



H2020-INFRAEOSC-2018-3

NI4OS-Europe

National Initiatives for Open Science in Europe

Deliverable D1.14

Data Management Plan

Lead beneficiary(s): ATHENA RC

Author(s): Eleni Toli, Elli Papadopoulou, Christos Liatas (editors), All partners

Status –Version: Final – d

Date: February 28, 2020

Dissemination Level: Public

Abstract: Deliverable D1.14 – Data Management Plan (DMP) outlines how data will be handled during and after the project, describes what data will be collected, processed or generated and following what methodology and standards, how these data will be shared and/or made open, and how they will be curated and preserved.

© Copyright by the NI4OS-Europe Consortium



The NI4OS-Europe project is funded by the European Commission under the Horizon 2020 European research infrastructures grant agreement no. 857645.

The NI4OS-Europe Consortium consists of:

GRNET SA	Coordinator	Greece
ATHENA RC	Beneficiary	Greece
CYI	Beneficiary	Cyprus
UCY	Beneficiary	Cyprus
IICT	Beneficiary	Bulgaria
SRCE	Beneficiary	Croatia
RBI	Beneficiary	Croatia
KIFU	Beneficiary	Hungary
DE	Beneficiary	Hungary
ICI BUCURESTI	Beneficiary	Romania
UEFISCDI	Beneficiary	Romania
ARNES	Beneficiary	Slovenia
UMUKM	Beneficiary	Slovenia
IPB	Beneficiary	Serbia
UOB	Beneficiary	Serbia
RASH	Beneficiary	Albania
UNI BL	Beneficiary	Bosnia-Herzegovina
UKIM	Beneficiary	North Macedonia
UOM	Beneficiary	Montenegro
RENAM	Beneficiary	Moldova (Republic of)
IIAP NAS RA	Beneficiary	Armenia
GRENA	Beneficiary	Georgia

This document contains material, which is the copyright of certain NI4OS-Europe beneficiaries and the European Commission, and may not be reproduced or copied without permission. The information herein does not express the opinion of the European Commission. The European Commission is not responsible for any use that might be made of data appearing herein. The NI4OS-Europe beneficiaries do not warrant that the information contained herein is capable of use, or that use of the information is free from risk, and accept no liability for loss or damage suffered by any person using this information.

Document Revision History

Date	Issue	Author/Editor/Contributor	Summary of main changes
31/01/2020	a	Eleni Toli, Elli Papadopoulou, Electra Sifakaki, Christos Liatas	First Draft
12/02/2020	b	Eleni Toli, Elli Papadopoulou, Electra Sifakaki, Christos Liatas	Updates from partners' input
24/02/2020	c	Eleni Toli, Elli Papadopoulou, Electra Sifakaki, Christos Liatas	Refinements and updates from partners' input
28/02/2020	d	Eleni Toli, Elli Papadopoulou, Electra Sifakaki, Christos Liatas, Ognjen Prnjat	Layout and context corrections

Table of contents

1. Introduction	9
2. Data summary	10
2.1. WHAT IS THE PURPOSE OF THE DATA COLLECTION/GENERATION AND ITS RELATION TO THE OBJECTIVES OF THE PROJECT?	10
2.2. WHAT TYPES AND FORMATS OF DATA WILL THE PROJECT GENERATE/COLLECT?	11
2.3. WILL YOU REUSE ANY EXISTING DATA AND HOW?	12
3. FAIR data.....	14
3.1. MAKING DATA FINDABLE, INCLUDING PROVISIONS FOR METADATA	14
3.2. MAKING DATA OPENLY ACCESSIBLE.....	15
3.3. MAKING DATA INTEROPERABLE.....	16
3.4. INCREASE DATA RE-USE (THROUGH CLARIFYING LICENCES)	16
4. Allocation of resources	18
5. Data security	18
6. Ethical aspects.....	19
7. Other issues.....	19
8. Conclusions.....	20

References

- [1] Guidelines on FAIR Data Management in Horizon 2020, v3.0 – 26 July 2016:
https://docs.google.com/viewer?url=https%3A%2F%2Fec.europa.eu%2Fresearch%2Fparticipants%2Fdata%2Fref%2Fh2020%2Fgrants_manual%2Fhi%2Ffoa_pilot%2Fh2020-hi-oa-data-mgt_en.pdf
- [2] Template for the Data Management Plan. (Annotated version for participants under Societal Challenge 1), v2.0 – 15 February 2018:
https://docs.google.com/viewer?url=https%3A%2F%2Fec.europa.eu%2Fresearch%2Fparticipants%2Fdata%2Fref%2Fh2020%2Fother%2Fgm%2Freporting%2Fh2020-tpl-oa-data-mgt-plan-annotated_en.pdf
- [3] Disciplinary Metadata: <http://www.dcc.ac.uk/resources/metadata-standards>
- [4] CERIF (Common European Research Information Format):
<http://rd-alliance.github.io/metadata-directory/standards/cerif.html>
- [5] VI-SEEM receives funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 675121:
<https://vi-seem.eu/overview/>

List of Tables

TABLE 1: PROJECT OBJECTIVES WITH RELATED DATASET COLLECTION/GENERATION	11
TABLE 2: DATASETS COLLECTED/GENERATED THROUGH NI4OS-EUROPE.....	11
TABLE 3: DATA REUSED OR INTENDED TO BE REUSED BY NI4OS-EUROPE	12
TABLE 4: NI4OS-EUROPE ORIGIN, EXPECTED SIZE, DATA UTILITY OF DATASETS COLLECTED/GENERATED.....	13
TABLE 5: FINDABLE DATA INCLUDING PROVISIONS FOR METADATA	14
TABLE 6: PROVISIONS FOR OPENLY ACCESSIBLE DATA	16
TABLE 7: NI4OS-EUROPE PROVISIONS FOR MAKING DATA INTEROPERABLE.....	16
TABLE 8: NI4OS-EUROPE DATA LICENCES AND PROVISIONS FOR DATA RE-USE.....	17
TABLE 9: RESOURCES ALLOCATED FOR DATA MANAGEMENT AND MAKING DATA FAIR IN NI4OS-EUROPE	18
TABLE 10: DATA SECURITY PROVISIONS IN NI4OS-EUROPE	18
TABLE 11: ETHICAL ASPECTS RELATED TO DATA SHARING IN NI4OS-EUROPE	19

List of Acronyms

NOADs	National Open Access Desks
CERIF	Common European Research Information Format
DOI	Digital object identifier
GDPR	General Data Protection Regulation
EOSC	European Open Science Cloud
OpenAIRE	Open Access Infrastructure for Research in Europe
FAIR	Findable Accessible Interoperable Reusable
ORDM	Open Research Data Management
OS	Open Science
OA	Open Access
WP	Work package
PID	Persistent Identifier
PMB	Partner member board
ERA	European Research Area
RoP	Rules of Participation

Executive summary

What is the focus of this Deliverable?

The D1.14 - Data Management Plan (DMP) is an initial version of a living document, intended to describe the data management life cycle for all data sets that will be collected, processed or generated by NI4OS-Europe. It also intends to formalize the means of production of software as well as the goals and structures that will ensure the software produced by NI4OS-Europe remains accessible and reusable in the short, medium and long term.

What are the deliverable contents?

NI4OS-Europe deals with the provision of tools and support for the onboarding of open science services to EOSC. Thus, it deals with any type of data and datasets generated or aggregated through its various activities. These types of data and datasets will be made available, in line with the guidelines for the Extended Pilot on Open Research Data in Horizon 2020 [1]. The DMP will also include the dataset metadata specification that will be used in the data registry, following an appropriate relevant standard [2]. It will specify the recommended licensing schemes as suggested by H2020. In this first version of the DMP both existing and planned data sets are described.

Conclusions and recommendations

The NI4OS-Europe DMP as depicted in D1.14 constitutes the basic tool that will be used to manage data in the NI4OS-Europe project.

1. Introduction

NI4OS-Europe participates in the EU Extended Open Research Data Pilot and as such will deliver a Data Management Plan (D1.14). Open access to research data refers to the right to access and re-use digital research data under the terms and conditions set out in the Grant Agreement. Openly accessible research data can typically be accessed, mined, exploited, reproduced and disseminated under defined and clearly specified terms and conditions (either free of charge or at a cost for the user).

The DMP is not a fixed document; it evolves and gains more precision and substance during the lifespan of the project.

This first version of the DMP is delivered in month 6 of the project. As the project evolves, updates of this plan would be useful at mid-way (month 18) and at the end of the project (in preparation for exploitation).

2. Data summary

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

The overall objective of NI4OS-Europe is to support the development and inclusion of the national Open Science Cloud initiatives in 15 Member States and Associated Countries in the overall scheme of EOSC governance; spread the EOSC and FAIR principles in the community and train it; and provide technical and policy support in on-boarding of the existing and future service providers into EOSC, including generic services (compute, data storage, data management), thematic (domain-specific) services, repositories and data sets – thus introducing new services that cover the whole spectrum of services related to Open Science, data and publications.

The project objectives associated with data collection/generation are depicted in the following table:

<p>01. Support EOSC governance framework by building national Open Science Cloud (OSC) initiatives for open research data, infrastructure and services and enabling them to support the overall EOSC governance and the related EOSC coordination structure.</p>	<p>Stakeholders data was initially collected within the scope of the landscaping activity conducted under the NI4OS-Europe project (WP2). The idea was to contact stakeholders, classified into five groups, as potential respondents in the landscaping survey. The dataset will also be used to create an interactive online map showing stakeholders in the 15 partner countries. The Survey dataset was collected within the scope of the landscaping activity conducted under the NI4OS-Europe project (WP2). The survey was conducted in the 15 partner countries with the aim of collecting information about Open Science initiatives, infrastructures, services, policies, stakeholders and topics, in line with the objectives of NI4OS-Europe, but also to contribute to the final mapping in the overall EOSC Landscape Activity. The collected information will be used by all WPs in the NI4OS project, primarily as in initial input in designing further activities. Selected data from the dataset will be used to create visualizations.</p>
<p>03. Enable the EOSC-relevant, non-commercial services to be accessed through the EOSC portal.</p>	<p>Horizontal service delivery and platform provisioning, will provide the final integrated production-level platform, consisting of EOSC on-boarded regional resources from the previous tasks. A dataset will be associated with this objective and WP5. The dataset will contain all relevant information about regional services onboarded and offered through the</p>

	EOSC, which is directly connected to the project's objective O3, and indirectly to all others.
O4. Provide the necessary technical, organizational and legal guidelines, tools, mechanisms and certification schemes, to support Open Research Data Management (ORDM) and its implementation in a harmonised and coordinated fashion.	To cover the increased demand for providing technical solutions to address the needs for researchers to publish in FAIR/open modes, we are focusing on the aggregation, improvement and development of tools. The processes for the tools' development will be separate datasets.
O5. Ensure the engagement of the targeted communities and validate the project solutions	Training data to understand the OS and EOSC capacity building needs in the area

Table 1: Project objectives with related dataset collection/generation

2.2. What types and formats of data will the project generate/collect?

Dataset	Type	Format
WP2 (T2.1) – Survey data	Text, numeric data	.csv
WP2 (T2.1) - Stakeholders in Open Science (in the 15 partner countries involved in the project NI4OS-Europe)	Text, numeric data	.csv
WP4 (T4.2) Policies and Guidelines	Textual descriptors, documents	CMS, wiki catalogue records, PDF and Office documents
WP4 (T4.3) – IPR Tool	Software tool (Web App)	Many different types
WP4 (T4.4) - Self-assessment Tool	Software tool (Web App)	Many different types
WP5 (T5.5) - NI4OS-Europe services	On-line Platform	In the first phase of the project, it will be implemented as a spreadsheet document (.csv) and later on as a relational database.

Table 2: Datasets collected/generated through NI4OS-Europe

2.3. Will you reuse any existing data and how?

NI4OS-Europe uses and will reuse existing data where needed throughout the project life. Existing data will be used in various tasks and could be integrated with new data where appropriate. In such cases references to authors and institutions will be made available. Table 3 below shows datasets (including literature/journals) used or intended to be used:

WP & Task	Data sources reused
WP2 (T2.1)	Literature review, datasets review from OpenAIRE, GEANT and others
WP2 (T2.2)	Literature review, National Data Statistics, Eurostat
WP2 (T2.4)	Reuse of data from T2.1, as well as complementary data from EOSCPilot, EOSC-Hub projects and other INFRAEOSC-5a funded projects.
WP2 (T2.5)	Reuse of data from T2.1, T2.2, T2.4 and available data from EOSC Governance Board and EOSC Executive Board
WP2 (T2.6)	Literature review, data reuse from OpenAIRE with worldwide OS actors information and collaboration status
WP3 (T3.1)	Literature review, related data sources (EOSC-Hub, eInfraCentral, EOSCpilot, Research Data Alliance, EOSC governance) integration and harmonization. Reuse of common metadata.
WP3 (T3.3)	Reuse data related to tools of the EOSC-Hub project and VI-SEEM project.
WP4 (T4.1)	Statistics from the NOADs network. Literature research in OSPP Recommendations for Rewards and Incentives, the Data Stewardship Skills and Capability Framework or the EOSCpilot's and DG-RTD's Open Science Monitor.
WP4 (T4.2)	Literature research on existing FAIR data practices, guidelines and certifications
WP5 (T5.1)	Reuse service information datasets previously collected within the framework of the VI-SEEM[5] project
WP5 (T5.4)	Reuse national repositories datasets and related metadata from OpenAIRE.

Table 3: Data reused or intended to be reused by NI4OS-Europe

- *What is the origin of the data?*
- *What is the expected size of the data?*
- *To whom might it be useful ('data utility')?*

Answers to the above questions are integrated in Table 4:

Dataset	Origin	Expected Size	Data Utility
WP2 (T2.1) - Survey	Data from the answers to the survey organized by the consortium,	A few MB	Researchers, research performing organizations, policy makers, public

	focused interviews, f2f meetings		sector, funders, innovation agencies, local businesses
WP2 (T2.1) - Stakeholders	The data was provided by partner institutions in the project.	up to 1 MB	All NI4OS-Europe partners, EOSC Secretariat, EU institutions and bodies and all organizations and individuals interested in Southeast Europe. The dataset can also be used by third parties to create visualizations.
WP4 (T4.2) Policies and Guidelines	OS projects and communities, meta-repositories, literature research	Unknown yet	Researchers, research performing organizations, policymakers, service providers, public sector, funders, innovation agencies, local businesses
WP4 (T4.3) IPR Tool	OS projects and communities, meta-repositories, specifications, charts code, manuals, user-provided data recorded by tools	Unknown yet	Researchers, research performing organizations, service providers, innovation agencies, local businesses
WP4 (T4.4) Self-assessment Tool	OS projects and communities, meta-repositories, specifications, charts, code, manuals, user-provided data recorded by tools	Unknown yet	Researchers, research performing organizations, service providers, innovation agencies, local businesses
WP5 (T5.5) - NI4OS-Europe services	Data will be collected from the resource providers.	Dataset will contain at least 20 generic service instances, 20 thematic services, and 15 repositories.	To wide range of stakeholders, in particular, to researchers as end-users (including business organizations), service providers, and research funders.

Table 4: NI4OS-Europe origin, expected size, data utility of datasets collected/generated

3. FAIR data

3.1. Making data findable, including provisions for metadata

Mission and vision of the project are to be a core contributor to European Open Science Cloud service portfolio, commit to EOSC governance and ensure inclusiveness on the European level. Within the project's major objectives is to spread the EOSC and FAIR principles in the community and train it. As, such NI4OS-Europe will not only provide FAIR data where applicable (data should be "as open as possible and as close as necessary") but will mostly support the community in the uptake of research data sharing and practices, in alignment with FAIR principles. The following table provides information for making NI4OS-Europe data findable, including provisions for metadata to promote reuse.

<i>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)?</i>	YES – DOI, unique and persistent URI. Collected services' information will be available at the dedicated portfolio/catalogue system URL. Other aspects of the persistent identifiers could be implemented by periodical snapshots of the database that contains the data.
<i>What naming conventions do you follow?</i>	NI4OS-[WP]-[title]-[ver]-[year]-[month]-[day].[ext]
<i>Will search keywords be provided that optimize possibilities for reuse?</i>	YES
<i>Do you provide clear version numbers?</i>	YES – accessed with unique and persistent URI
<i>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.</i>	<p>Metadata for General Research Data will follow the Dublin Core and DataCite Metadata Schema.</p> <p>For WP2 Survey and Stakeholders datasets the following metadata will be provided: title, description, keywords, controlled vocabulary terms, data provenance, time when the dataset was collected, spatial coverage, funding information (relation to the project), version, PID (DOI).</p> <p>For the WP5 – NI4OS-Europe services dataset a fully EOSC-compatible service description template (that structures the collected information) will be used.</p>

Table 5: Findable data including provisions for metadata

3.2. Making data openly accessible

By default, NI4OS-Europe will openly provide data produced following the principle “as open as possible, as closed as necessary”, to comply with ethical or security requirements and avoid related conflicting issues. The following table summarizes current considerations.

<i>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.</i>	<p>Anonymised WP2 survey data WP2 Stakeholders data Anonymised WP6 Training Capacity data Self-assessment tool repository IPR tool repository</p> <p>NI4OS services dataset: A predefined portion of all available metadata will be openly accessible, while authorization will be needed for full access.</p>
<i>How will the data be made accessible (e.g. by deposition in a repository)?</i>	Zenodo for all other datasets. WP5 NI4OS services data will be made available for human and machine access through the projects' service portfolio/catalogue system.
<i>What methods or software tools are needed to access the data?</i>	<p>WP2 – Survey dataset can be accessed with Microsoft Excel or similar open office tools.</p> <p>WP5 NI4OS Services Dataset: The project will deploy and host a service portfolio/catalogue system that will provide full human and machine access to the database that stores information about regional services. Datasets for the software tools will require text editors, pdf viewers, more information will be specified in a following DMP version.</p>
<i>Is documentation about the software needed to access the data included?</i>	NO, for other datasets apart from NI4OS services data.
<i>Is it possible to include the relevant software (e.g. in open source code)?</i>	NO, unless open source alternative is not provided – many software choices are available online.
<i>Where will the data and associated metadata, documentation and code be deposited?</i>	Dedicated repository in Zenodo (or else)
<i>Have you explored appropriate arrangements with the identified repository?</i>	For the WP5 NI4OS Services Dataset, repository is run by project coordinator GRNET
<i>If there are restrictions on use, how will access be provided?</i>	Access will be given using user authentication and authorization, handling users' verification and access levels.
<i>Is there a need for a data access committee?</i>	YES – there are GDPR issues associated

<i>Are there well described conditions for access (i.e. a machine-readable license)?</i>	Conditions for access will be described when the access policies are finalized.
<i>How will the identity of the person accessing the data be ascertained?</i>	Project's AAI integrated with the eduGAIN federation.

Table 6: Provisions for openly accessible data

3.3. Making data interoperable

<i>Are the data produced in the project interoperable, that is allowing data exchange and reuse 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 recombinations with different datasets from different origins)?</i>	Yes, metadata interoperability is ensured by following metadata standards. Widely used standard formats and protocols such as OAI-PMH and .odf will ensure datasets exchange and reusability between researchers.
<i>What data and metadata vocabularies, standards or methodologies will you follow to make your data interoperable?</i>	Standard vocabularies will be used, such as the ASIS&T Thesaurus of Information Science, Technology, & Librarianship. For WP2 Survey and Stakeholders datasets the standardized vocabulary EU Country Named Authority List (https://data.europa.eu/euodp/en/data/dataset/country) will be used.
<i>Will you be using standard vocabularies for all data types present in your data set, to allow inter-disciplinary interoperability?</i>	Yes, although it may evolve dynamically during the project lifetime to ensure an ontology alignment within the EOSC.
<i>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?</i>	YES, if uncommon or project specific ontologies or vocabularies are used.

Table 7: NI4OS-Europe provisions for making data interoperable

3.4. Increase data re-use (through clarifying licences)

<i>How will the data be licensed to permit the widest re-use possible?</i>	Probably one of Creative Commons license options, but this will be decided when a full picture of the data will be available. The license for each dataset will be one providing the widest re-use possible.
--	--

	For software, possibly Apache or GNU. More information will be specified in a following DMP version.
<i>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.</i>	The data will be made available for re-use as soon as the final/publishable version of the data is available.
<i>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.</i>	This will also be made clear as soon as there is a good view of the datasets achieved.
<i>How long is it intended that the data remains re-usable?</i>	Data is intended to remain re-usable for as long as it is allowed by project resources and infrastructure. This will be reviewed as the project progresses. NI4OS services data will be integrated into the EOSC portfolio system, which will ensure the reuse of collected information.
<i>Are data quality assurance processes described?</i>	Processes are not described, but effort will be put in providing quality data. The following data quality metadata will be provided: accuracy, relevance and consistency for the Stakeholders and Survey datasets. For the Services dataset collected services' information will be available at the dedicated portfolio/catalogue system URL.

Table 8: NI4OS-Europe data licences and provisions for data re-use.

Further to the FAIR principles, DMPs should also address:

4. Allocation of resources

<i>What are the costs for making data FAIR in your project?</i>	Unknown yet
<i>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).</i>	This is under consideration and further information will be specified in a following DMP version.
<i>Who will be responsible for data management in your project?</i>	Project Director
<i>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)?</i>	Resources for long term preservation have not yet been discussed, as this involves the general progress of EOSC and related projects.

Table 9: Resources allocated for data management and making data FAIR in NI4OS-Europe

5. Data security

<i>What provisions are in place for data security (including data recovery as well as secure storage and transfer of sensitive data)?</i>	Data collected/generated are stored in a NI4OS-Europe cloud workspace, where every file stored is maintained and encrypted using AES 256-bit encryption in geographically diverse areas. HTTPS protocol is used for secure communication between endpoints as a standard. It is the usual HTTP which runs on top of encrypted sockets (SSL/TLS) on the transport layer of the network stack (TCP/IP). Data will be moved to Zenodo repositories for long term preservation.
<i>Is the data safely stored in certified repositories for long term preservation and curation?</i>	YES – For the Zenodo repositories all files uploaded to Zenodo are stored in CERN’s EOS service in an 18 petabytes disk cluster. Each file copy has two replicas located on different disk servers. For each file they store two independent MD5 checksums. One checksum is stored by Invenio, used to detect changes to files made from outside of Invenio. The other checksum stored by EOS, is used for automatic detection and recovery of file corruption on disks. For NI4OS services data periodical database snapshots will be made and stored independently.

Table 10: Data security provisions in NI4OS-Europe

6. Ethical aspects

<p><i>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).</i></p>	<p>No ethical or legal issues foreseen for any of the generated datasets.</p>
<p><i>Is informed consent for data sharing and long-term preservation included in questionnaires dealing with personal data?</i></p>	<p>YES – NI4OS is requesting compliance for collection and reuse of data. WP2 Survey included a full privacy policy for the protection of natural persons concerning the processing of personal data and on the free movement of such data (GDPR). Describes the policies and procedures in place by NI4OS to protect the privacy of users, how the confidentiality of such information is ensured, laws, rights of data subjects and a communication path for further clarifications, should there be needed.</p>

Table 11: Ethical aspects related to data sharing in NI4OS-Europe

7. Other issues

Do you make use of other national/funder/sectorial/ departmental procedures for data management? If yes, which ones?

- NO

8. Conclusions

This deliverable is the first version of the NI4OS-Europe Data Management Plan and contains an initial description of datasets collected/generated in this first stage of the project. The described datasets may be of value for the project and will be exploited by the different tasks through the course of the project. The document will be updated as the list of datasets is enriched with new information or datasets. Datasets use, sharing, preservation and dissemination aspects will be specified in all cases. All this updated information will be included in the future versions and revisions of the current document.