Document Number**:** H2020-EUK-815323/5G-ALLSTAR/D6.12

Project Name:
5G AgiLe and fLexible integration of SaTellite And cellulaR (5G-ALLSTAR)

Deliverable D6.12

Data Management Plan

Date of delivery:	30/09/2018	Version:	1.0
Start date of Project:	01/07/2018	Duration:	36 months

Deliverable D6.12
Data Management Plan

Project Number:	H2020-EUK-815323
Project Name:	5G AgiLe and fLexible integration of SaTellite And cellu-laR

Document Number:	H2020-EUK-815323/5G-ALLSTAR/D6.12
Document Title:	Data Management Plan
Editor(s):	Nicolas Cassiau (CEA)
Authors:	Nicolas Cassiau (CEA), Stephan Jaeckel (FhG)
Dissemination Level:	PU
Contractual Date of Delivery:	30/09/2018
Security:	Public
Status:	Final
Version:	1.0
File Name:	5G_ALLSTAR_D6.12.docx

Abstract

This document describes the data management life cycle for the data to be collected, processed and/or generated by the Horizon 2020 project 5G-ALLSTAR.

Keywords

Data Management Plan, Findable, Accessible, Interoperable and Reusable Data

Versions

Version	Date	Content
1.0	30th September 2018	Initial information on data management in the 5G-ALLSTAR project.

Contents

1 Introduction

This deliverable describes the data management life cycle for the data to be collected, processed and/or generated by the Horizon 2020 project 5G-ALLSTAR. As part of making research data findable, accessible, interoperable and reusable (FAIR), it includes information on:

- the handling of research data during and after the end of the project
- what data will be collected, processed and/or generated
- which methodology and standards will be applied
- whether data will be shared/made open access and
- how data will be curated and preserved (including after the end of the project).

This deliverable is a living document that will updated continuously during the project.

List of Abbreviations

DMP	Data Management Plan
FAIR	Findable, Accessible, Interoperable and Reusable

2 Data summary

<u>What is the purpose of the data collection/generation and its relation to the objectives of the project?</u>

Data collection/generation pursue several goals. It will first be a support for exchanges between partners (meeting reports, emails, spreadsheets…). Data will also be used as a mean of recording project results (mainly thanks to deliverables) for possible future use. They will be then used for the demonstration of the project results (simulation and experiment). Finally, data will be a basis for dissemination of the project outcomes (publications, slidesets). All data also aims at demonstrating to the EU and to the Korean government that 5G-ALLSTAR as reached its objectives.

<u>What types and formats of data will the project generate/collect?</u>

The project will generate text, spreadsheets, emails, slidesets, software and algorithms.

<u>Will you re-use any existing data and how?</u>

A software (Quadriga) previously developed by Fraunhofer will be used for channel simulation. This software will be enhanced during the project to meet the project requirements. Aside this, the 5G-ALLSTAR project aims at producing new results, therefore no re-use of existing data is planned. Obviously, existing literature will be used to compile the state of the art in each scientific field studied in the project.

<u>What is the origin of the data?</u>

All data will be generated during the project. Text will come from deliverables, reports, publications and press releases. Spreadsheets would collect for example simulation and experiment results. Slidesets will be generated by physical and phone meetings but also by external presentations. Software and algorithms will be developed during the project to answer 5G-ALLSTAR problematics.

<u>What is the expected size of the data?</u>

At this early stage of the project, the total volume of data and the number of files cannot be evaluated. This section will be iteratively updated during the project.

<u>To whom might it be useful ('data utility')?</u>

Data will be first useful as a mean of exchanges between 5G-ALLSTAR partners. Data will then be used by project reviewers, to evaluate the progress of the project. Each partner will also use the data produced by the project to serve its company's objectives (for example normalization). Finally, data will be used by the scientific community as a basis for future studies or by industries for validating development objectives.

3 FAIR data

3.1 Making data findable, including provisions for metadata

<u>Are the data produced and/or used in the project discoverable with metadata, identifiable and locatable by means of a standard identification mechanism (e.g. persistent and unique identifiers such as Digital Object Identifiers)?</u>

No Digital Object Identifiers are used in the project.

<u>What naming conventions do you follow?</u>

For deliverables, the name must follow the pattern:

5G-ALLSTAR_Dx.y.docx(pdf)

with x the work package number and y deliverable number in the work package.

Other documents must start with 5G-ALLSTAR_. This must be followed by a date and a place (if from a meeting), the WP number if relevant, and the type of document (minutes, agenda, etc.)

<u>Will search keywords be provided that optimize possibilities for re-use?</u>

Keywords will be provided in deliverables.

<u>Do you provide clear version numbers?</u>

The data will be stored in a shared space, using the BSCW (Basic Support for Cooperative Work) tool provided by Fraunhofer FIT. The BSCW system is based on the notation of a shared workspace, a joint storage facility that may contain various kinds of objects such as documents, tables, graphics, spreadsheets or links to other Web pages. A workspace can be set up and objects stored, managed, edited or downloaded with any Web browser. The BSCW system will keep the members of a group informed about each other's relevant activities in a shared workspace.

This tool provides a versioning capability that will be used in the 5G-ALLSTAR project.

<u>What metadata will be created? In case metadata standards do not exist in your discipline, please outline what type of metadata will be created and how.</u>

No metadata will be created.

3.2 Making data openly accessible

<u>Which data produced and/or used in the project will be made openly available as the default? If certain datasets cannot be shared (or need to be shared under restrictions), explain why, clearly separating legal and contractual reasons from voluntary restrictions.</u>

<u>Note that in multi-beneficiary projects it is also possible for specific beneficiaries to keep their data closed if relevant provisions are made in the consortium agreement and are in line with the reasons for opting out.</u>

All public deliverables will be made openly available as the default. Quadriga software will be made openly available before the end of the project. Inside the project, all data produced by the project will be shared between partners.

How will the data be made accessible (e.g. by deposition in a repository)?

The data will be stored in a shared space, using the BSCW (Basic Support for Cooperative Work) tool provided by Fraunhofer FIT. The BSCW system is based on the notation of a shared workspace, a joint storage facility that may contain various kinds of objects such as documents, tables, graphics, spreadsheets or links to other Web pages. A workspace can be set up and objects stored, managed, edited or downloaded with any Web browser. The BSCW system will keep the members of a group informed about each other's relevant activities in a shared work-space.

Public documents will be made available in the public section of the 5G-ALLSTAR project.

What methods or software tools are needed to access the data?

Please see previous answer.

Is documentation about the software needed to access the data included?

The BSCW tool includes a very detailed documentation on how to use it. Furthermore, good practice rules have been provided to all partners.

Is it possible to include the relevant software (e.g. in open source code)?

Aside the channel simulation software, the software produced during the project will not be openly accessible. (This answer may be revised during the project).

Where will the data and associated metadata, documentation and code be deposited? Preference should be given to certified repositories which support open access where possible.

Please see the answer to the question "How will the data be made accessible?".

Have you explored appropriate arrangements with the identified repository?

The data in the repository is arranged in a way that makes it easy to access.

If there are restrictions on use, how will access be provided?

For project members, no restriction on use are foreseen.

Is there a need for a data access committee?

There is no need for a data access committee.

Are there well described conditions for access (i.e. a machine readable license)?

The conditions of use of the BSCW tool are available on the web.

How will the identity of the person accessing the data be ascertained?

Each person is assigned a login and a password to access the data in the BSCW tool.

3.3   Making data interoperable

Are the data produced in the project interoperable, that is allowing data exchange and re-use between researchers, institutions, organisations, countries, etc. (i.e. adhering to standards for formats, as much as possible compliant with available (open) software applications, and in particular facilitating re-combinations with different datasets from different origins)?

Data for project internal usage will use the Windows Office format (i.e. .docx, .xlsx and .pptx). Text data that will be made available outside the project will use the pdf (portable document format) format that can be read with open software.

What data and metadata vocabularies, standards or methodologies will you follow to make your data interoperable?

Common data vocabularies, standards or methodologies will be used.

Will you be using standard vocabularies for all data types present in your data set, to allow inter-disciplinary interoperability?

Yes.

In case it is unavoidable that you use uncommon or generate project specific ontologies or vocabularies, will you provide mappings to more commonly used ontologies?

No uncommon or project specific ontologies or vocabularies will be used.

3.4   Increase data re-use (through clarifying licences)

How will the data be licensed to permit the widest re-use possible?

Most of deliverables will be public, and therefore freely available on the project website. The software Quadriga, which will be enhanced during the project, will be made freely available before the end of the project. Conference and journal publications will be available on the web, with license depending on the type of publication (IEEE,…).

When will the data be made available for re-use? If an embargo is sought to give time to publish or seek patents, specify why and how long this will apply, bearing in mind that research data should be made available as soon as possible.

Public deliverables will be made available as soon as accepted by the reviewers.

Channel model emulator will be made available on month 18.

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

Please see previous answer.

How long is it intended that the data remains re-usable?

No time restriction is planned.

Are data quality assurance processes described?

No data quality assurance processes is available.

4   Allocation of resources

What are the costs for making data FAIR in your project?
During the project, for partners:

- BSCW is operated by Fraunhofer HHI free of charge for project partners. BSCW will be shut down 3 months after project ends. All contained data has to be archived at the partners premises.

- Domain registration for project website (5g-allstar.eu) costs 25 € annually and will be covered by Fraunhofers expenses. Website will remain online for at least 3 years after the end of the project.

How will these be covered? Note that costs related to open access to research data are eligible as part of the Horizon 2020 grant (if compliant with the Grant Agreement conditions).
Who will be responsible for data management in your project?
Please refer to the answer to the previous question.

Are the resources for long-term preservation discussed (costs and potential value, who decides and how what data will be kept and for how long)?
This question will be discussed during the project, and this deliverable will be updated accordingly.

5   Data security

What provisions are in place for data security (including data recovery as well as secure storage and transfer of sensitive data)?

Access to the BSCW server is password protected on a per-person level. The server is located at the Fraunhofer HHI premises in Berlin, Germany. No third party has access to the stored data without permission. All connections to the server are encrypted (Certified SSL connection). Weekly incremental backups are in place in case of hardware failure.

Is the data safely stored in certified repositories for long-term preservation and curation?

BSCW will be shut down 3 months after project ends. All contained data has to be archived at the at the project partners premises for long-term preservation.

6   Ethical aspects

Are there any ethical or legal issues that can have an impact on data sharing? These can also be discussed in the context of the ethics review. If relevant, include references to ethics deliverables and ethics chapter in the Description of the Action (DoA).
There are no ethical or legal issues that can have an impact on data sharing.

Is informed consent for data sharing and long term preservation included in questionnaires dealing with personal data?
No personal data will be used in the project.

7 Other issues

Do you make use of other national/funder/sectorial/departmental procedures for data management? If yes, which ones?
We do not make use of other national/funder/sectorial/departmental procedures.