



## D2.3 Developed & Integrated DM components, v3

M31/MARCH 2026



**Funded by  
the European Union**

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

Acronym	SoilWise
Project Full Title	An open access knowledge and data repository to safeguard soils
GA number	101112838
Topic	HORIZON-MISS-2022-SOIL-01-01
Type of Action	HORIZON Innovation Actions
Project Duration	48 months
Project Start Date	1-Sep-23
Project Website	<a href="http://www.soilwise-he.eu">www.soilwise-he.eu</a>
Deliverable Title	D2.3 Developed & Integrated DM components, v3
Delivery Time (DOA)	M31
Deliverable Submission Date	30/04/2026
Status	V3
Dissemination Level	PU - Public
Deliverable Lead	Dajana Snopková (MU)
Author(s)/Organisation(s)	Dajana Snopková (MU), Tomáš Řezník (MU), Paul van Genuchten (ISRIC), Wannes De Man (EV ILVO)
Contributor(s)	MU, WU, ZALF, EV ILVO, DOMG – VL O, CREA, ISRIC, WE, NP
Peer-Reviewers	Céline Blitz-Frayert (CIRAD), Giovanni L'Abate (CREA)
Contact	<a href="mailto:dajana.snopkova@geogr.muni.cz">dajana.snopkova@geogr.muni.cz</a>
Work Package	WP2
Dissemination level	Public
Keywords	data management components, prototype, technical documentation
Abstract	The purpose of this deliverable is to describe the process of co-design, development, validation, and integration of Data Management components of the SoilWise Catalogue and the current version of Technical documentation, comprising a description of the component's functionalities after the second integration and validation cycle.

## Disclaimer

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union. Neither the European Union nor the granting authority can be held responsible for them.

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

AI	Artificial Intelligence
API	Application Programming Interface
CI-CD	Continuous Integration and Continuous Delivery / Development
CIRAD	CENTRE DE COOPERATION INTERNATIONALE EN RECHERCHE AGRONOMIQUE POUR LEDEVELOPPEMENT
CREA	CONSIGLIO PER LA RICERCA IN AGRICOLTURA E L'ANALISI DELL'ECONOMIA AGRARIA
CSW	Catalog Service for the Web
DM	Data Management
EJP	European Joint Programme
ESDAC	European Soil Data Centre
ETS	Executable Test Suite
EU	European Union
EUSO	EU Soil Observatory
EV ILVO	EIGEN VERMOGEN VAN HET INSTITUUT VOOR LANDBOUW- EN VISSERIJONDERZOEK
FAIR	Findable, Accessible, Interoperable and Reusable
INSPIRE	Infrastructure for Spatial Information in Europe
ISO	International Organization for Standardization
ISRIC	STICHTING INTERNATIONAL SOIL REFERENCE AND INFORMATION CENTRE

---

KM	Knowledge Management
M	Month
ML	Machine Learning
MU	Masaryk Univerzity
NP	NEUROPUBLIC AE PLIROFORIKIS & EPIKOINONION
OGC	Open Geospatial Consortium
PSI	public sector information
PU	Public
RDF	Resource Description Framework
SOS	Sensor Observation Service
STAC	Spatio Temporal Asset Catalogue
SWC	SoilWise Catalogue
T	Task
DOMG – VL O	VLAAMSE GEWEST
WE	WETRANSFORM GMBH
WCS	Web Coverage Service
WFS	Web Feature Service
WMS	Web Map Service
WP	Work Package

WPS	Web Processing Service
WR	STICHTING WAGENINGEN RESEARCH
WU	WAGENINGEN UNIVERSITY
ZALF	LEIBNIZ-ZENTRUM FUER AGRARLANDSCHAFTSFORSCHUNG

# Table of Contents

<b>EXECUTIVE SUMMARY .....</b>	<b>8</b>
<b>1 INTRODUCTION.....</b>	<b>9</b>
1.1 PROJECT SUMMARY.....	9
1.2 DOCUMENT SCOPE.....	9
1.2.1 <i>Summary of deliverable version updates.....</i>	<i>11</i>
1.3 DOCUMENT STRUCTURE .....	11
1.4 RELATIONSHIP TO OTHER PROJECT DELIVERABLES.....	11
1.5 RELATIONSHIP TO PROJECT TASKS.....	12
<b>2 METHODOLOGY .....</b>	<b>13</b>
2.1 SOILWISE PROJECT APPROACH.....	13
2.2 CO-DESIGN OF DATA MANAGEMENT COMPONENTS .....	13
2.3 DEVELOPMENT OF DATA MANAGEMENT COMPONENTS.....	16
2.3.1 <i>First development cycle and lessons learned .....</i>	<i>16</i>
2.3.2 <i>Second development cycle .....</i>	<i>16</i>
2.4 VALIDATION OF DATA MANAGEMENTCOMPONENTSS.....	20
2.4.1 <i>First development cycle and lessons learned .....</i>	<i>20</i>
2.4.2 <i>Second development cycle .....</i>	<i>20</i>
2.4.3 <i>Validation reports .....</i>	<i>21</i>
2.5 INTEGRATION OF DATA MANAGEMENT COMPONENTS.....	29
<b>3 DATA MANAGEMENT COMPONENTS .....</b>	<b>30</b>
<b>4 TECHNICAL DOCUMENTATION .....</b>	<b>35</b>
<b>ANNEX I – SOILWISE-HE ROAD MAP .....</b>	<b>36</b>
<b>ANNEX II - FULFILMENT OF THE SYSTEM REQUIREMENTS (DATA).....</b>	<b>37</b>
<b>REFERENCES.....</b>	<b>43</b>

---

## List of Tables and Figures

Figure 1 SoilWise project approach to the delivery of the SoilWise repository based on three development cycles, each comprising four phases ..... 13

Figure 2 High-level architecture diagram of the SoilWise Catalogue (light-colored elements are under development and not part of 3<sup>rd</sup> prototype) ..... 15

Figure 3 The agile process flow applied for the development of the Data Management components ..... 17

Figure 4 Illustration of the SoilWise 2<sup>nd</sup> Iteration Sprint Backlog board containing 24 User Stories (second column) and associated tasks the development team plans to work on or is working on during each sprint (other columns). The backlog also gives an overview of possible blockers and their status. .... 18

Figure 5 Gantt Diagram of third prototype development iterations, version numbers are for reference only ... 19

Figure 6 First round of validation and testing of the SoilWise Finder, comments were provided on top of mock-up designs through a shared Miro Board ..... 23

Figure 7 Second round of validation and testing of the SoilWise Finder, comments and notes corresponding to bugs and issues were provided on top of screenshots of the current version through a shared Miro Board ..... 24

Figure 8 Example of directly entering spatial and observational data for visualization on maps via QGIS interfaces. .... 27

Figure 9 SoilWise Tabular Soil Data Annotation tool ..... 28

Figure 9 SoilWise Road Map. Normal text describes functionality developed in the 1<sup>st</sup> iteration, bold text describes functionality foreseen to be developed in the 2<sup>nd</sup> iteration..... 36

---

---

## Executive Summary

The SoilWise project, under the Horizon 2020 Innovation Actions, aims to develop an open-access soil knowledge and data repository to support users in safeguarding soil across Europe. With 60-70% of European soils currently considered unhealthy, i.e. are losing their capacity to support food production, biodiversity, and the regulation of water, nutrient, and carbon cycles. The project addresses the urgent need for reliable and harmonized data to support informed decision-making at various levels, aligning with the EU Mission 'A Soil Deal for Europe' and other related strategies. SoilWise, spanning 48 months and involving 15 partners, is designed to create a scalable, modular solution that leverages AI and ML technologies to make soil data Findable, Accessible, Interoperable, and Reusable (FAIR).

This deliverable, D2.3 Developed & Integrated DM Components, presents the development, integration, and validation of Data Management (DM) components within the SoilWise Catalogue. This deliverable is already a third version reporting on the development works after the second development cycle out of three foreseen. It builds upon the previous version delivered after the first development cycle (M13). Developments of the SoilWise Catalogue follow architectural and technical specifications, aligning with the broader project objectives and stakeholder requirements. The report details the agile methodology used, emphasizing iterative development cycles, stakeholder co-design, and robust validation processes. Technical documentation available at <https://prototype-3-0.soilwise-documentation.pages.dev/> accompanies this deliverable to provide even more detailed information on the functionality, data flows, and interfaces of delivered components.

The third prototype of the SoilWise integrates a dozen key DM components, including:

- **Harvester:** Collects metadata and knowledge resources.
- **Metadata harmonisation:** Harmonises collected metadata to a common model.
- **Repository Storage:** Provides secure and scalable storage solutions.
- **SoilWise Finder (Metadata Catalogue):** supported by Apache Solr Engine, facilitates querying, metadata downloads, and visualization.
- **Metadata Validation:** Ensures metadata completeness and structural integrity.
- **Metadata Augmentation:** Enhances metadata records with missing spatial information, additional links, and keywords.
- **DOI Resolution widget:** Facilitates validation of DOI and verification of the existence of this record within SoilWise. Enables the suggestion of a new DOI to be harvested.
- **Tabular soil data annotation:** Supports FAIR data publishing by AI-suggested metadata.
- **Transformation and Harmonization:** Makes metadata interoperable and coherent from the users' perspective.
- **Authentication & Authorisation:** Ensuring user management and access control.
- **System Usage & Monitoring:** Supporting operational tasks and administration of the system.
- **Pycsw:** Providing standardised APIs for further reuse of the SoilWise Catalogue harvested content.

Developed through an agile, user-centric approach, these components underwent a rigorous validation process involving technical assessments and stakeholder feedback. Integration followed ISO 19100 standards, progressing from “loose integration” to “tight integration” for enhanced scalability and flexibility.

This deliverable marks a significant milestone in the SoilWise project, laying a robust foundation for the final development phase. The next deliverable version (M46) will focus on deepening integration and expanding functionalities to further align with EU soil health objectives.

# 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 repositories 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 data repository 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 EU SO evolution accordingly**. The SoilWise repository 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 deliverable describes the methodology used in the process of co-designing SoilWise Catalogue (SWC) Data Management (DM) components, as well as their implementation, integration, validation, and delivery of their source code. As such, it documents the developments of the SWC prototype within the second project cycle out of the three planned. The first (described in D2.1) and second prototypes (described in D2.2) have been designed as a proof-of-concept, partly still building on existing (open source) components and integration of independent developments of SoilWise partners. The third prototype was designed taking into account the feedback of our users, JRC, and EC to the previous prototypes, focused primarily on improving the search performance of the SoilWise Catalogue, shifting the emphasis to knowledge from metadata and data, and on a stronger integration of individual solutions into a comprehensive solution. Nevertheless, all the technical development achievements are tightly coupled with the SoilWise mission, scope, and deliverables that have been published so far. Among others, this deliverable (1) comprises achievements that are framed by the architecture defined in D1.3, (2) respects and follows stakeholders’ demands/suggestions described in D1.1, (3) has been shaped by a number of

in-depth discussions with the European Commission, DG REA/JRC/AGRI, and (4) dozens of Mission Soil Horizon Europe projects.

Initially, a high-level architectural **design** was established, which was detailed in deliverable D1.3 and further refined through collaborative efforts in development clusters and monthly meetings. The **implementation** phase adopted an agile strategy (see Chapter **Error! Reference source not found.**). The development process was broken down into iterative cycles (sprints), and along the timeline adaptations were implemented further align the process to the dynamics of the project.

Delivery of the third SoilWise prototype comprises the following Data Management components: (1) **Harvester** component responsible for harvesting metadata and knowledge resources; (2) **Metadata harmonisation** component harmonizing metadata to a common model, and merging multiple instances of a record to single entity; (3) **SoilWise Finder** as the Catalogue UI, powered by **Apache Solr Search Engine**, allowing users to query the harvested knowledge (and data) resources, download their metadata, and provides links to knowledge assets; (4) **Repository Storage** component responsible for the storage of raw and augmented metadata, indexed metadata, user-enhanced content, and knowledge graphs; (5) **pycsw** serving several standardised APIs and enabling access to SoilWise content for third parties; (6) **Metadata Validation** component focusing on evaluating metadata completeness, structure validation, and link liveness assessment; (7) **Metadata augmentation** component powered by AI and LM suited to enrich source metadata records, e.g. by adding new links, extracting spatial information and keywords; (8) **Transformation and Harmonization** component handling tasks such as metadata standards transformation, CRS transformation, data restructuring, and format conversion, including a standalone tool to store data and INSPIRE metadata as Geopackage format; (9) **DOI Resolution widget** enabling DOI validation and suggestion of records to be harvested; (10) **Tabular soil data annotation** component enabling data providers to easily create metadata from their soil measurements; (11) **Authentication and Authorisation** component enabling log in and verification of user access rights, (12) **System and Usage Monitoring** components taking care of system health monitoring, with additional tools to monitor usage statistics. Detailed documentation, source code, and access points are provided for each component in Chapter 3.

The first version of this deliverable (D2.1 Developed & Integrated DM components) was submitted in M13 after the first two phases (co-design and development) of the first project cycle, and the second version (submitted in M18) reported on the development efforts after the additional integration and validation phase. For more information, see Chapter 2.1.

This (third) version reports on the development efforts after the co-design and development phase of the second project cycle. It was created hand-in-hand with the other SoilWise technical Work Packages to seamlessly document all the technical developments and achievements so far. As such, this D2.3 document, together with D3.3 Developed & Integrated KM components (M31) and D4.3 Repository infrastructure, components, and APIs (M31), delivers the third Soilwise prototype. The delivery of the third SoilWise prototype comprises (1) source code of developed technical components (DM + KM), (2) functionality description in the form of external Technical documentation, (3) Release notes (listed in D4.3), (4) three separate accompanying methodological and management overview documents.

The fourth (final) version of this deliverable will be delivered two months before the end of the project in M46 (end of the third project iteration cycle). All the mentioned deliverable versions will describe the technical developments of the SoilWise Data Management components. This technical level will be accompanied by a complex mid- and long-term strategy, within the D2.5 Report on strategy for FAIRness on soil data in two iterations (M27 and M42).

### 1.2.1 Summary of deliverable version updates

The **second version** of this deliverable was, in comparison to the first version, enriched with two new chapters in the Methodology section. Specifically, Chapter 2.1 SoilWise project approach, where the phases of the project cycles are explained in more detail, and Chapter 2.4 Validation of Data Management Components. This chapter was added to describe the activities that took place within the third phase of the first project cycle, as the first version of this deliverable was submitted before the start of this phase. A second version of the SoilWise prototype (its source codes) was released together with this deliverable, thus the Chapter 3 Data Management components was updated correspondingly. Additionally, minor updates were performed in the chapters: 1.2 Document scope, 1.3 Document structure, and 1.4 Relationship to other project deliverables.

The **third version** of this deliverable contains minor changes in the chapters: 1.2 Document scope and 1.4 Relationship to other project deliverables. Methodology chapters were updated according to the approach taken in the second development cycle, while for the development and validation phase, we also added a summary and lessons learned from the previous cycle. Detailed validation reports are provided in the chapter 2.4.3. The third version of the SoilWise prototype (its source codes) was released together with this deliverable, thus the Chapter 3 Data Management components was updated correspondingly. Additionally, in line with the expected Grant agreement amendment, the designation of the SoilWise Repository was changed to the SoilWise Catalogue.

## 1.3 Document structure

This document comprises the following chapters:

- **Chapter 1** provides an introduction to the project and the document
- **Chapter 2** explains the methodology used in the process of designing, implementing, integrating, and validating SoilWise Repository Data Management components. Validation reports for developments considering DM components are included.
- **Chapter 3** presents a list of Data Management components that are present in the first delivered prototype. It includes an external link to the Technical documentation that further details the component's functionality and technical specifications.
- **Chapter 4** details the Technical Documentation

Technical documentation is available at <https://prototype-3-0.soilwise-documentation.pages.dev/>, and a PDF exported version is also available on demand as a non-editable version saved at the date of the deliverable. The documentation includes links to the current versions of the components' code on Github, or version published on Zenodo, when available.

## 1.4 Relationship to other project deliverables

This deliverable relates to and complements the following deliverables:

- D2.1, D2.2, D2.4 – Developed & Integrated DM components, v1, v2, v4 (M13, M18, M47)
- D3.1, D3.2, D3.3, D3.4 – Developed & Integrated KM components, v1, v2, v3, v4 (M13, M18, M31, M47)

- D4.1, D4.2, D4.3, D4.4 – Repository infrastructure, components and APIs, v1, v2, v3, v4 (M13, M18, M31, M47)
- D1.3 – Repository architecture, v1, v2 (M08, M42)
- D1.1, D1.2 – Usage Scenarios, Requirements, v1, v2 (M6, M36)
- D1.5, D1.6 – Repository GM, v1, v2 (M21, M42)
- D4.5 – Repository Data and Knowledge Resources, v1, v2, v3 (M21)
- D4.6, D4.7 – User Manual, v1, v2 (M34, M46)
- D5.3, D5.4, D5.6 – Deployment and Evaluation Report, v1, v2, v3 (M21, M34, M46)
- 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:

- T1.3 Requirements, Validation framework and Rolling plan – will feed and update the design and implementation of DM components (T2.1 and T2.2)
- T1.4 Define SoilWise Architectural Design – will be considered in the design and implementation of DM components (T2.1 and T2.2)
- T1.5 Define SoilWise Multi-Stakeholder governance model – will be considered in the design and implementation of DM components (T2.1 and T2.2)
- **T2.1 Design of the data technology components**
- **T2.2 Implementation and deployment of data components**
- T2.3 AI and ML for data findability and accessibility – will extend the design and implementation of DM components (T2.1 and T2.2)
- T2.4 Strategy for FAIRness on soil data – will be followed in the design and implementation of DM components (T2.1 and T2.2)
- T3.1 Design of the Knowledge Management components – will consider and align with the design of DM components (T2.1)
- T3.2 Implementation and deployment of Knowledge Management components – will consider and align with the implementation of DM components (T2.2)
- T4.1 Repository digital infrastructure for deployment and delivery – will integrate implementation of DM components (T2.2) and KM components (T3.2)
- T4.2 Interfaces for access, sharing, population and integration with EUSO – will be considered in the design and implementation of DM components (T2.1 and T2.2)
- T4.3 Solutions & repository validation and population – will validate designed and implemented DM components (T2.1 and T2.2)
- T5.2 User Cases implementation and demonstration – will demonstrate the functionality of implemented DM components (T2.2)
- T7.2 Technical and Scientific Management – prepares an Open Science and Data Management Plan (D7.2 – D7.4) describing the data management level of handled (meta)data resources.

## 2 Methodology

### 2.1 SoilWise project approach

The development of the SoilWise data and knowledge repository follows a user-centred, agile, and iterative approach, fostering collaboration among stakeholders, including data owners, reusers, managers, technology experts, policymakers, academics, and regional stakeholders. This four-phase process, illustrated on Figure 1, unfolds across three development cycles. **Phase 1** involves co-design and stakeholder engagement to identify innovation needs and align the repository with practical governance frameworks, emphasizing open innovation, co-creation, and multi-disciplinary problem-solving. **Phase 2** builds on these insights, developing technological and knowledge components through adaptive sprints, with feedback loops ensuring synchronization and continuous improvement. **Phase 3** integrates these components in a shared repository, where solutions are tested and tailored to user needs, combining existing datasets with project advancements to achieve the desired technical and usability level. In **Phase 4**, the solutions are demonstrated and evaluated within user cases, ensuring they address diverse stakeholder requirements, enhance data sharing and governance models, and provide added value for target groups. The evaluation process involves feedback from users, consortium partners, and external evaluators to ensure the repository's broad applicability and effectiveness for real-world challenges.



Figure 1 SoilWise project approach to the delivery of the SoilWise repository based on three development cycles, each comprising four phases

The following chapters describe these phases in more detail from the Data Management components' perspective.

### 2.2 Co-design of Data Management components

The co-design phase is a crucial step preceding the technical development. Co-design phase results in the list and detailed description of desired functionalities to be developed, incl. acceptance criteria. In the **first development cycle**, the co-design phase was driven by vision scenarios and functional requirements (referred to as functionality points), based on main requirements. In the **second development cycle**, **24 user stories** were formulated to provide a base for functionality description and demonstration scenario. More information about

the process and their outputs can be found in Deliverables D1.1. and D1.2 (Usage scenarios, requirements, v1, v2).

Additionally, a high-level architectural design is produced (within the framework of T1.4 Define SoilWise Architectural Design). Its first version was the subject of deliverable D1.3 Repository architecture, v1., while some of the individual functions of the components were already detailed in the Technical documentation. After the submission of D1.3, work on the more detailed level of architecture design continued separately in the so-called “development clusters”, each having its dedicated technical component. Meanwhile, the architectural design and the Technical documentation were updated to reflect the recent design suggestions and development activities. Figure 2 displays the up-to-date version of the High-level architectural design corresponding to the release of the 3<sup>rd</sup> SoilWise prototype. An interactive and full version can be found at: <https://prototype-3-0.soilwise-architecture.pages.dev/> under Views tab.

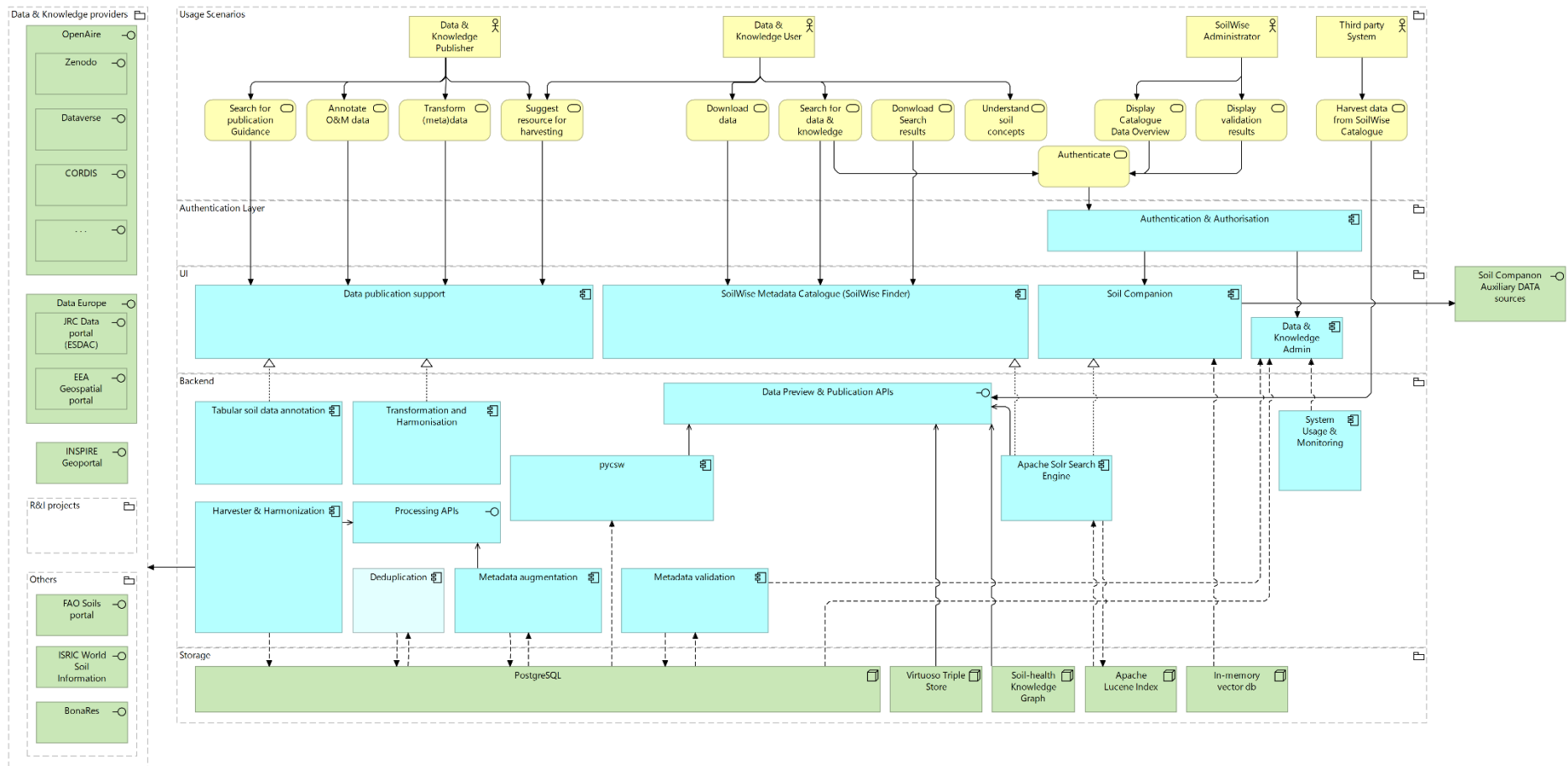


Figure 2 High-level architecture diagram of the SoilWise Catalogue (light-colored elements are under development and not part of 3<sup>rd</sup> prototype)



## 2.3 Development of Data Management Components

### 2.3.1 First development cycle and lessons learned

During the first development cycle, the Data Management components were developed following a structured agile methodology based on **one-month sprints**. Development activities were organised through a Product Backlog and Sprint Backlogs, with tasks defined as issues with Definition of Done (DoD) criteria, directed by user stories and user and functional/technical requirements. Coordination was managed via dedicated GitHub repositories and sprint boards, and six sprints were completed to deliver the first prototype comprising 11 core functionalities. More information can be found in the previous version of this deliverable (D2.2 Developed & Integrated DM components, v2).

While the sprint-based approach provided a clear framework and transparency for planning, implementation, review, and retrospective evaluation, several drawbacks became apparent. Aligning development work across multiple teams from different consortium partners proved challenging, particularly when coordinating shared deadlines within fixed sprint cycles. In addition, the one-month sprint structure was sometimes insufficiently flexible to accommodate newly emerging feedback, evolving requirements, or releases planned for specific testing sessions or demonstrations. As a result, the rigid sprint framework occasionally limited the teams' ability to adapt quickly to changing priorities.

### 2.3.2 Second development cycle

An important difference compared with the first iteration is the creation of a set of User Stories based on requirements coming from (i) the SoilWise Grant Agreement, (ii) User Cases, and (iii) JRC. To maintain a high-level management overview, a "SoilWise Roadmap" was created, including all requirements from each iteration. A full overview of this roadmap is available in Annex I.

In the first development cycle the development of Data and Knowledge Management (KM) components occurred in independent environments, with **integration** achieved through open APIs, resulting in a "loose integration" per ISO 19100 standards. The second project cycle focused on migrating distributed development activities to a single test and production environment. From there, integration depth will be gradually enhanced, progressing from "mixed integration" to a "hybrid model" balancing tightly integrated core components with loosely integrated modular elements, ensuring scalability and flexibility of the SoilWise Catalogue (SWC).

The development process for the Data Management components in the second iteration was based on an agile project management practice. This time, however, we did not follow a joint approach for all developments, but rather approached each user story individually. In the co-design phase, each User Story received an appointed Problem owner and Technical lead, responsible for following up with further clarification of requirements, planning the work, and taking care of reporting for monitoring purposes.

This approach, illustrated in Figure 3, again follows an iterative workflow:

1. Adding User Stories, requirements, acceptance criteria, and functionality descriptions to the Product Backlog<sup>1</sup>. For the second iteration, the Product Backlog was populated with User Stories.
2. Reviewing and prioritising tasks in the Product Backlog, identification of blockers.
3. Developing technical component functionality according to the tasks.
4. Verifying acceptance criteria.
5. Sprint Review and Retrospective (conducted during monthly WP2+3+4 meetings).

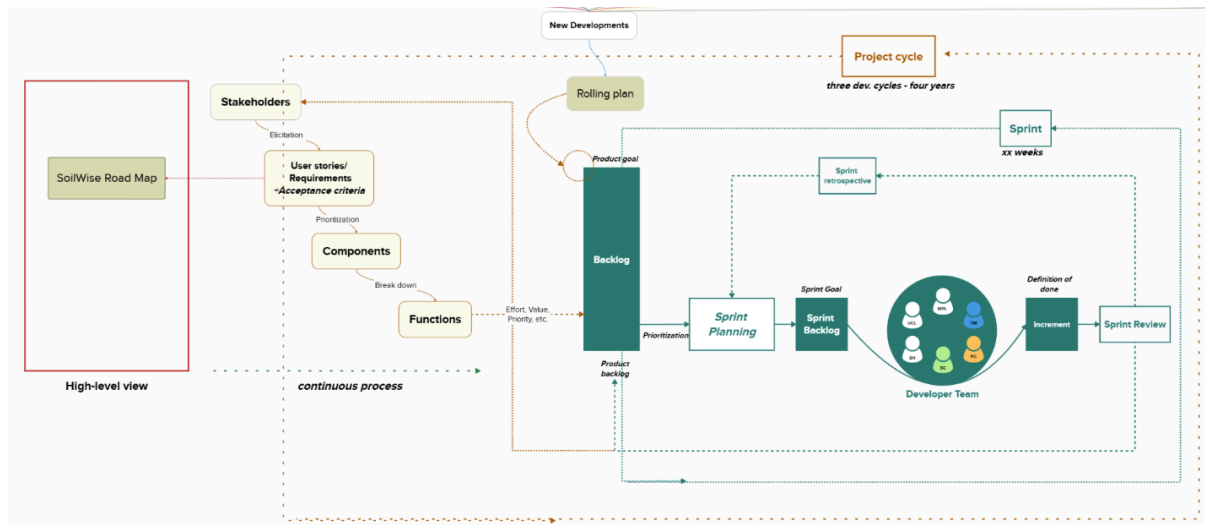


Figure 3 The agile process flow applied for the development of the Data Management components

Coordination of developments was managed using GitHub and a dedicated GitHub environment at <https://github.com/soilwise-he>. Separate repositories were used for each technical component, containing both the developed code and tasks (GitHub issues). A GitHub project *SoilWise Sprint Backlog – second iteration* (Figure 4) was used to plan and track the developments of all User Stories. Problem owners, developers, and collaborators involved in developing a User Story were also identified, forming the developmental team for that User Story.

<sup>1</sup> a prioritized list of functionality which a product should contain and that drives the developmental work

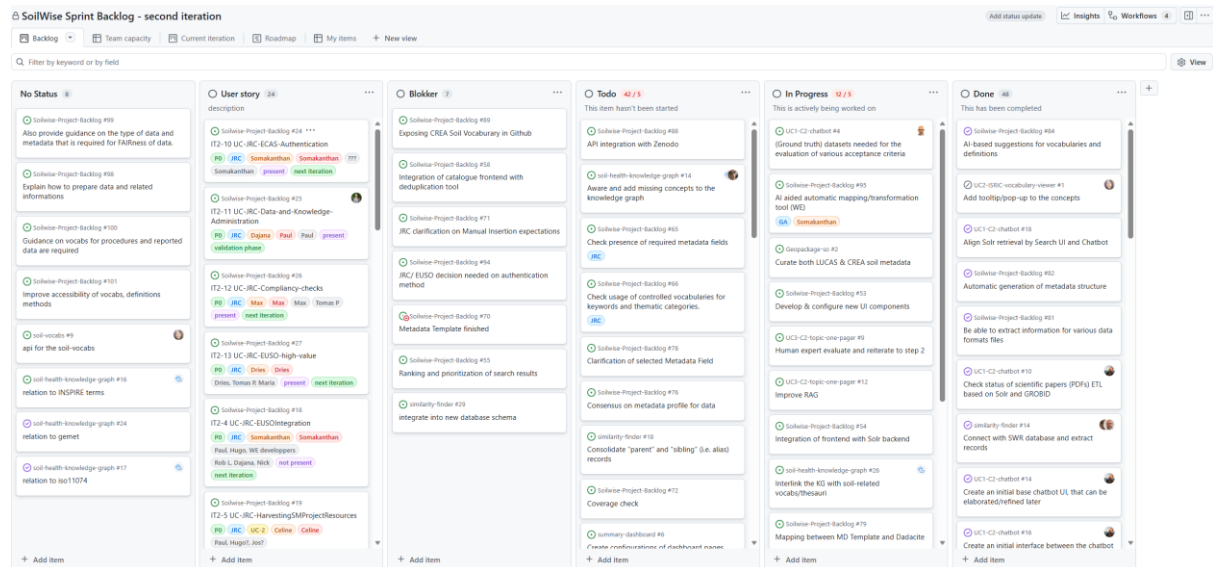


Figure 4 Illustration of the SoilWise 2<sup>nd</sup> Iteration Sprint Backlog board containing 24 User Stories (second column) and associated tasks the development team plans to work on or is working on during each sprint (other columns). The backlog also gives an overview of possible blockers and their status.

In order to monitor the progress of the development work in all User Stories, bi-weekly **User Story meetings** were held between WP2-3-4 leads, User Story problem owners, and EV ILVO as technical manager (M25-M27). Meetings were held in two major groups, one working on metadata, Grant Agreement, and JRC requirements, and another one on search functionality and chatbots. These meetings focused on monitoring the progress of all User Stories, identifying possible blockers or issues, and facilitating communication and collaboration between User Story groups. In addition to the User Story meetings, **development standups** occurred every 2 weeks, allowing developers to share issues and questions or request support. These standups encouraged co-development with developers from different teams in the consortium. Monthly **retrospective & planning meetings** were held between WP2, WP3, and WP4 leads and the Technical Coordinator to monitor the general progress in respective WPs.

As mentioned earlier, the development in the second iteration followed an agile and continuous approach in order to accommodate newly emerging feedback, evolving requirements, or releases planned for specific testing sessions or demonstrations. Instead of applying a rigid schedule where all User Stories would follow the same scheme, small User Story-specific iterations of co-design, development, and validation were implemented. This proved beneficial, as different functionalities require different needs and timelines for validation, testing, and feedback (see also Section 2.4.2). In addition, User Stories that were not fully developed could already collect tailored feedback on functionalities that were completed. In total, the second development phase consisted of 24 User Stories, which served as the primary organizational unit for tracking work and structuring development

efforts.

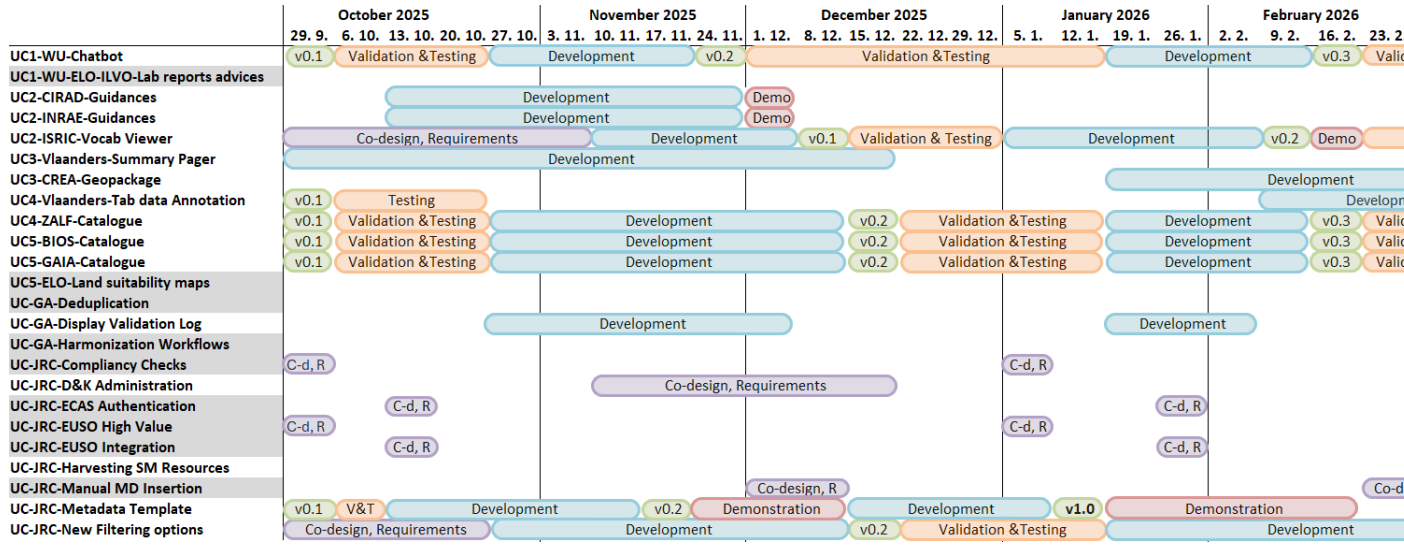


Figure 5 provides an overview of the internal iteration cycles for each User Story.

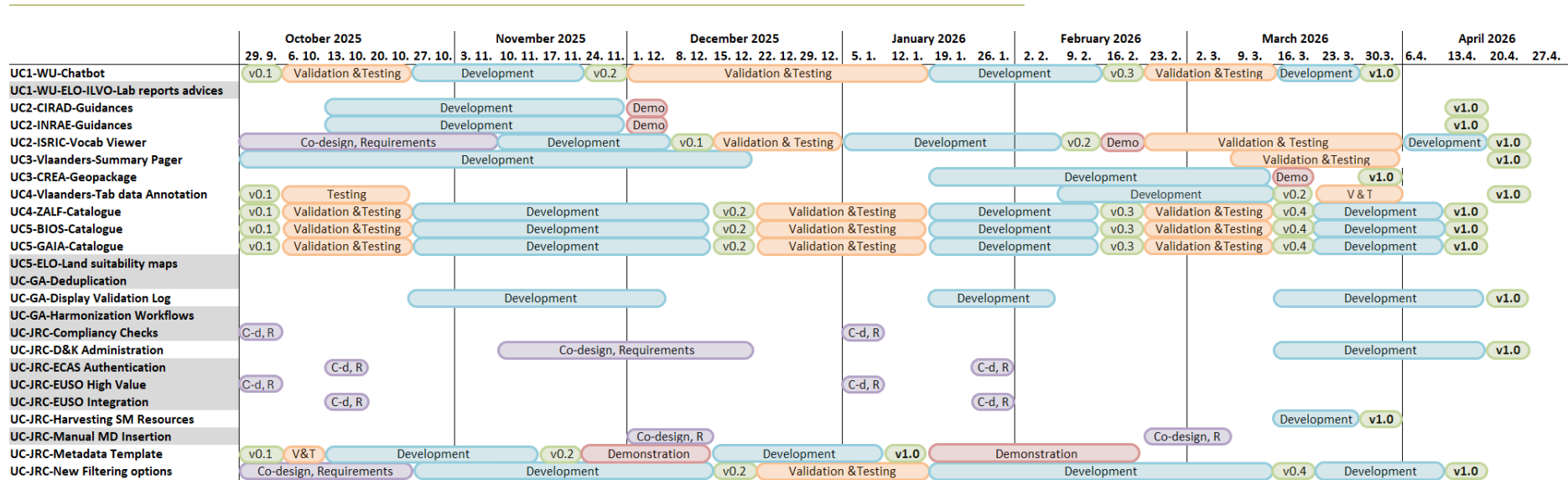


Figure 5 Gantt Diagram of third prototype development iterations, version numbers are for reference only

While the progress of the development phase was monitored closely, the following issues resulted in a delay of developmental work:

- Several User Stories were affected by the **restructuring of the database**, which was needed to enable distinguishing between original and augmented metadata, and to change the target harvesting metadata schema. This caused a delay in the developmental work foreseen, especially the Metadata augmentation component, and the adaptation of the search engine configuration and Catalogue functions for the newly defined database structure.
- **Migration of distributed developments** to a single hosting environment was thus also delayed, since we wanted to avoid duplicate efforts and waited until the development was finished in the original environment.
- While User Stories were defined during the co-design phase, **additional requirements from JRC** were added throughout the development phase and also changed throughout the cycle. These unforeseen requirements needed to be matched to dedicated User Story groups while development was already ongoing. Although challenging, the flexibility of consortium partners and the agility of the development framework allowed us to include some JRC requirements as User Stories in the current and next cycle.

As several issues resulted in a delay of developmental work, developmental work in the majority of User Stories continued during the validation phase. Bi-weekly User Story meetings were prolonged to monitor both ongoing development and validation of User Stories.



## 2.4 Validation of Data Management Components

### 2.4.1 First development cycle and lessons learned

The validation phase of the first iteration was characterized by adaptive and corrective developmental activities driven by (i) lower priority requirements from the prior phase and (ii) feedback originating from the technical validation activities carried out by the User Case (UC) groups in task T4.3 Solutions & repository validation and population. Feedback from the validation activities was gathered through a general online form, centralized in an Excel document, regularly reviewed, and translated into technical issues. These issues, which were classified as feature requests or bug reports critical to showcasing the demonstrators, were uploaded to the SoilWise GitHub repository and assigned to developers for further resolution.

To foster collaboration, biweekly meetings and workshops were organized between developers and UC partners. These sessions encouraged discussions on SoilWise shortcomings and identified functionality requirements for the next project iteration. Feedback from these interactions was also gathered via the online form, processed, and addressed in the same structured manner as described in the paragraph above. More detailed information on the validation of Data Management components in the first iteration can be found in D4.5 Repository Data and Knowledge Resources, v1, and D5.4 Deployment and Evaluation Report, v1.

Although the general feedback form provided a good base for validation of components during the first iteration, there was an increasing need for tailored feedback in the second iteration. In addition, the general form required manual assignment of feedback to relevant components. This was manageable for the number of components in the first iteration, but it would have been too cumbersome in the second iteration. While the first iteration focused on the development of components on the catalogue itself, the developments in the second iteration were more user-oriented and thus allowed a more specific approach to testing and validation by users. Considering these needs, the approach for validation was modified to a more agile and continuous solution, focusing on User Story-specific tools and timelines.

### 2.4.2 Second development cycle

Phase 3 of the second project iteration, **Integration and Validation** (M28-M31), was characterized by a continuous and agile development approach combined with monitoring and identifying functionalities ready for validation and testing by User Case (UC) groups. The validation framework will be described in *Deliverable D1.2 Usage Scenarios, Requirements v2*, providing detailed information on the approach for validating the functionalities developed in the second iteration.

Development activities were driven by (i) requirements from the prior phase and (ii) by feedback originating from the technical **validation** activities carried out by the UC groups in the task *T4.3 Solutions & repository validation and population*, and (iii) new and updated requirements defined by the JRC within both bilateral SoilWise – JRC meetings and related developments of EUSO & Mission Soil. As the development phase of the second iteration cycle was challenged by several blocking issues, the validation phase required an agile and continuous approach. Efforts for both development and validation, therefore, partly occurred simultaneously, allowing additional time for development without significantly impeding validation of newly developed functionalities.

To foster such agile and continuous development, biweekly meetings between User Story problem owners and developers were continued (M28-31) to monitor the progress of remaining developmental work from phase 2 while also identifying functionalities ready for validation by UC groups. These sessions thus focused on general updates, timing, blockers, and status for validation. Once a User Story was identified to be ready for validation by UC groups, the User Story Problem Owner and/or developer gave an internal demonstration to UC groups during upcoming WP5 meetings. These internal demonstrations focused on the newly developed functionality and how feedback could be provided.

The collection of feedback from UC groups was tailored to the functionalities developed within each User Story. Depending on the User Story, feedback was gathered through online forms or directly in the tools developed. Rather than centralizing all feedback received on the SWC as a whole (cf. first iteration), User Story Problem Owners directly received feedback in the most appropriate form and could directly implement additional changes. Internal demonstrations fostered collaboration between UC groups and developers, stimulating interaction and a clear understanding of how feedback should be provided. Feedback directly flowed from UC groups to developers, who could regularly review and translate feedback into technical issues in the SoilWise GitHub repository and assign them to developers for further resolution.

Validation activities were conducted in preparation for the individual demonstrators for each of the five UC groups and were coordinated under the scope of WP5. *Deliverable D5.4 Deployment and Evaluation Report, v2 (M34)*, will provide a comprehensive overview of the demonstrations of the second iteration and their outcomes. The demonstrations will be presented to stakeholders and evaluated in Phase 4 of the project cycle. A detailed overview of USs covering **Data Management components** and the status of their acceptance criteria is provided in Annex II of this deliverable.

### 2.4.3 Validation reports

The following paragraph outlines the validation reports of the different User Stories, focusing on data components developed during the second development cycle. The reports describe the process of validation and collection of feedback for each of the User Stories. Validation reports covering knowledge components are described in D3.3 Developed & Integrated KM components, v3.

Table 1 gives an overview of all User Stories part of the second development cycle and whether they are reported in the Knowledge or Data Management deliverable. In the following paragraphs, a short description of the used validation method(s) is provided for the Data Management related user stories.

*Table 1 Overview of SoilWise User Stories identified the second development cycle.*

User Story name*	Described in D2.3	Described in D3.3
IT2-2-UC-GA: Display MD Validation Results	x	
IT2-20-UC-UC1-WU: SoilWise Knowledge (conversational) chatbot		x
IT2-21-UC-UC2-ISRIC: Vocab viewer		x
IT2-18-UC-UC2-CIRAD2: Guidance on sustainable websites publication		x
IT2-19-UC-UC2-INRAE: Guidance for FAIR soil data publication	x	
IT2-17-UC-UC3-VLAANDERS: Summary pager		x
IT2-22-UC-UC3-CREA: Geopackage	x	
IT2-14-UC-UC4-ZALF: SWC search performance improvement, download	x	
IT2-23-UC-UC4-VLAANDERS: Tabular soil data annotation, metadata creation	x	
IT2-15-UC-UC5-GAIA: SWC UI improvement	x	
IT2-16-UC-UC5-BIOS: SWC search performance improvement	x	
IT2-7-UC-JRC: Metadata template	x	
IT2-8-UC-JRC: New filtering options	x	
IT2-11-UC-JRC: Data and knowledge administration console		x
IT2-5-UC-JRC: Harvesting Mission Soil Project Resources	x	
IT2-3-UC-GA: Harmonization Workflows	-	-
IT2-1-UC-GA: Deduplication	-	-

IT2-24-UC-UC1-WU-ELO-ILVO: Advice based on in-situ soil measurements and lab reports	-	-
IT2-25-UC-UC5-ELO: Land suitability maps	-	-
IT2-4-UC-JRC: EUSO integration	-	-
IT2-6-UC-JRC: Manual Insertion	-	-
IT2-10-UC-JRC: ECAS authentication	-	-
IT2-12-UC-JRC: Compliance checks	-	-
IT2-13-UC-JRC: EUSO high-value datasets	-	-

\* Codenames for the User Stories are built as follows: “ITx-y-UC-Requirement: Title”, with x the iteration number; y the number assigned to the specific User Story; requirement the Grant Agreement (GA), User Case (UC), or JRC; and title the actual name of the functionality developed. User Stories marked in grey are postponed to the next iteration.

#### 2.4.3.1 Metadata Catalogue front-end (SoilWise Finder)

The development of the SoilWise Finder was spread over four User Stories:

- IT2-14-UC-UC4-ZALF: SWC search performance improvement, download
- IT2-15-UC-UC5-GAIA: SWC UI improvement
- IT2-16-UC-UC5-BIOS: SWC search performance improvement
- IT2-8-UC-JRC: New filtering options

A **first round of feedback** was organized during the SoilWise consortium meeting in October 2025. Use Case partners provided feedback on the general design of the user interface in the form of post-its on printed mock-up screens. They could also try out the first released prototype of the catalogue search interface. The collected

feedback was transferred to a Miro Board (Figure 6) to enable further continuation of feedback collection, among other during dedicated meetings with user stories partners.

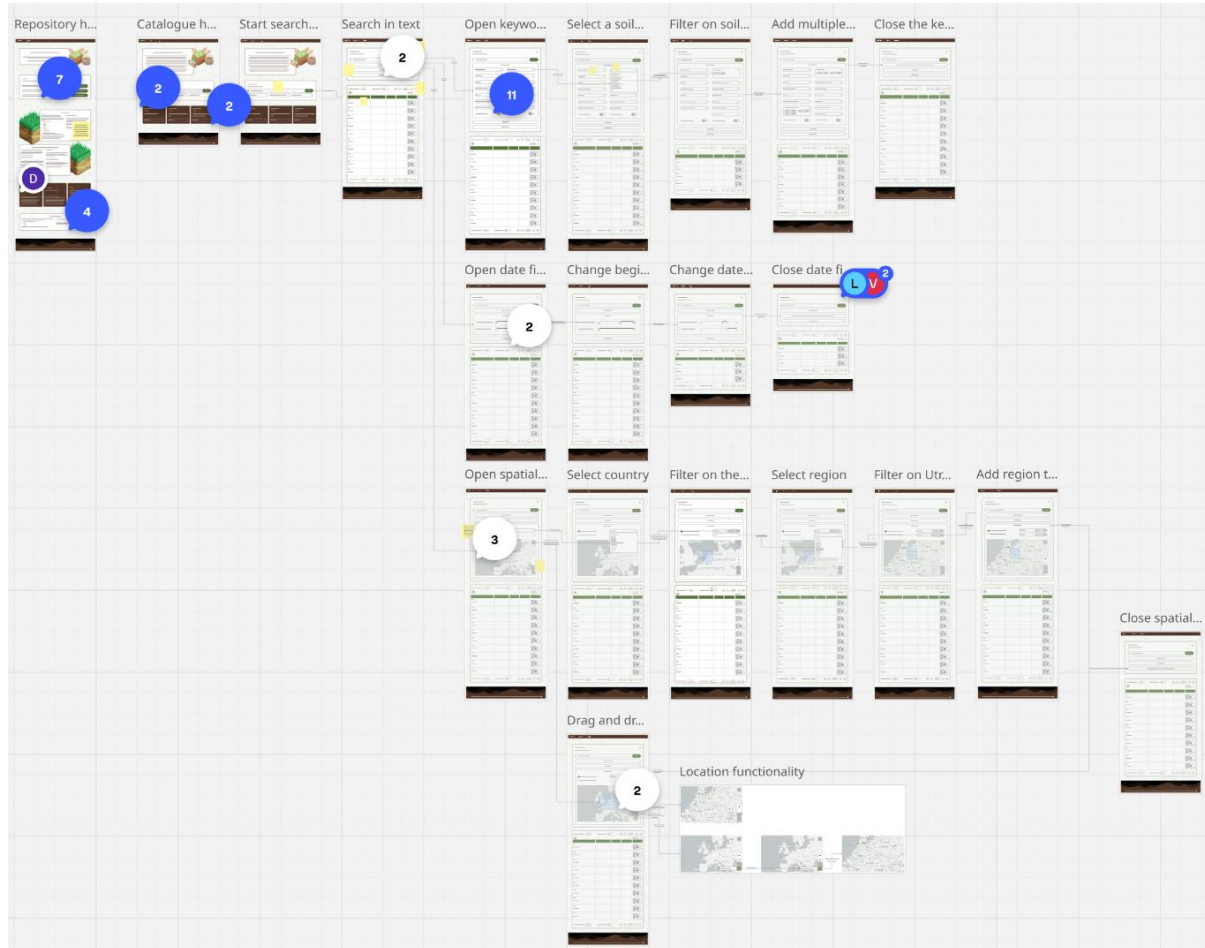


Figure 6 First round of validation and testing of the SoilWise Finder, comments were provided on top of mock-up designs through a shared Miro Board

A **second round of feedback** was organized from December 2025 to January 2026. This time, the new SoilWise Finder was presented to the UCs to start a first round of validation and feedback focusing on the search functionality (fulltext, spatial, term filters) and a common frontpage (general design, responsivity, names, titles, tooltips, understanding of individual functions, filters, etc). Feedback was again collected by using the Miro board, containing various screenshots of the current SoilWise Finder prototype, where UCs partners could share their comments or indicate if they experienced any bugs or items that did not work intuitively (Figure 7).

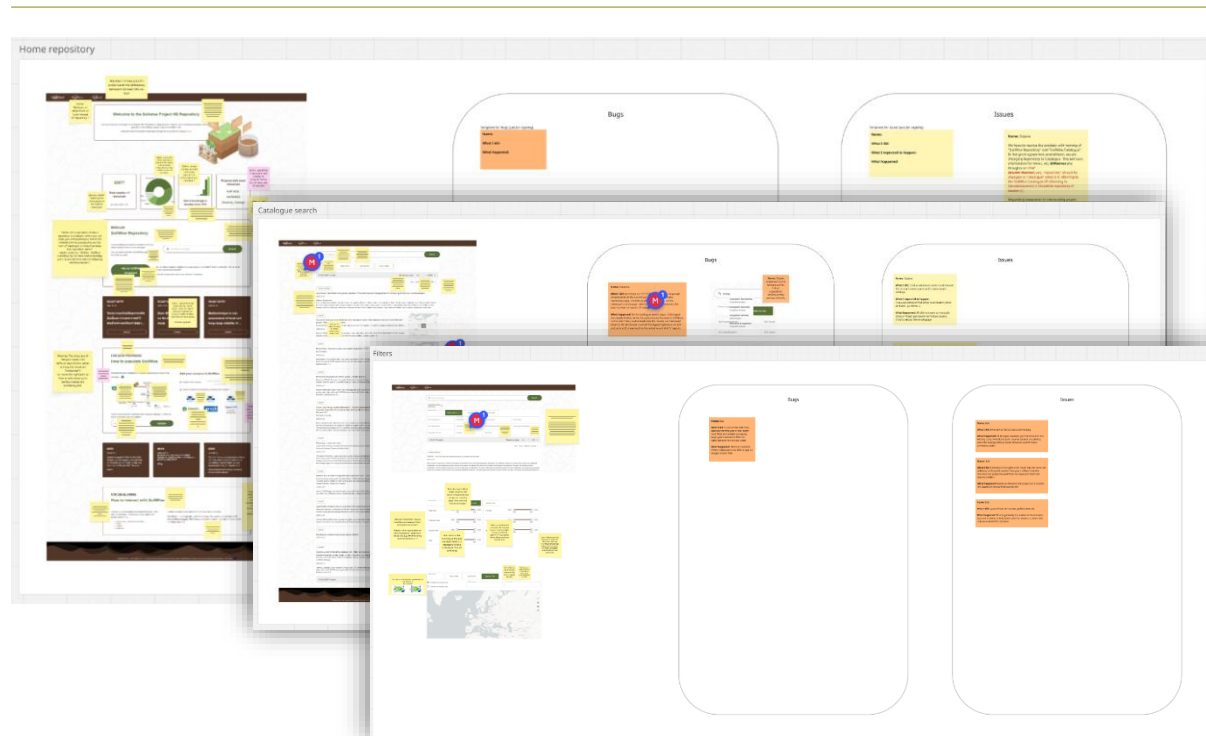


Figure 7 Second round of validation and testing of the SoilWise Finder, comments and notes corresponding to bugs and issues were provided on top of screenshots of the current version through a shared Miro Board

A **third round of feedback** was organized in February 2026, where UCs were invited to provide input on the record's detail page. As before, feedback was collected through a Miro board, containing mock-ups of the SoilWise Finder. UCs were specifically instructed to provide feedback on the new side-by-side layout and whether the information provided was complete.

For demonstration purposes, feedback collection continues via **Hotjar** Game and Feedback surveys. The Game survey is used throughout the demonstrations to interactively present the Metadata Catalogue to a new audience. The Feedback survey is used between the demonstrations to collect more general feedback and to handle reporting of potential bugs.

#### 2.4.3.2 Metadata template

The development of the Metadata template for data providers was part of the IT2-7-UC-JRC: Metadata template.

The development and validation of the Metadata template were conducted in cooperation with JRC, as the primary recipient and future adopter of the template.

The **first round** of feedback was organized as a series of facilitated discussion sessions during the SoilWise consortium meeting in October 2025. The feedback session with SoilWise project partners focused on reviewing the metadata fields and dissemination possibilities. The main goal was to gain initial feedback, gather recommendations, and solve open questions raised during the development of the template. The results of the sessions shaped subsequent developments and expectations.

The **second round** of feedback was held in November 2025 with a total of 8 representatives of three selected Mission Soil projects, namely projects ARAGORN, BENCHMARKS, and ISLANDR. During the online meetings, the concept and purpose of the Metadata template were presented, and the feedback was collected through a Mentimeter questionnaire, as well as through facilitated discussions. The primary goal was to gather feedback

and insights from Mission Soil data providers. The focus was on the understandability, content, and importance of metadata fields, as well as the expected content and requirements of the Metadata template in general. The feedback was then incorporated into the Metadata template and demonstrated shortly after at the Data & Knowledge Cluster meeting.

The **third round** of feedback was collected from Mission Soil projects and JRC representatives during the Data & Knowledge Cluster meeting in Brussels on December 2, 2025. The structure of these sessions was similar to the online meetings executed during the second round of feedback, i.e., several rounds of facilitated discussions with a Mentimeter questionnaire to collect the feedback. However, the audience of these live sessions included not only data providers but also project managers, data stewards, and JRC/EC representatives, providing more detailed and broader feedback than the online sessions.

The notes and Mentimeter results from the second and third feedback rounds were valuable inputs incorporated into the Metadata template. The primary changes from these sessions included field renaming, refinements to soil-specific fields, and also major adjustments regarding field multiplicity. At the same time, several fields proved to be unnecessary, whereas others (e.g., Data Collection Procedures) were mentioned as a beneficial addition.

Several open topics remain, such as codelist specifications, means and willingness to fill out the template by data providers, and fitness of the Metadata template within the scope of the SoilWise project. However, these issues were not the primary focus during the feedback sessions and are the subject of ongoing discussions with JRC.

#### *2.4.3.3 Guidance for FAIR soil data publication*

The development of guidances for data providers was part of the following two User Stories:

- IT2-18-UC-UC2-CIRAD2: Guidance on sustainable websites publication
- IT2-19-UC-UC2-INRAE: Guidance for FAIR soil data publication

Both User Stories were created based on user needs identified during dedicated stakeholder interviews conducted between February and June 2025. Stakeholders involved were mainly part of Mission Soil projects (e.g., Sus-Soil, Curiosoil, Soilguard, Loess, Edaphos, Seacure, Bin2Bean, Echo, Gov4All), but also researchers outside the Mission Soil program.

From July 2025 until December 2025, the practical guidances were developed for FAIR publications based on Zenodo, applied to observational data and websites. In November 2025, external validation with Mission Soil projects was executed to collect additional feedback that has been incorporated into the first version of the guidances.

Finally, the first version of the guidances has been published and deposited on Zenodo in November 2025 and is accessible via the following link: <https://doi.org/10.5281/zenodo.17693227>. The published guidances have been demonstrated to Mission Soil projects during the Data & Knowledge Cluster meeting in Brussels on December 2, 2025.

In January 2026, the practical guidances were demonstrated internally to SoilWise UCs for collecting additional feedback that can be incorporated into the next version of the guidances. Feedback is continuously being collected in the guidance documents as comments. Examples of items that could be added in the next version are:

- including other formats for observational data (e.g., geotiff, geopackages);
- performing tests on the tool to transform Excel files into machine-readable files;
- re-organizing the practical guidances document to match SoilWise website organization;
- adding separate mandatory criteria to make a record discoverable through SoilWise from the good practice.

Since the first version of the guidances was published in November 2025, additional feedback collected will be implemented in the next version of the guidances.

#### 2.4.3.4 *GeoPackage*

The development of the SoilWise GeoPackage is part of IT2-22-UC-UC3-CREA: GeoPackage.

The SoilWise GeoPackage is a novel approach to soil data sharing and storage. Merging the power of a database with the simplicity of a file, this approach enables storing relational data on soil, including observational meta-information such as procedures (both field and laboratory methodologies). The GeoPackage can be pre-filled with information on observable properties, together with applicable measurement methodologies. This single file stores both data and metadata, to facilitate maintenance and sharing of data, also by means of repositories. The data model utilized in the GeoPackage is aligned with relevant international standards (INSPIRE, OGC, ISO, W3C).

The SoilWise GeoPackage supports two different modes of interaction:

- Direct interaction via QGIS: In this usage scenario, the user can directly enter both spatial and observational data via QGIS interfaces. As a first step, the object of measurement (usually a soil site, soil profile, layer, or horizon) is defined, both via location on the map as well as additional attributes. Once the spatial objects have been defined, observations or measurements can be added to these spatial objects. For each observation or measurement, the user can indicate what property this value describes, in addition to the utilized methodology and unit of measure. This data can then be visualized on maps in QGIS (Figure 8). For field surveys, a QField version is being developed.
- Extract, Transform, and Load (ETL) from databases: In cases where soil data is already stored in local databases, the GeoPackage can be a useful exchange format, allowing data holders to provide excerpts from their database. In this usage scenario, existing data can be copied into the GeoPackage utilizing ETL tools, whereby the structure of the GeoPackage assures that all relevant measurement meta-information can be conveyed.

In January 2026, the work on the GeoPackage was presented to the UCs to initiate the collection of feedback. In order to enable diverse users to test the usefulness of this solution, the current list of observable properties and their methodologies (116 properties, each including a written description and a persistent identifier (IRI) linking each property to its formal definition within the Global Soil Information System (GloSIS) ontology) must be extended to include those properties for which data is to be encoded. Stakeholders from different soil subdomains are requested to provide information on observable properties being surveyed at their sites, together with methodologies used to help create a joint vocabulary with these concepts. This information will be prefilled into the GeoPackage as required for stakeholder testing.

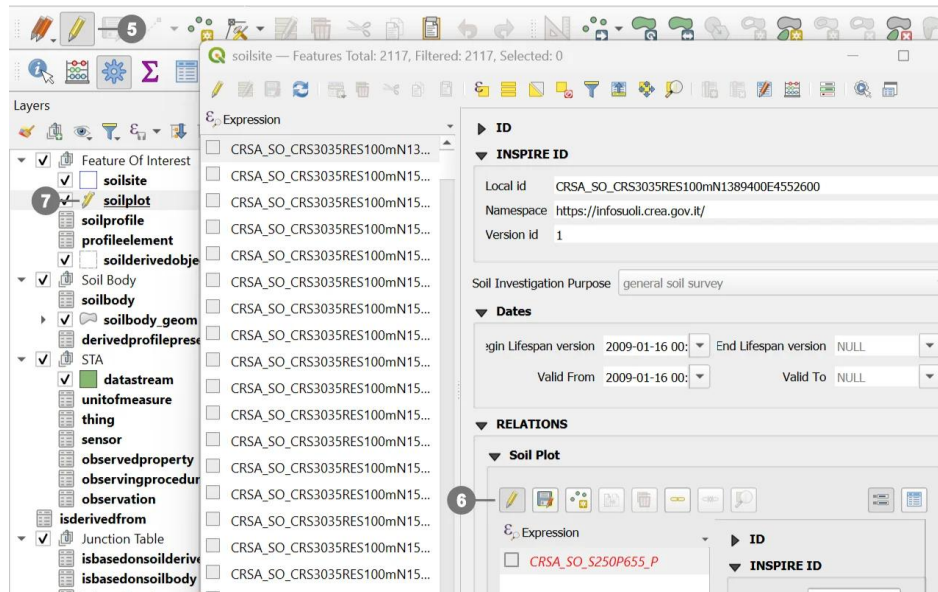


Figure 8 Example of directly entering spatial and observational data for visualization on maps via QGIS interfaces.

A follow-up demonstration to UCs was done in April 2026, where the completed work on the Geo-package will be presented and feedback collected.

#### 2.4.3.5 Tabular soil data annotation, metadata creation

A **first prototype** of the SoilWise tabular soil data annotation tool (<https://dataannotator-swr.streamlit.app/>) was demonstrated to all UCs in March 2026, after which feedback was collected through a tailored questionnaire. Relevant UC partners were asked to upload a soil dataset together with existing metadata and/or contextual documents. The tool then allowed validators to perform appropriate vocabulary matching (i.e., mapping column headers to standardised soil science terminology), where column metadata could be edited and corrected manually (data types, units, methods, etc.). The resulting annotated metadata was then collected through the questionnaire for further analysis. Figure 9 shows the landing page of the prototype as presented to validators.

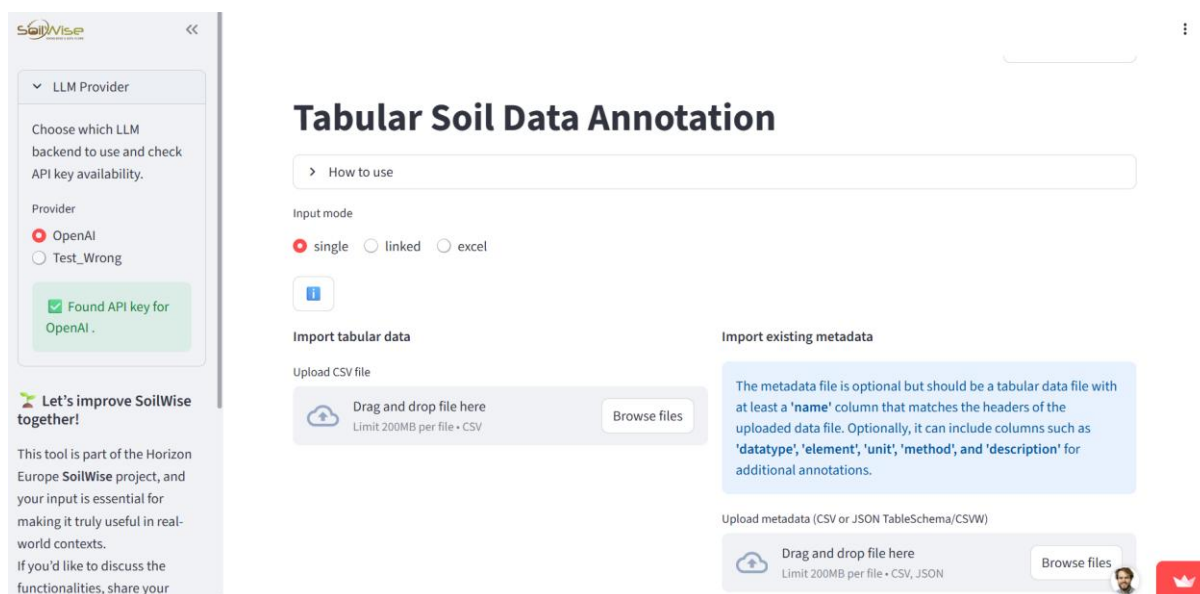


Figure 9 SoilWise Tabular Soil Data Annotation tool

#### 2.4.3.6 Harvesting Mission Soil project resources

This User Story is highly influenced by the SoilWise harvester and governance and has a direct link to the following other User Stories through which validation has been performed indirectly:

- IT2-8-UC-JRC: New filtering options
- IT2-18-UC-UC2-CIRAD2: Guidance on sustainable websites publication
- IT2-19-UC-UC2-INRAE: Guidance for FAIR soil data publication

In addition, screening forms have been prepared to identify data and knowledge that will be published by Mission Soil projects. This work is done in the framework of the Data & Knowledge Cluster of the Mission Soil Platform and in close collaboration with REA, JRC, DG AGRI, and the Mission Secretariat. Screening forms are distributed among Mission Soil projects to identify which type of data and knowledge will be produced and where they are planned to be repositied. This information will then be taken into account for the SoilWise harvester and governance components. Screening forms were developed and tested with selected Mission Soil projects. In December 2025, the first large wave of screening forms will be sent out, and a second wave is scheduled in spring 2026. The screening forms are updated regularly to reflect the information required. Starting and ending projects will be targeted specifically in the following waves to ensure all information is registered in due time and can be analyzed accordingly. An updated version of the screening form has already been sent (March 30, 2026) to 10 projects that are ending in 2026.

#### 2.4.3.7 Display MD validation results

The US *Display Metadata Validation Results* concerns backend infrastructure and internal system components that are not yet exposed at the user interface level. As such, functional validation from an end-user perspective could not be performed at this stage. The implemented work has been reviewed and confirmed at the technical level, and validation by UCs will be carried out once the relevant interface components are accessible.

## 2.5 Integration of Data Management Components

Developments conducted within WP2 (Design and Development of Data Management components) and WP3 (Design and Development of Knowledge Management components) have begun in independent environments of the involved developer teams to manage resources efficiently, i.e., to save time for the iterative development itself rather than focusing on the integration of every component version at all costs. **Integration** of the independently developed components has been performed at the end of the first integration and validation phase (see Figure 1), where applicable, through open APIs, mostly the ones standardized by the World Wide Web Consortium and Open Geospatial Consortium (OGC). As a result, the SWC and its components are characterized by "loose integration" at the end of the **first development cycle**, as defined by the ISO 19100 series of standards.

The **second development cycle** focused on "tight" integration of distributed developments and their subsequent migration to Wetransform premises. A set of components, which are not an integral part of the SoilWise Catalogue, but represent tools to support FAIR data and knowledge publication, is foreseen to remain at the "loose/mixed integration" levels, as defined by the ISO 19100 series of standards. More details can be found in D4.3 Repository infrastructure, components and APIs, v3.

The **third development cycle** will focus on the migration to EUSO. Since SoilWise developments follow a modular and scalable architecture design, partial migration of the system will be supported.

### 3 Data Management components

This chapter lists the Data Management components that are subject to the delivery of the second SoilWise prototype, including the links to their source code (GitHub), the latest release (Zenodo), and an access point for users (if applicable). Knowledge Management components, which represent another part of the delivered prototype, are listed in Chapter 3 of the parallel deliverable D3.3, Developed & Integrated KM components. Since data and knowledge in the SoilWise Catalogue are closely linked, in some cases, the components cannot be clearly assigned to only one Work Package, as the developments have been shared by both WP2 and WP3. Therefore, in the list of Data Management component functions, we indicate that they belong to Knowledge Management components and vice versa.

A more detailed description of the components' functionality is described in the Technical documentation, see more details in Chapter 4. A summary in the form of release notes is also available in Deliverable D4.3 - Repository infrastructure, components, and APIs. These components are also visible in the high-level architecture schema depicted in Figure 2.

<b>Name:</b>	<b>Harvester</b>
<b>Version:</b>	0.3.0
<b>Functions:</b>	Harvest metadata resources Harvest knowledge resources ( <i>part of KM</i> ) Metadata RDF Turtle serialization RDF to Triple Store ( <i>part of KM</i> ) Harvest Soil Mission news feed ( <i>part of KM</i> )
<b>Documentation:</b>	<a href="https://prototype-3-0.soilwise-documentation.pages.dev/technical_components/ingestion/">https://prototype-3-0.soilwise-documentation.pages.dev/technical_components/ingestion/</a>
<b>Source code:</b>	<a href="https://github.com/soilwise-he/harvesters/releases/tag/v0.3.0">https://github.com/soilwise-he/harvesters/releases/tag/v0.3.0</a>
<b>Release:</b>	<a href="https://doi.org/10.5281/zenodo.14923562">https://doi.org/10.5281/zenodo.14923562</a>
<b>Access point:</b>	not applicable
<b>Name:</b>	<b>Metadata Harmonisation</b>
<b>Version:</b>	0.1.0
<b>Functions:</b>	ETL Metadata harmonization Duplication identification ( <i>part of KM</i> )
<b>Documentation:</b>	<a href="https://prototype-3-0.soilwise-documentation.pages.dev/technical_components/metadata_harmonisation/">https://prototype-3-0.soilwise-documentation.pages.dev/technical_components/metadata_harmonisation/</a>
<b>Source code:</b>	<a href="https://github.com/soilwise-he/md-harmonization/releases/tag/0.1.0">https://github.com/soilwise-he/md-harmonization/releases/tag/0.1.0</a>
<b>Release:</b>	<a href="https://zenodo.org/records/19692782">https://zenodo.org/records/19692782</a>
<b>Access point:</b>	not applicable

<b>Name:</b>	<b>Repository Storage</b>
<b>Version:</b>	PostgreSQL release 16.11 Virtuoso release 07.20.3239 Apache Lucene 9.11.1
<b>Functions:</b>	Storage of raw harvested metadata – PostgreSQL / vector DB Storage of harmonised augmented metadata – PostgreSQL / vector DB Storage of augmented, linked metadata, knowledge graph - Virtuoso ( <i>part of KM</i> ) Storage of indexed metadata and knowledge content – Apache Lucene Index ( <i>part of KM</i> )
<b>Documentation:</b>	<a href="https://prototype-3-0.soilwise-documentation.pages.dev/technical_components/storage/">https://prototype-3-0.soilwise-documentation.pages.dev/technical_components/storage/</a>
<b>Source code:</b>	not applicable – available as independent Open Source components
<b>Access point:</b>	PostgreSQL: not applicable Triple Store (Virtuoso): <a href="https://sparql.soilwise.wetransform.eu/sparql">https://sparql.soilwise.wetransform.eu/sparql</a> Apache Lucene Index: not applicable
<b>Name:</b>	<b>SoilWise Finder (Metadata Catalogue)</b>
<b>Version:</b>	2.0.0
<b>Functions:</b>	Metadata & knowledge NLP transformations and indexation (part of KM) Search API on Solr Solr configuration (part of KM) Catalogue User Interface, offering the following functions <ul style="list-style-type: none"> <li>• Fulltext search</li> <li>• Thematic filtering (licence, language, project, source, keywords, Soil Mission programme)</li> <li>• Spatial filtering</li> <li>• Temporal filtering</li> <li>• Ranking / sorting</li> <li>• Spatial extent preview</li> <li>• Display record details</li> <li>• Data download (AS IS)</li> <li>• Download search results</li> </ul>
<b>Documentation:</b>	<a href="https://prototype-3-0.soilwise-documentation.pages.dev/technical_components/catalogue/#search-engine-index-and-search-strategies">https://prototype-3-0.soilwise-documentation.pages.dev/technical_components/catalogue/#search-engine-index-and-search-strategies</a> <a href="https://prototype-3-0.soilwise-documentation.pages.dev/technical_components/catalogue/#user-interface">https://prototype-3-0.soilwise-documentation.pages.dev/technical_components/catalogue/#user-interface</a>
<b>Source code:</b>	<a href="https://github.com/soilwise-he/SoilWiseFinder/releases/tag/v2.0.0">https://github.com/soilwise-he/SoilWiseFinder/releases/tag/v2.0.0</a>
<b>Release:</b>	<a href="https://doi.org/10.5281/zenodo.19691525">https://doi.org/10.5281/zenodo.19691525</a>
<b>Access point:</b>	<a href="https://catalogue.soilwise-he.eu/search">https://catalogue.soilwise-he.eu/search</a>

<b>Name:</b>	<b>Apache Solr Search Engine</b>
<b>Version:</b>	9.7.0
<b>Functions:</b>	Search platform
<b>Documentation:</b>	<a href="https://prototype-3-0.soilwise-documentation.pages.dev/technical_components/catalogue/#search-engine-index-and-search-strategies">https://prototype-3-0.soilwise-documentation.pages.dev/technical_components/catalogue/#search-engine-index-and-search-strategies</a>
<b>Source code:</b>	Not applicable – available as independent Open Source component
<b>Access point:</b>	<a href="https://solr.soilwise.wetransform.eu/">https://solr.soilwise.wetransform.eu/</a> login needed
<b>Name:</b>	<b>Metadata Validation</b>
<b>Version:</b>	0.4.1
<b>Functions:</b>	ETS INSPIRE validation Metadata completeness check
<b>Documentation:</b>	<a href="https://prototype-3-0.soilwise-documentation.pages.dev/technical_components/metadata_validation/">https://prototype-3-0.soilwise-documentation.pages.dev/technical_components/metadata_validation/</a>
<b>Source code:</b>	ETS INSPIRE validation: <a href="https://github.com/soilwise-he/metadata-validator/releases/tag/v0.4.1">https://github.com/soilwise-he/metadata-validator/releases/tag/v0.4.1</a> Metadata completeness check: <a href="https://github.com/soilwise-he/metadata-validator/releases/tag/v0.4.1">https://github.com/soilwise-he/metadata-validator/releases/tag/v0.4.1</a>
<b>Release:</b>	Metadata validator: <a href="https://doi.org/10.5281/zenodo.14924543">https://doi.org/10.5281/zenodo.14924543</a>
<b>Access point:</b>	<a href="https://superset.soilwise.wetransform.eu/login/?next=https://superset.soilwise.wetransform.eu/superset/dashboard/p/P52OgRVBGo9/">https://superset.soilwise.wetransform.eu/login/?next=https://superset.soilwise.wetransform.eu/superset/dashboard/p/P52OgRVBGo9/</a> login needed
<b>Name:</b>	<b>Metadata Augmentation</b>
<b>Version:</b>	0.3.0
<b>Functions:</b>	Translation module ( <i>part of KM</i> ) Link liveliness assessment Keyword matcher ( <i>part of KM</i> ) Element matcher ( <i>part of KM</i> ) Spatial metadata extractor
<b>Documentation:</b>	<a href="https://prototype-3-0.soilwise-documentation.pages.dev/technical_components/metadata_augmentation/">https://prototype-3-0.soilwise-documentation.pages.dev/technical_components/metadata_augmentation/</a>
<b>Source code:</b>	<a href="https://github.com/soilwise-he/metadata-augmentation/releases/tag/v0.3.0">https://github.com/soilwise-he/metadata-augmentation/releases/tag/v0.3.0</a> Link liveliness assessment: <a href="https://github.com/soilwise-he/link-liveliness-assessment/releases/tag/1.1.8">https://github.com/soilwise-he/link-liveliness-assessment/releases/tag/1.1.8</a>
<b>Release:</b>	<a href="https://doi.org/10.5281/zenodo.14924181">https://doi.org/10.5281/zenodo.14924181</a> Link liveliness assessment: <a href="https://doi.org/10.5281/zenodo.14923789">https://doi.org/10.5281/zenodo.14923789</a>
<b>Access point:</b>	not applicable

<b>Name:</b>	<b>DOI Resolution Widget</b>
<b>Version:</b>	1.2.0
<b>Functions:</b>	Validate DOI Suggest DOI for harvesting
<b>Documentation:</b>	<a href="https://prototype-3-0.soilwise-documentation.pages.dev/technical_components/data_publication_support/#doi-resolution-widget">https://prototype-3-0.soilwise-documentation.pages.dev/technical_components/data_publication_support/#doi-resolution-widget</a>
<b>Source code:</b>	<a href="https://github.com/soilwise-he/soilwise-util/releases/tag/v1.2.0">https://github.com/soilwise-he/soilwise-util/releases/tag/v1.2.0</a>
<b>Release:</b>	<a href="https://doi.org/10.5281/zenodo.19708262">https://doi.org/10.5281/zenodo.19708262</a>
<b>Access point:</b>	<a href="https://soilwise-he.eu/repository/">https://soilwise-he.eu/repository/</a>
<b>Name:</b>	<b>Tabular soil data annotation</b>
<b>Version:</b>	1.0.1
<b>Functions:</b>	(Meta)data upload NLP-based annotation suggestions Annotated (meta)data download
<b>Documentation:</b>	<a href="https://prototype-3-0.soilwise-documentation.pages.dev/technical_components/data_publication_support/#tabular-soil-data-annotation">https://prototype-3-0.soilwise-documentation.pages.dev/technical_components/data_publication_support/#tabular-soil-data-annotation</a>
<b>Source code:</b>	<a href="https://github.com/soilwise-he/TabularDataAnnotator-streamlit/releases/tag/v1.0.1">https://github.com/soilwise-he/TabularDataAnnotator-streamlit/releases/tag/v1.0.1</a>
<b>Release:</b>	<a href="https://doi.org/10.5281/zenodo.19569572">https://doi.org/10.5281/zenodo.19569572</a>
<b>Access point:</b>	<a href="https://dataannotator-swr.streamlit.app/">https://dataannotator-swr.streamlit.app/</a>
<b>Name:</b>	<b>Transformation and Harmonization - Geopackage</b>
<b>Version:</b>	1.0.0
<b>Functions:</b>	Unified Data Storage Standardized Conceptual Mapping (INSPIRE model) Semantic Harmonisation & Interoperability Time-Series and Sensor Integration (STAC2) Native GIS support, Guided Data Entry via Custom QGIS Forms
<b>Documentation:</b>	<a href="https://prototype-3-0.soilwise-documentation.pages.dev/technical_components/data_publication_support/#soilwise-geopackage">https://prototype-3-0.soilwise-documentation.pages.dev/technical_components/data_publication_support/#soilwise-geopackage</a>
<b>Source code:</b>	<a href="https://github.com/soilwise-he/Geopackage-so/releases/tag/1.0.0">https://github.com/soilwise-he/Geopackage-so/releases/tag/1.0.0</a>
<b>Release:</b>	<a href="https://doi.org/10.5281/zenodo.19465866">https://doi.org/10.5281/zenodo.19465866</a>
<b>Access point:</b>	not applicable

<b>Name:</b>	<b>Authentication &amp; Authorisation</b>
<b>Version:</b>	Keycloak 7.1.9
<b>Functions:</b>	Authentication (Login) Authorisation
<b>Documentation:</b>	<a href="https://prototype-3-0.soilwise-documentation.pages.dev/technical_components/user_management/">https://prototype-3-0.soilwise-documentation.pages.dev/technical_components/user_management/</a>
<b>Source code:</b>	not applicable
<b>Release:</b>	not applicable
<b>Access point:</b>	<a href="https://id.soilwise.wetransform.eu/">https://id.soilwise.wetransform.eu/</a> login needed
<b>Name:</b>	<b>System Usage &amp; Monitoring</b>
<b>Version:</b>	Grafana 12.1.1 / Prometheus v2.55.1 / Prometheus LTS v2.18.1 / nginx 1.0.2
<b>Functions:</b>	System Health & Status Monitoring Usage statistics
<b>Documentation:</b>	<a href="https://prototype-3-0.soilwise-documentation.pages.dev/technical_components/monitoring/">https://prototype-3-0.soilwise-documentation.pages.dev/technical_components/monitoring/</a>
<b>Source code:</b>	<a href="https://github.com/grafana/grafana/releases/tag/v12.1.1">https://github.com/grafana/grafana/releases/tag/v12.1.1</a> <a href="https://github.com/prometheus/prometheus/releases/tag/v2.18.1">https://github.com/prometheus/prometheus/releases/tag/v2.18.1</a> <a href="https://github.com/prometheus/prometheus/releases/tag/v2.55.1">https://github.com/prometheus/prometheus/releases/tag/v2.55.1</a> <a href="http://nginx.org/download/nginx-1.0.2.tar.gz">http://nginx.org/download/nginx-1.0.2.tar.gz</a>
<b>Access point:</b>	<a href="https://grafana.soilwise.wetransform.eu/">https://grafana.soilwise.wetransform.eu/</a> login needed
<b>Name:</b>	<b>pycsw</b>
<b>Version:</b>	3.0.0.beta1
<b>Functions:</b>	CSW API OGC API-Records STAC API OAI-PMH API OpenSearch
<b>Documentation:</b>	<a href="https://prototype-3-0.soilwise-documentation.pages.dev/technical_components/catalogue/#integrations-interfaces">https://prototype-3-0.soilwise-documentation.pages.dev/technical_components/catalogue/#integrations-interfaces</a>
<b>Source code:</b>	<a href="https://github.com/geopython/pycsw/releases/tag/3.0.0-beta1">https://github.com/geopython/pycsw/releases/tag/3.0.0-beta1</a>
<b>Access point:</b>	<a href="https://api.soilwise.wetransform.eu/records/openapi">https://api.soilwise.wetransform.eu/records/openapi</a>

## 4 Technical documentation

Technical documentation was built in the form of an external web page accompanying this deliverable to capture the complexity of the SoilWise achievements. This documentation consists of a description of functionalities, detailed technical specifications, and interfaces between all modular components that will be employed in the SWC. This fits with one of the work package (WP) 2 goals to detail the design and develop the DM components of SWC, more specifically, within the task (T) 2.1 Design of the data technology components (M7-M37) and T2.2 Implementation and deployment of data components (M7-M46).

For this deliverable, a stable release of the SWC technical documentation is published at <https://prototype-3-0.soilwise-documentation.pages.dev/>. A PDF exported version is also available on demand as the non-editable version made at the date of deliverable submission. The same documentation is maintained as a living environment in the public GitHub repository at <https://github.com/soilwise-he/SoilWise-documentation>, where it is continuously updated during the development process. It comprises a description of functionality, detailed technical specifications, and interfaces between all modular components, and is structured according to the main technical components.

# Annex I – SoilWise-he Road Map

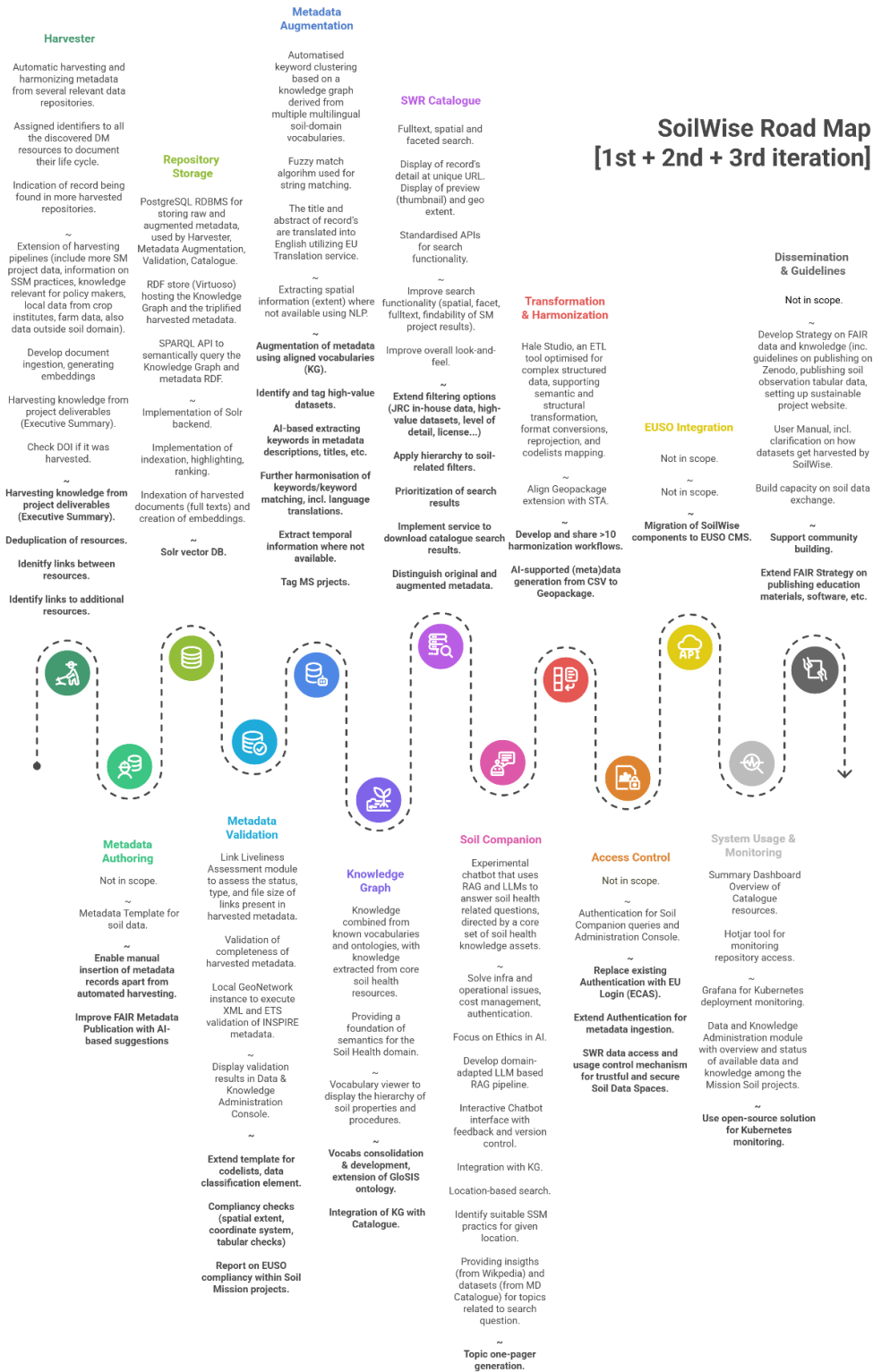


Figure 10 SoilWise Road Map. Normal text describes functionality developed in the 1<sup>st</sup> iteration, bold text describes functionality foreseen to be developed in the 2<sup>nd</sup> iteration

## Annex II - Fulfilment of the system requirements (Data)

The SoilWise core product and its functionalities delivered in the second iteration were driven by the **User Stories** (US) developed during the co-design phase (see Deliverable 1.2). Each US describes the need for a functionality based on a problem statement, which forms the basis for defining the US-specific acceptance criteria. During the 2<sup>nd</sup> iteration, USs were created for requirements defined by (i) the SoilWise Grant Agreement (i.e. business requirements), (ii) User Cases, and (iii) JRC. A total of 24 USs has been identified during the second iteration, of which 9 will be developed during the third iteration. The following tables report on the acceptance criteria derived from the USs for **Data** and indicate their current status.

### User Stories from Grant Agreement requirements

IT2-2-UC-GA: Display MD Validation Results	
<b>Problem statement</b>	As a data user, I want to see the results of ETS validation for selected INSPIRE metadata record, so I can see whether it is INSPIRE compliant or what the validation issues are.
<b>GitHub link</b>	<a href="https://github.com/soilwise-he/Soilwise-Project-Backlog/issues/16">https://github.com/soilwise-he/Soilwise-Project-Backlog/issues/16</a>
<b>Acceptance criteria</b>	<ul style="list-style-type: none"> <li>× I can open the validation log from every INSPIRE metadata record's detail.</li> <li>× I can immediately see if the given metadata record is INSPIRE compliant.</li> <li>× Validation is performed on the latest metadata version.</li> <li>× Validation is returning the same results as INSPIRE Reference Validator: <a href="https://inspire.ec.europa.eu/validator/home/index.html">https://inspire.ec.europa.eu/validator/home/index.html</a></li> </ul>

### User Stories from User Case requirements

IT2-19-UC-UC2-INRAE: Guidance for FAIR soil data publication	
<b>Problem statement</b>	As a researcher I want to receive proper guidance on how to publish a dataset on Zenodo, so I that the data can be easily reused
<b>GitHub link</b>	<a href="https://github.com/soilwise-he/Soilwise-Project-Backlog/issues/33">https://github.com/soilwise-he/Soilwise-Project-Backlog/issues/33</a>
<b>Acceptance criteria</b>	<ul style="list-style-type: none"> <li>× Guidance explains how to prepare data and related information.</li> <li>× Provide guidance on the type of data and metadata that is required for FAIRness of data.</li> <li>× Guidance on vocab for procedures and reported data is required.</li> <li>× Improve accessibility of vocab, definitions methods.</li> </ul>

IT2-22-UC-UC3-CREA: Geopackage	
<b>Problem statement</b>	As a data aggregation and developer organization I want to adopt a standard template for the digital exchange of soil data between data producers (researchers, surveyors and laboratories) and/or Public Authorities. My organization last aim is to advise policymakers about Europe's environment. The requirements are 1) Portable database (allowing digital exchange); 2) Open-source RDBMS; 3) INSPIRE compliance; 4) Extensible for every soil subdomain subject (Hydrological data, soil fauna, soil microbiology, biochemical, mineralogical). So, I searched online for structured soil database and could find the INSPIRE SO GeoPackage which seems to align to all the requirements but also to be quite difficult to populate. Difficulty is due both on the number of metadata required to be populated for describing data and the adoption of vocabularies to align metadata.

<b>GitHub link</b>	<a href="https://github.com/soilwise-he/Soilwise-Project-Backlog/issues/36">https://github.com/soilwise-he/Soilwise-Project-Backlog/issues/36</a>
<b>Acceptance criteria</b>	<ul style="list-style-type: none"> <li>✓ Given the new requirement to simplify population of the database maintaining of all the other requirements, when the metadata documentation is already compiled or reduced in terms of number of metadata, then populating the database will result easier.</li> <li>✓ Moreover, a reason of complexity is related to the observational data structure implemented (O&amp;M), updating data structure to the new standard OGC SensorThings API, will result on a reduction of complexity and an easier population of the database.</li> <li>× The development of Knowledge Graph, Metadata Authoring, and Metadata Validation component will result on an easier management of metadata.</li> <li>× The Transformation and Harmonization component will contribute to the easy population of the database.</li> </ul>

**IT2-14-UC-UC4-ZALF: SWC search performance improvement, download**

<b>Problem statement</b>	<p>As a geographical information scientist demanding soil and environmental data in spatial modelling studies, someone working at a research institution, I want user-friendly data discovery and download tools. An interface providing an overview of what SWC can provide.</p> <ul style="list-style-type: none"> <li>• User-friendly interface with a clear overview of available datasets, publications, and experiments. — Smart search &amp; filters to quickly find relevant knowledge.</li> <li>• Customizable downloads in interoperable formats, with options to select spatial extent, temporal range, variables of interest, coordinate reference system, and preferred file format (e.g., CSV, GeoTIFF, JSON).</li> <li>• Seamless access to knowledge that supports both research and policy use.</li> </ul> <p>So, I can get data easily, efficiently use time, improve research ideas with insights, and validate a climate model focused on agricultural impacts.</p>
<b>GitHub link</b>	<a href="https://github.com/soilwise-he/Soilwise-Project-Backlog/issues/28">https://github.com/soilwise-he/Soilwise-Project-Backlog/issues/28</a>
<b>Acceptance criteria</b>	<ul style="list-style-type: none"> <li>✓ When I refine my search by selecting a specific geographic area or a time range, then the results update to show only data from that chosen location and period</li> <li>✓ When I use a very specific keyword or phrase in the search bar, i.e. when I search for a term like “soil organic carbon”, then the system should return results that exactly match this term. It should not include results for broader or loosely related terms (for example, “soil organic matter”)</li> <li>× I can easily understand the types of resources.</li> <li>× Pdfs are not returned if I search for datasets.</li> <li>✓ I can use multiple keywords to search at the same time: For example, carbon sequestration, peatland, filter only contents with the mentioned keywords, that means C sequestration only in peatland not in grasslands or forests</li> <li>✓ There is guidance on search available in the UI.</li> <li>× Embargoed, or not available to download, datasets are lower in search results.</li> <li>✓ When needed, the search will perform an exact match with keywords available in metadata.</li> <li>× Using plural vs singular should not affect the number of returned search results (i.e. long term field experiment/s).</li> <li>✓ A combination of full-text and facet filters is possible. Facet filters will narrow down the full-text search results.</li> </ul>

	<ul style="list-style-type: none"> <li>✓ A combination of multiple facet filters is possible to make the search more specific. (AND condition)</li> <li>✓ An option to download the search results returned after customized search in .csv format</li> </ul>
--	---

#### IT2-23-UC-UC4-VLAANDERS: Tabular soil data annotation, metadata creation

<b>Problem statement</b>	As a researcher in Living Labs with minimal knowledge of FAIR principles I want to publish raw observation data in an interoperable way with minimal effort so my results can efficiently be reused.
<b>GitHub link</b>	<a href="https://github.com/soilwise-he/Soilwise-Project-Backlog/issues/37">https://github.com/soilwise-he/Soilwise-Project-Backlog/issues/37</a>
<b>Acceptance criteria</b>	<ul style="list-style-type: none"> <li>✓ Without an extensive knowledge of metadata profiles, I get based on the information provided an acceptable proposal for a metadata file that I can use for a better publication</li> <li>✓ A researcher should be able to import my data in an interoperable way, ingest the data in a triple store or similar and query it efficiently</li> <li>× The perception among a large portion (70%) of the test group (&gt;5 participants) is that added value (time savings, better effort–benefit balance) is being created within their own process. This added value may lie in improved documentation of the dataset and enhanced interoperability.</li> <li>× Possible input files; csv, geotiff, (excel), geopackage, sqlLite</li> </ul>

#### IT2-15-UC-UC5-GAIA: SWC UI improvement

<b>Problem statement</b>	As a representative organization for Greek farmers and agricultural cooperatives, I want to improve and modernize the layout/design/template of the SoilWise hub, so it is more inviting and user-friendly for our stakeholders, and to attract a wider group of users and with a little patience to navigate a system.
<b>GitHub link</b>	<a href="https://github.com/soilwise-he/Soilwise-Project-Backlog/issues/29">https://github.com/soilwise-he/Soilwise-Project-Backlog/issues/29</a>
<b>Acceptance criteria</b>	<ul style="list-style-type: none"> <li>✓ The spatial filter is not overwhelmed by displaying BBOXes of the current search results.</li> <li>× avoid duplicates including versioning of datasets</li> </ul>

#### IT2-16-UC-UC5-BIOS: SWC search performance improvement

<b>Problem statement</b>	As researchers working on yield predictive models based on soil, weather and satellite data we want to search for alternative data, primarily related to soil, so we can improve our models.
<b>GitHub link</b>	<a href="https://github.com/soilwise-he/Soilwise-Project-Backlog/issues/30">https://github.com/soilwise-he/Soilwise-Project-Backlog/issues/30</a>
<b>Acceptance criteria</b>	<ul style="list-style-type: none"> <li>✓ We can open spatial filter on a bigger portion of the screen, or in a new tab.</li> <li>✓ We can spatially filter data according to a target country, or region.</li> <li>✓ We can sort the dates in an ascending and descending way.</li> <li>× Spatial search is returning relevant results also for records that do not have spatial information included in the metadata.</li> </ul>

## User Stories from JRC requirements

#### IT2-7-UC-JRC: Metadata template

<b>Problem statement</b>	Data provider of Mission Soil Project wanting access to a standardized metadata structure that is able to handle soil-specific metadata elements
<b>GitHub link</b>	<a href="https://github.com/soilwise-he/Soilwise-Project-Backlog/issues/21">https://github.com/soilwise-he/Soilwise-Project-Backlog/issues/21</a>

<b>Acceptance criteria</b>	<ul style="list-style-type: none"> <li>✓ Metadata Template reviewed and accepted by JRC and selected MS projects</li> <li>× The user can easily find his/her resources with metadata filled in the schema provided by the template</li> </ul>
----------------------------	---

IT2-8-UC-JRC: New filtering options	
<b>Problem statement</b>	As a JRC data analyst I want to easily find Mission Soil project results so I can monitor the outcomes of this program.
<b>GitHub link</b>	<a href="https://github.com/soilwise-he/Soilwise-Project-Backlog/issues/22">https://github.com/soilwise-he/Soilwise-Project-Backlog/issues/22</a>
<b>Acceptance criteria</b>	<ul style="list-style-type: none"> <li>✓ I can set up a BBOX to filter data by relevant spatial extent. When I confirm the BBOX, I only see records which spatial extent matches the one I provided.</li> <li>✓ I can perform a thematic query. When I type in the desired keywords, I only see records that match them.</li> <li>× I can filter out only Mission Soil Projects contributions within SoilWise Catalogue record's.</li> <li>× I can filter out only high-value datasets within SoilWise Catalogue record's.</li> </ul>

IT2-5-UC-JRC: Harvesting Mission Soil Project Resources	
<b>Problem statement</b>	Improve harvesting of existing SM project data & knowledge. As a JRC stakeholder, I want to be able to find data and knowledge productions from Mission Soil projects so that the Mission Soil productions become sustainable and FAIR.
<b>GitHub link</b>	<a href="https://github.com/soilwise-he/Soilwise-Project-Backlog/issues/19">https://github.com/soilwise-he/Soilwise-Project-Backlog/issues/19</a>
<b>Acceptance criteria</b>	<ul style="list-style-type: none"> <li>× D&amp;K Mission Soil productions that SoilWise is aware of should be found on SoilWise Catalogue.</li> </ul>

## User Stories identified for next iteration

IT2-3-UC-GA: HarmonizationWorkflows	
<b>Problem statement</b>	As a soil data scientist, I want to use the preset data harmonization workflow and mapping functions provided by SWC so that I can automatically convert my unstructured soil measurement data into the INSPIRE Soil format, eliminating the need for manual mapping.
<b>GitHub link</b>	<a href="https://github.com/soilwise-he/Soilwise-Project-Backlog/issues/17">https://github.com/soilwise-he/Soilwise-Project-Backlog/issues/17</a>
<b>Acceptance criteria</b>	<ul style="list-style-type: none"> <li>× The SWC has a drop-down function with the item "Transformation preset," which takes me to a list of presets.</li> <li>× 10+ workflow and mapping presets created and displayed in the list, covering state-of-the-art target formats in the soil domain.</li> <li>× I can select a workflow/mapping preset from a list that appears when I click on the corresponding drop-down menu.</li> <li>× I can upload my dataset.</li> <li>× I can perform the transformation based on the selected preset</li> <li>× I can download the transformed database</li> </ul>

IT2-1-UC-GA: Deduplication	
<b>Problem statement</b>	As a catalogue manager, I want to automatically detect potential duplicate products in my catalogue so I can maintain data quality and prevent customer confusion from seeing the same item listed multiple times cases:

	<ul style="list-style-type: none"> <li>• All combinations of DOI - content</li> <li>• Pairwise comparison DOI: same DOI – different DOI – no DOI (?)</li> <li>• Content = title, description and keywords (authors?) -&gt; same content, different content, partial content. Possible causes different content: missing title-description-keywords, partial content, use of subtitles, different languages, preprints vs peer-reviewed, different versions, corrected/retracted versions, conference paper vs proceedings, dataset vs article, supplementary materials vs main article, individual article vs special issue.</li> </ul>
<b>GitHub link</b>	<a href="https://github.com/soilwise-he/Soilwise-Project-Backlog/issues/15">https://github.com/soilwise-he/Soilwise-Project-Backlog/issues/15</a>
<b>Acceptance criteria</b>	<ul style="list-style-type: none"> <li>✓ Display suspected duplicates in a review interface showing side-by-side comparison</li> <li>✓ Show a confidence score or percentage for each potential duplicate match</li> <li>✓ All records in the catalogue should be mutually compared and assessed for duplication</li> <li>× The system must maintain an audit trail of all duplicate resolution actions</li> <li>× The system must generate reports showing total duplicates found, resolved, (and pending review)</li> </ul>

#### IT2-24-UC-UC1-WU-ELO-ILVO: Advice based on in-situ soil measurements and lab reports

<b>Problem statement</b>	As an advisory service, landowner, farmer and land manager I want to receive advice based on my soil data (analyses of my soil samples) in order to develop the set of inputs and practices I need to improve the soil health* of my farm/estate/land. So I can improve the resilience of my farm in order to increase profitability, income and sustainability. Therefore, monitoring my spendings and see my eligibility for policy measures.
<b>GitHub link</b>	<a href="https://github.com/soilwise-he/Soilwise-Project-Backlog/issues/40">https://github.com/soilwise-he/Soilwise-Project-Backlog/issues/40</a>
<b>Acceptance criteria</b>	<ul style="list-style-type: none"> <li>• tbc</li> </ul>

#### IT2-25-UC-UC5-ELO: Land suitability maps

<b>Problem statement</b>	As a company active in the processing agricultural raw materials (products in first transformation such as sugar beets, potatoes, vegetables, cereals), I want to view the land suitability maps for present and future for certain crop and/or elements for certain region taking changes in climate into account, for example: wines in Switzerland, oranges and almonds in Sicilia (example of Lodovica), .. so I can sustain and continue my business, making smart investments decisions for example on what crops to choose and where to locate my processing capacity
<b>GitHub link</b>	<a href="https://github.com/soilwise-he/Soilwise-Project-Backlog/issues/41">https://github.com/soilwise-he/Soilwise-Project-Backlog/issues/41</a>
<b>Acceptance criteria</b>	<ul style="list-style-type: none"> <li>• tbc</li> </ul>

#### IT2-4-UC-JRC: EUSO integration

<b>Problem statement</b>	EUSO Integration, Monitoring & Access Control
<b>GitHub link</b>	<a href="https://github.com/soilwise-he/Soilwise-Project-Backlog/issues/18">https://github.com/soilwise-he/Soilwise-Project-Backlog/issues/18</a>
<b>Acceptance criteria</b>	<ul style="list-style-type: none"> <li>• tbc</li> </ul>

IT2-6-UC-JRC: Manual Insertion	
<b>Problem statement</b>	As a data provider, I want to fill metadata directly in the SoilWise Catalogue so my dataset is findable through SoilWise Catalogue
<b>GitHub link</b>	<a href="https://github.com/soilwise-he/Soilwise-Project-Backlog/issues/20">https://github.com/soilwise-he/Soilwise-Project-Backlog/issues/20</a>
<b>Acceptance criteria</b>	<ul style="list-style-type: none"> <li>× Prerequisite: Structured Metadata Template</li> <li>× The user can fill manually metadata via module accessible from the SWC</li> <li>× The Metadata editor supports all Metadata Template elements, including all relations and links to codelists, vocabularies etc.</li> <li>× The user can automatically see invalid inputs in the editor</li> <li>× Only the authorized user can submit created MD record to the catalogue</li> <li>× Metadata created via Metadata editor are findable in SWC</li> </ul>

IT2-10-UC-JRC: ECAS authentication	
<b>Problem statement</b>	As SWC Board-appointed Administrator, I want my EU login (formerly ECAS) to be authorized by the SWC super administrator upon my request so that I can access the SWC admin section for non-public functionalities. So that I can access non-public functions, review, edit, and save relevant components, functions, etc., e.g., editing metadata templates to comply with a changed target schema. This is important for proper operations/maintenance of SWC.
<b>GitHub link</b>	<a href="https://github.com/soilwise-he/Soilwise-Project-Backlog/issues/24">https://github.com/soilwise-he/Soilwise-Project-Backlog/issues/24</a>
<b>Acceptance criteria</b>	<ul style="list-style-type: none"> <li>× I can register/sign-up with SWC using my EU login (formerly ECAS) and request authorization and assignment of my role (by the SWC super administrator).</li> <li>× After my registration has been approved, I can log in with my EU login and access non-public SWC functionalities, such as metadata templates, via the admin user interface.</li> <li>× I can edit and save components and functionalities and submit them for review and approval.</li> </ul>

IT2-12-UC-JRC: Compliance checks	
<b>Problem statement</b>	As a data governance officer from the EUSO, I want to assess all datasets coming from the Mission Soil projects so I can quickly respond to the relevant projects that datasets do not meet the standards.
<b>GitHub link</b>	<a href="https://github.com/soilwise-he/Soilwise-Project-Backlog/issues/26">https://github.com/soilwise-he/Soilwise-Project-Backlog/issues/26</a>
<b>Acceptance criteria</b>	<ul style="list-style-type: none"> <li>× Verify mandatory elements: Identification, Title, Abstract, File Format, Spatial Extent, Licensing, Contact, Reference system, Citation, File / service Location. --&gt; to be updated with the definition of the metadata template.</li> <li>× Use controlled vocabularies for keywords and thematic categories.</li> <li>× Data formats must belong to a geospatial format (e.g., GPKG, NetCDF, GeoJSON, GML, GeoTIFF).</li> <li>× Coordinate Reference System must be valid, coming from a control list (e.g., QGIS list of CRS).</li> </ul>

IT2-13-UC-JRC: EUSO high-value datasets	
<b>Problem statement</b>	EUSO high-value data sets will be identified, labelled and discoverable upon a filter in the SWC. However, SWC will not store any original EUSO high-value data.
<b>GitHub link</b>	<a href="https://github.com/soilwise-he/Soilwise-Project-Backlog/issues/27">https://github.com/soilwise-he/Soilwise-Project-Backlog/issues/27</a>

---

<b>Acceptance criteria</b>	x The user can filter the catalogue on High Value Datasets as defined by the requirements
----------------------------	---

## 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.