



## D2.5 Report on strategy for FAIRness on soil data, v1

M27/NOV 2025



**Funded by  
the European Union**

This project has received funding from the Horizon Europe research and innovation programme under Grant Agreement No 101112838.

<b>Acronym</b>	SoilWise
<b>Project Full Title</b>	An open access knowledge and data repository to safeguard soils
<b>GA number</b>	101112838
<b>Topic</b>	HORIZON-MISS-2022-SOIL-01-01
<b>Type of Action</b>	HORIZON Innovation Actions
<b>Project Duration</b>	48 months
<b>Project Start Date</b>	1-Sep-23
<b>Project Website</b>	www.soilwise-he.eu
<b>Deliverable Title</b>	Report on strategy for FAIRness on soil data, v1
<b>Delivery Time (DOA)</b>	M27
<b>Deliverable Submission Date</b>	28/11/2025
<b>Status</b>	draft
<b>Dissemination Level</b>	PU - Public
<b>Deliverable Lead</b>	Tomáš Pavelka (MU)
<b>Author(s)/Organisation(s)</b>	Tomáš Pavelka (MU), Céline Blitz Frayret (CIRAD), Tomáš Řezník (MU), Paul van Genuchten (ISRIC), Dajana Snopková (MU), Rob Lokers (WR), Radu Giurgiu (EV ILVO), Jan Oprchal (MU)
<b>Contributor(s)</b>	MU, CIRAD, ISRIC, EV ILVO, WR
<b>Peer-Reviewers</b>	Cenk Dönmez (ZALF), Maria Fantappiè (CREA)
<b>Contact</b>	tomas.pavelka@geogr.muni.cz
<b>Work Package</b>	WP2
<b>Dissemination level</b>	Public
<b>Keywords</b>	FAIR data principles; soil data; metadata; Mission Soil; SoilWise;

**Abstract**

This deliverable presents the first version of the strategy developed within SoilWise to improve the FAIRness of soil data and knowledge across Europe. The document outlines the conceptual foundations of FAIR principles, reviews the relevant European policy landscape, and analyses current gaps in findability, accessibility, interoperability and reusability of soil-related information. It describes the methodological process used to co-design both general and data type-specific FAIR Guidelines in close collaboration with Mission Soil projects and the Data and Knowledge Management Cluster.

The resulting guidelines provide practical recommendations for data providers, supporting metadata preparation, repository selection, vocabulary use, provenance documentation, and licensing. The strategy also clarifies how SoilWise contributes to improving long-term discoverability and sustainable reuse of soil data. This first version forms the foundation for subsequent iterations that will expand and refine the guidelines as SoilWise progresses.

## Disclaimer

The content of this deliverable represents the views of the author(s) only and does not necessarily reflect the official opinion of the European Union. The European Union institutions and bodies or any person acting on their behalf are not responsible for any use that may be made of the information it contains.

In this document, the acronym 'DOMG – VL O' is used to refer to the Department of the Environment and Spatial Development, Flanders, Belgium, as per the partner's request for clarification. It's noted that in the grant agreement, the partner is identified by the acronym VL O (Vlaamse Gewest).

## List of Abbreviations

<b>AI</b>	Artificial Intelligence
<b>API</b>	Application Programming Interface
<b>CC</b>	Creative Commons
<b>CSV</b>	Comma-Separated Values
<b>D</b>	Deliverable
<b>DGA</b>	Data Governance Act
<b>DMP</b>	Data Management Plan
<b>DOI</b>	Digital Object Identifier
<b>EC</b>	European Commission
<b>EEA</b>	European Environment Agency
<b>EOSC</b>	European Open Science Cloud
<b>ESDAC</b>	European Soil Data Centre
<b>EU</b>	European Union
<b>EUSO</b>	European Soil Observatory
<b>FAIR</b>	Findable, Accessible, Interoperable, Reusable
<b>GEMET</b>	General Multilingual Environmental Thesaurus
<b>GM</b>	Governance model
<b>HE</b>	Horizon Europe
<b>INSPIRE</b>	Infrastructure for Spatial Information in the European Community
<b>ISO</b>	International Organization for Standardization
<b>JSON</b>	JavaScript Object Notation
<b>KM</b>	Knowledge management
<b>LUCAS</b>	Land Use/Cover Area Frame Survey
<b>ML</b>	Machine Learning
<b>MS</b>	Mission Soil
<b>OAI-PMH</b>	Open Archives Initiative Protocol for Metadata Harvesting
<b>OGC</b>	Open Geospatial Consortium
<b>PID</b>	Persistent Identifier
<b>ROR</b>	Research Organization Registry
<b>SOC</b>	Soil Organic Carbon
<b>SWC</b>	SoilWise Catalogue
<b>T</b>	Task
<b>URI</b>	Uniform Resource Identifier
<b>WBR</b>	World Reference Base for Soil Resources
<b>WCS</b>	Web Coverage Service
<b>WFS</b>	Web Feature Service

PROJECT NAME – SoilWise

Project Number 101112838

---

<b>WP</b>	Work Package
<b>WMS</b>	Web Map Service
<b>XML</b>	Extensible Markup Language

# Table of Contents

<b>1</b>	<b>INTRODUCTION.....</b>	<b>8</b>
1.1	PROJECT SUMMARY.....	8
1.2	DOCUMENT SCOPE.....	8
1.3	DOCUMENT STRUCTURE.....	9
1.4	RELATIONSHIP TO OTHER PROJECT DELIVERABLES.....	9
1.5	RELATIONSHIP TO PROJECT TASKS.....	9
<b>2</b>	<b>FAIR PRINCIPLES.....</b>	<b>11</b>
2.1	OVERVIEW OF FAIR PRINCIPLES .....	11
2.1.1	<i>Findability</i> .....	11
2.1.2	<i>Accessibility</i> .....	11
2.1.3	<i>Interoperability</i> .....	12
2.1.4	<i>Reusability</i> .....	12
2.2	FAIRNESS IN THE CONTEXT OF SOIL DATA: WHY CHALLENGING AND NECESSARY .....	12
<b>3</b>	<b>POLICY FRAMEWORK .....</b>	<b>14</b>
3.1	EUROPEAN POLICY AND INSTITUTIONAL FRAMEWORK INFLUENCING FAIR SOIL DATA .....	14
3.1.1	<i>INSPIRE Directive (2007/2/EC)</i> .....	14
3.1.2	<i>Open Data Directive (2019/1024)</i> .....	14
3.1.3	<i>European Open Science Policies &amp; Horizon Europe Open Research Data Mandate</i> .....	15
3.1.4	<i>European Data Act &amp; Data Governance Act (DGA)</i> .....	15
3.1.5	<i>Soil Monitoring and Resilience Directive (Soil Monitoring Law)</i> .....	15
3.2	CURRENT GAPS IN FAIR IMPLEMENTATION OF SOIL DATA .....	15
3.2.1	<i>Findability: Metadata Gaps and Limited Discoverability</i> .....	16
3.2.2	<i>Accessibility: Legal and Technical Barriers</i> .....	16
3.2.3	<i>Interoperability: Fragmented Standards and Semantic Inconsistencies</i> .....	16
3.2.4	<i>Reusability: Lack of Provenance, Licensing and Long-term Preservation</i> .....	17
3.2.5	<i>Organizational and Cultural Barriers</i> .....	17
<b>4</b>	<b>GUIDELINES DEVELOPMENT PROCESS.....</b>	<b>19</b>
4.1	METHODOLOGY .....	19
4.1.1	<i>General Guidelines Methodology</i> .....	20
4.1.2	<i>Data Type-specific Guidelines Methodology</i> .....	20
<b>5</b>	<b>GENERAL GUIDELINES ON DATA AND KNOWLEDGE FAIRNESS .....</b>	<b>22</b>
<b>6</b>	<b>DATA TYPE SPECIFIC GUIDELINES ON DATA AND KNOWLEDGE FAIRNESS .....</b>	<b>27</b>
<b>7</b>	<b>ADDITIONAL SOURCES .....</b>	<b>27</b>
<b>8</b>	<b>CONCLUSION.....</b>	<b>28</b>

## Executive Summary

This deliverable presents the initial strategy for improving the FAIRness of soil data and knowledge within Europe. It explains why FAIR principles are essential for research, monitoring and policy making, and outlines the methodological approach used within SoilWise to develop general guidelines and data type-specific recommendations. The strategy is grounded in European policy needs, the objectives of the Mission Soil initiative and insights gathered from the wider soil data community. This collaborative process involved close engagement with Mission Soil research projects, fostering peer-to-peer interactions and leveraging insights from the Data and Knowledge Management Cluster.

The document is intended for experts working with soil data, organisations responsible for soil monitoring systems, data service providers, research teams, Mission Soil projects and policy makers who require harmonised information to support planning and decision making. It is also relevant for technical specialists developing or managing repositories, catalogues and infrastructures related to soil information.

The main activities focused on analysing existing policy and technical frameworks, identifying key obstacles to achieving FAIR principles and defining a methodological process for creating both general and data type-specific guidelines. This work combined stakeholder consultations, a review of relevant standards, an assessment of current repositories and an examination of user requirements. Additional effort was dedicated to preparing a structure that will support future guideline development in upcoming project cycles.

The results provide a clear overview of how FAIR principles can be applied in the soil domain. A 'cookbooks' on step-by-step soil data 'FAIRification' is provided as the major result, are available in chapters 5 and 6. The document highlights major limitations that restrict the usability of soil information, including inconsistent metadata practices, weak interoperability across national and European classification systems, unclear or missing licences and insufficient documentation of data provenance. At the same time, it proposes a framework for subsequent versions of the strategy, which will be refined as SoilWise develops new components and collaborates with other Mission Soil projects.

The findings bring benefits for both research and practice. For research teams, the strategy clarifies how to create and publish data that remain understandable, findable and reusable in the long term. For practitioners, including authorities responsible for soil monitoring or policy development, it offers a foundation for more reliable indicators, better understanding of soil trends and more informed decision making. Improved harmonisation and the adoption of shared standards will also support agricultural advisors, land managers and organisations working on soil conservation.

There are also relevant policy implications. The documented shortcomings indicate which areas require targeted attention from the European Commission, national bodies and other institutions. The strategy supports future data coordination within the European Soil Observatory and helps prioritise investments into interoperable standards, open access and long-term preservation. It directly contributes to the objectives of the Soil Monitoring Law and the Mission Soil implementation plan.

In summary, this deliverable establishes a solid basis for the development of more detailed FAIRification methodologies in future iterations. It encourages a coordinated European approach to soil data management and strengthens SoilWise as a trusted, sustainable entry point for integrating soil-related information across Europe.

# 1 Introduction

## 1.1 Project summary

Now more than ever, soil health is an issue that needs to be addressed urgently, as recent assessments state that 60-70% of European soils can be considered unhealthy (Bouman, 2022). The EU Mission ‘A Soil Deal for Europe’, the EU Soil Strategy and the proposal for a Soil Monitoring and Resilience Directive (5 July 2023), aims to have 75% of EU soils healthy or significantly improved by 2030 and all soils healthy in 2050. Reaching such an ambition requires, among others, access to reliable, harmonised existing and new data and knowledge collected at local, national and EU levels to allow **informed decision-making at all scales to support the proposed Soil Monitoring and Resilience Directive and the EU Soil Strategy**.

The SoilWise project will provide an integrated and actionable access point to scattered and heterogeneous soil data and knowledge in Europe, making them FAIR (Findable, Accessible, Interoperable and Reusable) and improve trust, willingness, and ability to share and re-use soil data and knowledge. In three project development cycles, **co-creation and co-validation by multi-stakeholder groups are at the centre of project activities**. SoilWise recognises existing workflows and public and permanent repositories, such as Zenodo which is suggested to be used by European Commission for permanent data storage, for specific user needs and aims to work with them to enhance their discoverability, approachability, and interconnection. An open, modular, scalable, and extensible knowledge and metadata catalogue building on existing and new technologies will be provided while respecting data ownership, access policies and privacy. AI- and ML- techniques will be employed to interlink scattered data and knowledge, automatise the processes, infer new knowledge and increase FAIRness. **SoilWise applies infrastructure thinking instead of project thinking to design a repository for at least a decade to support EUSO evolution accordingly**. The SoilWise metadata catalogue and community are designed to be a joint starting point and common ground for countries, the European Commission, and other stakeholders to jointly guide soil and related spatial policy and informed decision-making towards the 2030 goals of the Green Deal, achieve healthy soils in 2050 and ensure broad uptake and implementation by land managers, policy, research, and industry.

All personal data acquired through SoilWise is processed in strict accordance with the relevant EU privacy regulations, highlighting our dedication to uphold to the highest standards of data privacy and security for our users.

## 1.2 Document scope

This document presents the first version of the SoilWise strategy for improving the FAIRness of soil data and knowledge, with a scope that reflects both the needs of the SoilWise project and the broader objectives of Mission Soil. It establishes a structured basis for understanding how FAIR principles can be operationalised within the soil domain and provides the conceptual and methodological framework that underpins the development of the SoilWise FAIR Guidelines, including both general and data type-specific guidelines.

The scope includes the conceptual framing of FAIRness in the soil domain, as well as an overview of the current challenges affecting soil data management across Europe, including gaps in metadata quality, accessibility, semantic consistency, licensing and long-term preservation. It also identifies the organizational and cultural obstacles that influence the ability of data providers and institutions to implement FAIR practices. By clarifying these challenges, the document frames the need for harmonised guidelines that can support data providers working in different scientific, monitoring and policy contexts.

The document further explains the iterative and participatory approach used to develop the FAIR Guidelines. This includes the co-design process carried out with Mission Soil projects, insights gathered from the broader Data and Knowledge Management Cluster and feedback from stakeholder engagement activities. The scope therefore covers not only conceptual elements but also the practical processes through which FAIR practices can be strengthened in real operational settings.

Given the strong links between soil-related data and European policy frameworks, the scope also outlines the strategic alignment of this FAIRness strategy with the Soil Monitoring and Resilience Directive, Horizon Europe open science requirements and the future role of the European Soil Observatory. This ensures that the strategy supports current legal obligations while also preparing the ground for long-term integration with European infrastructures and future SoilWise development cycles.

## 1.3 Document structure

This document comprises the following chapters to provide an overview of the FAIRness strategy for soil data and knowledge.:

- **Chapter 1** introduces the scope of the document and its relation to the project.
- **Chapter 2** summarises the FAIR principles and explains their relevance for soil-related information.
- **Chapter 3** presents the main European policies and initiatives shaping requirements for data sharing and interoperability.
- **Chapter 4** describes the methodological process for preparing general and data type-specific guidelines.
- **Chapter 5** presents the first version of general FAIR Guidelines, which offer practical recommendations for improving FAIRness across soil-related data and knowledge.
- **Chapter 6** introduces the data type-specific guidelines, which provide tailored recommendations for publication of specific data types, tabular soil data and project websites respectively.
- **Chapter 7** lists additional sources that supported the development of the strategy.
- **Chapter 8** concludes the document and summarises the key messages and next steps.

## 1.4 Relationship to other project deliverables

This deliverable relates to and complements the following deliverables:

- D2.6 – Report on strategy for FAIRness on soil data, v2 (M42)
- D3.5, 3.6 – Report on strategy for effective soil KM, v1, v2 (M27, M42)
- D7.2, D7.3, D7.4 – Open Science and Data Management plan, v1, v2, v3 (M6, M27, M48)

## 1.5 Relationship to project tasks

This deliverable relates to the following project tasks:

- T2.4 Strategy for FAIRness on soil data – Develops Strategy for FAIRness on soil data.
- T3.4 Strategy for efficient KM – Develops Strategy for efficient KM, which links and overlaps with Strategy for FAIRness on soil data.

- T5.2 User Cases implementation and demonstration – will demonstrate the developed Data-type specific Guidelines.
- T6.1 Communication and dissemination – Strategy for FAIRness on soil data will be promoted and disseminated to relevant stakeholders.
- T6.3 Exploitation and capacity Building – Strategy for FAIRness on soil data will contribute to stakeholder capacity building regarding data publication.
- T7.2 Technical and Scientific Management – prepares an Open Science and Data Management Plan (D7.2 – D7.4) which feeds the recommendations in the Strategy for FAIRness on soil data.

## 2 FAIR principles

The concept of FAIR data, i.e. ensuring that data are Findable, Accessible, Interoperable, and Reusable, has become a central pillar of European research and innovation policy (see section 3 for policy insights). The FAIR framework aims to enhance the value, visibility, and sustainability of data by ensuring that they can be efficiently discovered, accessed, understood, and reused across different scientific and policy domains.

In the context of soil data, the FAIR principles are particularly relevant. Soil information is inherently diverse, multi-scale, and heterogeneous, encompassing field measurements, laboratory analyses, spatial data, and model outputs. These data are often produced by various organizations under different standards, formats, and levels of openness. As a result, ensuring their FAIRness is not only a matter of technical implementation but also of governance, coordination, and community alignment. By incorporating FAIR principles, such as common standards, interoperable vocabularies, and transparent metadata practices, the FAIR approach helps bridge these gaps, making soil data more comparable, discoverable, and usable across projects, domains, and governance levels. It provides a foundation for integrating fragmented soil information into a coherent, accessible European knowledge base.

This chapter introduces the FAIR principles, discusses their specific relevance to soil data, and presents the key frameworks that guide FAIR implementation in the European context.

### 2.1 Overview of FAIR Principles

The FAIR principles were first articulated in 2016 (Wilkinson et al.) as a response to the growing need for systematic and machine-actionable data management. They provide a guiding framework for improving the discovery, accessibility, and reusability of digital resources throughout their lifecycle. Importantly, FAIRness does not imply that all data must be open, but that the rules for discovery and access are transparent and technically enforceable.

#### 2.1.1 Findability

To be findable, data and metadata must be uniquely and persistently identified, well-documented, and indexed in searchable repositories.

For soil data, this means that individual datasets (such as soil profiles, monitoring time series, or land management data) should have Persistent Identifiers (PIDs) (e.g. DOI, URI) and be described using standardized metadata (e.g. ISO 19115, Dublin Core or INSPIRE). Data should be deposited in trustworthy and permanent repositories, and metadata in persistent catalogues, enabling discovery through both human and machine queries.

Enhanced findability facilitates cross-project visibility, allows linking between datasets and publications, and supports reproducibility of scientific and policy analyses.

#### 2.1.2 Accessibility

Accessibility ensures that data and metadata can be retrieved via standardized and open protocols, regardless of their access conditions. For soil data, this typically involves web services (e.g. OGC WMS/WFS/WCS, OAI-PMH, REST APIs) that allow automated data exchange and integration.

Crucially, accessibility also covers the clarity of access rights and conditions, indicating whether data are open, restricted, or available upon request, and under which license (e.g. Creative Commons). Note that data and metadata can have different access conditions. In many cases, soil data are sensitive (e.g. farm-level, environmental monitoring), so ensuring transparent and machine-readable access policies is as important as the technical mechanisms themselves.

### 2.1.3 Interoperability

Interoperability enables data to be integrated, exchanged, and interpreted across systems, domains, and scales.

For soil data, this principle requires the use of common vocabularies, ontologies, and data models that define soil properties, sampling methods, and land-use categories in a consistent way. Examples include INSPIRE Soil Data Specification, World Reference Base (WRB), AGROVOC or GEMET.

Interoperability also means that data can be opened and used without the need for specialized or proprietary software. This is achieved through the use of widely accepted and transparent data structures that can be read by many systems and shared across different platforms. In practice, this ensures that soil information collected by different organizations or projects can be easily combined, compared, and reused in broader environmental and policy contexts.

### 2.1.4 Reusability

Reusability ensures that data can be reliably and meaningfully reused beyond their original purpose.

This requires comprehensive metadata, explicit licensing, provenance information, and, ideally, quality indicators describing data accuracy, completeness, and validation status. For soil data, provenance includes information on e.g. sampling location, date, methods, laboratory protocols, and processing steps.

Reusable soil data enable long-term applications such as soil health monitoring, modelling, and policy evaluation, and help build trust in derived indicators and decision-support tools.

Moreover, data should be stored in repositories committed to long-term preservation, ensuring future accessibility and traceability in line with FAIR and open science principles.

## 2.2 FAIRness in the Context of Soil Data: Why Challenging and Necessary

Soil data are intrinsically complex, heterogeneous and fragmented. They originate from a wide range of sources, e.g. field surveys for soil mapping, field surveys for agronomic evaluation, laboratory analyses, remote sensing, national monitoring networks, agronomic models, legacy maps and related datasets, research projects, and citizen science initiatives. Each of these sources operates with different methodologies, spatial and temporal resolutions, file formats, and terminologies.

Unlike many other agri-food domains, soil data are produced across multiple scientific, environmental, and policy fields such as agriculture, climate, biodiversity, land use, or pollution monitoring. Each of these domains applies its own standards, vocabularies, and objectives, leading to variations in how soil is measured, described,

and classified. This cross-domain diversity makes it particularly difficult to integrate soil information into a coherent system and to ensure its long-term usability.

This diversity makes soil data particularly vulnerable to inconsistencies, lack of standardisation, and limited usability outside the context in which they were originally produced. Ensuring FAIRness is therefore not only desirable, but essential to make these data understandable, comparable and reusable across regions, sectors and disciplines.

The implementation of FAIR principles in the soil domain faces several key **challenges**:

- **Fragmented data ownership and governance**

Soil data are often collected and managed by multiple institutions, such as research organizations, national agencies, private companies, or farmers, each with their own data policies, formats, and access rules. This leads to limited interoperability and difficulty integrating datasets across borders or administrative levels.

- **Lack of harmonized vocabularies and standards**

Different classification systems (e.g., WRB, USDA Soil Taxonomy), laboratory methods, or metadata conventions make it difficult to align datasets. In some cases, metadata are missing or insufficient, hindering both discoverability and reuse. Several soil ontologies exist (as reported e.g. by EJP Soil project in Deliverable [6.1 Report on harmonized procedures for creation of databases and maps](#), chapter 3.4 Controlled vocabularies and ontologies on Soil), partly overlapping, without consolidation and clear governance. A GLOSIS working group has been recently established which has plan to work towards a common world soil ontology definition. Ontologies depicting soil data structure are fragmented among data sources or are not used.

- **Spatial and temporal heterogeneity**

Soil properties vary significantly across small distances and time periods (created by soil properties or different analytical methods). Combining data from different campaigns or monitoring systems requires detailed contextual information, quality indicators, and provenance documentation.

- **Licensing and sensitivity issues**

Some soil-related data (e.g., from farm plots or private land) may be restricted due to privacy, commercial, or legal constraints. Clear licensing and access policies are often missing or ambiguous, which creates uncertainty and reduces willingness to share.

- **Legacy and analogue data**

Many valuable soil datasets exist in non-digital formats, offline (individual) backups or in outdated systems. Without digitization, metadata enhancement, and persistent identifiers, these legacy resources remain largely invisible to modern infrastructures such as EUSO.

Despite these challenges, improving FAIRness in soil data management is **necessary** for several reasons:

- It enables **evidence-based policymaking** on soil health, climate adaptation, land degradation, and sustainable agriculture.
- It ensures **long-term findability, accessibility, and preservation of soil data and project results**, preventing the loss of valuable outputs after projects end and supporting their continuity, citation, monitoring, and reuse by future initiatives.

- It enhances the **impact and reusability of research outputs**, avoiding duplication of efforts and ensuring the long-term value of data generated through publicly funded projects.
- It contributes to the development of the **European Soil Observatory (EUSO)** as a central gateway for soil knowledge, where data from multiple sources can be integrated, harmonized, and used to produce indicators and decision-support tools.
- It supports the broader objectives of **Mission Soil**, the **European Green Deal Data Space**, and the **European Open Science Cloud (EOSC)**, where FAIR data forms a foundational element.

In summary, achieving FAIRness in soil data is both a technical challenge and a strategic necessity. It requires coordinated actions on standards, governance, infrastructure, and community practices, all of which are further developed in the guidelines presented in the following chapters.

## 3 Policy Framework

The development of FAIR soil data does not take place in isolation, but within a rapidly evolving landscape of European policies, regulations, and strategic initiatives. These frameworks define legal obligations, shape technical requirements, and create incentives for more open, interoperable, and reusable data. For soil-related information, they play a crucial role in harmonising national practices, supporting cross-border data sharing, and ensuring that soil knowledge can effectively inform research, innovation, and policy. This chapter provides an overview of the key EU policies that influence FAIR data practices and outlines how they collectively shape the environment in which SoilWise operates.

Note that the policy context of key EU policies and frameworks focusing on soil health and facilitating the sharing of spatial data across Europe is also mentioned in deliverable 1.5 Repository GM, v1. The national and EU regulations on agricultural soil data sharing and national monitoring activities were also summarised within the EJP Soil project, reported in their [Deliverable 6.2](#).

### 3.1 European Policy and Institutional Framework Influencing FAIR Soil Data

FAIR management of soil data is no longer just a technical or research-oriented issue. It is increasingly embedded in European legislation and policy frameworks that require transparent, interoperable and reusable environmental information. Several key policies shape how soil data must be collected, managed and shared across EU Member States:

#### 3.1.1 INSPIRE Directive (2007/2/EC)

The [INSPIRE Directive 2007/2/EC](#) (2007) requires EU countries to make spatial environmental data themes, including soil-related layers, available in interoperable formats and with common metadata, cataloguing and web-services. While it supports machine-readability and access across borders, it does not fully address research-data reusability, persistent identifiers or detailed semantic standardization. For soil-data FAIRness, INSPIRE provides a strong foundation for metadata and discoverability but needs to be complemented by FAIR-oriented practices. Note, the ongoing revision of the INSPIRE Directive (November 2025) tends to eliminate the burden imposed on data providers concerning data harmonization.

#### 3.1.2 Open Data Directive (2019/1024)

The [Open Data Directive](#) (2019/1024) emphasizes the importance of making publicly funded data, including high-value datasets like soil data, available for reuse under open access principles. High-value soil data, which often includes critical information on soil health, composition, and environmental impact, can be more effectively leveraged for research, policy-making, and environmental management when it adheres to FAIR standards. The Open Data Directive is also being understood as a follow-up of the INSPIRE Directive that has narrowed the scope on the one hand while it is more demanding in terms of FAIRification steps for such narrowed scope than the original INSPIRE Directive. Soil data are within such narrowed scope of the Open Data Directive.

### 3.1.3 European Open Science Policies & Horizon Europe Open Research Data Mandate

Under the Horizon Europe framework, all funded projects are required to deposit research data in trusted repositories, ensure open access where possible and apply the FAIR principles. Data Management Plans (DMPs) are mandatory, and projects are expected to provide metadata, persistent identifiers (e.g., DOIs) and clear licensing. For the soil domain, this means that dataset outputs from research must be prepared for reuse and integration, aligning with the wider data infrastructures such as EUSO.

### 3.1.4 European Data Act & Data Governance Act (DGA)

These two regulations support the broader [European Data Strategy](#) and indirectly influence soil and environmental data management:

- **[Data Governance Act \(DGA, 2022\)](#)**: establishes frameworks for data intermediaries, data altruism, and re-use of public-sector data, including environmental data. Encourages secure data spaces (e.g. Green Deal Data Space, Agriculture Data Space).
- **[Data Act \(2024\)](#)**: regulates data access and reuse obligations, particularly for data generated by devices, sensors, and public-private collaboration. It promotes interoperability, data sharing between sectors, and common standards – relevant for sensor-based soil monitoring and agri-environmental data.

These Acts push for cross-sector FAIR practices, including sharing privately or sensor-generated soil data under fair conditions.

### 3.1.5 Soil Monitoring and Resilience Directive (Soil Monitoring Law)

The EU has adopted the [Soil Monitoring and Resilience Directive](#) (2025), also known as Soil Monitoring Law, to establish a harmonised system of soil health monitoring across Member States. Member States must set up monitoring systems based on a common methodology, with comparable soil-health descriptors, and report to the European Environment Agency (EEA) and the European Commission (EC). Its relevance for FAIRness lies in the legal push towards standardised, interoperable, and publicly accessible soil data, aligning closely with the FAIR principles and offering a clear mandate for the EUSO and the SoilWise project to build data systems accordingly.

## 3.2 Current Gaps in FAIR Implementation of Soil Data

Despite increasing political and institutional support for open and FAIR data in Europe, the implementation of FAIR principles in soil-related data systems remains inconsistent and fragmented. The challenges described in

Chapter 2.2 manifest clearly in current infrastructures, national monitoring systems and project-based data outputs. The following subsections summarize the most significant gaps in terms of each FAIR dimension.

### 3.2.1 Findability: Metadata Gaps and Limited Discoverability

Soil-related datasets are dispersed across multiple platforms, including ESDAC, national soil institutes and national geoportals, INSPIRE geoportal, research and scholar repositories, and project-specific portals, with no unified discovery mechanism. As a result, users often rely on personal networks or ad hoc searches to locate relevant data.

Although INSPIRE mandates metadata, metadata completeness and quality vary significantly between Member States and organizations. Mandatory elements are often filled minimally or inconsistently, which reduces their utility for data discovery. This situation often arises because metadata creation is perceived as an administrative obligation rather than as an integral part of data management. Data providers frequently lack training or dedicated tools to produce high-quality metadata, and existing templates can be complex or poorly aligned with soil data specifics. As a result, metadata is completed only to satisfy formal requirements, without ensuring that they enable effective discovery or reuse of datasets.

Another critical limitation is the lack of Persistent Identifiers (PIDs) such as DOIs. Without them, datasets cannot be reliably cited, tracked across versions, or linked to related resources, limiting both visibility and reproducibility.

Finally, many project outputs are frequently published only within deliverables or PDFs, without structured metadata or repository deposition, reducing visibility and reusability. This often happens because FAIR compliance is treated as a reporting requirement rather than an active process. Researchers tend to focus on delivering scientific results, while repository deposition and metadata creation require additional time, tools, and expertise. Moreover, some projects lack clear guidelines on where and how to store datasets after completion. Without proper metadata, it is often impossible to know what kind of data or knowledge are contained within these documents, under what conditions they were collected, or whether they could be reused for other analyses.

### 3.2.2 Accessibility: Legal and Technical Barriers

Although the open data landscape in Europe is improving, restricted or unclear access conditions remain common, especially for national soil monitoring data or commercially collected soil sampling data.

Licensing represents another major barrier. Even when data are technically available for download, the lack of a clear license, or the presence of ambiguous terms, makes it uncertain whether and how the data can be reused. Many datasets are labelled as “open access” but without specifying what this allows, i.e. whether users can modify, redistribute, or use the data commercially. As a result, potential users often avoid reusing such data out of legal caution, which undermines the very purpose of making them accessible in the first place.

Technical accessibility is also limited. Many legacy datasets exist only in outdated or proprietary formats (e.g. Excel spreadsheets, shapefiles without metadata, or local databases), often lacking open download links or API interfaces. In other cases, access requires bilateral agreements or manual requests, making automated or large-scale integration impossible.

### 3.2.3 Interoperability: Fragmented Standards and Semantic Inconsistencies

The ability of data to be integrated and understood across systems — remains one of the most significant weaknesses in the soil data domain. Multiple soil classification systems (e.g. WRB, USDA national classifications), analytical procedures, and sampling protocols coexist, often without mappings between terms, translation layers, or validated transfer function between analytical procedures. This diversity reflects the long history of national soil science traditions but complicates data comparison and integration.

Semantic inconsistencies are also widespread. Soil attributes are often described using different terms such as “SOC,” “organic carbon,” or “C<sub>org</sub>,” sometimes with incompatible units or reference depths, and even the same term is used for different properties: e.g. soil organic carbon stock (which is a soil organic carbon mass content in a specific measure, that is, a density) is referred as soil organic carbon, confusing it with the soil organic carbon content, which is the content soil organic carbon content as mass percentage. Without a shared vocabulary or ontology, for soil properties, soil analytical procedures and the respective units of measure, automated data processing and cross-domain analysis become unreliable.

Although soil data standards exist at European and national levels, they are usually only partially implemented and do not cover all variables relevant to research or policy (e.g., soil health indicators, biological parameters). Many datasets also lack machine-readable metadata structures and controlled vocabularies, which are essential prerequisites for interoperability and automated integration across systems.

### 3.2.4 Reusability: Lack of Provenance, Licensing and Long-term Preservation

Reusability depends on transparency, documentation, and long-term availability. However, provenance of information, describing how, when, and by whom data were created and processed, is often incomplete or entirely missing. Without it, users cannot assess data quality or suitability for secondary analysis.

Versioning and citation mechanisms are rarely implemented, especially within national or institutional databases. As a result, users cannot trace corrections, updates, or methodological changes over time.

Licensing problems are again a major obstacle. Missing or overly restrictive licenses prevent legal reuse and integration, even when data are technically available, hindering scientific collaboration, cross-border data reuse or commercial application.

Many datasets produced by EU-funded projects are not stored in trustworthy and permanent repositories after the project ends, resulting in data loss or inaccessible supplementary materials. The lack of sustainable preservation strategies further increases the risk of data loss, where data remains on institutional servers or private storage without guarantees of sustainability.

### 3.2.5 Organizational and Cultural Barriers

Beyond technical challenges, organizational and cultural factors play a role in FAIR implementation. Many data providers lack the incentives, resources, or training needed to produce FAIR data. The benefits of data sharing are often unclear or overshadowed by concerns over data misuse, loss of control, or lack of recognition.

Soil data collection is also highly decentralized. It is performed by various actors, such as research projects, local authorities, environmental agencies, and private laboratories, each operating under different mandates and expectations. This fragmentation makes coordination difficult and results in duplicated efforts and inconsistent standards.

Limited collaboration between national monitoring systems, European infrastructures, and research initiatives further intensifies the problem. Without a shared vision or operational framework, data remain locked within institutional silos, and opportunities for integration and innovation are lost.

Current FAIR implementation in soil data is constrained by a combination of technical, semantic, legal and organisational gaps. Many of these deficiencies do not stem from missing data, but from missing structure, policies, identifiers and long-term infrastructure. These gaps emphasise the need for harmonised guidelines and shared vocabularies.

## 4 Guidelines Development Process

Ensuring that soil data and knowledge produced across Europe are FAIR is essential for achieving the sustainability of results and objectives of the EU Soil Mission. Despite numerous initiatives promoting open science and data sharing practices (see chapter 7), soil-related information remains highly fragmented, without clear documentation and with limited metadata (as described in chapter 3.2).

The SoilWise project acknowledges that the world of metadata is complex, hard to understand, and out of scope for many data providers. The SoilWise FAIR Guidelines address these gaps by providing a practical, standardised framework that translates high-level FAIR and policy principles into actionable procedures for data providers, researchers, and policymakers. These guidelines thus form a core pillar of the SoilWise strategy. Moreover, the guidelines provide the information on how to publish compliant to SoilWise harvesting strategies.

This chapter explains how the SoilWise FAIR Guidelines are being developed, including methodology, sources of input, stakeholder involvement, and alignment with EU standards. It ensures that the guidelines are transparent, reproducible, and can be updated during the project's lifecycle and beyond. Note that the guidelines development is an ongoing iterative process, and the guidelines mentioned in this deliverable are not a final product.

The methodological approach for developing guidelines towards FAIR soil data and knowledge handling was shaped over two years of intensive co-design, co-creation, and co-validation with a diverse range of user groups. This collaborative process involved close engagement with Mission Soil research projects, fostering peer-to-peer interactions and leveraging insights from the Data and Knowledge Management Cluster. Through a series of workshops, feedback loops, and iterative testing, the guidelines were refined to ensure they address the unique challenges of soil data management, from data standardization to interoperability. This participatory approach ensured that the resulting guidelines are not only scientifically robust but also aligned with the practical needs and priorities of soil researchers, data managers, and policymakers. The outcome is a set of actionable, user-driven guidelines that will enhance the FAIRness of soil data and knowledge, supporting greater collaboration and reuse across the soil science community.

### 4.1 Methodology

The guidelines have been developed in several steps, in line with the process described in the previous paragraph. First, to establish a baseline for the SoilWise guidelines and possible approaches towards their development, the related policies, standards, and already existing guidelines were reviewed. The development process was designed to translate the requirements of the relevant policies (see chapter 3.1) and various already established guidelines and frameworks (see chapter 7), to build upon them, and to fill the gaps relevant for SoilWise, soil data, and Mission Soil projects in general.

Moreover, the metadata requirements from soil data providers were collected during the stakeholder engagement activities. It was identified that the main gaps during the metadata provision are:

- Insufficient knowledge of FAIR principles in general,
- Not understanding the purpose of metadata provision,
- Insufficient knowledge of expected content of some specific metadata fields,
- No soil-specific metadata fields in most of the common repositories (e.g., lack of descriptors for soil properties, sampling depth, analytical method),

- Decision making regarding the vocabularies or repositories,
- Insufficient knowledge of after-project data/knowledge management practices,
- Problematics regarding different data types.

Among the most common soil-specific metadata fields are soil properties, soil indicators and used methodologies. Without this information, the reusability of soil data is limited.

The outputs of these discussions was a set of user requirements focused on practical guidelines about how to publish FAIR datasets (vocabularies, repositories to choose, etc.). These findings demonstrated that guidelines must not only define what metadata should be provided but also explain why it matters and how to implement it in practice. After the summarization, the most suitable form and content of the guidelines were discussed. It was decided that there is a need for data provider support at two levels: the general, universally oriented level with basic information on FAIR principles, and second, more detailed level, describing best practices and recommendations for different data types.

#### 4.1.1 General Guidelines Methodology

The first level will be covered by the general guidelines (see chapter 5), as an explanation of basic terms and purposes, and a manual for basic metadata provision, with a focus on the specifics of soil data. These general, more universal guidelines should serve as a baseline for soil-related metadata providers to understand the context, concept, and expected content of metadata. The goal is to fill the gaps related to insufficient or incorrect metadata fulfilment, stemming from a lack of knowledge on the provider.

The first version of these guidelines was developed as a list of 6 steps of “How to FAIRify data and knowledge”, with each step divided into more detailed recommendations, accompanied by other basic information. These steps covered both common FAIR recommendations (such as “Use community standards to ensure consistency and interoperability”) and soil-specific recommendations (such as listing standardised vocabularies containing soil-domain terms).

This first version was internally tested to gather initial feedback on the guideline’s content, form, understandability, and level of provided detail. The general approach and direction were considered useful and valid. However, the biggest issues were seen in the level of detail, as the presented guidelines were focusing on “what” should be done to support data FAIRness, but lacked more practical information on “how” it should be done.

Reflecting the feedback, the second version of the guidelines was developed. The second version reordered and supported the initial 6-step approach with more practically oriented examples and descriptions of “How to” implement the principles.

Although most of these guidelines are still universal across stakeholder groups, some examples are inclining towards data providers from academia, and/or on the European level, as stakeholders from these environments are foreseen to be primary users of these guidelines.

#### 4.1.2 Data Type-specific Guidelines Methodology

The second level should focus on specifics of data types (see chapter 6), since different data types have different structures on conceptual and logical data model levels, as well as they may vary in recommendations,

possibilities, and strategies towards metadata management. These guidelines are foreseen to be more detailed, with more tailored recommendations for data providers, dealing with various data and knowledge types.

This approach covers wide range of possible directions and recommendations, due to variety of target groups with different requirements, needs and expectations, as well as number of data types to be published. The methodology is based on requirements collected from interaction with Mission Soil stakeholders.

Based on the outputs of the stakeholder interviews, it was decided to develop a proof-of-concept guidelines on i) soil data in tabular format, and ii) project's websites, and their publication to Zenodo, as a pilot for further extension to other data types. The tabular data were chosen due to their specific nature, complexity, and wide base of data providers among scientists on one side, and limited possibilities for suitable metadata population on the other. Websites archiving was selected as several MS projects (CURIOSOIL, LOESS) indicated concerns regarding long-term preservation of information on their websites. Zenodo was chosen since it is a common repository for various scientific outputs and recommended to Mission Soil projects.

The second step focused on developing the guidelines. For this part, we tested the draft FAIR data strategy (as foreseen in the grant agreement) with a Mission Soil stakeholder. The test showed that more practical explanations were needed. As a result, practical guides for both data types are now being developed and will be presented at the Mission Soil D&K cluster on 2 December 2025.

Moreover, it is foreseen that guidelines for more data types will be developed in the future, because various data and knowledge sources might have different specifications, requirements, metadata annotations or recommended processes for their persistent preservation. For example, in case of soil data in tabular format, it is important to note also soil properties and used methodologies, or in case of website storage, the issue lies in storing various formats and interactivity (html, css, javascripts and other materials interconnected for full functionality). The final list of specific data types and level of guidelines detail were not yet chosen, but the following data types (or their sub selection) are considered:

Considered types of data:

- Raster data
- Vector data
- Geodatabase
- Service
- 3D model
- Tabular data

Considered types of knowledge:

- Publication or journal papers
- Guideline or manual
- Methodology
- Ontology
- Repository
- Models
- Deliverables
- Text
- Websites

- Video
- Audio
- Image
- Presentation
- Best practices and tools
- Interview
- Policy documents
- Software
- Validation results

Moreover, guidelines on following supporting materials are considered to be included:

- Manual on How to fill metadata fields
- Sensitive metadata handling
- Versioning
- Vocabularies
- List of SWC harvested repositories

The final selection of data types for which the guidelines will be developed will be decided upon following interviews with stakeholders and described in the following iteration of this deliverable. The relevance regarding soil data and knowledge will be considered.

## 5 General Guidelines on data and knowledge FAIRness

This chapter presents the current version (v0.2) of the general guidelines. Later versions of these guidelines will be published on Zenodo, as a standalone document. Note that the guidelines are still a work in progress and may change in future iterations:

### 6 Steps Towards FAIR Data and Knowledge in Mission Soil

Soil-related data and knowledge are essential for understanding and improving soil health, guiding sustainable land management, informing policy, and enabling scientific progress. Yet, these resources are often scattered, poorly documented, hard to access, or locked in incompatible formats. This fragmentation hampers their reuse, reduces their impact, and creates inefficiencies across projects and disciplines.

This strategy, developed within the SoilWise project, provides guidance to ensure that all data and knowledge generated, collected, or reused within Mission Soil projects are managed according to the FAIR principles – Findable, Accessible, Interoperable, and Reusable.

By applying the FAIR principles, we can increase the visibility and value of soil-related data and knowledge. FAIRness enables:

- **Transparency** – allowing others to understand how data were produced and under what conditions.
- **Collaboration** – making it easier to combine datasets across projects, regions, and domains.
- **Efficiency** – reducing duplication of efforts and enabling the reuse of existing data.

- **Impact** – enhancing the potential for innovation, decision support, and long-term use of project outputs.

It is intended not only as a conceptual framework, but also as a practical tool to support project teams in developing their own Data Management Plans (DMPs). Each project or work package should describe how its data and knowledge will be handled, shared, and sustained in a FAIR-compliant way.

Soil health knowledge is inherently diverse – ranging from field measurements and lab results to local experiences and spatial models. This strategy promotes coherence and openness while respecting contextual and disciplinary differences.

**Metadata** play a central role in this process: without clear and rich metadata, data cannot be truly findable or reusable. Metadata describe the context, origin, structure, and conditions of use of the data or knowledge asset, acting as the bridge between producers and users — whether human or machine.

**Approach: How to FAIRify Data and Knowledge**

We recommend the following stepwise process for implementing FAIR principles during metadata population:

**Step 1: Inventory & Typology**

Principles	How to
<ul style="list-style-type: none"> <li>• <b>Identify what data and knowledge will be created or reused.</b></li> <li>• <b>Classify them by type, format, origin, and reuse potential.</b></li> </ul>	<ul style="list-style-type: none"> <li>• During project planning or Data Management Plan (DMP) preparation, make a list of all expected outputs: datasets, reports, maps, models, code, workflows, protocols, websites, and other knowledge products.</li> <li>• For each resource, note whether it is newly created, collected from partners, or reused from external sources.</li> <li>• Record the data type and format (e.g., raster, tabular, text, time series).</li> </ul>

**Step 2: Choose Trustworthy Repositories**

Principles	How to
<ul style="list-style-type: none"> <li>• <b>Use trustworthy and permanent platforms for archiving and sharing</b></li> <li>• <b>Prefer repositories that support persistent identifiers (e.g., DOI), machine-readable metadata, and open licenses.</b></li> </ul> <p><b>Remember that FAIR doesn't necessarily mean open. In most</b></p>	<ul style="list-style-type: none"> <li>• Choose a repository before your project outputs are finalized, so you can prepare metadata accordingly.</li> <li>• For academic outputs, use <a href="#">Zenodo</a> (linked to OpenAIRE). For other contexts, explore <a href="#">re3data.org</a> to identify certified or domain-specific repositories.</li> <li>• Verify that the repository supports persistent identifiers (PID, DOI) and offers clear license options (e.g., Creative Commons).</li> <li>• Check repository policy: does it allow embargo periods or restricted access for sensitive data?</li> </ul>

<p><b>repositories, you can also safely store resources with restricted access.</b></p>	
---	--

**How to be discoverable by Soil Wise harvesting strategies?**

For outputs of Horizon Europe research, verify that the selected repository is harvested by [OpenAire](#) and that the Grant number of the project can be added as part of the funding information. SoilWise and Cordis use this information to select relevant publications.

For non-academic repositories, verify that the repository is harvested by [data.europa.eu](#) and add a keyword soil to the metadata record (in the relevant language of the record). Notice that data.europa.eu includes the contents of [ESDAC](#) as well as [INSPIRE Geoportal](#).

**Step 3: Choose Open and Standardized Formats**

Principles	How to
<ul style="list-style-type: none"> <li>• <b>Use non-proprietary, open formats whenever possible</b></li> <li>• <b>Ensure that data formats are machine-readable and suitable for long-term access.</b></li> </ul>	<ul style="list-style-type: none"> <li>• If possible, when preparing data for sharing, export or convert them into open formats:                         <ul style="list-style-type: none"> <li>• CSV, TSV → for tabular data</li> <li>• GeoJSON, GML → for vector spatial data</li> <li>• GeoTIFF → for raster data</li> <li>• PDF/A or TXT → for documents</li> <li>• SQLite, PostgreSQL dumps → for databases</li> </ul> </li> <li>• Avoid file formats that depend on proprietary software (e.g., XLSX, MDB, DOCX).</li> <li>• If necessary for reusability or additional description of your data, include a README.txt file in the dataset or archive describing:                         <ul style="list-style-type: none"> <li>• data structure, variable descriptions, units</li> <li>• coordinate system (for spatial data)</li> <li>• software used to create or open files</li> <li>• known limitations or assumptions.</li> </ul> </li> <li>• Compress related files into a ZIP archive before uploading, ensuring all dependencies are included.</li> </ul>

**Step 4: Metadata Design and Management**

Note that, in most cases, you can create metadata directly in the trustworthy repository during the resource upload.

Principles	How to
<ul style="list-style-type: none"> <li>• <b>Populate as much metadata as you can and describe all the information needed to understand and use your data. You will support the Reusability of the resource.</b></li> <li>• <b>Assign persistent identifiers (PIDs) to data objects, documents, and authors whenever possible.</b></li> <li>• <b>Use metadata to interlink related resources.</b></li> </ul>	<ul style="list-style-type: none"> <li>• Create metadata at the time of upload or using metadata templates provided by your repository.</li> <li>• Fill in all key fields: title, creators (with ORCID if possible), description, keywords, publication date, funding, related resources, license, and version.</li> <li>• Use clear, descriptive titles and structured abstracts that explain what the dataset contains, how it was produced, and its spatial/temporal coverage.</li> <li>• Add persistent identifiers (PIDs): <ul style="list-style-type: none"> <li>• Dataset → DOI</li> <li>• People → ORCID</li> <li>• Projects → Grant number or ROR (Research Organization Registry)</li> </ul> </li> <li>• Interlink related resources, such as datasets, publications, methodologies, software or reports. Most repositories allow “Related identifiers”.</li> <li>• Keep a local copy of metadata (e.g., in JSON, XML, or CSV format) as a backup for institutional archiving.</li> </ul>

#### Step 5: Use Controlled Vocabularies and Ontologies

Principles	How to
<ul style="list-style-type: none"> <li>• <b>Apply standardized vocabularies, code lists or established domain terms to describe data content and context.</b></li> <li>• <b>If possible, avoid creating new vocabularies and ontologies in your project. When local or project-specific terms are used, provide crosswalks or mappings to widely used vocabularies.</b></li> </ul>	<ul style="list-style-type: none"> <li>• Identify the relevant vocabulary or established domain terms source early in the project: <ul style="list-style-type: none"> <li>• <a href="#">AGROVOC</a> for agricultural concepts</li> <li>• <a href="#">GEMET</a> for environmental terms</li> <li>• <a href="#">INSPIRE themes</a> for spatial data</li> <li>• <a href="#">WRB soil classification</a></li> <li>• <a href="#">LUCAS codes</a></li> <li>• <a href="#">EUNIS habitat codes</a></li> </ul> </li> <li>• Use the suggested keywords from your repository’s predefined list — this ensures linking to standard vocabularies automatically.</li> <li>• For attributes or variables in datasets, provide code lists or lookup tables that explain categorical values (e.g., soil texture codes).</li> <li>• When using project-specific terminology, create a mapping table linking local terms to standard equivalents (crosswalk). Store it with your dataset.</li> </ul>

	<ul style="list-style-type: none"> <li>For ontologies or semantic models, reference them by URI where possible (e.g., from <a href="#">INSPIRE Registry</a>, <a href="#">EcoPortal</a>, <a href="#">AgroPortal</a>, <a href="#">BioPortal</a>, or <a href="#">GLOSIS</a>).</li> </ul>
--	---

### Step 6: Define Licensing and Access Conditions

Principles	How to
<ul style="list-style-type: none"> <li>Clearly specify the terms under which data and knowledge can be reused. Always use already existing, standardised licences.</li> <li>Provide machine-readable licensing information in metadata.</li> <li>FAIR doesn't necessarily mean open. In most repositories, you can also safely store resources with restricted access.</li> </ul>	<ul style="list-style-type: none"> <li>Before publication, review whether your data contains personal, confidential, or sensitive information (e.g., farmer names, coordinates of private land).</li> <li>If necessary, anonymize or aggregate data before release.</li> <li>Choose a <a href="#">standard license</a> compatible with your data type (If possible, choose one of the <a href="#">Creative Commons licenses</a>. They are widely used, and easy to understand), such as:                         <ul style="list-style-type: none"> <li><a href="#">CC-BY 4.0</a> – requires attribution (recommended default)</li> <li><a href="#">CC0</a> – public domain</li> <li><a href="#">CC-BY-NC</a> – non-commercial use</li> <li><a href="#">ODC-BY</a> – for databases</li> </ul> </li> <li>Include the license text or identifier (e.g., <a href="https://creativecommons.org/licenses/by/4.0/">https://creativecommons.org/licenses/by/4.0/</a>) in your metadata record.</li> <li>Choose the most open license possible, while respecting privacy, ethical and legal constraints.</li> <li>Ensure the repository captures the license as machine-readable metadata.</li> <li>If access restrictions apply, specify the access level (open, embargoed, restricted) and provide contact information for requests.</li> <li>Mention licensing and access terms clearly in your DMP and README file.</li> <li>When in doubt, consult your institutional data steward or legal office before choosing the license.</li> </ul>

If you are part of a Mission Soil project and you want your resources to be findable through SoilWise Catalogue, the easiest way is to upload them to [Zenodo](#) and link the assets to your project (in the *Funding section*). It will be automatically harvested and logged in SoilWise Catalogue in a few days.

For more information on FAIR principles, you can visit guides at e.g. [OpenAIRE](#) or [go-fair.org](#), or you can test your resource FAIRness at [FAIR Data Self-Assessment](#) or [FAIRsFAIR](#) tools. See also the related official Horizon 2020 documents, [Guidelines on FAIR Data Management in Horizon 2020](#), and [Guidelines to the Rules on Open Access to Scientific Publications and Open Access to Research Data in Horizon 2020](#).

## 6 Data type specific Guidelines on data and knowledge FAIRness

This first version of a practical guidelines displays how to make FAIR data and knowledge discoverable on SoilWise. It considers the use of Zenodo source repository and emphasizes the observational data and websites specificities. The guidelines are published on Zenodo, DOI: [10.5281/zenodo.17693227](https://doi.org/10.5281/zenodo.17693227). The first version is in .pptx format, as the primary mean of dissemination of this version of these guidelines is via presentation. The following iterations of these guidelines will be transformed into other formats as well.

## 7 Additional Sources

This chapter provides selected additional sources that are particularly useful for understanding, implementing, and evaluating FAIR principles in data. Each resource offers guidance, practical tools, or official policy references to support data providers, researchers, and policymakers. However, none of these are focusing on soil data publication specifics. The SoilWise FAIR Guidelines are built upon and aligned with these sources, drawing on their recommendations and frameworks, to ensure consistency with European and international FAIR initiatives. By referencing these materials, the guidelines provide a coherent and practical approach to FAIRification of soil data while remaining connected to established standards and community practices.

### Horizon 2020 Official Documents

The Horizon 2020 [Guidelines on FAIR Data Management](#) and the [Guidelines to the Rules on Open Access to Scientific Publications and Open Access to Research Data](#) provide the official EU framework for research data management in H2020 projects. They cover requirements for data management plans, metadata standards, licensing, and access conditions. For soil-related projects, these documents serve as authoritative references to ensure that datasets meet EU requirements, are legally reusable, and can be integrated into broader European data infrastructures. These guidelines should be the baseline for metadata management for data providers within Mission Soil projects.

### OpenAIRE Guides

[OpenAIRE](#) provides comprehensive [guides](#) on open science and FAIR principles in the European context. The platform offers practical recommendations on metadata, data management plans, and repository use. For data providers, OpenAIRE guides help clarify how to structure metadata, select appropriate repositories, and ensure compliance with European policies, thereby improving discoverability and reusability.

### GO FAIR Resources

[GO FAIR](#) is an international initiative promoting the adoption of FAIR principles across scientific domains. It provides conceptual frameworks, FAIR implementation guides, and community-driven best practices. Data providers can use GO FAIR materials to understand FAIR concepts at a high level, plan implementation strategies, and align soil datasets with broader European and global FAIR initiatives. The strength is in a detailed explanation of each step.

### FAIR Data Self-Assessment Tools

The [FAIR Data Self-Assessment](#) tool allows data providers to evaluate the FAIRness of their datasets through structured questionnaires. By identifying gaps and weaknesses in metadata, accessibility, interoperability, and reuse potential, soil data managers can prioritize improvements, ensure consistency, and support long-term usability of project outputs.

### **FAIRsFAIR Tools**

[FAIRsFAIR](#) offers practical tools, checklists, and workflows to support the FAIRification of data. The resources include automated assessment tools, training materials, and recommendations for repository selection. These tools help soil data providers implement FAIR principles systematically, from metadata creation to data deposition and monitoring of FAIR compliance.

## 8 Conclusion

This first version of the FAIRness strategy provides a structured starting point for strengthening the findability, accessibility, interoperability and reusability of soil data and knowledge across Europe. By analyzing current challenges and assembling a clear methodological framework, it outlines the conceptual foundations, policy context and methodological processes that guide the development of both general and data type-specific FAIR Guidelines. It highlights the diverse obstacles faced by data providers, researchers and public authorities, and formulates a coherent approach for addressing them through harmonized practices, improved metadata quality, stronger semantic alignment and clearer licensing.

The strategy contributes to a shared vision within the Mission Soil community by offering a unified framework that supports collaboration among projects, infrastructures and institutions. The work presented here enhances the capacity of SoilWise to act as a central gateway for soil-related resources and aligns its development with the needs of EUSO, the Soil Monitoring and Resilience Directive and European open science initiatives. As the project advances, the insights and structures outlined in this document will guide technical development, stakeholder engagement and the integration of new data workflows.

The strategy acknowledges the complexity and diversity of soil information and responds to it by proposing practical, user-oriented guidelines that translate high-level FAIR principles into actionable steps. The creation of general guidelines and the development of the first data type-specific guidelines represent an important milestone, as they address the concrete needs of data providers dealing with different formats, data structures and disciplinary requirements. These outputs support better metadata quality, clearer licensing, stronger semantic alignment and improved long-term preservation across the soil-data landscape.

Looking ahead, these deliverable forms the basis for the next version of the strategy, which will incorporate new findings, feedback from stakeholders and lessons learned from the first deployment cycles of the SoilWise Catalogue and associated components. The next version will offer more detailed and comprehensive guidelines, including new data types, expanding the scope of data that can be published, discovered and reused through SoilWise and supporting the long-term sustainability of FAIR soil data practices.

In summary, this deliverable establishes a clear and forward-oriented framework that strengthens the FAIRness of soil data and knowledge, supports more efficient collaboration and contributes to informed decision making in soil health, land management and environmental policy. It represents a significant step toward a robust and interconnected European soil data ecosystem and provides the basis for ongoing development in the next phase of the SoilWise project.

---

## References

Bouman, J., & Veerman, C. P. (2022). Developing Management Practices in: “Living Labs” That Result in Healthy Soils for the Future, Contributing to Sustainable Development. *Land*, 11(12), 2178.

EJP Soil project, Deliverable 6.1 Report on harmonized procedures for creation of databases and maps, available at: <https://edepot.wur.nl/563842>

European Commission (2016). Guidelines on FAIR Data Management in Horizon 2020. Available at: [https://ec.europa.eu/research/participants/data/ref/h2020/grants\\_manual/hi/oa\\_pilot/h2020-hi-oa-data-mgt\\_en.pdf](https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf)

European Commission (2017). Guidelines to the Rules on Open Access to Scientific Publications and Open Access to Research Data in Horizon 2020. Available at: [https://ec.europa.eu/research/participants/data/ref/h2020/grants\\_manual/hi/oa\\_pilot/h2020-hi-oa-pilot-guide\\_en.pdf](https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-pilot-guide_en.pdf)

European Union (2007). Directive (EU) 2019/1024 of the European Parliament and of the Council of 20 June 2019 on open data and the re-use of public sector information (recast). Official Journal of the European Union. Available at: <https://eur-lex.europa.eu/eli/dir/2007/2/oj/eng>

European Union (2019). Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007 establishing an Infrastructure for Spatial Information in the European Community (INSPIRE). Official Journal of the European Union. Available at: <https://eur-lex.europa.eu/eli/dir/2019/1024/oj/eng>

European Union (2020). COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT, THE COUNCIL, THE EUROPEAN ECONOMIC AND SOCIAL COMMITTEE AND THE COMMITTEE OF THE REGIONS A European strategy for data. Official Journal of the European Union. Available at: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52020DC0066>

European Union (2022). Regulation (EU) 2022/868 of the European Parliament and of the Council of 30 May 2022 on European data governance and amending Regulation (EU) 2018/1724 (Data Governance Act) (Text with EEA relevance). Official Journal of the European Union. Available at: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32022R0868>

European Union (2023). Regulation (EU) 2023/2854 of the European Parliament and of the Council of 13 December 2023 on harmonised rules on fair access to and use of data and amending Regulation (EU) 2017/2394 and Directive (EU) 2020/1828 (Data Act) (Text with EEA relevance). Official Journal of the European Union. Available at: <https://eur-lex.europa.eu/eli/reg/2023/2854>

European Union (2025). Directive (EU) 2025/2360 of the European Parliament and of the Council of 12 November 2025 on soil monitoring and resilience (Soil Monitoring Law). Official Journal of the European Union. Available at: [https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=OJ:L\\_202502360](https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=OJ:L_202502360)

Fantappiè, M., Peruginelli, G., Conti, S., Rennes, S., Le Bas, C., van Egmond, F., Smreczak, B., Wetterlind, J., Chenu, C., Bispo, A., Oorts, K., & Bulens, J. (2021). Report on the national and EU regulations on agricultural soil data sharing and national monitoring activities. Zenodo. Available at: <https://doi.org/10.5281/zenodo.10014912>

PROJECT NAME – SoilWise

Project Number 101112838

---

Wilkinson, M., Dumontier, M., Aalbersberg, I. et al. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, 3, 160018. Available at: <https://doi.org/10.1038/sdata.2016.18>