



Data management plan

D6.2

The DETERMINISTIC6G project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement no 1010965604.



Data management plan

Grant agreement number:	101096504
Project title:	Deterministic E2E communication with 6G
Project acronym:	DETERMINISTIC6G
Project website:	Deterministic6g.eu
Programme:	EU JU SNS Phase 1
Deliverable type:	Report
Deliverable reference number:	D24
Contributing workpackages:	WP6
Dissemination level:	Public
Due date:	M06
Actual submission date:	30-06-2023
Responsible organization:	ERI-DE
Editor(s):	Dhruvin Patel
Version number:	V1.0
Status:	Final
Short abstract:	This deliverable describes data management plan of the DETERMINISTIC6G project.
Keywords:	FAIR, data, security and DMP

Contributor(s):	James Gross, Gourav Prateek Sharma Frank Duerr Raheeb Muzaffar
-----------------	---

Revision History

V0.1	First draft
V0.2	Second draft with complete details
V0.3	Third draft for the first project review
V0.4	Fourth draft for the second project

Disclaimer

This work has been performed in the framework of the Horizon Europe project DETERMINISTIC6G co-funded by the EU. This information reflects the consortium's view, but the consortium is not liable for any use that may be made of any of the information contained therein. This deliverable has been submitted to the EU commission, but it has not been reviewed and it has not been accepted by the EU commission yet.

Contents

Revision History	1
Disclaimer.....	2
1 Introduction	4
1.1 Data Summary.....	4
1.1.1 Purpose of the data collection	4
1.1.2 Types of data.....	4
1.1.3 Re-use of existing data.....	5
1.1.4 Expected data size.....	5
1.1.5 Data utility.....	5
1.2 Findable, accessible, interoperable, and re-usable (FAIR) data.....	6
1.2.1 Making data findable, including provisioning of metadata	6
1.2.2 Making data openly accessible	6
1.2.3 Making data interoperable	8
1.2.4 Increase data re-use.....	8
1.3 Allocation of resources	8
1.4 Data security	8
1.4.1 Shared project data.....	8
1.4.2 Data within each partner institution.....	9
References	9
List of abbreviations.....	9

1 Introduction

The data management plan (DMP) describes how data is collected/generated and managed during the project run time. The DMP should clarify who is responsible for the data management in different stages. The present document describes the DMP, detailing the management life cycle for the data to be collected, processed, and generated by the DETERMINISTIC6G project. The DMP is based on the format provided by the ARGOS platform [AOS201] answering the questions essential for the data management plan. This DMP is planned to be updated in the context of the periodic evaluation of the projects if required.

1.1 Data Summary

1.1.1 Purpose of the data collection

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

Across the project, partners of the consortium will generate data in the form of simulation results, presentations pertaining towards project objectives. The collection and sharing of the information across the project are essential to keep the partners up to date and aligned to ensure the successful execution of the project. Data across the consortium will be shared internally via collaboration space provided by the project coordinator. The collaboration space is accessible only to the project partners.

In addition, a dissemination and communication plan, documented in D6.1 [D6G23-D61] will ensure the outcome of the DETERMINISTIC6G project reaching the target audiences.

1.1.2 Types of data

What type and formats of data will be project generated/collected?

During the run time of the project, different types and formats of the **open access data** will be generated and collected which will be shared via public website of the project, social media platforms, Zenodo [ZEDO23] platform.

Deliverables: The project will develop a series of the open public deliverables in form of technical reports and open-source software packages. All the public deliverables will be released in the public website. Open-source software packages will be available in Zenodo [ZEDO23] platform.

Research outcomes: All the partners will publish their outcomes (e.g., journal papers, conference papers, white papers, presentations, tutorials etc.) both in the project repository and on the public website.

Standardization of technical contributions: All the partners will also publish the standards contributions in the official webpages.

Additional information: Including newsletter, press release will be published through the public website and the social networks.

The DETERMINISTIC6G consortium has adapted the review procedure which allows all the above data/information being shared with the consortium members before it gets published for the external audience. Additionally, meeting minutes, project internal reports will be shared internally among the project consortium members via collaboration space hosted by the project coordinator.

1.1.3 Re-use of existing data

Will you re-use any existing data and how? What is the origin of the existing data?

DETERMINISTIC6G will utilize the data developed and validated by the partners outside the framework of the project. The public specifications across various standardization organizations such as 3GPP, OPC UA and IEEE 802.1 TSN working group will be used as starting point for further development of the concepts. Additionally, existing scientific publications, tutorials, concepts will be used as baseline to further develop DETERMINISTIC6G technical enablers. The DETERMINISTIC6G validation platform builds upon NetSTInG [NST23]. It is based on the OMNeT++ [OMT23] discrete event simulation platform. To evaluate DETERMINISTIC6G technical enablers for latency characterization, the testbed based on OpenAirInterface (OAI) [OAI23] will be a baseline for benchmarking the developed solution.

1.1.4 Expected data size

What is the expected size of the project data?

DETERMINISTIC6G will develop new technological concepts for E2E deterministic communication services. The output of the project will be in the form of

- 26 public deliverables including an open-source validation framework package
- Measurement data and open-source packages will be in the range of Gigabytes

Report D5.1 [D6G23-D51] further details the expected dissemination and communication activities.

1.1.5 Data utility

To whom might it be useful ('data utility')?

Data generated from the project is planned to be disseminated towards the identified target audience. All the partners within the consortium have already made a dedicated dissemination plan, details of which are available in the deliverable D5.1 "Impact, exploitation and dissemination plan for academic research, awareness of potential benefits to industry, standardization synergies" [D6G23-D51] and is available on the project webpage.

Industrial community: The project results will be promoted towards the industrial community through dissemination at various industry fora and standardization bodies such as 3GPP, IEEE 802.1, IETF, NGMN, and OPC Foundation. Standardization activities will mainly be driven by industry partners namely Ericsson, Orange, and B&R. Ericsson is actively involved in 3GPP and IEEE 802.1, IEC/IEEE 60802 standardization working groups, whereas B&R is active in OPC UA FLC. Moreover, participation in industrial workshops, and global industrial alliances such as NGMN, and 5G-ACIA is targeted.

Scientific community: The scientific dissemination will mainly be driven by academic partners but highly supported by industry partners and SMEs. The scientific dissemination will be in the form of peer-reviewed scientific journals, conferences, workshops, and white papers in addition to organizing scientific workshops, tutorials, and special sessions. All publications will be available as open access, however, according to the guidelines of the publication venues.

Social networks: The scientific findings and the industrial advances of the project will be integrated into lectures, seminars, and summer schools conducted by the R&D partners. Summer schools and info days will be organized by academic partners to have in-depth discussions of the project and to disseminate the knowledge to young PhD students and researchers. Moreover, awareness to a wider community will be created through digital media such as the project website, social media channels, and through professional networks.

1.2 Findable, accessible, interoperable, and re-usable (FAIR) data

1.2.1 Making data findable, including provisioning of metadata

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)?

The internal project repository is defined with specific structure and format for tracking all the internal shared documents. Externally towards target audience, several metadata tracking methods will be utilized depending upon the type of the data. For instance, DOI (Digital Object Identifiers) mechanism used for publication, measurement data and open-source software packages (provided through Zenodo platform).

What naming conventions do you follow? Do you provide clear version numbers

Deliverables are assigned unique IDs along with full title, making it easily accessible via search engines and on the project website. The nomenclature agreed within the consortium is explained in the project management handbook D6.1 [D6G23-D61] as follows,

1. Deliverables have names according to the following template “DETERMINISTIC6G-DX.Y-vN.M”, were
 - X.Y is the deliverable number
 - N.M is the version number of the document

For journals, articles and standard contributions, the naming is as follows

`<event>_<yyyy>_<aut>_<Words>_<N.M>`, where

- `<event>` - abbreviation of the event name such a magazine, journal, conference e.g., 3GPP, IEEE, etc.
- `<aut>` -- first three letters of the last name of the author; in case of several authors to the paper, indicate only the last name of the first author and append the rest with “etal”, e.g. if the author is Tom Smith, the abbreviation would be Smi; in case of several authors, e.g., Tom Smith and Robert Brown, the abbreviation would be Smi_etal
- `<yyyy>` - year of the publication
- `<vN.M>` for version handling as explained in the following subsection.

Will search keywords be provided that optimize possibilities for re-use?

Keywords are also provided in all public deliverables to optimize possibilities for re-use.

1.2.2 Making data openly accessible

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.

All the deliverables providing various results will be made public. Some of the results produced will be shared (or shared under certain restrictions) to the open public. All the internal communication, reports, tutorials, presentations and discussions within the consortium leading to the project results will not be made public. The restricted data due to contractual reasons will not be shared to open public.

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

All the data collection and sharing essential to project tasks will be shared internally via collaboration space hosted by the project coordinator. Data is only accessible to selected colleagues of each organization within the consortium. The project will also provide main results, updates by making use of the official project website. The website is open and accessible to the public. Innovation and dissemination team members will also ensure results are disseminated towards social platforms such as Twitter and LinkedIn. Zenodo will be used as platform to share the open-source software packages and measurement data.

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

Public deliverables and presentations will be uploaded on the public website using Portable Document Format (PDF) format. Guidelines to utilize the DETERMINISTIC6G validation framework to reproduce the results and to perform further experiments will be provided along with the open-source software packages.

Is documentation about software needed to access the data included? Is it possible to include the relevant software (e.g., in an open-source code?)

The DETERMINISTIC validation framework software package is planned to be made open-source. Guidelines on how to run the validation framework will be provided.

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.

All the deliverables and results including the metadata will be made available in the Download area of the official project website. Internally, a specific project structure is defined to ensure proper organization of the data at work package level. Measurement data and open-source packages will be deposited in the Zenodo platform, which is a European Union supported research data repository.

Have you explored appropriate arrangements with the identified repository?

The project coordinator has ensured that the repository is properly organized and all the project partners fulfil the appropriate arrangements with the internal repository.

If there are restrictions on use, how will access be provided? How will the identity of the person accessing the data be ascertained.

The project coordinator has ensured that the collaboration space is accessible only by registered users from the consortium. No registration is required to visit the project webpage which means also public deliverables are open and accessible to the public without any restriction.

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.

The timeline of the DETERMINISTIC6G deliverables is openly available on the project website. All the data related to deliverables will be available in accordance with the timeline. Different embargo guidelines for different publication venues (such as IEEE, Elsevier, Springer, ACM etc.) will be considered before making it public on the website.

1.2.3 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)?

It is estimated that data produced within the project is interoperable. All the reports, deliverables, and public presentations will be documented and presented in English. The final version of all the documents will be provided in PDF format, making it readable on a wide range of devices/platforms. Where it is not possible, the project will endeavor to provide data in formats which are open, widely accepted and are accessible to the wider public through open-source utilities.

1.2.4 Increase data re-use

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

Project output including public deliverables in form of the technical reports, software packages, and the project results are available on the project website and can be reused by other projects. Some of the contributions provided by the project will also be made public. Selected organizations such as 3GPP, IETF DetNet allow open access, while others allow access only to members (e.g., IEEE).

Specific confidential material may require direct licensing from the originating company.

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

Data produced within the DETERMINISTIC6G project and openly published on the website will be usable during and after the project.

1.3 Allocation of resources

What are the costs for making data FAIR in your project?

There is a dedicated work package (WP5) within the project working towards dissemination, standardization, and exploitation of the DETERMINISTIC6G results and research produced within the project. WP5 is operational throughout the project and has in total 59 Person Month (PM) allocated within a minimum of 1 person month assigned from each project partner.

Who is responsible for data management in your project?

The innovation manager of the project together with the project coordinator will be responsible for the data management within the project.

Is the resource for long term preservation discussed (cost and potential value, who decide and how what data will be kept and for how long)?

Data is planned to be preserved after the end of the project. Internal and confidential reports are expected to be available in the collaboration space in a static copy for at least 5 years.

1.4 Data security

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

1.4.1 Shared project data

The sharing of all non-public data within the project is done over a collaboration space provided by the project coordinator. Access to the collaboration space requires username and password. The

platform provided by the project coordinator is protected by the state-of-the-art security mechanisms.

1.4.2 Data within each partner institution

The consortium members have measures in place to preserve and protect data. Each partner has agreed to comply with EU data protection directive 95/96/EC.

References

- [AGOS23] ARGOS online tool for creating DMP, 2023, [online]: Argos - About (openaire.eu)
- [D6G23-D51] DETERMINISTIC6G, Deliverable 5.1, “impact, exploitation, and dissemination plan for academic research, awareness of benefits to industry, standardization synergies” March 2023
- [D6G23-D61] DETERMINISTIC6G, Deliverable 6.1, “Project Management Handbook”, January 2023
- [ZEDO23] Zenodo platform, [online]: <https://zenodo.org/>
- [NST23] NetSTInG framework, [online]: <https://gitlab.com/ipvs/nesting>
- [OMT23] OMNeT++ simulation tool, [online]: OMNeT++ Discrete Event Simulator (omnetpp.org)
- [OAI23] OpenAirInterface, [online]: OpenAirInterface – 5G software alliance for democratising wireless innovation

List of abbreviations

Table 1: List of abbreviations

DMP	Data Management Plan
FAIR	Findable, accessible, interoperable, and re-usable
EU	European Union