



SCORPION

Cost effective robots for smart
precision spraying

Cost effective robots for smart precision spraying

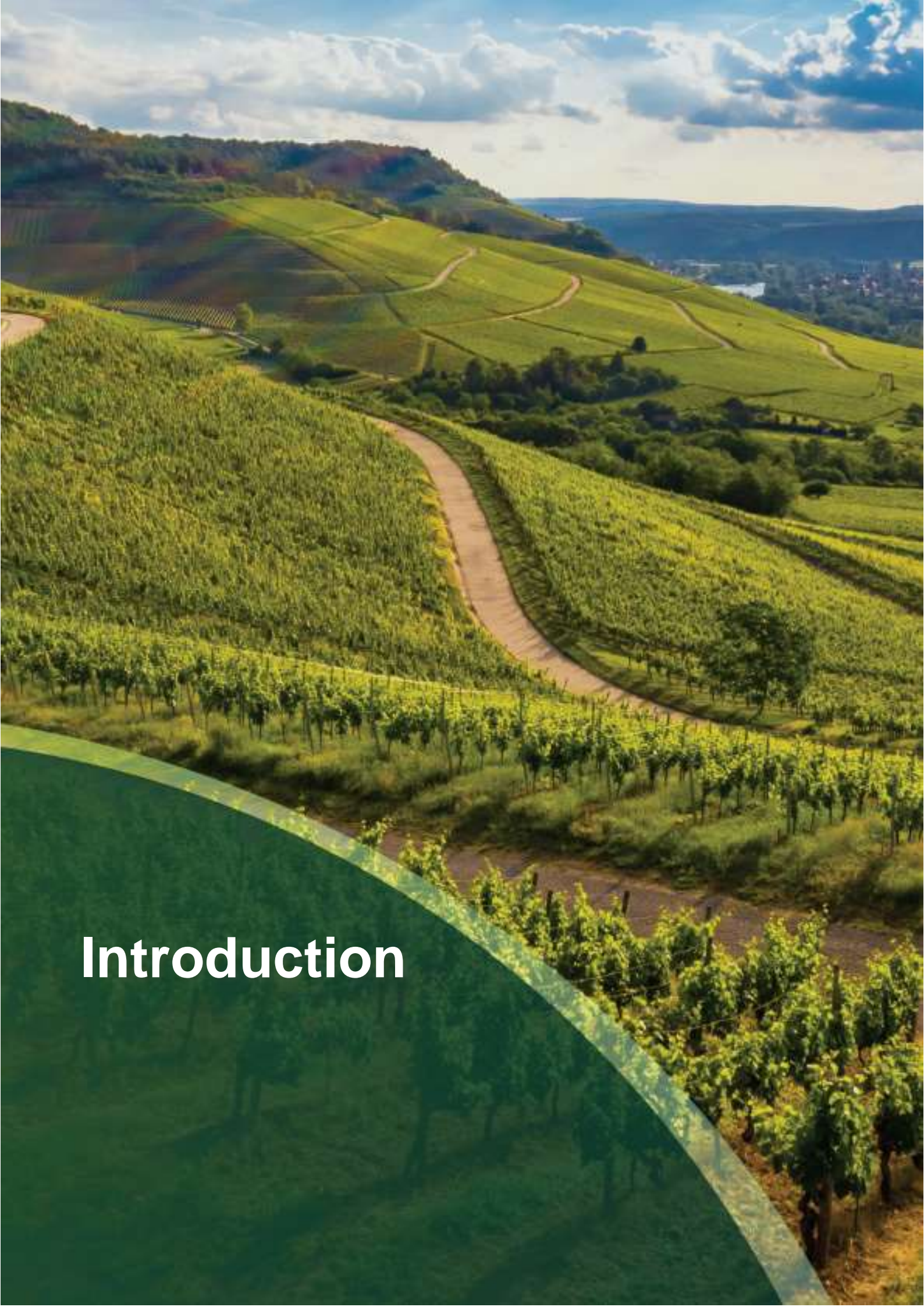
D6.2 First version of the data Management Plan



Project Title	Cost effective robots for smart precision spraying			
Horizon 2002 Call	SU-SPACE-EGNSS-3-2019-2020 - EGNSS applications fostering societal resilience and protecting the environment			
Type of Action	Innovation Action			
Project Start date	1st January 2021			
Project Duration	36 months			
Project URL	http://scorpion-h2020.eu			
Document URL	-			
EU Project Officer	Joaquin Reyes Gonzalez			
Project Coordinator	Filipe Neves dos Santos (INESC TEC)			
Deliverable	D6.2 First version of the data Management Plan			
Work Package	WP6 - Pilots, Dissemination and Communication Activities			
Date of Delivery	Contractual Date:	M4	Actual date:	M10
Type	Other			
Dissemination level	Public			
Lead Beneficiary	INESC TEC			
Lead Authors	Filipe Santos	Email:	fbsantos@inesctec.pt	
		Phone:	+ 351 91 61 56 516	
Other Authors	Marta Godinho (SPI)			
Reviewer (s)	Danilo Rabino (STEMS) Tatiana Martins (INESC TEC) Ana Vaz (IPN) Ard Nieuwenhuizen (WR) Victor Tirado (INNOVI) Danilo Rabino (STEMS)			
Keywords	Data Management; Dissemination			

Index

1. Introduction.....	5
1.1 Scope of the document.....	6
1.2 Deliverable’s structure	6
2. Data summary	8
2.1 Data Management Plan (DMP) guiding principles.....	13
3. FAIR Data	16
3.1 Making data findable, including provisions for metadata	16
3.2 Making data openly accessible	18
3.3 Making data interoperable	20
4. Allocation of resources	23
5. Data security	25
6. Ethical aspects.....	27
Annex 1 - Data Inventory Table.....	29



Introduction

1. Introduction

The Open Research Data Pilot of the European Commission enables open access and reuse of research data generated by the Horizon 2020 projects. The main pillars of the Open Research Data Pilot are the following: developing a Data Management Plan (DMP) and providing open access to research data from funded projects.

In order to adhere to the Pilot, the research consortia under H2020 have to:

- Develop (and keep up to date) a Data Management Plan (DMP);
 - Deposit data in a research data repository.
 - Ensure that third parties can freely access, mine, exploit, reproduce and disseminate these data.
 - Provide related information and identify (or provide) the tools needed to use the raw data to validate the research.
- The Pilot applies to:
 - The data (and metadata) needed to validate results in scientific publications.
 - Other curated and/or raw data (and metadata).

The DMP will function as a guiding document to ensure good data management throughout the lifecycle of the project to make the information collected and produced Findable, Accessible, Interoperable and Re-usable (FAIR).

1.1 Scope of the document

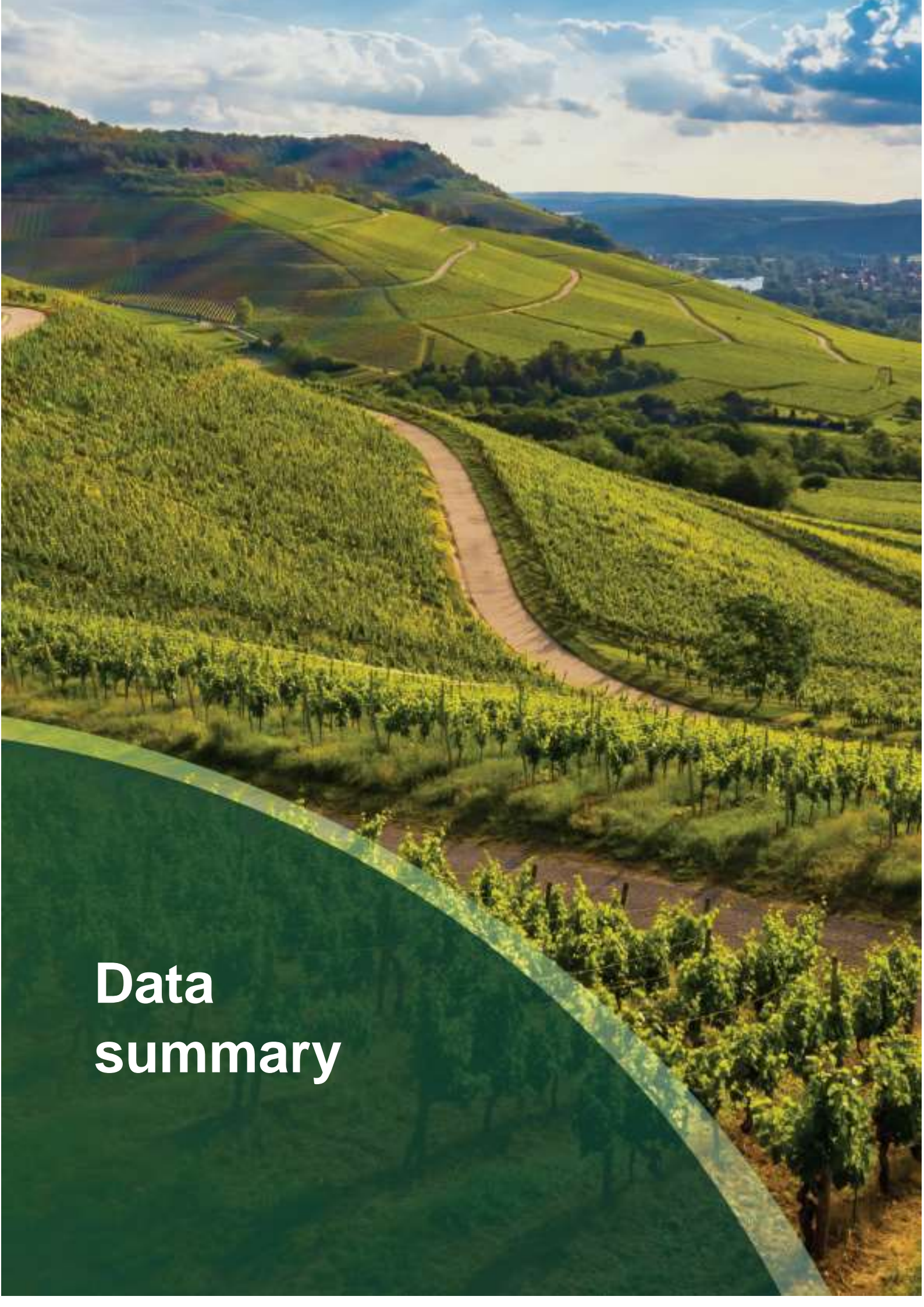
The Data Management Plan (DMP) describes the data management life cycle for the data to be collected, processed and/or generated by SCORPION project, as a Horizon 2020 project. It identifies procedures and minimum requirements to collect, store, analyse, and publish data generated during the project with the purpose to making them findable, accessible, interoperable and re-usable (FAIR), both throughout the lifecycle of SCORPION and after the end of the project.

The DMP is a living document and will be regularly updated, when necessary, by informing all project partners about the changes made.

1.2 Deliverable's structure

The document is structured following the guidelines of H2020 programme on FAIR Data Management in Horizon 2020, including the following information:

- ✓ Data Management Plan (DMP) guiding principles and strategy;
- ✓ Description of SCORPION type of data;
- ✓ Description of FAIR DATA characteristics including DMP Review Process & data inventory;
- ✓ Allocation of resources;
- ✓ Data Security;
- ✓ Ethical Aspects.



Data summary

2. Data summary

The project will collect different types of data:

- Questionnaires and inquiries,
- papers,
- technical reports,
- pictures,
- graphs,
- figures,
- videos,
- In addition, raw data from sensors will also be collected.

These data are important to replicate and compare results with other projects in the field. Valuable information produced by researchers in SCORPION project will:

**Be shared freely on Zenodo repository set up by the
OpenAIRE project after a total agreement among the
entire Consortium.**

This includes the system testing data, scientific publications as well as the technical data needed to validate the results presented in the deposited scientific publications, known as “underlying data”.

This will lead to better and more efficient science and improve transparency for citizens and society. It will also contribute to economic growth through open innovation.

It aims to improve and maximize access and re-use of research data generated by SCORPION project for the benefit of society and the economy.

Besides Zenodo repository, project data will also be collected and shared in the SCORPION website. Data will be available during and after the end of the project on the SCORPION website. The website will be hosted and updated by SPI until the data

becomes obsolete. Additionally, after the end of the project, collected data will be available in the platform Zenodo repository free of charge. The ownership and access to key knowledge (IPR, data, etc.) is defined by the rules established in the consortium agreement.

All European funded projects must try to disseminate as much information as possible and on top of that the SCORPION project was signed up to the “Open Research Data Pilot” which means that we are committed to give open access to data generated during the project unless it goes against our legitimate interests. In this regard, the main purpose of the DMP is to ensure the accessibility and intelligibility of the data generated during the SCORPION project to comply with the Guidelines of the “Open Research Data Pilot”. Each dataset created during the project will be assessed and categorized as open, embargo or restricted by the owners of the content of the data set.

All the datasets, regardless of their categorization, will be stored in the Microsoft One Drive folder created by the coordinator as internal database of the partners. In addition, those categorized as open or embargo will be publicly shared (in the case of embargo, after the embargo period is over) through the platform Zenodo repository (<https://zenodo.org/>).

Location - <https://zenodo.org/communities/scorpion-h2020>

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

Type	Work packages	Partners	Storage
Raw Data			
ROS BAG	WP2, WP3, WP4, WP5	INESC TEC, CNR, EURECAT, DEIMOS, TEYME	ZENODO WEBSITE
RINEX	WP2, WP3, WP4, WP5		

RTCM	WP2, WP3, WP4, WP5	INESC TEC, CNR, EURECAT, DEIMOS, TEYME	ZENODO WEBSITE
ISOBUS	WP2, WP3, WP4, WP5	INESC TEC, CNR, EURECAT, DEIMOS, TEYME	ZENODO WEBSITE
Coding C, C++, Python	WP2, WP3, WP4, WP5	INESC TEC, CNR, EURECAT, DEIMOS, TEYME	GITLAB GITHUB
General Data			
Inquires	WP1, WP6	ALL	OneDrive / ShareFolder of SCORPION project WebSite
Scientific Papers	ALL	ALL	OneDrive / ShareFolder of SCORPION project Zenodo WebSite
Technical Documents	ALL	ALL	OneDrive / ShareFolder of SCORPION project Zenodo WebSite
Pictures	WP5 and 6	ALL	OneDrive / ShareFolder of SCORPION project Zenodo WebSite
Videos	WP5 and 6	ALL	OneDrive / ShareFolder of SCORPION project Zenodo WebSite
Technical reports	ALL	ALL	OneDrive / ShareFolder of SCORPION project Zenodo

WebSite

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

Purpose relation to the objectives of the SCORPION project?	Types	Project Objective
Collect data to explore the use Galileo Signals on Robotics Navigation solutions Publish scientific results	ROSBAG RINEX Files RTCM Files	<i>O1: Bring Multi Frequency EGNSS (fully Galileo-based services) benefits to permanent crops</i>
Collect data for robotic algorithms development Publish scientific results Algorithm's validation	ROSBAG RINEX Files RTCM Files Photos	<i>O2: Localization and navigation system development;</i>
Collect data for sprayer controllers' development Publish scientific results Algorithm's validation	ROSBAG RINEX Files RTCM Files Photos	<i>O3: Advanced sprayer tool with Variable Rate Technologies (VRT) development;</i>
Collect data for safety systems development Publish scientific results Algorithm's validation	ROSBAG RINEX Files RTCM Files ISOBUS Photo Video	<i>O4: Safety systems</i>
Collect data for validation of system interoperability and compliance Publish scientific results Algorithm's validation	ROSBAG RINEX Files RTCM Files ISOBUS	<i>O5: Modularity, compliance and interoperability;</i>
Publish scientific results	ROSBAG	<i>O6: SCORPION demonstration and validation.</i>

Dissemination and Communication of project results	Technical Documents Photos Video Questionnaires and inquiries
--	--

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

Type	Size per unit	End-users for data
ROS BAG	1-20GB	SMEs and Researchers
RINEX	1-200MB	SMEs and Researchers
RTCM	1-200MB	SMEs and Researchers
ISOBUS	1-200MB	SMEs and Researchers
Inquires	1-200MB	SMEs, End-users
Scientific Papers	1-50MB	Scientific community
Technical Documents	1-20MB	Consortium members
Pictures	1-50MB	General audience and Media
Videos	1-2GB	General audience and Media
Technical reports	1-20MB	Consortium members

2.1 Data Management Plan (DMP) guiding principles

The Data Management Plan of SCORPION project is encompassed within the Work Package 6.

The SCORPION project data management plan follows the principle of Open Access according to the Horizon 2020 guideline summarized in the following diagram (Figure 1 **Erro! A origem da referência não foi encontrada.**).

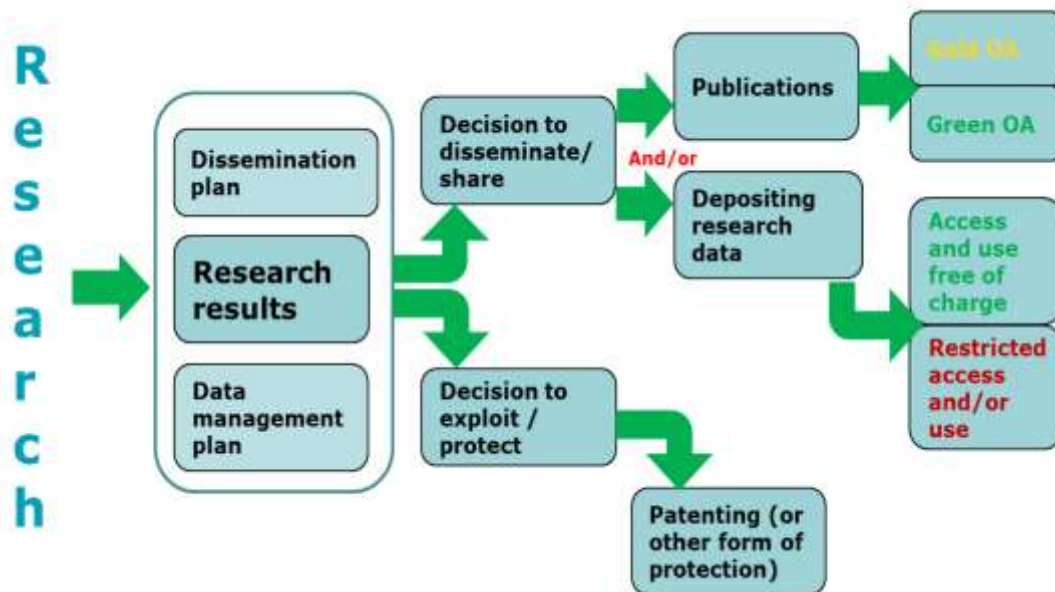


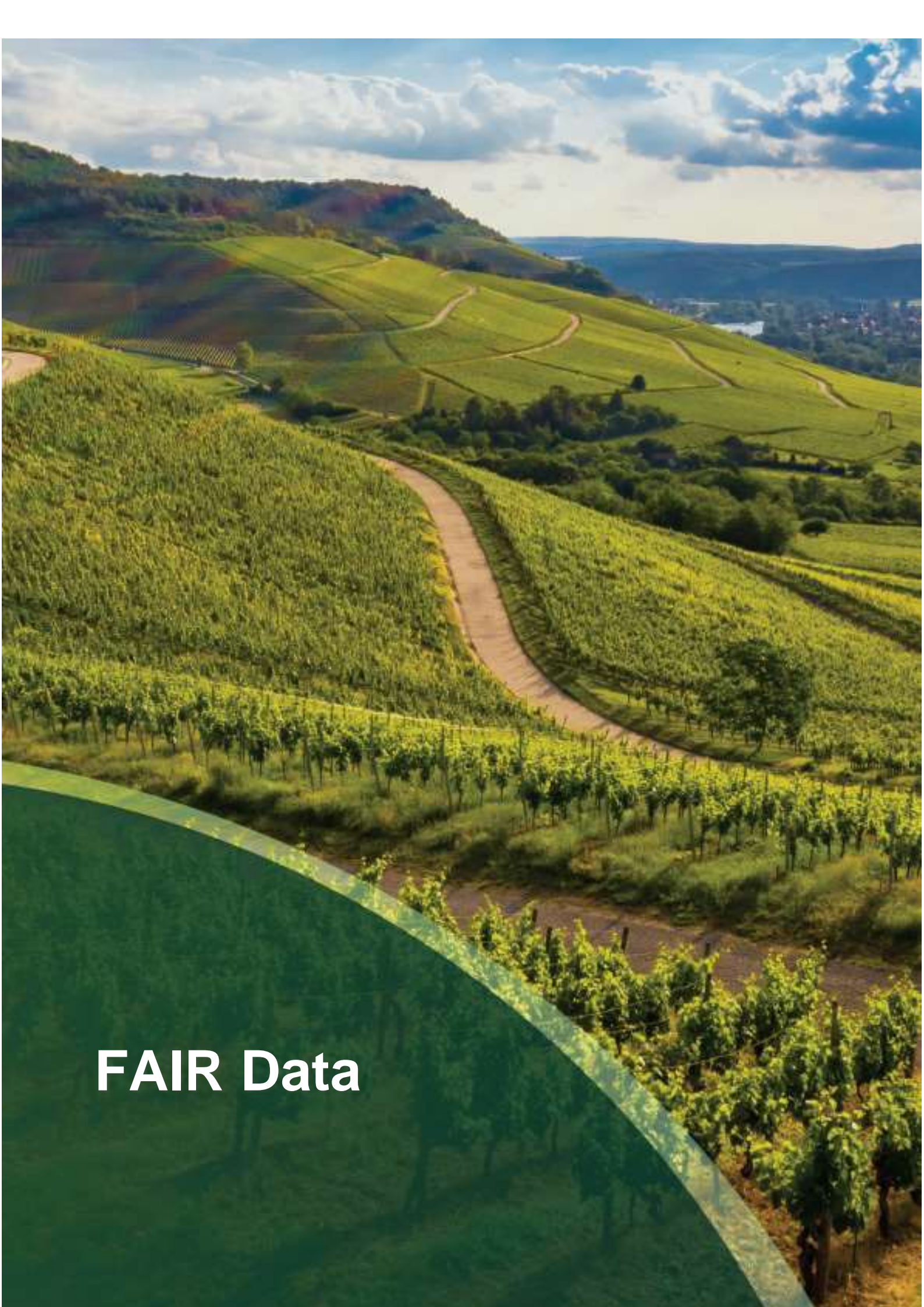
Figure 1 - Open access to research data and publication decision diagram (from Guidelines to the Rules on Open Access to Scientific publications and Open Access to Research Data in Horizon 2020).

The others main principles for the SCORPION project DPM are the following:

1. This Data Management Plan (DMP) has been prepared by taking into account the template of the “Guidelines on Data Management in Horizon 2020” http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf
2. The DMP is an official project Deliverable (D6.6 due in Month 36 – December 2023), but it will be updated throughout the project. The first initial version (D6.2

due in Month 4 - April 2021) will evolve depending on significant changes arising and periodic reviews at relevant reporting stages of the project.

3. The consortium complies with the requirements of Regulation (EU) 2016/679 and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data and repealing Directive 95/46/EC (General Data Protection Regulation). Guidance on how these regulations interact with open-access data policy can be found here: <https://www.openaire.eu>.
4. Type of data, storage, confidentiality, ownership, management of intellectual property and access: procedures that will be implemented for data collection, storage, access, sharing policies, protection, retention and destruction will be in line with EU standards as described in the Grant Agreement and the Consortium Agreement.



FAIR Data

3. FAIR Data

3.1 Making data findable, including provisions for metadata

Metadata are data on the research data themselves. They enable other researchers to find data in an online repository and are, as such, essential for the reusability of the datasets. Once added rich and detailed metadata, other researchers can better determine whether the dataset is relevant and useful for their own research. Metadata (type of data, location, etc.) will be uploaded in a standardized form. This metadata will be kept separate from the original raw research data.

As described in the project Grant Agreement (Article 29.2), the bibliographic metadata include all of the following:

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

SCORPION open data will be collected in an open online research data repository: ZENODO. Its repository structure, facilities and management follow FAIR data principles. Zenodo is an OpenAIRE that allows researchers to deposit both publications and data, providing tools to linking them to these through persistent identifiers and data citations. Zenodo is set up to facilitate the finding, accessing, re-using and interoperating of data sets, which are the basic principles that ORD projects must comply with. Zenodo repository is provided by OpenAIRE and hosted by CERN. Zenodo is a catch-all repository that enables researchers, scientists, EU projects and institutions to:

- Share research results in a wide variety of formats including text, spreadsheets, audio, video, and images across all fields of science.
- Display their research results and get credited by making the research results citable and integrating them into existing reporting lines to funding agencies like the European Commission.
- Easily access and reuse shared research results.
- Integrate their research outputs with the OpenAIRE portal.

Search keywords

Zenodo allows to perform simple and advanced search queries on Zenodo using the keywords. Zenodo also provides a user guide with easy-to-understand examples.

Location - <https://zenodo.org/communities/scorpion-h2020>

Naming conventions

Files and folders at data repositories will be versioned and structured by using a name convention consisting as follow: SCORPION_Dx.y_YYYYMMDD_Vzz.doc

Version numbers

Individual file names will contain version numbers that will be incremented at each revision (vx.y).

Metadata

This will defined along the project and presented in the final DMP document.

3.2 Making data openly accessible

Most of produced data during the project activities will be openly available as the default, except for technical documents with detailed information about the solution implementation. Selected data and results will be shared with the scientific community and other stakeholders through publications in scientific journals and presentations at conferences, as well as through open access data repositories, promoting the main objective of maximize the impact of the research data by sharing the results within and beyond the consortium.

As before said, all the SCORPION data will be stored on the Zenodo repository after delivery. Before that moment, the data will be stored in cloud folders, or on the institutional secure server, and made accessible to all consortium members.

Zenodo assigns a DOI to all documents uploaded and allows the storage of metadata.

Most files will be accessible with general use software (word, excel, power point and their open versions). Next versions of the Data Management Plan will be more specific regarding other types of data and their accessibility.

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

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

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

The SCORPION project will consider open standards and formats readable by open-source software. This project will privilege the following formats:

- **RINEX Files**
- **ROSBAG Files**
- **RTCM Files**
- **ISOBUS files**
- **Photos in standard formats (JPG, PNG, TIFF)**
- **Video (MP4 and AVI)**

Besides all SCOPRION documents will be published in PDF format. For example, text that contains quantitative data of the market where SCORPION robots will be applied in future and data from anonymised results from questionnaires will be published in PDF and in CVS files.

3.3 Making data interoperable

All the project partners will observe OpenAIRE guidelines for online interoperability of generated data, including OpenAIRE Guidelines for Literature Repositories, OpenAIRE Guidelines for Data Archives, OpenAIRE Guidelines for CRIS Managers based on CERIF-XML.

These guidelines can be found at <https://guidelines.openaire.eu/en/latest/>.

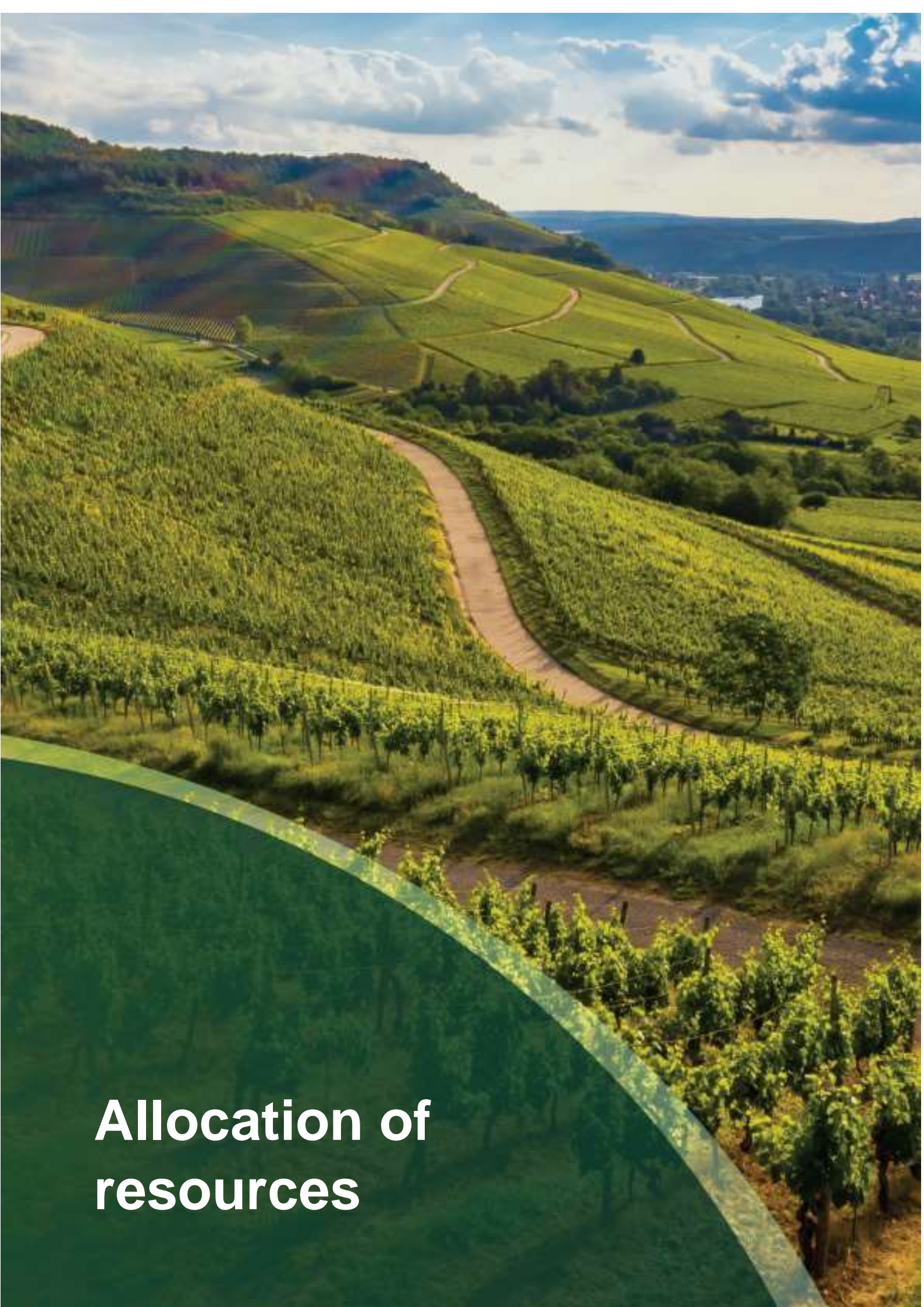
Partners will also ensure that SCORPION generated data observes FAIR data principles under H2020 open-access policy: http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-datamgt_en.pdf.

In order to ensure the interoperability, all datasets will use the same standards for data and metadata capture/creation.

As the project progresses and data are identified and collected, further information on making data interoperable will be outlined in subsequent versions of the present Data Management Plan. More in specific, information on data and metadata vocabularies, standards or methodology to follow to facilitate interoperability and whether the project uses standard vocabulary for all data types present to allow interdisciplinary interoperability.

The SCORPION project will consider open standards and formats readable by open-source software. This project will privilege the following raw data formats:

- **ROSBAG**
- **RINEX Files**
- **RTCM Files**
- **ISOBUS**



Allocation of resources

4. Allocation of resources

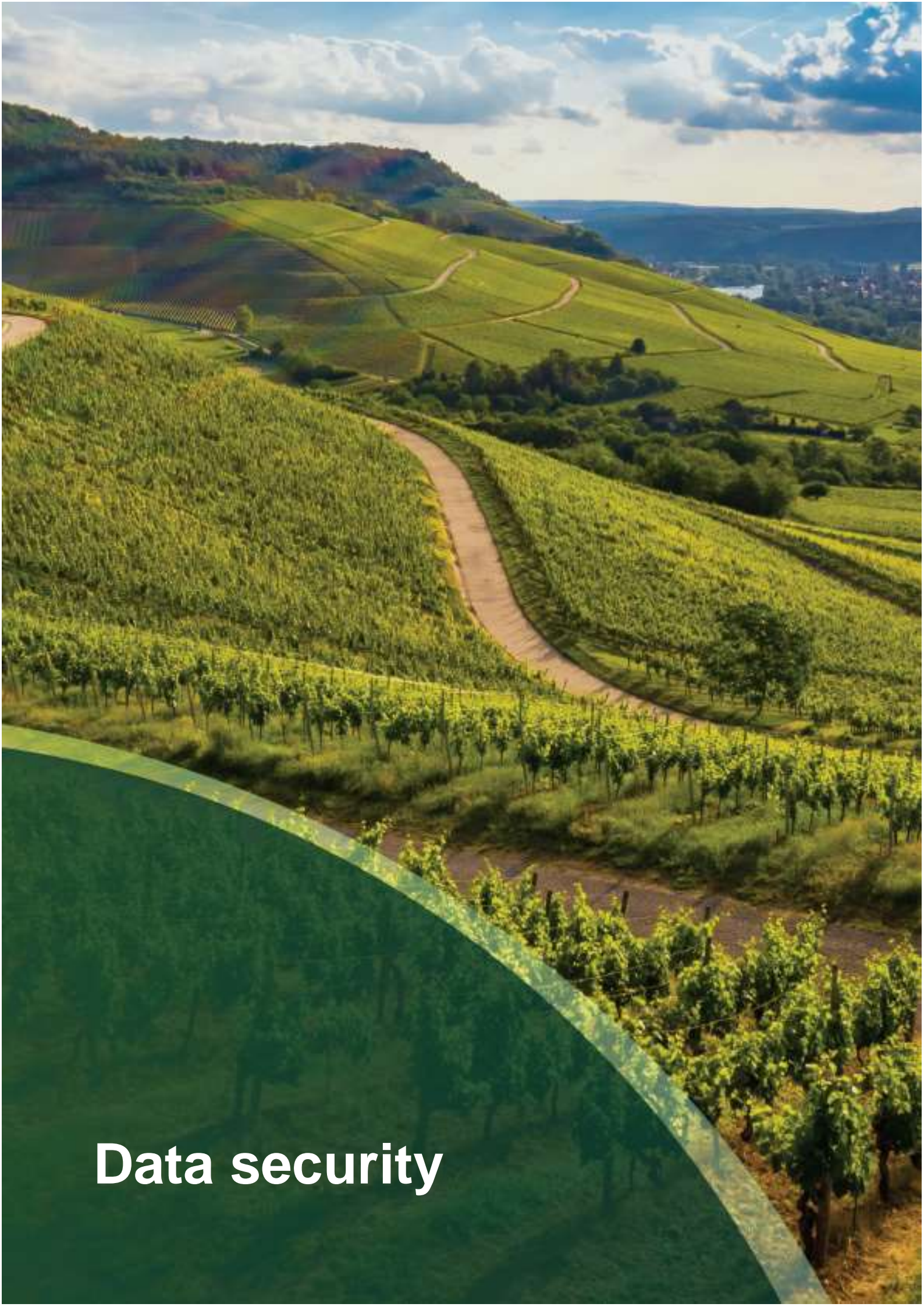
All the costs related to open-access to research data provided during a Horizon 2020 project are eligible for reimbursement under the conditions defined in the H2020 Grant Agreement, in particular Article 6 and Article 6.2.D.3, but also other articles relevant for the cost category chosen. Project beneficiaries will be responsible for applying for reimbursement for costs related to making data accessible to others beyond the consortium.

The costs for making data FAIR includes:

- Fees associated with the publication of scientific articles containing project's research data in "Gold" Open access journals. The cost sharing, in case of multiple authors, shall be decided among the authors on a case-by-case basis.
- Project Website operation: **to be determined**.
- Data archiving at ZENODO and on other on line data base: free of charge.

Regarding the budget allocated for open access to research data, the expected amount for SCORPION is 3000€.

Each partner is responsible for the data they produce. Any fee incurred for Open Access through scientific publication of the data will be the responsibility of the data owner (authors) partner(s).



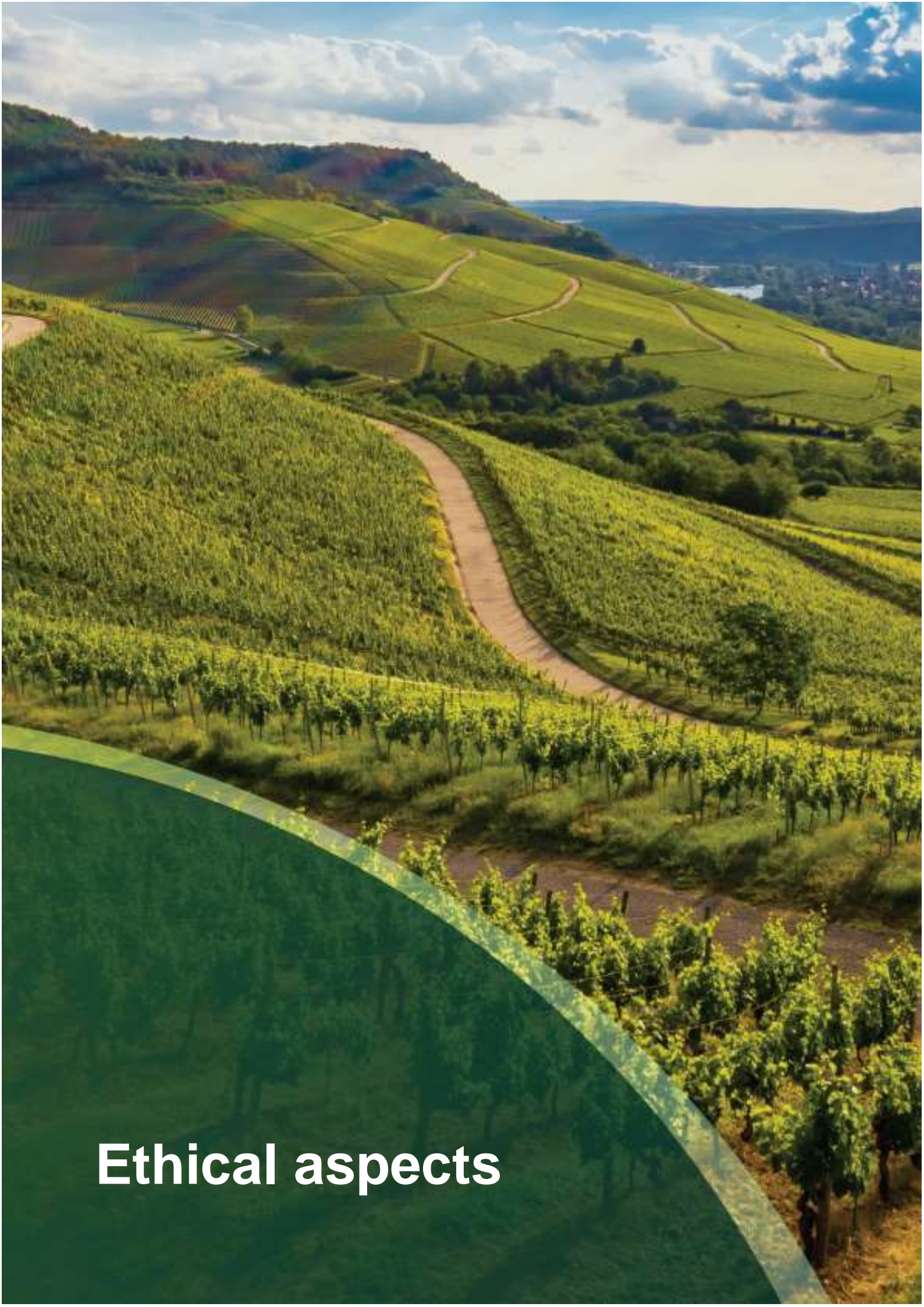
Data security

5. Data security¹

Most SCORPION data will be stored in Zenodo repository. Zenodo is provided by OpenAIRE and hosted by CERN. During the project will be founded other alternatives to store this data in a secondary place.

Technical data and documents will be stored in INESC TEC servers with redundancy systems that can unsure the long term data preservation.

¹ *What provisions are in place for data security (including data recovery as well as secure storage and transfer of sensitive data)? Is the data safely stored in certified repositories for long term preservation and curation?*



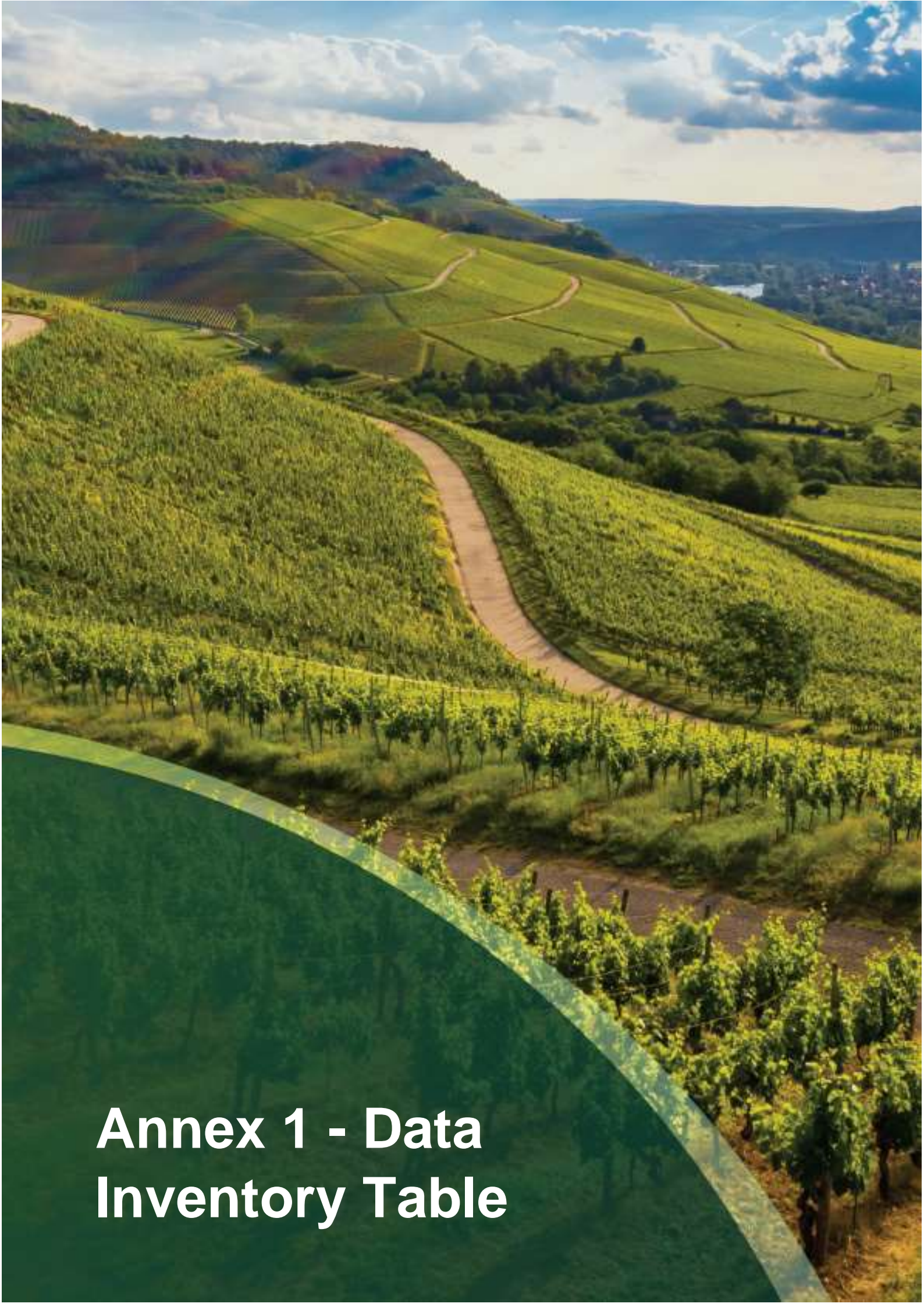
Ethical aspects

6. Ethical aspects²

During the SCORPION prototypes developing, testing and validation the robotic sensors may record images and videos of persons. This information may be collected and selected for data sharing. Being this a sensitive issue, the procedure before any data publication will required a data revision to identify if persons are identifiable on a particular dataset and if they are, one of two procedures will be followed:

1. Remove information from dataset to make these persons not identifiable/recognizable.
2. Ask to the involved persons, the permission to publish that dataset with that information.

² 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). Is informed consent for data sharing and long term preservation included in questionnaires dealing with personal data?



Annex 1 - Data Inventory Table

Annex 1 - Data Inventory Table

To be presented on the final DMP.

