname	Beneficiary	Contact email	
Anthony	Dupont	SIGMA	anthony.dupont@sigma-orionis.com	
Co-authors				
#	First Name	Surname	Beneficiary	Contact email
1	Pietro Alberto	Cultrona	COMAU	pietro.cultrona@comau.com
2	Stephanie	Dos Santos	SIGMA	stephanie.dos-santos@sigma-orionis.com
3	Dimosthenis	Ioannidis	CERTH	djoannid@iti.gr
4	Marc	Jentsch	FIT	marc.jentsch@fit.fraunhofer.de
5	Ifigeneia	Metaxa	ATLANTIS	metaxa@abe.gr
6	Symeon	Parcharidis	SUNLIGHT	s.parcharidis@sunlight.gr
7	Apostolos	Perdikakis	EPFL	apostolos.perdikakis@epfl.ch
8	Antonio	Serra	REGOLA	a.serra@regola.it
9	Francesco	Sottile	ISMB	sottile@ismb.it

REVIEWERS LIST

List of Reviewers				
#	First Name	Surname	Beneficiary	Contact email
1	Pietro Alberto	Cultrona	COMAU	pietro.cultrona@comau.com
2	Stelios	Krinidis	CERTH	krinidis@iti.gr

REVISION CONTROL

Version	Author	Date	Status
0.1	SIGMA	February 17, 2015	Initial Draft
0.2	SIGMA	June 3, 2015	Draft
0.3	COMAU, CERTH	June 10, 2015	Quality Check
0.4	SIGMA	June 19, 2015	Revised draft
0.5	SIGMA	June 29, 2015	Second revised draft
1.0	CERTH	June 30, 2015	Submission to the EC

TABLE OF CONTENTS

List of Definitions & Abbreviations	6
Executive Summary	7
1. Introduction	8
2. Dataset List	9
3. General principles	11
Participation in the Pilot on Open Research Data	11
IPR management and Security	11
Personal Data Protection	11
4. Data Management Plan	13
Conclusion	42

LIST OF DEFINITIONS & ABBREVIATIONS

Abbreviation	Definition
API	Application Programming Interface
DMP	Data Management Plan
EC	European Commission
EU	European Union
IEEE	Institute of Electrical and Electronics Engineers
PAKE	Password-authenticated key agreement
RAID	Redundant array of independent disks
RFID	Radio-frequency identification
SOP	Standard operating procedure
SRP	Secure Remote Password
UWB	Ultra-wideband, a radio technology

EXECUTIVE SUMMARY

The present document is a deliverable of the SatisFactory project, funded by the European Commission's Directorate-General for Research and Innovation (DG RTD), under its Horizon 2020 Research and Innovation programme (H2020).

It presents the first version of the project Data Management Plan (DMP). This first version lists, at a preliminary stage, the various datasets that will be produced by the project, the main exploitation perspectives for each of those datasets, and the major management principles the project will implement around them. It paves the way for the specification of a Data Management Portal, to be created by the project in the next period.

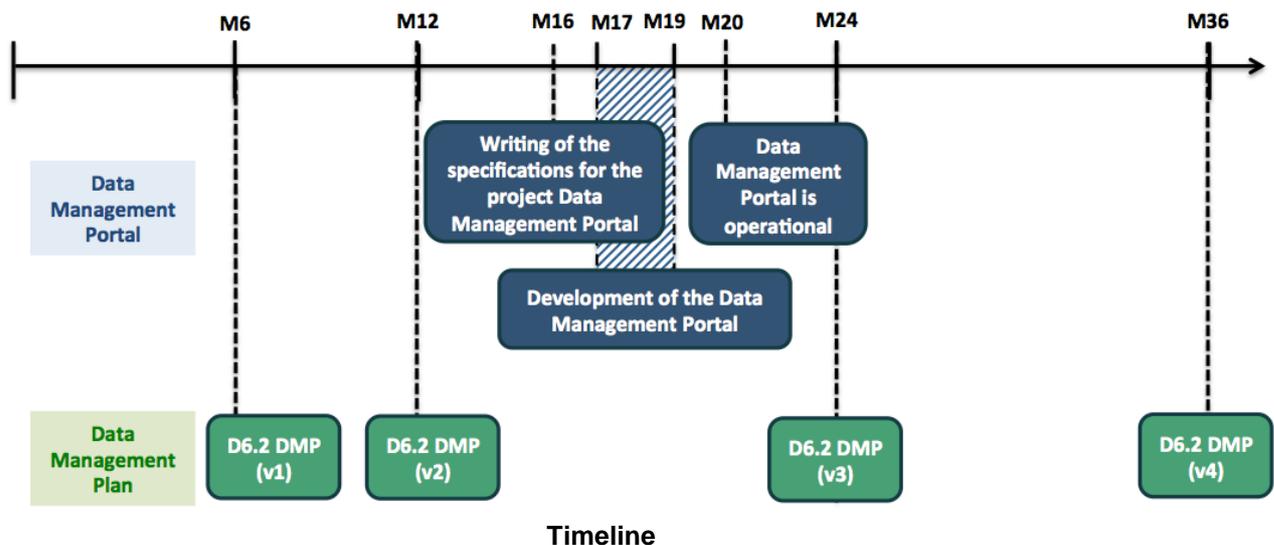
1. INTRODUCTION

The SatisFactory project aims to enhance and enrich the manufacturing working environment towards attractive factories of the future that encompass key enabling technologies such as augmented reality, wearable and ubiquitous computing as well as and customised social communication platforms coupled with experience design and gamification techniques for the efficient transfer of knowledge and experience among employees.

The purpose of the Data Management Plan (DMP) is to provide an analysis of the main elements of the data management policy that will be used by the consortium with regard to all the datasets that will be generated by the project. The DMP is not a fixed document, but will evolve during the lifespan of the project. This first version of the DMP includes an overview of the data sets to be produced by the project, and the specific conditions that are attached to them. The next version of the DMP, to be published at M12, will describe more in detail the practical data management procedures implemented by the SatisFactory project.

The activities related to data management along the SatisFactory project are planned as follows:

- M6: Preliminary analysis and production of the first version of the Data Management Plan (contained in this document);
- M12: Refined analysis based on the progress in the development of the tools and the definition of the case studies, described in the second version of the Data Management Plan;
- M16: Writing of the specifications for the project Data Management Portal;
- M17-M19: Development of the Data Management Portal (to be carried out by CERTH);
- M20: the Data Management Portal is operational;
- M24: Third version of the Data Management Plan, describing actual, proven procedures implemented by the project in the pilot demonstrators, and preparing the sustainability of the data storage after the end of the project;
- M36: Final Data Management Plan, reflecting on the lessons learnt through the project, and describing the plans implemented by SatisFactory for sustainable storage and accessibility of the data.



2. DATASET LIST

All SatisFactory partners have identified the data that will be produced in the different project activities. The list is provided below, while the nature and details for each dataset are given in the subsequent sections.

This list is indicative and allows estimating the data that SatisFactory will produce – it may be adapted (addition/removal of datasets) in the next versions of the DMP to account for the progress of the project activities.

#	Dataset Name
1	DS.CERTH.01.IncidentDetection
2	DS.CERTH.02.ProcessField
3	DS.COMAU.01.Accelerometer_jacket
4	DS.COMAU.01.Gyroscope_jacket
5	DS.COMAU.01.Cardio_jacket
6	DS.COMAU.01.Temp_jacket
7	DS.COMAU.02.RFID_torque_wrench_01
8	DS.COMAU.02.Work_bench_camera
9	DS.COMAU.03.Glasses_Accelerometer
10	DS.COMAU.03.Glasses_Gyroscope
11	DS.COMAU.03.Glasses_Indoor_pos
12	DS.COMAU.03.Glasses_temp
13	DS.COMAU.03.Glasses_camera
14	DS.ISMB.01.IncidentDetection
15	DS.ISMB.01.VirtualFencing
16	DS.ISMB.02.CameraBased_VirtualFencing
17	DS.ISMB.02.UWB_VirtualFencing
18	DS.ATLANTIS.01.FeedbackEngine
19	DS.FIT.01.UserRequirements
20	DS.Regola.01.ARModels
21	DS.Sunlight.01.MotiveBatteriesAssembly

22	DS.Sunlight.02.JarFormationDataAnalysis
23	DS.Sunlight.03.TempMonitoringInJarFormation
24	DS.Sunlight.04.MalfunctionIncidentManagement
25	DS.Sunlight.05.ClimateMonitorAndControl
26	DS.Sunlight.06.WaterConductivityMonitor

3. GENERAL PRINCIPLES

Participation in the Pilot on Open Research Data

SatisFactory participates in the Pilot on Open Research Data launched by the European Commission along with the Horizon2020 programme. The consortium believes firmly in the concepts of open science, and the large potential benefits the European innovation and economy can draw from allowing reusing data at a larger scale. Therefore, all data produced by the project may be published with open access – though this objective will obviously need to be balanced with the other principles described below.

IPR management and Security

As an innovation action close to the market, SatisFactory covers high-TRL technologies and aims at developing marketable solutions. The project consortium includes many partners from the private sector, in particular technology developers (namely Atlantis, GlassUp, and Regola) and end-users (namely COMAU and SUNLIGHT). Those partners obviously have Intellectual Property Rights on their technologies and data, on which their economic sustainability is at stake. Consequently, the SatisFactory consortium will protect that data and crosscheck with the concerned partners before every publication of data.

Another consequence of this aspect is that – with the data collected through SatisFactory being of high value – all measures should be taken to prevent them to leak or being hacked. This is another key aspect of SatisFactory data management, and all data repositories used by the project will include effective protection.

A holistic security approach will be followed, in order to protect the pillars of information security (confidentiality, integrity, availability). The security approach will consist of a methodical assessment of security risks followed by their impact analysis. This analysis will be performed on the personal information and data processed by the proposed system, their flows and any risk associated to their processing.

Security measures will include the implementation of PAKE protocols, such as the SRP protocol, and protection about bots such as CAPTCHA technologies. Moreover, the industrial demo sites apply monitored and controlled procedures related to the data collection, their integrity and protection. The data protection and privacy of personal information will include protective measures against infiltration as well as physical protection of core parts of the systems and access control measures.

Personal Data Protection

SatisFactory's activities will involve human participants, as the pilots will be conducted in real shop floors with actual workers. For some of the activities to be carried out by the project, it may be necessary to collect basic personal data (e.g. name, background, contact details), even though the project will avoid collecting such data unless really necessary. Such data

will be protected in accordance with the EU's *Data Protection Directive 95/46/EC*² “on the protection of individuals with regard to the processing of personal data and on the free movement of such data”. National legislations applicable to the project will also be strictly applied, such as the *Italian Personal Data Protection Code*³. The industrial pilot sites also implement health and safety management standards (BS OHSAS 18001:2007).

All personal data, or data directly related to shop floor workers, to be collected by the project will be done after giving data subjects full details on the experiments to be conducted, and obtaining from them a signed informed consent form.

² <http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:31995L0046&from=en>

³ <http://www.privacy.it/privacycode-en.html>

4. DATA MANAGEMENT PLAN

DS.CERTH.01.IncidentDetection	
Data Identification	
Data set description	Dataset for incident detection, along with high-level activities and business processes monitoring (e.g. activities occurring at the shop-floor, etc.), obtained with thermal and depth cameras mounted at specific locations in the shop floor. The dataset will also include human monitoring and tracking and activity tracking during different affective states, to be used for incident analysis and detection.
Source (e.g. which device?)	The dataset will be collected using thermal and depth cameras located at the areas under interest.
Partners activities and responsibilities	
Partner owner of the device	The device will be owned to the industry (CERTH/CPERI, COMAU, SUNLIGHT), where the data collection is going to be performed.
Partner in charge of the data collection (if different)	CERTH
Partner in charge of the data analysis (if different)	CERTH
Partner in charge of the data storage (if different)	CERTH
WPs and tasks	The data are going to be collected within activities of WP3 and more specifically within activities of T3.3.
Standards	
Info about metadata (Production and storage dates, places) and documentation?	The dataset will be accompanied with a detailed documentation of its contents. Indicative metadata include: (a) description of the experimental setup (e.g. location, date, etc.) and procedure that led to the generation of the dataset, (b) annotated incident, activity, business process, state of the monitored activity and the involved humans per time interval, etc.
Standards, Format, Estimated volume of data	The data will be stored at XML format and are estimated to be 80 GB per camera per day.
Data exploitation and sharing	

Data exploitation (purpose/use of the data analysis)	<p>The collected data will be used for the development of the activities analysis and incident detection methods of the SatisFactory project and all the tasks, activities and methods that are related to it.</p> <p>Furthermore, the different parts of the dataset could be useful in the benchmarking of a series of human detection and tracking methods, activity detection focusing either on pose and gestures analysis and tracking, on high-level activity recognition, on affect related human activity analysis and on incident analysis and detection.</p>
Data access policy / Dissemination level (Confidential, only for members of the Consortium and the Commission Services) / Public	<p>The full dataset will be confidential and only the members of the consortium will have access on it.</p> <p>Furthermore, if the dataset or specific portions of it (e.g. metadata, statistics, etc.) are decided to become of widely open access, a data management portal will be created that should provide a description of the dataset and link to a download section. Of course these data will be anonymized, so as not to have any potential ethical issues with their publication and dissemination.</p>
Data sharing, re-use and distribution (How?)	<p>The full dataset will be shared using a data management portal that is going to be created and maintained by CERTH. The public version of the data will be shared within the portal as well. Of course, the data management portal will be equipped with authentication mechanisms, so as to handle the identity of the persons/organizations that download them, as well as the purpose and the use of the downloaded dataset.</p>
Embargo periods (if any)	None
Archiving and preservation (including storage and backup)	
Data storage (including backup): where? For how long?	<p>Both full and public versions of the dataset will be accommodated at the data management portal created and maintained by CERTH, while links to the portal will exist at the SatisFactory website. Furthermore, in order to avoid data loses, RAID and other common backup mechanism will be utilized ensuring data reliability and performance improvement.</p>

DS.CERTH.02.ProcessField	
Data Identification	
Data set description	Dataset for shop-floor information related to the status and condition of the involved machinery and production-process system (e.g. field data related to the status of a device (pump, motor) that workers interact at the shop-floor, etc.). The dataset will also include the human actions and the logging of commands and activity during different conditions and states, to be used for the decision support system and procedures.
Source (e.g. which device?)	The dataset will be collected through the automation systems that acquire the signals from the respective field network of interest. The device managers will communicate with the automation systems in order to transfer the selected data to the Satisfactory repository.
Partners activities and responsibilities	
Partner owner of the device	The device will be owned by the CERTH/CPERI, where the data collection is going to be performed.
Partner in charge of the data collection (if different)	CERTH
Partner in charge of the data analysis (if different)	CERTH
Partner in charge of the data storage (if different)	CERTH
WPs and tasks	Initially the static and persistent data along with a set of dynamic data will be collected within activities of WP3 and more specifically within activities of T3.3 and T3.5. The dynamic data will be updated during WP5 and more specifically at T5.3.
Standards	
Info about metadata (Production and storage dates, places) and documentation?	The dataset will be accompanied with the respective documentation of its contents. Indicative metadata include: (a) description of the experimental setup (e.g. process system, date, etc.) and procedure which is related to the dataset (e.g. proactive maintenance action, unplanned event, nominal operation. etc.), (b) scenario related procedures, state of the monitored activity and involved workers, involved system etc.
Standards, Format, Estimated volume of data	The data will be stored at XML format and are estimated to be 200-1000 MB per month.

Data exploitation and sharing	
Data exploitation (purpose/use of the data analysis)	<p>The raw field layer data will be enriched with semantics by the middleware components and will be used for the sharing of information between the process system and the involved actors at the shop floor.</p> <p>The collected data will be used by the integrated decision support system and the event manager in order to represent the input for the operation and the maintenance procedures along with the identification of unplanned incidents at the shop floor and to analyse the response and behaviour of the workers through their interaction with the Satisfactory platform.</p> <p>The initial set of dynamic data will be analysed during the development of the Satisfactory platform and the nominal dynamic data will be used during the deployment phase at the shop floor.</p>
Data access policy / Dissemination level (Confidential, only for members of the Consortium and the Commission Services) / Public	<p>The full dataset will be confidential and only the members of the consortium will have access on it.</p> <p>Furthermore, if specific portions of it (e.g. metadata, statistics, etc.) are decided to become of widely open access, a data management portal will be created that should provide a description of the dataset and link to a download section. Of course these data will be anonymized, so as not to have any potential correlation and identification of the ethical issues with their publication and dissemination.</p>
Data sharing, re-use and distribution (How?)	<p>The created dataset will be shared using a data management portal that is going to be created and maintained by CERTH. The public version of the data will be shared within the portal as well. Of course, the data management portal will be equipped with authentication mechanisms, so as to handle the identity of the persons/organizations that download them, as well as the purpose and the use of the downloaded dataset.</p>
Embargo periods (if any)	None
Archiving and preservation (including storage and backup)	
Data storage (including backup): where? For how long?	<p>Both full and public versions of the dataset will be accommodated at the data management portal created and maintained by CERTH, while links to the portal will exist at the SatisFactory website. Furthermore, in order to avoid data losses, RAID and other common backup mechanism will be utilized ensuring data reliability and performance improvement.</p>

DS.COMAU.01.Jacket	
Data Identification	
Data set description	Data for ergonomic parameters monitoring, coming from sensors installed on a jacket worn by the operators. This dataset will comprise several distinct sub-datasets corresponding to each type of sensor, in order to simplify its maintainability. Those sub-datasets, named DSCOMAU.01.Jacket_accelerometer, DSCOMAU.01.Jacket_gyroscope, etc., will follow similar procedures but will be managed independently.
Source (e.g. which device?)	The dataset will be collected by different sensors installed on jackets worn by operators, namely: an accelerometer, a gyroscope, and a temperature sensor.
Partners activities and responsibilities	
Partner owner of the device	ISMB
Partner in charge of the data collection (if different)	COMAU with ISMB
Partner in charge of the data analysis (if different)	ISMB
Partner in charge of the data storage (if different)	ISMB
WPs and tasks	The data are going to be collected within activities of WP3 and WP4
Standards	
Info about metadata (Production and storage dates, places) and documentation?	Indicative metadata include worker's posture, e.g., trunk bending forward/backward and time stamp.
Standards, Format, Estimated volume of data	The data format has not been defined yet. Approximately, the estimated volume of data is less than 2MB per day per worker.
Data exploitation and sharing	
Data exploitation (purpose/use of the data analysis)	Benchmarking of a series of human detection and tracking methods, activity detection focusing either on pose and gestures analysis and tracking

Data access policy / Dissemination level (Confidential, only for members of the Consortium and the Commission Services) / Public	The full dataset will be confidential and only the members of the consortium will have access on it. Furthermore, if the dataset or specific portions of it (e.g. metadata, statistics, etc.) are decided to become of widely open access, a data management portal will be created that should provide a description of the dataset and link to a download section. Of course these data will be anonymized, so as not to have any potential ethical issues with their publication and dissemination.
Data sharing, re-use and distribution (How?)	The created dataset could be shared by using open APIs through the middleware.
Embargo periods (if any)	None
Archiving and preservation (including storage and backup)	
Data storage (including backup): where? For how long?	All information belongs to the industrial partner that owns the shop floor. All data will respect the partner policies.

DS.COMAU.02.RFID_torque_wrench	
Data Identification	
Data set description	Sensors installed on the workbench where the operator normally works
Source (e.g. which device?)	RFID installed on the torque wrenches used by the operator.
Partners activities and responsibilities	
Partner owner of the device	REGOLA
Partner in charge of the data collection (if different)	COMAU with REGOLA
Partner in charge of the data analysis (if different)	REGOLA
Partner in charge of the data storage (if different)	REGOLA
WPs and tasks	The data are going to be collected within activities of WP3 and WP4
Standards	
Info about metadata (Production and storage dates, places) and documentation?	Metadata is yet to be defined by the solution provider.
Standards, Format, Estimated volume of data	The data format has not been defined yet.
Data exploitation and sharing	
Data exploitation (purpose/use of the data analysis)	Production process recognition and help during the different production phases, avoiding mistakes

Data access policy / Dissemination level (Confidential, only for members of the Consortium and the Commission Services) / Public	The full dataset will be confidential and only the members of the consortium will have access on it. Furthermore, if the dataset or specific portions of it (e.g. metadata, statistics, etc.) are decided to become of widely open access, a data management portal will be created that should provide a description of the dataset and link to a download section. Of course these data will be anonymized, so as not to have any potential ethical issues with their publication and dissemination.
Data sharing, re-use and distribution (How?)	The sharing of this data is yet to be decided in accordance to COMAU policies and other partners' requirements.
Embargo periods (if any)	None
Archiving and preservation (including storage and backup)	
Data storage (including backup): where? For how long?	All information belongs to the industrial partner that owns the shop floor. All data will respect the partner policies.

DS.COMAU.03.Work_bench_camera	
Data Identification	
Data set description	Sensors installed on the workbench where the operator normally works
Source (e.g. which device?)	Camera installed on the workbench where the operator works. The type of camera will be defined at a later stage based on the use-cases to be developed.
Partners activities and responsibilities	
Partner owner of the device	REGOLA
Partner in charge of the data collection (if different)	COMAU with REGOLA
Partner in charge of the data analysis (if different)	REGOLA
Partner in charge of the data storage (if different)	REGOLA
WPs and tasks	The data are going to be collected within activities of WP3 and WP4
Standards	
Info about metadata (Production and storage dates, places) and documentation?	Metadata is yet to be defined by the solution provider.
Standards, Format, Estimated volume of data	The data format has not been defined yet.
Data exploitation and sharing	
Data exploitation (purpose/use of the data analysis)	Production process recognition and help during the different production phases, avoiding mistakes

Data access policy / Dissemination level (Confidential, only for members of the Consortium and the Commission Services) / Public	The full dataset will be confidential and only the members of the consortium will have access on it. Furthermore, if the dataset or specific portions of it (e.g. metadata, statistics, etc.) are decided to become of widely open access, a data management portal will be created that should provide a description of the dataset and link to a download section. Of course these data will be anonymized, so as not to have any potential ethical issues with their publication and dissemination.
Data sharing, re-use and distribution (How?)	The sharing of this data is yet to be decided in accordance to COMAU policies and other partners' requirements.
Embargo periods (if any)	None
Archiving and preservation (including storage and backup)	
Data storage (including backup): where? For how long?	All information belongs to the industrial partner that owns the shop floor. All data will respect the partner policies.

DS.COMAU.04.Glasses	
Data Identification	
Data set description	Sensors installed on the glasses developed by GlassUp
Source (e.g. which device?)	The dataset will be collected by different sensors installed on the GlassUp glasses worn by operators, namely: an accelerometer, a gyroscope, an indoor localisation sensor, a temperature sensor, and a camera.
Partners activities and responsibilities	
Partner owner of the device	GLASSUP
Partner in charge of the data collection (if different)	COMAU with GLASSUP
Partner in charge of the data analysis (if different)	GLASSUP
Partner in charge of the data storage (if different)	GLASSUP
WPs and tasks	The data are going to be collected within activities of WP3 and WP4
Standards	
Info about metadata (Production and storage dates, places) and documentation?	Metadata is yet to be defined by the solution provider.
Standards, Format, Estimated volume of data	The data format has not been defined yet.
Data exploitation and sharing	
Data exploitation (purpose/use of the data analysis)	Benchmarking of a series of human detection and tracking methods, activity detection focusing either on pose and gestures analysis and tracking + for the cameras: Support for contextual data related on the machine/devices looked, remote support for maintenance

Data access policy / Dissemination level (Confidential, only for members of the Consortium and the Commission Services) / Public	The full dataset will be confidential and only the members of the consortium will have access on it. Furthermore, if the dataset or specific portions of it (e.g. metadata, statistics, etc.) are decided to become of widely open access, a data management portal will be created that should provide a description of the dataset and link to a download section. Of course these data will be anonymized, so as not to have any potential ethical issues with their publication and dissemination.
Data sharing, re-use and distribution (How?)	The sharing of this data is yet to be decided in accordance to COMAU policies and other partners' requirements.
Embargo periods (if any)	None
Archiving and preservation (including storage and backup)	
Data storage (including backup): where? For how long?	All information belongs to the industrial partner that owns the shop floor. All data will respect the partner policies.

DS.ISMB.01.IncidentDetection	
Data Identification	
Data set description	Data from cameras are processed to detect specific gestures in case of emergency, in order to signal an alert. A fall prevention system is also in scope. The identification of the most suitable type of cameras is yet to be done in the scope of WP3.
Source (e.g. which device?)	Gesture & content recognition manager component.
Partners activities and responsibilities	
Partner owner of the device	Stored data and devices will be owned by the industrial partner.
Partner in charge of the data collection (if different)	
Partner in charge of the data analysis (if different)	
Partner in charge of the data storage (if different)	
WPs and tasks	WP3: T3.3
Standards	
Info about metadata (Production and storage dates, places) and documentation?	Either XLM or JSON format will be used.
Standards, Format, Estimated volume of data	This information cannot be predicted in advance of a real use of the technology at the shop floor level.
Data exploitation and sharing	
Data exploitation (purpose/use of the data analysis)	Stored data can be used to trigger specific alerts level based on the incident detection system.
Data access policy / Dissemination level (Confidential, only for members of the Consortium and the Commission Services) / Public	All information belongs to the industrial partner that owns the shop floor. All data will respect the partner policies.

Data sharing, re-use and distribution (How?)	Specific triggered will provide specific alerts and the actuators for these alerts are dependent on the re-adaptation of existing facilities.
Embargo periods (if any)	None
Archiving and preservation (including storage and backup)	
Data storage (including backup): where? For how long?	All information belongs to the industrial partner that owns the shop floor. All data will respect the partner policies.

DS.ISMB.02.VirtualFencing	
Data Identification	
Data set description	Data from the cameras will be processed and triggered events will be propagated. An historical database of events that have been triggered will be stored and processed. The identification of the most suitable type of cameras is in the scope of WP3.
Source (e.g. which device?)	Multimedia Manager and Gesture & content recognition manager components
Partners activities and responsibilities	
Partner owner of the device	Stored data and devices will be owned by the industrial partner.
Partner in charge of the data collection (if different)	
Partner in charge of the data analysis (if different)	
Partner in charge of the data storage (if different)	
WPs and tasks	WP3: T3.3
Standards	
Info about metadata (Production and storage dates, places) and documentation?	Either XLM or JSON format will be used.
Standards, Format, Estimated volume of data	This information cannot be predicted in advance of a real use of the technology at the shop floor level.
Data exploitation and sharing	
Data exploitation (purpose/use of the data analysis)	Stored data can be used to feed management toolkits, incident frequency benchmarking.
Data access policy / Dissemination level (Confidential, only for members of the Consortium and the Commission Services) / Public	All information belongs to the industrial partner that owns the shop floor. All data will respect the partner policies.
Data sharing, re-use and distribution (How?)	Data management tool will visualize processed information.
Embargo periods (if any)	None
Archiving and preservation (including storage and backup)	
Data storage (including backup): where? For how long?	All information belongs to the industrial partner that owns the shop floor. All data will respect the partner policies.

DS.ISMB.03.CameraBased_VirtualFencing	
Data Identification	
Data set description	Data from cameras are going to be processed in order to analyse if the personnel in the area wears the required safety equipment. The identification of the most suitable type of cameras is in the scope of WP3.
Source (e.g. which device?)	Multimedia Manager component.
Partners activities and responsibilities	
Partner owner of the device	Stored data and devices will be owned by the industrial partner.
Partner in charge of the data collection (if different)	
Partner in charge of the data analysis (if different)	
Partner in charge of the data storage (if different)	
WPs and tasks	WP3: T3.3
Standards	
Info about metadata (Production and storage dates, places) and documentation?	Either XLM or JSON format will be used.
Standards, Format, Estimated volume of data	This information cannot be predicted in advance of a real use of the technology at the shop floor level.
Data exploitation and sharing	
Data exploitation (purpose/use of the data analysis)	Stored data can be use for statistical evaluation and analysis of the safety equipment.
Data access policy / Dissemination level (Confidential, only for members of the Consortium and the Commission Services) / Public	All information belongs to the industrial partner that owns the shop floor. All data will respect the partner policies.
Data sharing, re-use and distribution (How?)	Alert system will be implemented accordingly with the use cases and requirements of the industrial partners.
Embargo periods (if any)	None
Archiving and preservation (including storage and backup)	
Data storage (including backup): where? For how long?	All information belongs to the industrial partner that owns the shop floor. All data will respect the partner policies.

DS.ISMB.04.UWB_VirtualFencing	
Data Identification	
Data set description	Dataset reporting potential incidents based on a virtual fencing approach (e.g., a worker is approaching a dangerous area, the worker is out of the dangerous area). Incident events are obtained by the Localization Manager by processing position data coming from UWB devices
Source (e.g. which device?)	The raw localization data will be collected by the UWB-tags carried out by the workers. In particular, the UWB-tag continuously performs ranging (i.e. distance) measurements from fixed anchors (named anchors) and estimates its position running a localization algorithm on the tag itself.
Partners activities and responsibilities	
Partner owner of the device	The UWB devices will be owned by ISMB. In particular, ISMB is going to develop wearable UWB tags based on the IEEE 802.15.4a standard.
Partner in charge of the data collection (if different)	ISMB
Partner in charge of the data analysis (if different)	ISMB
Partner in charge of the data storage (if different)	ISMB
WPs and tasks	The data are going to be collected within activities of WP3 and more specifically within activities of T3.3.
Standards	
Info about metadata (Production and storage dates, places) and documentation?	Indicative metadata include: state of the worker (e.g., inside or outside a dangerous area), name of the dangerous area, relative and absolute coordinates, and time stamp.
Standards, Format, Estimated volume of data	The geofencing events as well as localization data will be stored in a relational DB inside the Localization Manager component. These data can be accessed through REST APIs. Approximately, at most 5MB of events could be stored per day per worker.
Data exploitation and sharing	
Data exploitation (purpose/use of the data analysis)	The collected data will be used for the development of the proactive incident detection functionalities of the SatisFactory platform. Furthermore, localization data can be exploited by the other components of SatisFactory such as Augmented Reality and Collaboration Tools.

Data access policy / Dissemination level (Confidential, only for members of the Consortium and the Commission Services) / Public	The full dataset will be confidential and only the members of the consortium will have access on it. Furthermore, if the dataset or specific portions of it (e.g. metadata, statistics, etc.) are decided to become of widely open access, these could be shared through the Localization Manager. Of course these data will be anonymized, so as not to have any potential ethical issues with their publication and dissemination.
Data sharing, re-use and distribution (How?)	The UWB VirtualFencing dataset could be shared using the REST APIs of the Localization Manager component for debugging, demonstration and management purposes. The public version of the data could be shared through the Localization Manager by using a suitable authentication mechanism so as to handle the identity of the persons/organizations that access them.
Embargo periods (if any)	None
Archiving and preservation (including storage and backup)	
Data storage (including backup): where? For how long?	Data will be stored in a relational DB of the Localization Manager component that will be developed by ISMB. Typically, the Localization Manager server will be installed in the end-users' premises (i.e., COMAU and Sunlight) for the industrial pilot demonstrators.

DS.ATLANTIS.01.FeedbackEngine	
Data Identification	
Data set description	Dataset for shop floor feedback engine obtained by Smart Sensor Network and Thermal and Optical Sensors, made available through Device Manager, together with data stemming from Incident Management Tools.
Source (e.g. which device?)	Device Manager, Incident Management Tools
Partners activities and responsibilities	
Partner owner of the device	The device will be owned by the industry (COMAU, CERTH/CPERI, SUNLIGHT), where the data collection is going to be performed.
Partner in charge of the data collection (if different)	
Partner in charge of the data analysis (if different)	
Partner in charge of the data storage (if different)	
WPs and tasks	The data are going to be collected within activities of WP3 and WP4
Standards	
Info about metadata (Production and storage dates, places) and documentation?	Metadata is yet to be defined by the solution provider.
Standards, Format, Estimated volume of data	Data must be available at XML or JSON format. Estimation of the volume of data can be done only by the source
Data exploitation and sharing	
Data exploitation (purpose/use of the data analysis)	The collected data will be used for better understanding of the processes and activities evolving in the shop floor which in conjunction with pre-defined response policies and strategies will provide a actionable knowledge in the form of a set of recommendations regarding both maintenance and manufacturing operations.
Data access policy / Dissemination level (Confidential, only for members of the Consortium and the Commission Services) / Public	All information belongs to the industrial partner that owns the shop floor. All data will respect the partner policies.
Data sharing, re-use and distribution (How?)	The sharing of this data is yet to be decided along with the industrial partners.
Embargo periods (if any)	None
Archiving and preservation (including storage and backup)	
Data storage (including backup): where? For how long?	All information belongs to the industrial partner that owns the shop floor. All data will respect the partner policies.

DS.FIT.01.UserRequirements	
Data Identification	
Data set description	Data that is collected in user workshops with the goal of understanding the shop floor workers work environment and their needs
Source (e.g. which device?)	Semi-structured interview and other questioning techniques in user workshops
Partners activities and responsibilities	
Partner owner of the device	Data not collected by device
Partner in charge of the data collection (if different)	The interviews will be conducted by COMAU, SUNLIGHT, CERTH/CPERI and FIT
Partner in charge of the data analysis (if different)	FIT
Partner in charge of the data storage (if different)	FIT
WPs and tasks	The requirements engineering is focus of WP1
Standards	
Info about metadata (Production and storage dates, places) and documentation?	The data collection process will be described as well as there will be minutes of the workshops
Standards, Format, Estimated volume of data	Semi-structured interviews, Questionnaires, Shadowing, Think Aloud Prototypes, Velcro Modelling
Data exploitation and sharing	
Data exploitation (purpose/use of the data analysis)	The collected data builds the foundation for all activities in the project. The analysis will determine what the SatisFactory shall achieve and thus it will determine the actions for all WPs
Data access policy / Dissemination level (Confidential, only for members of the Consortium and the Commission Services) / Public	The full dataset will be confidential and only the members of the consortium will have access on it. The results of the analysis will be accessible in the public deliverables D1.1 and partly D1.2. For this, all data will be anonymized
Data sharing, re-use and distribution (How?)	The dataset will be stored in a restricted folder of the BSCW and only shared with the partners. It cannot be made available due to confidentiality agreements with the interviewees themselves.
Embargo periods (if any)	Forever
Archiving and preservation (including storage and backup)	
Data storage (including backup): where? For how long?	Included in the normal BSCW backup strategy

DS.Regola.01.ARModels	
Data Identification	
Data set description	Dataset containing information captured at low-level and representing human activities in the shop floor context; graphic model describing objects processed during the activities; video and audio recorded during actions (e.g. activities occurring at the shop-floor, etc.) obtained with AR cameras mounted on specific wearable devices; executable script containing step-by-step instructions, dynamic help visualized through the glasses, etc.
Source (e.g. which device?)	The dataset will be collected using cameras integrated in the wearable device and cameras located at the areas under interest. The recordings will be colour and HR.
Partners activities and responsibilities	
Partner owner of the device	The device will be owned by the industry (COMAU, CERTH/CPERI, SUNLIGHT), where the data collection is going to be performed.
Partner in charge of the data collection (if different)	Regola
Partner in charge of the data analysis (if different)	Regola
Partner in charge of the data storage (if different)	Regola
WPs and tasks	The data are going to be collected within activities of T2.5 and of T4.3.
Standards	
Info about metadata (Production and storage dates, places) and documentation?	The dataset will be accompanied with a detailed documentation of its contents. Indicative metadata include: (a) description of the working phase (e.g. location, date, etc.) SOP and procedure that led to the generation of the dataset, (b) description of involved objects

Standards, Format, Estimated volume of data	3D Formats: 3DM, 3DS, DXF, DWG, IGES, Collada DAE, FBX, OBJ, PLY, ASC, RAW, SKP, SLDPR2, STP, STEP, STL, WRL, VRML, SGF e SGP (proprietary scenegraph file formats). Image Formats: BMP, DIB, JPG, TGA, PNG, DDS, HDR Audio Formats: WAV, MID, MP3 Video Formats: AVP, MPG Motion Capture Formats: BVH, C3D, HTR, GTR Original SOP Formats: PDF, DOCX, XLS, XLSX, etc. R3D RT SOP Formats: RTS (proprietary XML-Based file format). The data will be stored at XML format and are estimated to be 40 GB per day.
Data exploitation and sharing	
Data exploitation (purpose/use of the data analysis)	The collected data will be used for analysis of operator's behaviour and the development of the scripts containing step-by-step instruction and the controls of correctness of activities
Data access policy / Dissemination level (Confidential, only for members of the Consortium and the Commission Services) / Public	The full dataset will be confidential and only the members of the consortium will have access on it. Furthermore, if the dataset or specific portions of it (e.g. metadata, statistics, etc.) are decided to become of widely open access, a data management portal will be created that should provide a description of the dataset and link to a download section. Of course these data will be anonymized, so as not to have any potential ethical issues with their publication and dissemination.
Data sharing, re-use and distribution (How?)	The full dataset will be shared using a data management portal. The data management portal will be equipped with authentication mechanisms, so as to handle the identity of the persons/organizations that download them, as well as the purpose and the use of the downloaded dataset.
Embargo periods (if any)	None
Archiving and preservation (including storage and backup)	
Data storage (including backup): where? For how long?	Both full and public versions of the dataset will be accommodated at the data management portal; RAID and other common backup mechanism will be utilized ensuring data reliability and performance improvement.

DS.Sunlight.01.MotiveBatteriesAssembly	
Data Identification	
Data set description	Technical data for battery assembly and working instructions for assembly and quality control.
Source (e.g. which device?)	Technical data for battery assembly will be provided from SAP. Working instructions will be available from a database, which will be accessed through the internal company network.
Partners activities and responsibilities	
Partner owner of the device	Sunlight will be the owner of the device.
Partner in charge of the data collection (if different)	
Partner in charge of the data analysis (if different)	
Partner in charge of the data storage (if different)	
WPs and tasks	Data will be collected for WP3, T3.4
Standards	
Info about metadata (Production and storage dates, places) and documentation?	Metadata that will be used are SAP codes, order numbers and drawing numbers
Standards, Format, Estimated volume of data	The data will be in text format including drawings and photos. The estimated total volume will not exceed 1TB
Data exploitation and sharing	
Data exploitation (purpose/use of the data analysis)	The data will be used only for the development of the Satisfactory application.
Data access policy / Dissemination level (Confidential, only for members of the Consortium and the Commission Services) / Public	Data will be confidential. Data that cannot be shared because it includes Customer order details and technical know-how details.
Data sharing, re-use and distribution (How?)	Data can be not shared
Embargo periods (if any)	None
Archiving and preservation (including storage and backup)	
Data storage (including backup): where? For how long?	Data are stored in SAP but an intermediate storage unit will be used in order to avoid data loses and provide a data backup

DS.Sunlight.02.JarFormationDataAnalysis	
Data Identification	
Data set description	Battery data (e.g. voltage, current, temperature, etc.) collected during battery formation
Source (e.g. which device?)	Data are stored in the Jar formation SCADA system
Partners activities and responsibilities	
Partner owner of the device	Sunlight will be the owner of the device.
Partner in charge of the data collection (if different)	
Partner in charge of the data analysis (if different)	
Partner in charge of the data storage (if different)	
WPs and tasks	Data will be collected for WP3, T3.5
Standards	
Info about metadata (Production and storage dates, places) and documentation?	Metadata that will be used are production dates and the Jar formation equipment code number.
Standards, Format, Estimated volume of data	The data will be in text format
Data exploitation and sharing	
Data exploitation (purpose/use of the data analysis)	The data will be used only for the development of the Satisfactory application.
Data access policy / Dissemination level (Confidential, only for members of the Consortium and the Commission Services) / Public	Data will be confidential because it includes technical know-how details.
Data sharing, re-use and distribution (How?)	Data will be shared by using a data management portal.
Embargo periods (if any)	None
Archiving and preservation (including storage and backup)	
Data storage (including backup): where? For how long?	Data are stored in the storage device of the SCADA system (computer). A back up will be stored in an external storage device.

DS.Sunlight.03.TempMonitoringInJarFormation	
Data Identification	
Data set description	Battery Temperature measurements
Source (e.g. which device?)	Thermal cameras
Partners activities and responsibilities	
Partner owner of the device	Sunlight will be the owner of the device.
Partner in charge of the data collection (if different)	
Partner in charge of the data analysis (if different)	
Partner in charge of the data storage (if different)	
WPs and tasks	Data will be collected for WP3, T3.3
Standards	
Info about metadata (Production and storage dates, places) and documentation?	Metadata that will be used are production dates and the Jar formation equipment code number.
Standards, Format, Estimated volume of data	The data will be in text format
Data exploitation and sharing	
Data exploitation (purpose/use of the data analysis)	The data will be used only for the development of the Satisfactory application.
Data access policy / Dissemination level (Confidential, only for members of the Consortium and the Commission Services) / Public	Data will be available only for members of the Consortium and the Commission Services
Data sharing, re-use and distribution (How?)	Data will be shared by using a data management portal.
Embargo periods (if any)	None
Archiving and preservation (including storage and backup)	
Data storage (including backup): where? For how long?	Data will be stored in the storage device of the developed system (computer). A back up will be stored in an external storage device.

DS.Sunlight.04.MalfunctionIncidentManagement	
Data Identification	
Data set description	Malfunction Incidents
Source (e.g. which device?)	Malfunction Incidents are logged manually in .XLS file
Partners activities and responsibilities	
Partner owner of the device	Sunlight will be the owner of the device.
Partner in charge of the data collection (if different)	
Partner in charge of the data analysis (if different)	
Partner in charge of the data storage (if different)	
WPs and tasks	Data will be collected for WP3, T3.5
Standards	
Info about metadata (Production and storage dates, places) and documentation?	Metadata that will be used are the incident details (date, hour, place, etc.)
Standards, Format, Estimated volume of data	The data will be in text format
Data exploitation and sharing	
Data exploitation (purpose/use of the data analysis)	The data will be used only for the development of the Satisfactory application.
Data access policy / Dissemination level (Confidential, only for members of the Consortium and the Commission Services) / Public	Data will be available only for members of the Consortium and the Commission Services
Data sharing, re-use and distribution (How?)	Data will be shared by using a data management portal.
Embargo periods (if any)	None
Archiving and preservation (including storage and backup)	

Data storage (including backup): where? For how long?	Data will be stored in the storage device of the developed system (computer). A back up will be stored in an external storage device.
---	---

DS.Sunlight.05.ClimateMonitorAndControl	
Data Identification	
Data set description	Ambient temperature and humidity measurements from the curing chambers
Source (e.g. which device?)	Temperature and humidity sensors
Partners activities and responsibilities	
Partner owner of the device	Sunlight will be the owner of the device.
Partner in charge of the data collection (if different)	
Partner in charge of the data analysis (if different)	
Partner in charge of the data storage (if different)	
WPs and tasks	Data will be collected for WP3, T3.3
Standards	
Info about metadata (Production and storage dates, places) and documentation?	Metadata that will be used are the chamber code numbers
Standards, Format, Estimated volume of data	The data will be in text format
Data exploitation and sharing	
Data exploitation (purpose/use of the data analysis)	The data will be used only for the development of the Satisfactory application.
Data access policy / Dissemination level (Confidential, only for members of the Consortium and the Commission Services) / Public	Data will be available only for members of the Consortium and the Commission Services
Data sharing, re-use and distribution (How?)	Data will be shared by using a data management portal.
Embargo periods (if any)	None
Archiving and preservation (including storage and backup)	

Data storage (including backup): where? For how long?	Data will be stored in the storage device of the developed system (computer). A back up will be stored in an external storage device.
---	---

DS.Sunlight.06.WaterConductivityMonitor	
Data Identification	
Data set description	De-ionized water conductivity measurements
Source (e.g. which device?)	Water conductivity sensor
Partners activities and responsibilities	
Partner owner of the device	Sunlight will be the owner of the device.
Partner in charge of the data collection (if different)	
Partner in charge of the data analysis (if different)	
Partner in charge of the data storage (if different)	
WPs and tasks	Data will be collected for WP3, T3.3
Standards	
Info about metadata (Production and storage dates, places) and documentation?	Metadata that will be used are the production date and the code number of the reverse osmosis unit
Standards, Format, Estimated volume of data	The data will be in text format
Data exploitation and sharing	
Data exploitation (purpose/use of the data analysis)	The data will be used only for the development of the Satisfactory application.
Data access policy / Dissemination level (Confidential, only for members of the Consortium and the Commission Services) / Public	Data can be public
Data sharing, re-use and distribution (How?)	Data will be shared by using a data management portal.
Embargo periods (if any)	None
Archiving and preservation (including storage and backup)	

Data storage (including backup): where? For how long?

Data will be stored in the storage device of the developed system (computer). A back up will be stored in an external storage device.

CONCLUSION

The first analysis contained in this report allows to anticipate the procedures and infrastructures to be implemented by SatisFactory to efficiently manage the data it will produce. In particular, this first version of the SatisFactory Data Management Plan leads the project to draft the specifications of one of the key tools to be implemented: the SatisFactory Data Management Portal.

Based on the information detailed in this report, this portal will need to manage a large number of datasets, collected by various devices: to data collected by sensors, must be added data manually inputted in IT systems, collected through direct interactions with employees (e.g. interviews), and data registered by cameras. This last item will be the most significant in terms of volume, since it is expected that this data will quickly go beyond the 1TB/day threshold – depending of course on the actual specifications of the pilot testing, which are still to be defined.

Another lesson from this report is that nearly all project partners will be owners or/and producers of data. Similarly, all technical Work Packages of the project will produce data – though the majority of data will be produced through WP3. This means that the Data Management Portal will need to allow for specific access to each project partner, and that editing / access rights will need to be managed accordingly.

The SatisFactory Data Management Plan will put a strong emphasis of the appropriate collection – and publication should the data be published – of metadata, storing all the information necessary for the optimal use and reuse of those datasets. This metadata will be managed by each data producer, and will be integrated in the Data Management Portal.

Finally, the Data Management Portal will need to be flexible in terms of the parts of datasets that are made public. Special care will be given to ensuring that the data made public violates neither IPR issues related to the project partners, nor the regulations and good practices around personal data protection. For this latter point, systematic anonymization of personal data will be made.

This draft Data Management Plan provides an overview of the data that SatisFactory will produce, and of the challenges and constraints that need to be taken into account for managing it. Once the solutions to be tested by the project and the content of this testing are specified, the next step – which will be described in the updated version of this report due in December 2015 – will be to finalize the specifications of the Data Management Portal of the project and to provide a detailed Management Plan for each data set.