



FIREURISK - DEVELOPING A HOLISTIC, RISK-WISE STRATEGY FOR EUROPEAN WILDFIRE MANAGEMENT

Grant Agreement Number: 101003890	
Call identifier: H2020-LC-CLA-2018-2019-2020	
Topic:	LC-CLA-15-2020 Forest Fires risk reduction: towards an integrated fire management approach in the E.U.
Instrument:	RIA

D4.2 – Data Management Plan

Deliverable Identifier:	D4.2
Deliverable Due Date:	31/08/2021
Deliverable Submission Date:	31/08/2021
Deliverable Version:	1.0
Lead partner:	AUTH
Authors:	AUTH: Christos Sotiropoulos, Dimitris Stavrakoudis, Ioannis Gitas TUD: Matthias Forkel, Christopher Marrs GMV: Marta Gómez-Giménez, Julia Yagüe-Ballester STICHTING VU: Peter Verburg, Cecilia Zagaria ADAI: Luís Mário Ribeiro UCILeR: Dulce Lopes



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003890.

	FESB: Damir Krstinić, Ljiljana Šerić
Work Package:	WP4 - Multi-level integration
Task:	Task 4.1 - Integration within the project
Dissemination Level:	<input type="checkbox"/> PU: Public <input checked="" type="checkbox"/> CO: Confidential, only for members of the Consortium (including the Commission Services)

Revision History

Version	Date	Edited by	Description
v.0.1	23/04/2021	AUTH	Document Content outline / structure
v.0.2	27/06/2021	AUTH	Draft version
v.0.3	22/06/2021	Christos Sotiropoulos (AUTH)	Draft version
v.0.4	28/07/2021	Ioannis Gitas (AUTH) and Dimitris Stavrakoudis (AUTH)	Complete first draft version
v.0.5	06/08/2021	AUTH, ADAI, UAH, TUD, VUA, FESB, ARBO, NCSRD, SATWAYS, UAVR, SU	Internal review from partners involved, including WP4 Leader
v.0.7	29/08/2021	S. Doerr (SU) & G. Eftychidis (KEMEA)	Internal review of the Complete version
v.0.8	29/08/2021	S. Doerr (SU) & J. Pausas (CSIC)	Internal approval review by the QRB
v.0.9	30/08/2021	Dimitris Stavrakoudis (AUTH)	Final version ready for submission
v.1.0	31/08/2021	Domingos Viegas (ADAI)	Final version submitted

Quality Control

Type	Date	Reviewed by	Approved/Comment
Internal & Peer	21/08/2021	Stefan Doerr	Quality & Risk Management Board: approved
Internal & Peer	29/08/2021	George Eftychidis	Approved

Executive Summary

FirEUrisk will collect, process and/or generate a vast amount of information from diverse sources, whereas several (intermediate or final) products will be developed and delivered. This necessitates the establishment of a coordinated approach to the production of outputs, sharing the data internally, and transferring selected data and products externally. This deliverable, namely, the Data Management Plan (DMP), describes the rules and suggestions for ensuring proper data interoperability and sharing procedures.

The DMP functions as a guiding document to ensure good data management and to facilitate the creation of a common understanding and—where possible—common practices. A key aspect of the DMP is to make the information collected and produced Findable, Accessible, Interoperable and Re-usable (FAIR). In particular, the DMP identifies the main data to be generated within FirEUrisk, outlining how data are used during the project as well as protocols and policies required for data exchange and sharing within and outside the consortium.

It is important to note that the DMP will be a living document, meaning that it will be further detailed, updated, amended and/or corrected during the whole lifespan of the project. Updated versions will be produced whenever significant changes arise, e.g., when new data formats are required to be defined or when new consortium policies on data exchange, sharing or security are enforced.

Disclaimer

The content of the publication herein is the sole responsibility of the publishers and it does not necessarily represent the views expressed by the European Commission or its services.

While the information contained in the documents is believed to be accurate, the authors(s) or any other participant in the FirEURisk consortium make no warranty of any kind with regard to this material including, but not limited to the implied warranties of merchantability and fitness for a particular purpose.

Neither the FirEURisk Consortium nor any of its members, their officers, employees or agents shall be responsible or liable in negligence or otherwise howsoever in respect of any inaccuracy or omission herein.

Without derogating from the generality of the foregoing neither the FirEURisk Consortium nor any of its members, their officers, employees or agents shall be liable for any direct or indirect or consequential loss or damage caused by or arising from any information advice or inaccuracy or omission herein.

Copyright message

© FirEURisk Consortium, 2021-2025. This deliverable contains original unpublished work except where clearly indicated otherwise. Acknowledgement of previously published material and of the work of others has been made through appropriate citation, quotation or both. Reproduction is authorised provided the source is acknowledged.

Table of Contents

1	Introduction	10
1.1	Purpose of the document	10
1.2	Structure of the document	10
2	Data Management Plan overview	11
3	Data summary	12
4	FAIR data	14
4.1	Making data findable, including provisions for metadata	15
4.1.1	File naming	15
4.1.2	Metadata	17
4.1.3	Data repository	17
4.1.4	Identification mechanisms	18
4.1.5	Search keywords	18
4.2	Making data openly accessible	18
4.3	Making data interoperable	20
4.4	Increase data reuse (through clarifying licenses)	20
5	Allocation of resources	22
6	Data security	22
7	Ethical aspects	22
8	Conclusions	23
9	References	23
10	Annex A: Product Inventory Form	25
11	Annex B: Dataset Inventory Form	28
12	Annex C: INSPIRE Metadata Example & Requirements	30
13	Annex D: Online open access resources	34

List of Figures

Figure 1: The scheme depicts the FAIRification process. The figure was reproduced from GO FAIR and focuses on data and metadata.15

List of tables

Table 1: List of Acronyms.....	9
Table 2: Dataset request form.....	13
Table 3: Points of contact for dataset requests for each FirEURisk region of interest	14
Table 4: Product Inventory Form.....	25
Table 5: Dataset Inventory Form	28
Table 6: INSPIRE metadata dataset example	30
Table 7: INSPIRE metadata requirements	32

List of Acronyms

Table 1: List of Acronyms

List of Acronyms	
DA	Demonstration Area
DMB	Data Management Board
DMP	Data Management Plan
DOI	Digital Object Identifier
ET	European Territory
FAIR	Findable, Accessible, Interoperable and Re-usable
INSPIRE	Infrastructure for Spatial Information in the European Community
N/A	Not Applicable
OGC	Open Geospatial Consortium
OpenAIRE	Open Access Infrastructure for Research in Europe
PS	Pilot Site
SEGLAB	Societal, Ethical, Gender and Legal Board
TBD	To Be Defined

1 Introduction

FirEURisk's goal is to develop, test and disseminate an integrated and science-based strategy for wildfire risk management in Europe. To achieve this goal, a vast amount of information will be collected, processed and/or generated from diverse sources, whereas several (intermediate or final) products will be developed and delivered. This necessitates the establishment of a coordinated approach to the production of outputs, sharing the data internally, and transferring selected data and products externally. The need was identified during the early stages of the proposal writing and led to the provision for a separate Data Management Board (DMB), which will assist the Consortium in all issues related to data management and security. One of DMB's primary objectives is to explore all interoperability aspects of the datasets that will be used or produced within the project and provide guidance to project partners so that the systems, data formats and relevant organisational aspects are harmonised and consistent with each other. The Data Management Plan (DMP) consolidates the results of this effort, detailing the rules and suggestions for ensuring proper data interoperability and sharing procedures.

1.1 Purpose of the document

This document represents the first version of the DMP, which should be followed by the partners for the datasets generated or collected throughout the lifecycle of the project. The DMP functions as a guiding document to ensure good data management and to facilitate the creation of a common understanding and—where possible—common practices. A key aspect of the DMP is to make the information collected and produced Findable, Accessible, Interoperable and Re-usable (FAIR). In particular, the DMP identifies the main data to be generated within FirEURisk, outlining how data are used during the project as well as protocols and policies required for data exchange and sharing within and outside the consortium.

It is important to note that the DMP will be a living document, meaning that it will be further detailed, updated, amended and/or corrected during the whole lifespan of the project. Updated versions will be produced whenever significant changes arise, e.g., when new data formats are required to be defined or when new consortium policies on data exchange, sharing or security are enforced. The final version will be produced towards the end of the project, provisionally on M46.

1.2 Structure of the document

The present document is structured in seven sections. Section 2 provides an overview of the DMP. Suggestions for properly identified datasets and reporting their specifications are provided in Section 3. Section 4 details the processes that need to be followed for making the datasets FAIR. Sections 5, 6 and 7 deal with resource allocation, data security and ethical aspect issues, respectively. Additionally, the document is accompanied by four annexes. Annex A and B provide the forms for reporting the necessary accompanying information for FirEURisk products and datasets, respectively. Annex C includes an INSPIRE (Infrastructure for Spatial Information in the European Community) metadata example for a dataset as well as the metadata requirements for INSPIRE-compliant datasets and systems. Finally, Annex D includes a list of open access resources to support beneficiaries in making their research data openly accessible.

2 Data Management Plan overview

The FirEURisk DMP aims to provide a description and analysis of FAIR data management lifecycle for the data to be collected, processed and/or generated. To assist the beneficiaries to make their research data FAIR, this DMP has been prepared considering the EC template of the Guidelines on Data Management in Horizon 2020 (H2020 Programme 2016)¹. The template consists of a set of questions that beneficiaries should answer with a level of detail appropriate to the project. If a requested information does not apply for a given dataset, then N/A (Not Applicable) will be used, whereas if not available yet, then TBD (To Be Defined) will be used. As DMP is intended to be a living document, further details on currently unavailable information will be updated during project implementation.

The DMP will address for each dataset collected, processed and/or generated in the project the following elements:

- Data handling during and after the end of the project
- Type and format of data to be generated/collected
- Data documentation and metadata
- Data access and re-use policy
- Data storage and preservation
- Ethics and legal compliance

The DMB has also identified interactions with other tasks and associated deliverables, which are relevant or complement some aspects of the DMP. These include:

- **Task T4.1 - Integration within the project.** This is the task the DMP deliverables belong to (specifically, activity 'A4.1.2 - Interoperability and data management'). In addition, the deliverable 'D4.3 - Internal cooperative platform of models, data and knowledge' is also relevant and, generally, activity 'A4.1.3 - FirEURisk internal cooperative platform of models, data and knowledge' that D4.3 belongs to. Strictly speaking, the present document serves as input to the latter, since several decisions made in the DMP influence the internal platform's design (e.g., naming conventions, metadata). At the same time, though, the physical implementation of the data security protocols will be reported in D4.3.
- **Task T5.1 - Selection of products for PS scalability demos.** Task's 5.1 objective is the selection of the most adequate products to apply them at PS and ET scales. In particular, the deliverable 'D5.2 - Portfolio of demonstration products for PS and pan-European demos' will produce a publishable portfolio of selected products for PS, consisting of a 1-page definition of the product(s), including the technical components, the methodology, proved accuracy at DA and validation/testimonial of local end-users. The collection of input from the partners developing the various FirEURisk products has already begun in the context of WP5. The reported technical specifications serve as base for DMP's Product Inventory Form (see [Annex A](#)).
- **Task T7.3 - Research Ethics, Legal and Societal Management.** This task is devoted to the assessment and management of ethical and gender issues throughout the project, among which are the ethical and legal

¹ https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf

restrictions imposed to data collection and processing. The associated deliverables ‘D7.1 - Initial Ethics, Legal and Societal management Report’ and ‘D7.2 - Final Ethics, Legal and Societal management Report’ supplement the DMP with respect to those issues.

Collectively, this document has been prepared considering the information provided in the following guidelines and documents:

- H2020 Annotated Model Grant Agreement (H2020 Programme 2019)
- Guidelines to the Rules on Open Access to Scientific Publications & Open Access to Research Data in Horizon 2020 (H2020 Programme 2017)
- Guidelines on FAIR Data Management in Horizon 2020 (H2020 Programme 2016)
- DMP template for H2020 projects (H2020 Programme 2016)
- OpenAIRE Research Data Management Briefing Paper (OpenAIRE2020 2017)
- DCC Checklist for a Data Management Plan (DCC 2013)

3 Data summary

DMB’s role is to assist project partners in maintaining harmonised and consistent systems, data formats and organisation of datasets that are used or produced within FirEURisk. Three groups of datasets are identified:

1. **FirEURisk products.** This category includes all final data products that will be created by FirEURisk partners during the project and will be used for achieving the project’s objectives. They will be made available to project partners through the internal platform, with some of them also being selected for distribution outside the project through the open platform as standalone products.
2. **Internal FirEURisk variables/datasets.** This category includes intermediate or temporal datasets that will be produced during the project implementation and need to be exchanged between partners. Examples include satellite imagery (or timeseries of imagery), LiDAR data acquired during FirEURisk, intermediate land cover classifications, results from the analysis of questionnaires, etc. These datasets are generally intended for use by project partners and will hence be made available through the internal platform (or equivalent mechanism, e.g., appropriate web interfaces), although some of them may be elected to be published publicly as well.
3. **External datasets.** This category includes all datasets created outside FirEURisk and supplied to the project either by the partners or third parties. Examples are national/regional fire statistics, field data, demographic statistics, etc. These datasets will only be available to the FirEURisk partners via the internal platform unless a more restrictive sharing policy is requested, in which case it will be available only to some partners. Sharing via the open platform is permitted only after the explicit consent of the dataset provider or if a permissive redistribution policy is in effect.

The first step to ensure the interoperability of the data used in FirEURisk (irrespective of category) is to provide a summary description of each dataset, which includes at least:

- a) the product name,

D4.2 – Data Management Plan

- b) a short description,
- c) the type/format of the dataset,
- d) the origin and accessibility of input data used (where applicable),
- e) the intended use of the dataset, and
- f) the quality assurance process.

These are subsequently complemented with information regarding the dataset’s technical specification, including:

- a) expected size of the data,
- b) pre-processing processes and/or methodology (where applicable),
- c) whether the partners intend to use the execution engine of the internal platform (for FirEURisk products), and
- d) whether beneficiaries intend to upload the data to the internal platform (and how they expect to be hosted) or shared via other means (e.g., appropriate web interfaces).

Finally, each dataset is accompanied by information regarding the FAIR principles that are followed (naming conventions, sharing policies, etc.) and legal/ethical considerations, which are detailed in the following sections of this document.

For the first category of datasets, the collection of the summary descriptions of the different FirEURisk products has already started in WP5 (led by beneficiary GMV), within the context of activity ‘A5.1.2 - Selection of Fire Risk Assessment, Reduction and Adaptation products for demos’ and the associated deliverable ‘D5.2 - Portfolio of demonstration products for PS and pan-European demos’ (led by beneficiary REEFMC). The rest of the information is being collected by the DMB. Eventually, each FirEURisk product will compile the information of the Product Inventory Form reported in [Annex A](#). For the other two categories of data, the DMB is coordinating the collection of information shown in the Dataset Inventory Form in [Annex B](#). Eventually, all this structured information will be available via the internal FirEURisk platform and, more specifically, through its Data Catalogue component.

Finally, to facilitate data requests and sharing for data among partners, the DMB has prepared a simple dataset request form (Table 2), which has been distributed to all partners by FirEURisk’s Steering Committee. The form captures the most essential information and is aimed at primary datasets of the third category (external datasets). The objective is to minimise interactions regarding data requests with third parties (e.g., public authorities) and to have a single point of contact handling the request (at least at a first level) per pilot site (PS) and the whole European territory (ET). For the latter, a single beneficiary and contact person has been assigned for each PS and the ET. For the PS level, the beneficiary responsible for the PS has been assigned, whereas the DMB manager serves as the point of contact for ET-level data requests (Table 3).

Table 2: Dataset request form

Information	Explanation	Example
Partner(s) requesting the dataset	<i>Use the short names / acronyms.</i>	SomePartner
Contact person(s)	<i>Name & email of the person(s) that can be contacted for further information on the request.</i>	SomePartner <some@partner.org>
Required dataset description	<i>A few lines or a short paragraph describing the dataset's main characteristics and features of</i>	Archive of fire ignition data from pilot site.

	<i>interest (e.g., limitations of the methodology employed).</i>	
Purpose of use	<i>Briefly describe why the dataset is necessary and where it will be used.</i>	Generation of wildfire ignitions models (Activity A1.1.3)
Dataset specifications	<i>Detail any technical specifications the dataset must have, e.g., existence of certain variables.</i>	Fire ignition location coordinates, fire date, and fire cause. Forested and agricultural areas. It should include all fires, irrespective of area burnt.
Temporal coverage (range & frequency)	<i>Time period the dataset should cover and frequency (if periodic). Both desired & necessary values can be specified.</i>	Desired: last 30 years Minimum: last 10 years
Spatial coverage	<i>ET (European territory), named PS (pilot site), DA (demonstration area); use a longer textual description in case of something more complex.</i>	Minimum: PS-05-EL Desired: Whole Greece
Notes	<i>Any other comment you may have.</i>	

Table 3: Points of contact for dataset requests for each FirEURisk region of interest

Region	Partner	Role	Contact person
ET (European Territory)	AUTH	Data Management Board leader	Ioannis Gitas
PS-01-SE (Northern Europe - Sweden)	RISE	Point of contact for PS-01	Johan Sjöström
PS-02-DE.CZ.PL (Central Europe - Germany/Czech Rep./Poland)	TUD	Point of contact for PS-02	Christopher Marrs
PS-03-PT (Portugal)	ADAI	Point of contact for PS-03	Jorge Raposo
PS-04-ES (Spain)	UdL	Point of contact for PS-04	Aitor Ameztegui
PS-05-EL (Greece)	KEMEA	Point of contact for PS-05	Chrysoula Papathanasiou

4 FAIR data

FAIR data are data, metadata, and supporting infrastructure (e.g., search engines) that meet principles of findability, accessibility, interoperability, and reusability. The FAIRification process (Figure 1) for making data FAIR starts from the moment we gain access to the data; either these are generated, collected or received data.

The first step of the FAIRification process is the analysis of data. In this step, beneficiaries should investigate the origin of input data needed to carry out the analysis—especially if they derive from non-open sources—and assess the licenses to ensure that input data can be used and reused by others without restrictions. For intermediate or final products, this means to investigate their FAIR status: whether best practices for file naming are used, the format, semantics and proper linking of all input data required to produce the data. Similarly, on metadata level, FAIR status should be investigated: if proper standards are used, if they have assigned licenses as well as the status of the information for reusing and accessing the data.

The next step is to assign a license to the dataset. The chosen licensing scheme determines data reusability. Choosing a truly open license makes data available to a larger audience and makes the widest range of uses possible. Subsequently, ensure data is described by proper and rich metadata. Standardised and machine-readable metadata

are essential for the automatic discovery of datasets and services, so this is a crucial component of the FAIRification process² (Wilkinson *et al.* 2016). Improving metadata requires including details such as identification and provenance, keywords, quality and validity, license, copyright, and description of use conditions and access of data.

The final step is to deposit or publish the FAIR data so that they can be indexed by search engines and be accessed even if authentication and authorisation are required³.

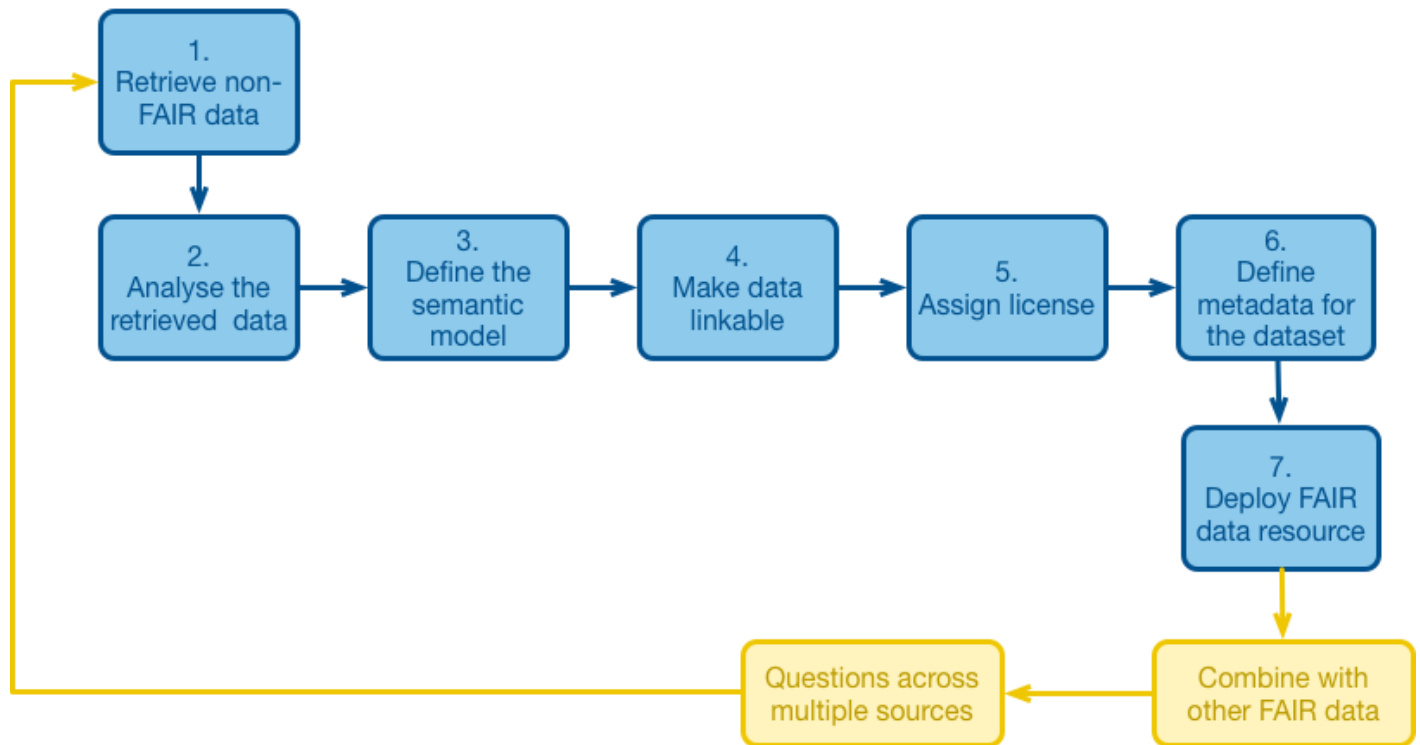


Figure 1: The scheme depicts the FAIRification process. The figure was reproduced from [GO FAIR](#) and focuses on data and metadata.

4.1 Making data findable, including provisions for metadata

This section of the dataset description forms includes a description of best practices for file naming, metadata and related standards, data repositories, keywords to be used and if a standard identification mechanism such as Digital Object Identifier (DOI) will be used.

4.1.1 File naming

The first step to identify, locate and use research data efficiently is file naming. A good file name functions as a unique identifier that provides descriptive information regarding the content, status and version of a file. This helps to

² <https://rwww.go-fair.org/fair-principles/>

³ <https://www.go-fair.org/fair-principles/fairification-process/>

D4.2 – Data Management Plan

organize, distinguish and share files effectively among different users. Proper file naming should be consistent, logical and predictable in order to be understood by everyone. Information to consider when developing a file name:

- Project name (here, FirEURisk).
- Deliverable identifier (e.g., D4.2) or task/activity identifier (e.g., T1.3, A1.1.4), which may be omitted if the dataset is intended for external use (i.e., the open FirEURisk platform or a public repository).
- Location/Study site (e.g., PS-01-SE for the Northern Europe / Sweden PS).
- Date (and time if needed); the ISO 8601-1:2019 basic formats for dates and times should be used (details provided below).
- Product short description (e.g., BurnedArea or BA).
- Data type short description (e.g., S2-20m).
- Version number (e.g., v01, v02 or v2021.01).
- Any other tag necessary for uniquely identifying the dataset from other datasets.

It is recommended to use dashes (-), underscores (_) or capitalize the first letter of each word to connect the above-mentioned information (tag elements) and avoid using spaces when developing a file name (since most operating systems require escape characters for correctly parsing filenames containing spaces in command line interfaces). Moreover, it is good practice to be consistent in the use of the tag element separation character. For example, if underscore (_) is used to connect the various tag elements, do not use underscore within any individual element (e.g., do not use 'S2_20m' for the data type description, use 'S2-20m' or 'S2.20m' instead). This facilitates computer programs splitting the filename string into its individual tags unambiguously. For dates and times, the basic formats of ISO 8601-1:2019 (ISO/TC 154 2019 p. 860) should be preferred, unless there is a valid reason to follow another scheme. The basic rules for that are:

- Years are represented by four digits as YYYY.
- Calendar dates are represented with two-digit months and day, as YYYYMMDD.
- Ordinal dates (year plus day of year) are represented by four-digit year followed by three-digit day of year, i.e., YYYYDDD.
- If time is necessary, use the basic format T[hh][mm][ss][.mmm][TZ], where:
 - [hh] is the zero-padded hour (00 to 23).
 - [mm] is the zero-padded minute (00 to 59).
 - [ss] is the zero-padded second (00 to 60).
 - [.mmm] is the decimal fraction (usually three decimal places, but can be any other number), applied to the lowest order time element present.
 - [TZ] is the zone designator. It is advised to always use Coordinated Universal Time (UTC), in which case [TZ] should be 'Z' (e.g., T0930Z). If local time should unavoidably be used, use the UTC offset designator (e.g., '-02' or '+0530'). Avoid omitting the [TZ] field, because local time is assumed in this case according to the standard.
 - All fields but [hh] are optional. For example, all the following are valid representations: T14Z (14h 00m 00s UTC), T1430Z (14h 30m 00s UTC), T12.5+02 (12h 30m 00s UTC+02:00), T113016.5Z (11h 30m 16.5s UTC), T235960Z (23h 59m 60s UTC, leap second addition date).

An example of a filename that respects the abovementioned rules is:

FirEURisk_A1.3.2_BA_S2-20m_PS-05-EL_20200724_v01.tif

In any case, the most important aspect is that the filenames adhere to the scheme that you report they follow.

4.1.2 Metadata

Data should be accompanied by proper and rich metadata, as they are essential for automatic discovery of datasets and services, therefore an essential component of the FAIRification process (Wilkinson *et al.* 2016). In order to be useful, the overall structure of the metadata needs to be standardised and machine-readable so that automated processes will be able to find, use, preserve and reuse data in the future.

One of the commonly used metadata standards for spatial resources (e.g., datasets, services, maps) is the ISO 19115/119/139. ISO 19139—with its latest version being ISO/TS 19139-1:2019 (ISO/TC 211 2019 p. 19139)—is an XML encoding of ISO 19115—current version being ISO 19115-1:2014 (ISO/TC 211 2014 p. 19115)— with additions from ISO 19119:2016 (ISO/TC 211 2016) to cover metadata for services that provide geospatial data. ISO 19115/119/139 provides information such as the identification, extent, quality, the spatial and temporal resolution, and the distribution of geographic data.

At European Union level and in order to ensure spatial data are compatible and usable, Member States follow the INSPIRE metadata guidelines⁴ that also comply with the ISO 19115/119/139 metadata standards. An example of an INSPIRE complied metadata scheme as well as an overview of the required metadata elements for dataset – dataset series and data services can be found in [Annex C](#).

The bibliographic metadata that identify scientific publications must be in a standard format and must include all of the following, according to FirEURisk’s Grant Agreement:

- a) the terms “European Union (EU)” and “Horizon 2020”;
- b) the name of the action, acronym and grant number;
- c) the publication date and length of embargo period if applicable, and
- d) a persistent identifier.

As the project progresses and data is identified and collected, information on metadata standards for non-spatial products will be outlined in subsequent versions of the DMP.

4.1.3 Data repository

Data will be deposited to the internal cooperative platform (activity A4.1.3 within WP4), planned to be delivered by M12. The platform will consist of a data portal, a code repository, and a wiki to exchange data, methods and models, and knowledge, respectively. This platform will be first developed and used within the project (internal cooperative platform) and then it will be exploited (public open platform) to disseminate results externally to European Commission Disaster Risk Management Knowledge Centre (DRMKC). The platform will use a multi-layered user interface providing different access levels (project-internal and open) and different frontends for scientific experts, end-users/authorities and general audience.

⁴ <https://inspire.ec.europa.eu/file/1557/download?token=UaQBcRvQ>

Partners seeking to share research with broader audiences, improve online discovery, increase citation rates and preserve data for the long-term, may additionally choose a certified online data repository. There are two general categories of repositories:

- a) a cost-free data repository (e.g., Zenodo, PANGAEA) that allows researchers to deposit both publications and data while providing tools to link them
- b) an institutional research data repository, featured in the project website.

The Registry of Research Data Repositories⁵ offers useful listings of repositories. The Open Access Infrastructure for Research in Europe (OpenAIRE)⁶ is the recommended entry point for researchers to determine what repository to choose and provides additional information and support on linking publications to underlying research data. Trusted online data repositories offer (a) safe storage, (b) possibility to assign a Digital Object Identifier (DOI) and make data citable and trackable, (c) assistance to create metadata and assign open licenses, (d) Open or closed/restricted access modes and (e) versioning / ability to easily update the dataset.

4.1.4 Identification mechanisms

A Digital Object Identifier (DOI) will be provided for each dataset so that it is findable, accessible and citable. Similar to the online data repositories, the internal cooperative platform is expected to assign a versioned DOI to the uploaded data, enabling users the possibility to update the shared material.

4.1.5 Search keywords

Search keywords to be provided once uploaded to internal platform or online data repository to increase findability and consequent use of each dataset.

4.2 Making data openly accessible

This section of the dataset description forms includes a description of how and which data will be made accessible and why some datasets cannot be made open due to possible legal, ethical or contractual considerations. A description of the potential software tools required to access the data is also included.

Article 29.2 of FirEURisk's Grant Agreement sets out detailed legal requirements on open access to scientific publications: *'each beneficiary must ensure open access (free of charge online access for any user) to all peer-reviewed scientific publications relating to its results'*. Beneficiaries must *'deposit a machine-readable electronic copy of the published version or final peer-reviewed manuscript accepted for publication in a repository for scientific publications. Moreover, the beneficiary must aim to deposit at the same time the research data needed to validate the results presented in the deposited scientific publications'*.

⁵ Registry of Research Data Repositories: <https://www.re3data.org/>

⁶ Open Access Infrastructure for Research in Europe (OpenAIRE): <https://www.openaire.eu/>

D4.2 – Data Management Plan

Openness of research data (data underlying publications, curated data and/or raw data) is addressed in article 29.3 of FirEURisk's Grant Agreement: Beneficiaries must '*a) deposit in a research data repository and take measures to make it possible for third parties to access, mine, exploit, reproduce and disseminate — free of charge for any user — the data, including associated metadata*', and '*b) provide information — via the repository — about tools and instruments at the disposal of the beneficiaries and necessary for validating the results (and — where possible — provide the tools and instruments themselves*'.

Based on the legal requirements of Articles 29.2–29.3, a number of options for sharing and accessing data through the cooperative platform or/and a data repository are foreseen:

- **Open Access.** This is the highly recommended option, which provides free access and rights to all data with no legal, ethical, or contractual considerations.
- **Partial Access:** This option refers to restrictions that only apply to a subset of the data, such as data that could jeopardize the achievement of the action's main objective. This option allows to openly share the remaining data, where access constraints do not apply. In this case, the DMP must contain the reasons for not giving access.
- **Embargoed Access.** This option refers to data underlying publications that, even if they have been deposited, open access will be provided once published in a scientific paper. In such case, information regarding the embargo period should be included in the metadata and the DMP.
- **Restricted Access.** This option—although not recommended—refers to legal, ethical, or contractual considerations to withhold access to research data such as personal or sensitive datasets. Access to these data needs to be requested and the depositor can decide whether and how the data will be accessed (this can be managed through a data repository). In this case, the DMP should contain information of how access is provided.
- **Closed Access.** This option may be suitable for highly sensitive data whose confidentiality is limited to consortium and EC services.

Different confidentiality levels are expected to be applied within FirEURisk:

- **Public:** This option applies to most FirEURisk (final) products.
- **Internal use:**
 1. **Confidential to partner:** This option applies when data are shared only between specific partners in the consortium (e.g., data involved in collaborative work, sharing of intermediate or incomplete products, data expected to be included in patent applications).
 2. **Confidential to consortium:** This option applies for data underlying publications that may not yet published in peer reviewed scientific papers or that are planned to be published before becoming publicly available.

As the project progresses and data is identified and collected, further information on making data openly accessible will be outlined in subsequent versions of the DMP. In specific, information on methods or software tools needed to access the data and how access will be provided in case there are restrictions.

4.3 Making data interoperable

This section of the dataset description forms details actions to facilitate data interoperability, including standards for formats, metadata vocabularies or ontologies of vocabularies.

Information related to the interoperability of FirEURisk datasets has been incorporated in [Annex A](#) (Product Inventory Form) and [Annex B](#) (Dataset Inventory Form). This includes documentation describing data generated within the project that will support the interpretation and reuse. In addition, proper and rich documentation of metadata in a standardised way will ensure that datasets can be correctly understood, interpreted, and re-used.

As the project progresses and data is identified and collected, further information on making data interoperable will be outlined in subsequent versions of the DMP. The use of a standard vocabulary will be investigated for all data types present in the various datasets to allow interdisciplinary interoperability.

4.4 Increase data reuse (through clarifying licenses)

This section of the dataset description forms describes how the data will be licensed to permit the widest reuse. It includes information regarding the period when data will be available for reuse and whether third parties will have the option to use the data, in particular after the end of the project. Description of quality assurance processes are also included.

Assigning a license is an important aspect of making sure data meet the reusable principle in FAIR data management. License assignment can be applied through the online data repositories when depositing data. Similarly, the cooperative platform is expected to automatically offer different licensing options to choose from. As stated in the Horizon2020 guidelines, one straightforward and effective way to enable third parties to access, mine, exploit, reproduce and disseminate (free of charge for any user) the research data is to attach either a Creative Commons Zero (CC0) or a Creative Commons Attribution (CC BY) license to the data (H2020 Programme 2017 p. 2). Creative Commons (CC) provides different license types among which CC0 and CC BY are truly open licenses:

- **Creative Commons Zero (CC0)**⁷. The most open license type, allowing data to be used by anyone, does not require to attribute the author, users can distribute, remix, adapt, and build upon the material in any medium or format and use the work commercially.
- **Creative Commons Attribution (CC BY)**⁸. Similar with CC0 but as stated in the name it requires to attribute the author.

⁷ <https://creativecommons.org/publicdomain/zero/1.0/>

⁸ <https://creativecommons.org/licenses/by/4.0/>

To make reuse as likely as possible it is recommended to choose one of the abovementioned licenses, which makes data available to a larger audience and permits the widest range of uses possible. In case partners wish to further limit access to the uploaded data, alternative license options provided by CC are:

- **Creative Commons Attribution Share-Alike (CC BY-SA)⁹**. Allows users to distribute, remix, adapt and build upon the material in any medium or format, and can be used commercially. However, it requires (a) giving credit to the author and (b) license the modified material under the same license as the original (or a license listed as compatible).
- **Creative Commons Attribution-NonCommercial (CC BY-NC)¹⁰**. Allows users to distribute, remix, adapt and build upon the material in any medium or format. However, (a) it requires giving credit to the author and (b) only non-commercial uses of the work are permitted.
- **Creative Commons Attribution-NonCommercial-ShareAlike (CC BY-NC-SA)¹¹**. Allows users to distribute, remix, adapt and build upon the material in any medium or format. However, (a) it requires giving credit to the author, (b) only non-commercial uses of the work are permitted and (c) you must license the modified material under the same license as the original (or a license listed as compatible).
- **Creative Commons Attribution-NoDerivatives (CC BY-ND)¹²**. Allows users to copy and distribute the material in any medium or format and can be used commercially. However, (a) it requires giving credit to the author and (b) forbids the distribution if you remix, transform, or build upon the material.
- **Creative Commons Attribution-NonCommercial- NoDerivatives (CCBY-NC-ND)¹³**. Allows users to copy and distribute the material in any medium or format. However (a) it requires giving credit to the author, (b) forbids the distribution if you remix, transform or build upon the material and (c) only non-commercial uses of the work are permitted.

To ease the selection of the right license to attach to a dataset or software, partners can use the License Selector Tool¹⁴. The tool was developed by a team led by Pawel Kamocki (researcher IT Law at IDS Mannheim, Germany) and is by no means mandatory to use, but purely aimed here as a support¹⁵. The link leads in a wizard-like interface with questions that will narrow down the possibilities of license choice to attach to your work.

⁹ <https://creativecommons.org/licenses/by-sa/4.0/>

¹⁰ <https://creativecommons.org/licenses/by-nc/4.0/>

¹¹ <https://creativecommons.org/licenses/by-nc-sa/4.0/>

¹² <https://creativecommons.org/licenses/by-nd/4.0/>

¹³ <https://creativecommons.org/licenses/by-nc-nd/4.0/>

¹⁴ <https://ufal.github.io/public-license-selector/>

¹⁵ <https://eudat.eu/services/userdoc/license-selector>

All data that will be stored in online data repositories can remain reusable for the lifetime of the repository. The lifetime of the FirEURisk cooperative platform will be defined during the course of the project. Data reusability by third parties during or after the end of the project, as well as data validation tools will be defined clearly through the project implementation. As the project progresses and data is identified and collected, further information on increasing data reuse will be outlined in subsequent versions of the DMP.

5 Allocation of resources

FAIR data management is part of the current deliverable D4.2 and all related costs that will occur during the project's implementation will be covered by the project budget. This involves a specific number of person-months needed to set up and perform the data collection and analysis of this activity, costs for storing and sharing the data, as well as costs for open access publications.

No further costs for data archiving are expected, since both the cooperative platform and data repositories offer free data archiving. Eventual long-term storage and preservation costs that may be incurred will be discussed among consortium members. Since such costs depend on data size—for datasets either collected or produced—it will be estimated towards the end of the project and will be reported at the final revision of the DMP.

6 Data security

Data deposited to the cooperative platform or/and a certified data repository will be protected with the server's security protocol. Each partner shall bear the responsibility of backing up data that will be openly shared through a data repository, whereas in the case of the cooperative platform a local backup system is expected during the project lifespan. The possibility for long-term preservation of data in the FirEURisk cooperative platform will be discussed among consortium members.

Special attention will be given to the security of sensitive data based on FirEURisk's Grant Agreement Articles related to security and confidentiality (Articles 36 and 37) and the processing of personal data (Article 39). FirEURisk has a dedicated work package (WP8) for privacy constraints on data, a Project Security Officer (PSO) to advise and review all security and sensitive material and matters of the project, and a dedicated board—the Societal, Ethical, Gender and Legal Board (SEGLAB)—to monitor the data protection of the project's research activities.

7 Ethical aspects

FirEURisk has a dedicated task 'T7.3 - Research Ethics, Legal and Societal Management' within WP7 to ensure that ethical requirements are met for all research undertaken in the project. Special attention will be paid by the FirEURisk Consortium to its compliance with the H2020 research ethics standards and EU/international codes of scientific conduct. Moreover, SEGLAB will provide guidance, advice and recommendations on the solutions and results of the

project and to monitor the data protection and ethical issues of the project's research activities. The Consortium will ensure the compliance of the performed activities with the basic ethical principles that are laid down in the European Charter of Fundamental Rights with a focus on protection of privacy and protection of personal data of the individual. The exact procedures to be followed will be detailed in the associated deliverables 'D7.1 - Initial Ethics, Legal and Societal management Report' (expected to be delivered on M12) and 'D7.2 - Final Ethics, Legal and Societal management Report' (expected to be delivered on M42).

8 Conclusions

Deliverable D4.2 serves as the first version of the DMP, intended to function as a guiding document to ensure good data management and to facilitate the creation of a common understanding and—where possible—common practices. The DMP will be a living document, which will further detailed, updated, amended and/or corrected during the whole lifespan of the project. At the very minimum, it will be updated in M18, M36 and M46 of the project, with potential additional updates whenever significant changes arise, e.g., when new data formats are required to be defined or when new consortium policies on data exchange, sharing or security are enforced. Eventually, the DMP aspires to integrate all data-related progress of the project and to further specify some aspects that are currently defined in general terms. For example, Open Geospatial Consortium (OGC) standards for data and metadata will most likely be adopted for most spatial datasets and products, to further increase interoperability with contemporary platforms within the INSPIRE framework. All these decisions will be reported in the updated versions of the DMP.

9 References

- DCC (2013) Checklist for a Data Management Plan, v4.0. Edinburgh: Digital Curation Centre, https://www.dcc.ac.uk/sites/default/files/documents/resource/DMP/DMP_Checklist_2013.pdf.
- H2020 Programme (2016) Guidelines on FAIR Data Management in Horizon 2020, Version 3.0. European Commission, Directorate-General for Research & Innovation, https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf.
- H2020 Programme (2017) Guidelines to the Rules on Open Access to Scientific Publications and Open Access to Research Data in Horizon 2020, Version 3.2. European Commission, Directorate-General for Research & Innovation, https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-pilot-guide_en.pdf.
- H2020 Programme (2019) AGA – Annotated Model Grant Agreement, Version 5.2. European Commission, Directorate-General for Research & Innovation, https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/amga/h2020-amga_en.pdf.

D4.2 – Data Management Plan

ISO/TC 154 (2019) ISO 8601-1:2019. Date and time — Representations for information interchange — Part 1: Basic rules. *International Organization for Standardization*.

<https://www.iso.org/cms/render/live/en/sites/isoorg/contents/data/standard/07/09/70907.html>.

ISO/TC 211 (2014) ISO 19115-1:2014. Geographic information — Metadata — Part 1: Fundamentals. *International Organization for Standardization*.

<https://www.iso.org/cms/render/live/en/sites/isoorg/contents/data/standard/05/37/53798.html>.

ISO/TC 211 (2016) ISO 19119:2016. Geographic information — Services. *International Organization for Standardization*.

<https://www.iso.org/cms/render/live/en/sites/isoorg/contents/data/standard/05/92/59221.html>.

ISO/TC 211 (2019) ISO/TS 19139-1:2019. Geographic information — XML schema implementation — Part 1: Encoding rules. *International Organization for Standardization*.

<https://www.iso.org/cms/render/live/en/sites/isoorg/contents/data/standard/06/72/67253.html>.

OpenAIRE2020 (2017) OpenAIRE Research Data Management Briefing paper: Understanding Research Data Management. OpenAIRE2020, <https://www.openaire.eu/briefpaper-rdm-infonoads/view-document>.

Wilkinson MD, Dumontier M, Aalbersberg IJJ, Appleton G, Axton M, Baak A, Blomberg N, Boiten J-W, da Silva Santos LB, Bourne PE, Bouwman J, Brookes AJ, Clark T, Crosas M, Dillo I, Dumon O, Edmunds S, Evelo CT, Finkers R, Gonzalez-Beltran A, Gray AJG, Groth P, Goble C, Grethe JS, Heringa J, 't Hoen PAC, Hooft R, Kuhn T, Kok R, Kok J, Lusher SJ, Martone ME, Mons A, Packer AL, Persson B, Rocca-Serra P, Roos M, van Schaik R, Sansone S-A, Schultes E, Sengstag T, Slater T, Strawn G, Swertz MA, Thompson M, van der Lei J, van Mulligen E, Velterop J, Waagmeester A, Wittenburg P, Wolstencroft K, Zhao J, Mons B (2016) The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data* **3**, 160018. doi:10.1038/sdata.2016.18.

10 Annex A: Product Inventory Form

Table 4 shows FirEURisk’ Product Inventory Form, which aims to collect all information necessary to define, maintain and share FirEURisk’s products, as defined in Section 3 of this document. The information of the first part (Product Description; light blue cells) is being collected from WP5 in the context of Activity ‘A5.1.2 - Selection of Fire Risk Assessment, Reduction and Adaptation products for demos’ (i.e., the Product Portfolio Template). The rest of the information is being collected by the DMB.

Table 4: Product Inventory Form

1. Product Description	
Product Name	< 20 words
Product Description	Objective / Methods to develop the product
Type of products	Indicate if the product is Geospatial / Document / Training, and provide legend units if applicable
Main applications	Prevention / Mitigation / Preparedness / Response / Recovery / Restoration
Update Frequency	Product: Days, Months, Year
Accuracy	Product
Output Format	.tif, .hdf, .docx, .pdf, video, .gdb, .db, etc.
Main Input data	Specify the datasets needed to carry out the analysis and create the product: satellite images HR, VHR, LiDAR, meteorological data, field data. In case of documents or training materials please refer to the pertinent sources.
Temporal coverage	Product
Archive Length	This item refers to geospatial products mainly. For temporal scalability please indicate if the main input data enjoys a lengthy archive in order to offer a broader temporal analysis.
Input data ownership: Public / Private	If input datasets are not publicly available, please indicate the name of the Institution holding the data.
Input data accessibility	Online / by request
Input data cost	Free / Non-free
Spatial resolution	Product: m, km, ha
Minimum Mapping Unit	m, km, ha
Validation	Summary of the producer's internal quality control / validation or with external stakeholders. Other products such as handbooks, training materials can be validated through satisfaction questionnaires
Local test site	Indicate the Local Demonstration Area if applicable
Local End-User	Details i.e. point of contact- email of potential users interested in your products
Regional Demonstration site for upscaling	Indicate the Pilot Site/s where your products will be implemented or indicate EU Territory
Regional data	Indicate specific data at regional scale that you do not have and if you need to contact local agencies
Regional End user	Details i.e. point of contact - email of potential users interested in your products

Associated Deliverable	DX.Y	
Product Overview	Example / image if available	
2. Data Specifications		
Input data description	Specify whether the data will be produced within the FirEURisk project (indicate the partner and project activity associated with data production) or existing data (generated outside the project) will be re-used (indicate the provider and whether data are already available).	
Expected size	Indicate the approximate expected size of the product, e.g., ~25MB, < 2 GB, etc.	
Execution engine	Do you intend to use the execution engine feature of FirEURisk's internal platform? If yes, please specify if your algorithm(s) relies on simple (potentially spatially) operations on the input data or if a dedicated application is required to run for obtaining the result.	
Programming language	Report which programming language(s) (if any) will be used for producing the dataset, especially if you intend to use the platform's execution engine.	
Internal platform data hosting	Will the data be uploaded to the internal platform?	Yes/No. If no, don't fill the next two rows
	In what way are they expected to be hosted?	e.g., interactive GIS map, passive dataset for download; state other way
	How do you expect they be uploaded to the platform?	e.g., manually through a platform-side interface, via FTP, through a dedicated web API, etc. Also include how the respective metadata are expected to be uploaded.
3. FAIR Data		
3.1. Making Data Findable & Interoperable		
Metadata	Format	e.g., ISO 19115, ISO/TS 19139, INSPIRE, etc. Use "TBD by DMB" if you don't have any metadata already and want the Data Management Board to investigate & propose the most fitting format for the dataset. Use "TBD" if you intend to determine yourself, but at a later stage.
	Vocabularies	Describe if any format metadata vocabulary is followed. Use "TBD by DMB" if you don't know what vocabularies to use and want the Data Management Board to investigate & propose the most fitting ones. Use "TBD" if you intend to determine yourself, but at a later stage.
	DOIs	Indicate how data are made identifiable, if a standard permanent identifier assignment scheme is used (i.e. ARK, DOI, PURL, URN, MODA ...)
Will you share data via a repository?	If yes, specify which data repository (e.g. Zenodo, PANGAEA)	
3.2. Making Data Accessible & Reusable		

D4.2 – Data Management Plan

Technologies / software required for accessing the data	<i>Report if specific software/technology is required for accessing the dataset (e.g. language: R/ software: Rstudio, GIS software, etc.)</i>
Access policy	<i>Identify any restrictions of access (e.g., Open Access, Partial Access, Embargoed Access, Restricted Access, Closed Access)</i>
Confidentiality level	<i>Specify confidentiality level, e.g., Internal use (consortium or specific members of it), Public - everyone can see the information, etc.</i>
When will the data become available for re-use?	<i>Identify any probation period that will be applied.</i>
Length of time data will remain re-usable or continue to be produced	<i>e.g., Permanently, 5 years following termination of the project</i>
License agreement type	<i>Please make a selection from the drop-down list. In case of other license agreement type, please select option "other" and fill in the type.</i>
Will/can the data be re-used by third parties?	<i>If yes, report any restrictions (e.g., specific type of third parties only)</i>
Quality assurance processes	<i>Identify tools required for data validation</i>
4. Allocation of Resources	
Cost for achieving FAIR data	<i>If possible, provide an estimate of the cost for making the data reusable, for the time period of expected reuse</i>
5. Ethical Aspects	
Ethical Issues	<i>Report any ethical issues that might arise from the use or distribution of the data</i>

11 Annex B: Dataset Inventory Form

Table 5 shows FirEURisk’ Dataset Inventory Form, which aims to collect all information necessary to define a dataset that will be stored or shared within FirEURisk. The datasets targeted here belong to the last two categories defined in Section 3, i.e., internal FirEURisk variables/datasets and external datasets. The information will be collected by the DMB to facilitate the design of the internal platform, ease available data exchange between partners and avoid unnecessary work by different partners.

Table 5: Dataset Inventory Form

Information	Explanation
Dataset name	<i>A short descriptive title of the dataset (<20 words, if possible)</i>
Dataset description	<i>A few lines or a short paragraph describing the dataset's main characteristics and features of interest (e.g., limitations of the methodology employed).</i>
Partner(s) responsible	<i>Use the short names / acronyms.</i>
Contact person(s)	<i>Name & email of the person(s) that can be contacted for further information on the dataset.</i>
Temporal coverage (range & frequency)	<i>Time period the dataset covers and frequency (if periodic).</i>
Spatial coverage	<i>ET (European territory), named PS (pilot site); use a longer textural description in case of something more complex.</i>
Data filetype	<i>Excel file, raster file(s), shapefile, etc.</i>
Layers / Variables	<i>Describe any layers (e.g., for multiband raster files or vector files with multiple layers) or variables (e.g., for tabular / database datasets), if needed. Leave empty if you instead provide a linked document describing the dataset structure below.</i>
Linked documents / resources	<i>Identify and briefly describe any document or other resource accompanying the dataset (e.g., a word document describing the dataset's structure & files, metadata, etc.).</i>
Accuracy / Validation	<i>Report any accuracy assessment result you may already have, even informal/qualitative one (e.g., "accuracy of coordinates questionable in certain cases). Leave empty if not relevant, unknown or in doubt.</i>
Restrictions on use / license	<i>By default, the dataset is assumed that can be used internally only, by all partners. In this case, leave this cell empty. Otherwise, report other restrictions of use (e.g., can only be used by certain partners after formal request, existing license, etc.)</i>
Redistribution policy	<i>By default, it is assumed that the dataset can be uploaded to the internal FirEURisk platform. In this case, leave this cell empty. Otherwise, describe the redistribution policy, e.g., if it should not be uploaded to the platform or if it can also be distributed through the public platform (identify any necessary attributions/restrictions in this case, if presently known).</i>
Metadata	<i>Is the dataset accompanied by metadata? If yes, what format and vocabulary is being used?</i>
Expected dataset size	<i>Report the estimated dataset storage size, if known (even if the estimate is just order of magnitude / rough estimate).</i>

D4.2 – Data Management Plan

Alternative hosting options	<i>Especially for large datasets, report if there is an option to host the data locally and expose them through a web interface (e.g., (S)FTP, web API), instead of storing them to the internal platform directly.</i>
Platform name	<i>The dataset name (and filename template) that will be used when creating the dataset entry to the internal FirEURisk platform. Leave empty and WP4 will assign an appropriate name. If you wish, you can suggest a name following similar rules with that shown in next column's example, but WP4 may eventually decide to change it.</i>

12 Annex C: INSPIRE Metadata Example & Requirements

This annex presents an example of the INSPIRE metadata for a spatial dataset (Table 6), as well as the required and more frequently used metadata elements when following the INSPIRE directive for defining metadata (Table 7).

Table 6: INSPIRE metadata dataset example¹⁶

Identification	
Resource Title	SPI: Standardized Precipitation
Resource Abstract	The Standardized Precipitation Index (SPI-n) is a statistical indicator comparing the total precipitation received at a particular location during a period of n months with the long-term rainfall distribution for the same period of time at that location. SPI is calculated on a monthly basis for a moving window of n months, where n indicates the rainfall accumulation period, which is typically 1, 3, 6, 9, 12, 24 or 48 months. The corresponding SPIs are denoted as SPI-1, SPI-3, SPI-6, etc. In order to allow for the statistical comparison of wetter and drier climates, SPI is based on a transformation of the accumulated precipitation into a standard normal variable with zero mean and variance equal to one. SPI results are given in units of standard deviation from the long-term mean of the standardized distribution. In 2010, WMO selected the SPI as a key meteorological drought indicator to be produced operationally by meteorological services.
Resource Type	dataset
Resource Locator	http://edo.jrc.ec.europa.eu/chm/
Resource Unique Identifier	e24425e1-b073-11e1-9105-0017085a97ab
Resource language	eng
Classification of data and services	
Topic category	climatologyMeteorologyAtmosphere
Keyword	
Keyword value	Atmospheric conditions
Geographic Location	
Geographic Location Bounding Box:	+ West:-15.00 + East:45.00 + North:35.00 + South:72.00
Temporal reference	
Date of publication	2012-02-20
Quality and validity	
Lineage	Computation of the SPI involves fitting a probability density function to a given frequency distribution of precipitation totals for a station or grid point and for an accumulation period. We use the gamma probability density function. The statistics for the frequency distribution are calculated on the basis of a reference period of at least 30 years. The parameters of the probability density function are then used to find the cumulative probability of the

¹⁶ [INSPIRE Metadata Implementing Rules: Technical Guidelines based on EN ISO 19115 and EN ISO 19119](#), ANNEX A, page 77

D4.2 – Data Management Plan

	78 observed precipitations for the required month and temporal scale. This cumulative probability is then transformed to the standardised normal distribution with mean zero and variance one, which results in the value of the SPI. The SPI values are computed using the so-called MARS weather stations as rainfall input data. Refer the MARS weather catalogue for characteristics of the quality and quantity of these data. We only rely on the rainfall data input.
Spatial Resolution	0.25 m
	Conformity
Specification: + title	+ COMMISSION REGULATION (EU) No 1089/2010 of 23 November 2010 implementing Directive 2007/2/EC of the European Parliament and of the Council as regards interoperability of spatial data sets and services
+ publication date:	+ 2010-12-08
Degree:	not evaluated
	Constraints related to access and use
Limitation on public access	no limitations
Condition applying to access and use	Reproduction for non-commercial purposes is authorised, provided the source is acknowledged. Commercial use is not permitted without prior written consent of the JRC. Reports, articles, papers, scientific and non-scientific works of any form, including tables, maps, or any other kind of output, in printed or electronic form, based in whole or in part on the data supplied, must contain an acknowledgement of the form: Data re-used from the European Drought Observatory (EDO) http://edo.jrc.ec.europa.eu . The SPI data were created as part of JRC's research activities. Although every care has been taken in preparing and testing the data, JRC cannot guarantee that the data are correct; neither does JRC accept any liability whatsoever for any error, missing data or omission in the data, or for any loss or damage arising from its use. The JRC will not be responsible for any direct or indirect use which might be made of the data. The JRC does not provide any assistance or support in using the data.
	Responsible Organisation
Responsible party + organisation: + e-mail + Responsible party role	+ European Commission, Joint Research Centre + ies-contact@jrc.ec.europa.eu + custodian
	Metadata on metadata
Metadata point of contact + organisation: + e-mail	+ European Commission, Joint Research Centre + ies-contact@jrc.ec.europa.eu
Metadata date	2012-02-20
Metadata language	eng

Table 7: INSPIRE metadata requirements¹⁷

INSPIRE Metadata Requirements		
Element Name	INSPIRE Obligation for dataset – dataset series	INSPIRE Obligation for dataset services
Resource title	Mandatory	Mandatory
Resource abstract	Mandatory	Mandatory
Resource type	Mandatory	Mandatory
Resource locator	Mandatory if a URL is available to obtain more information on the resources and/or access related services	Conditional, mandatory if linkage to service is available
Coupled resource	-	Conditional, mandatory if linkage to data sets on which the service operates are available.
Spatial data service type	-	Mandatory
Unique resource identifier	Mandatory	-
Resource language	Mandatory if the resource includes textual information.	-
Topic category	Mandatory	-
Keyword value	Mandatory	Mandatory
Originating controlled vocabulary	Conditional: Mandatory for each keyword if the keyword value originates from a controlled vocabulary	Conditional: Mandatory for each keyword if the keyword value originates from a controlled vocabulary
Geographic bounding box	Mandatory	Conditional, mandatory for services with an explicit geographic extent.
Temporal reference	At least one of Temporal extent, Date of publication, Date of last revision or Date of creation must be given	At least one of Temporal extent, Date of publication, Date of last revision or Date of creation must be given.
Temporal extent	Conditional	Conditional
Date of publication	Conditional	Conditional
Date of last revision	Conditional	Conditional
Date of creation	Conditional	Conditional
Lineage	Mandatory	-
Spatial resolution	Mandatory if an equivalent scale or a resolution distance can be specified	Mandatory when there is a restriction on the spatial resolution for this service.
Conformity	Mandatory	Mandatory
Specification	Mandatory for each conformity statement.	Mandatory for each conformity statement.
Degree	Mandatory for each conformity statement.	Mandatory for each conformity statement.

¹⁷ [Technical Guidance for the implementation of INSPIRE dataset and service metadata based on ISO/TS 19139:2007](#), ANNEX B, page 112 - 115

D4.2 – Data Management Plan

Conditions applying to access and use	Special values for unknown conditions or no applying conditions may be used.	Special values for unknown conditions or no applying conditions may be used.
Limitations on public access	Special value for no limitations may be used.	Special value for no limitations may be used.
Responsible organisation	Mandatory	Mandatory
Responsible party	Mandatory for each responsible organisation.	Mandatory for each responsible organisation.
Responsible party role	Mandatory for each responsible organisation.	Mandatory for each responsible organisation.
Metadata point of contact	Mandatory	Mandatory
Metadata date	Mandatory	Mandatory
Metadata language	Mandatory	Mandatory
Category	-	Conditional, mandatory for an Invocable Spatial Data Service
Coordinate Reference System	Mandatory	Mandatory if relevant for an Interoperable Spatial Data Service in order to comply with Annex V of [Regulation 1089/2010].
Temporal Reference System	Mandatory only if a non-default temporal reference system (i.e. Gregorian Calendar or the Coordinated Universal Time) is used	-
Encoding	Mandatory	-
Topological consistency	Mandatory only if the data set includes types from the Generic Network Model and does not assure centreline topology (connectivity of centrelines) for the network	-
Character Encoding	Mandatory only if the data set is not using UTF-8 encoding	-
Spatial representation type	Mandatory	-
Quality of service	-	Mandatory for an Interoperable Spatial Data Service. Three criteria for minimum quality of service shall be given to comply with Annex VI of [Regulation 1089/2010]: Availability, Performance and Capacity.
Invocation metadata	-	Mandatory for a Harmonised Spatial Data Service in order to comply with Annex VII of [Regulation 1089/2010]

13 Annex D: Online open access resources

This annex collects a list of open access resources to support beneficiaries in making their research data openly accessible. Most of them have been reversed in the rest of this document, but are collected here for convenience:

- ❖ Horizon 2020 Annotated Grant Agreement:
https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/amga/h2020-amga_en.pdf
- ❖ Guidelines to rules on Open Access to Scientific Publications & Open Access to Research Data in Horizon 2020:
https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-pilot-guide_en.pdf
- ❖ ZENODO Open-Access Data Repository:
<https://zenodo.org/>
- ❖ Registry of Research Data Repositories:
<https://www.re3data.org/>
- ❖ Directory of Open-Access Repositories:
<http://www.opendoar.org/>
- ❖ INSPIRE Metadata Implementing Rules: Technical Guidelines based on EN ISO 19115 and EN ISO 19119:
<https://inspire.ec.europa.eu/documents/inspire-metadata-implementing-rules-technical-guidelines-based-en-iso-19115-and-en-iso-1>
- ❖ INSPIRE metadata validator:
<https://inspire.ec.europa.eu/validator/about/>
- ❖ License Selector Tool:
<https://ufal.github.io/public-license-selector/>
- ❖ Creative Commons Licenses:
<https://creativecommons.org/about/ccllicenses/>
- ❖ FAQs on open access to scientific publications:
<https://ec.europa.eu/info/funding-tenders/opportunities/portal/screen/support/faq;categories=;programme=null;actions=;keyword=oapubsH2020>
- ❖ FAQs on research data management and open research data:
<https://ec.europa.eu/info/funding-tenders/opportunities/portal/screen/support/faq;categories=;programme=null;actions=;keyword=ordH2020>