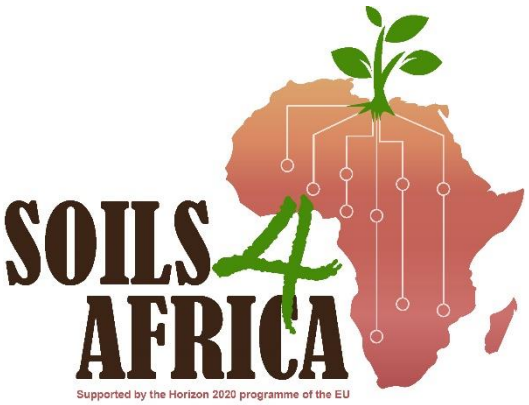



# Soil Information System Manuals

**Project No. 869200**

31 May 2025

**Deliverable: Soils4Africa D6.4**

Version 1.0

## 1. Introduction

The Soils4Africa project aims to develop a soil information system (SIS) that serves information on the prevalence and spatial distribution of soil quality indicators and constraints relevant to the sustainable intensification of agriculture in Africa. This information can serve as a baseline for monitoring changes in soil conditions across the continent's agricultural land in the future.

Soil samples were collected from approximately 15,000 locations across a wide range of African countries, at two soil depths. Uniform procedures were applied during both fieldwork and laboratory analysis to ensure consistency of the resulting soil dataset.

To support the use and hosting of the SIS, a set of manuals was developed as part of project deliverable D6.4. Two of these are included in this document:

- Part A: SIS User Manual
- Part B: SIS Installation Manual

Both manuals will also be made available through the SIS platform.

In addition to these, three complementary online resources can be accessed via the embedded hyperlinks embedded listed below:

- Digital Soil Mapping Tutorial
- Spectral Modelling Notebook
- Documentation of the Soils4Africa data model

The Digital Soil Mapping Tutorial provides a hands-on, reproducible digital soil mapping workflow implemented in the statistical software R that shows how to produce digital soil maps from the Soils4Africa soil data together with a set of covariate layers.

The Spectral Modelling Jupyter notebook offers guidance on modelling and predicting of soil properties from spectral measurements using Soils4Africa data. It covers:

i) visualization of spectral data,

ii) preprocessing techniques,

iii) training machine learning models for soil property prediction, and

iv) interpretability of model results.

The notebook introduces key concepts at three levels - beginner, intermediate and expert - making it available to a wide range of users. Instructions on how to run the notebook can be found here.

The *Soils4Africa Data Model Documentation* describes the data model that was developed for the Soils4Africa project and used to implement a relational database to store the field and analytical data collected during the project. The data model is rooted in the international standard for digital exchange of soil data (ISO 28258) that was further specialized to fit the requirements for storing soil data in the Soils4Africa database.



European Commission | Horizon 2020 European Union funding for Research & Innovation

This project has received funding from the
EU Horizon 2020 Research and Innovation Programme
under Grant Agreement 862900.

SOILS4AFRICA

# PART A
# Soil Information System
# User Manual

**Project No. 869200**

31 May 2025

**Deliverable: Soils4Africa_D6.4a_v01**

Version 1.0



European Commission | Horizon 2020 European Union funding for Research & Innovation

This project has received funding from the
EU Horizon 2020 Research and Innovation Programme
under Grant Agreement 862900.

SOILS4AFRICA

## Contact details

**Director of Coordinating Institute – ISRIC:** Rik van den Bosch

**Project Coordinator:** Mary Steverink-Mosugu

**Address:** Droevendaalsesteeg 3, 6708 PB Wageningen (Building 101), The Netherlands

**Postal:** PO Box 353, 6700 AJ Wageningen, The Netherlands

**Phone:** +31 317 48 7634

**Email:** mary.steverink-mosugu@isric.org

## Project details

| | |
|---|---|
| Project number | 862900 |
| Project acronym | Soils4Africa |
| Project name | Soil Information System for Africa |
| Starting date | 01/06/2020 |
| Duration In months | 48 |
| Call (part) identifier | H2020-SFS-2019-2 |
| Topic | SFS-35-2019-2020 Sustainable Intensification in Africa |
| CORDIS url | https://cordis.europa.eu/project/id/862900 |

## Document details

| | |
|---|---|
| Work Package | WP6 |
| Deliverable number | D6.4a |
| Version | V1.0 |
| Filename | Soils4Africa_D6.4a_SIS_user_manual.pdf |
| Type of deliverable | Report |
| Dissemination level | Public |
| Lead partners | ISRIC |
| Contributing partners | ISRIC |
| Author | Eric Asamoah, Bas Kempen |
| Contributors | Paul van Genuchten |
| Due date | 30/11/2024 |
| Submission date | 31/05/2025 |





This project has received funding from the
EU Horizon 2020 Research and Innovation Programme
under Grant Agreement 862900.

**Table of Contents**





European Commission | Horizon 2020 European Union funding for Research & Innovation

This project has received funding from the
EU Horizon 2020 Research and Innovation Programme
under Grant Agreement 862900.

SOILS4AFRICA

# 1. Introduction

The Soils4Africa project aims to develop a soil information system (SIS) that serves information on the prevalence and spatial distribution of soil quality indicators and constraints relevant to the sustainable intensification of agriculture in Africa. This information can serve as a baseline for monitoring changes in soil conditions across the continent's agricultural land in the future.

Soil samples were collected from approximately 15,000 locations across a wide range of African countries, at two soil depths. Uniform procedures were applied during both fieldwork and laboratory analysis to ensure consistency of the resulting soil dataset.

A SIS was developed to store and disseminate the field observations collected at the sampling locations and the results of the laboratory analyses conducted using mid-infrared (MIR) spectroscopy and wet chemistry. In addition to these core datasets, several other data products were created and collected that are also distributed through the SIS.

To support its adoption beyond the lifetime of the project, the SIS was designed with branding that is distinct from the Soils4Africa project. It is currently branded as "Soil Data of Africa" at the time of the writing of this manual.

The SIS is designed in a modular way, comprising various components that allow users to discover, explore and download soil data—all built using open source software solutions. These components are integrated into a central SIS homepage, which also provides access to a range of resources produced by the Soils4Africa project, including reports, manuals, protocols, and instructional videos.

This *Soil Information System user manual* provides a comprehensive explanation of the setup and functionality of the SIS. It offers a detailed overview of the system's various components and illustrates their functionality and practical use.

Chapter 2 introduces the design and structure of the SIS homepage. Chapters 3, 4 and 5 focus on the three main data-sharing components that allow users to discover, explore and download the data hosted within the SIS.



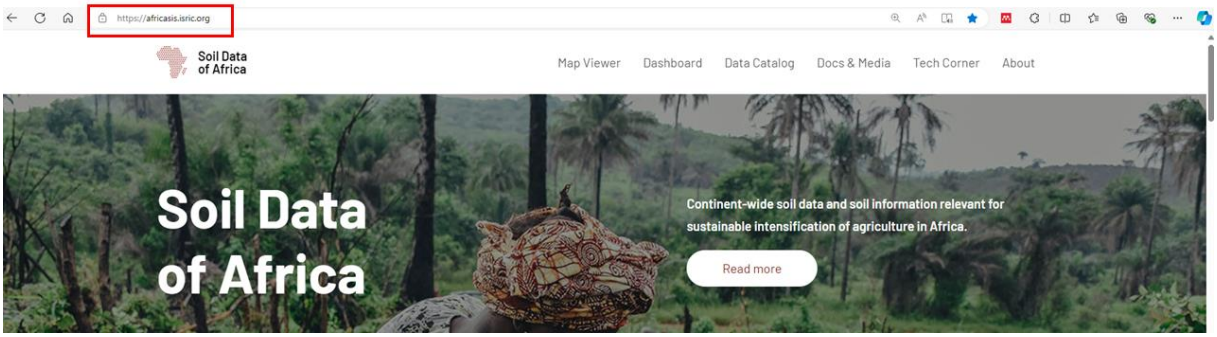
European Commission | Horizon 2020 European Union funding for Research & Innovation

This project has received funding from the
EU Horizon 2020 Research and Innovation Programme
under Grant Agreement 862900.

SOILS4AFRICA

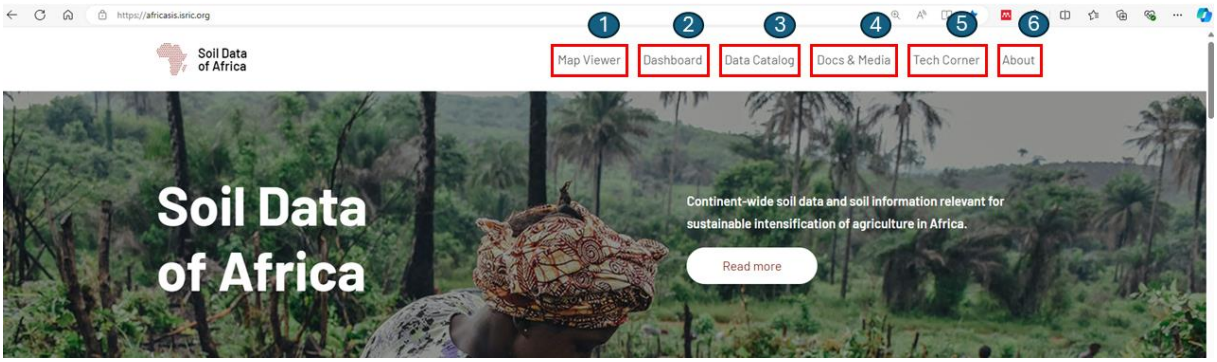
## 2. Soil Information System homepage

The homepage of the 'Soil Data of Africa' platform offers a comprehensive overview and access to various resources and tools related to soil data and information across the African continent. The 'Soil Data of Africa' platform is accessible at this URL: https://africasis.isric.org, where it is hosted until it is migrated to the final hosting institute in Africa.




The components of the soil information system are:

1. Map Viewer
2. Dashboard
3. Data Catalogue
4. Docs & Media
5. Tech Corner
6. About




The next sections describe the various components of the soil information system that are accessible through the system's homepage.



European Commission | Horizon 2020 European Union funding for Research & Innovation

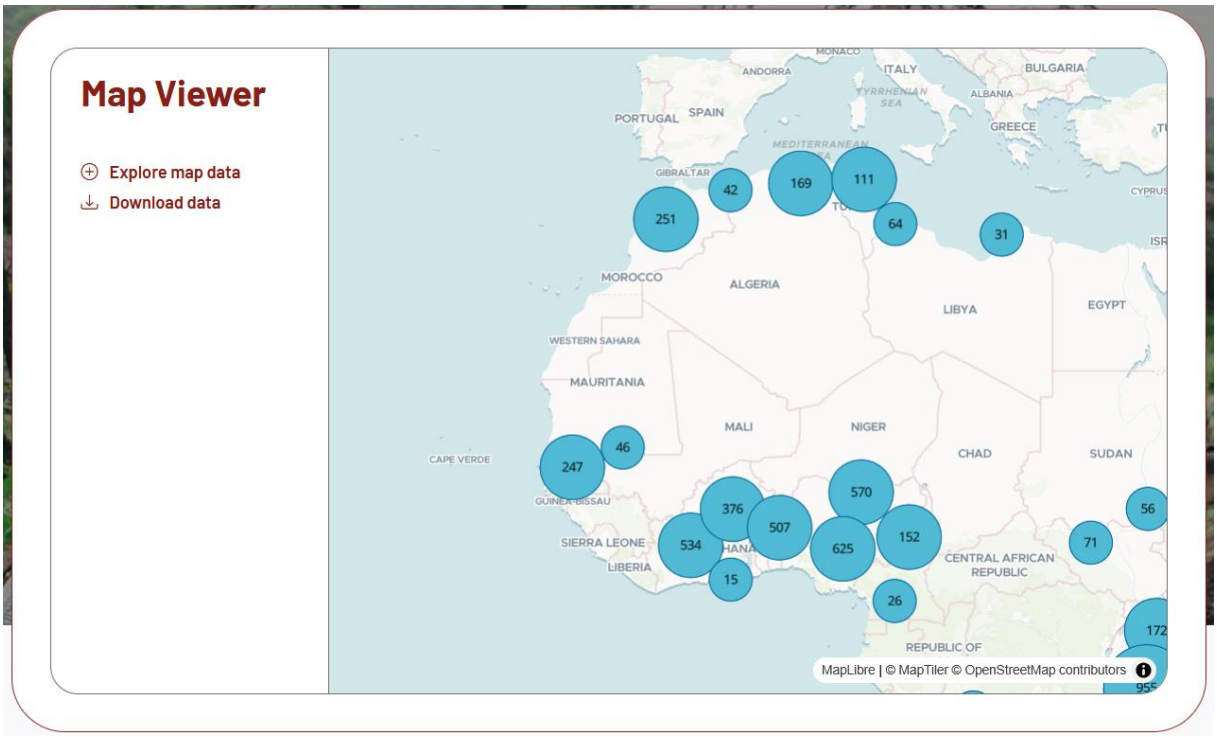
This project has received funding from the EU Horizon 2020 Research and Innovation Programme under Grant Agreement 862900.

SOILS4AFRICA

## Map Viewer

The *Map Viewer* provides an overview of the locations sampled during the Soils4Africa project. Users can zoom in on individual sampling sites and click on a sampling point to view some of the collected field data, including photos taken during the sampling. Clicking the 'Explore map data' button opens the full map viewer (Chapter 3) that allow exploration and interaction with the Soils4Africa data and many more geospatial data layers.

The 'Download data' button gives access to data dashboard from where the Soils4Africa data can be downloaded in different file formats.




## Dashboard

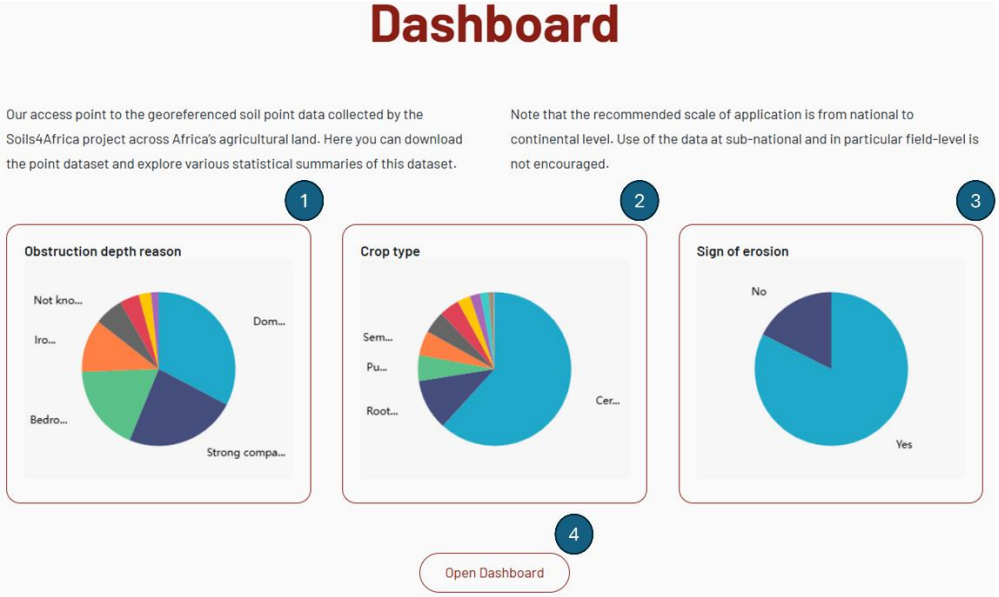
The *Dashboard* provides access to georeferenced soil point data collected by the Soils4Africa project across agricultural land in Africa. Users can download the point dataset and explore various statistical summaries of the datasets.

The *Dashboard has placeholders for graphical summaries of the data contained in the database of the SIS (1, 2, 3). These placeholders are interactive and provide more information by hovering the mouse pointer on top of these.*

A click on **Open dashboard (4) ,** takes you to the SOILS4AFRICA data dashboard.



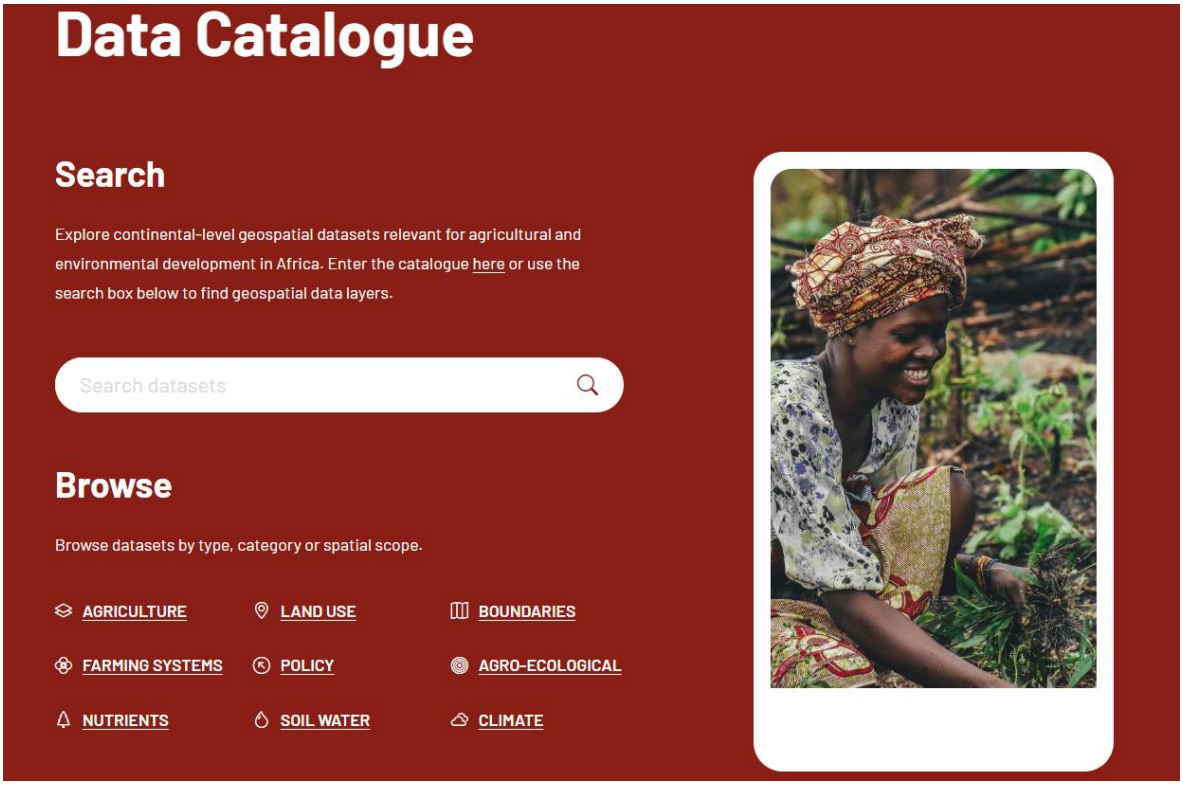
This project has received funding from the
EU Horizon 2020 Research and Innovation Programme
under Grant Agreement 862900.




## Data Catalogue

The *Data Catalogue* allows users to search through the system's geospatial data catalogue to find soil and other spatial datasets relevant for environmental and agricultural development in Africa.

One can enter the homepage catalogue service here or directly the catalogue via a search query.

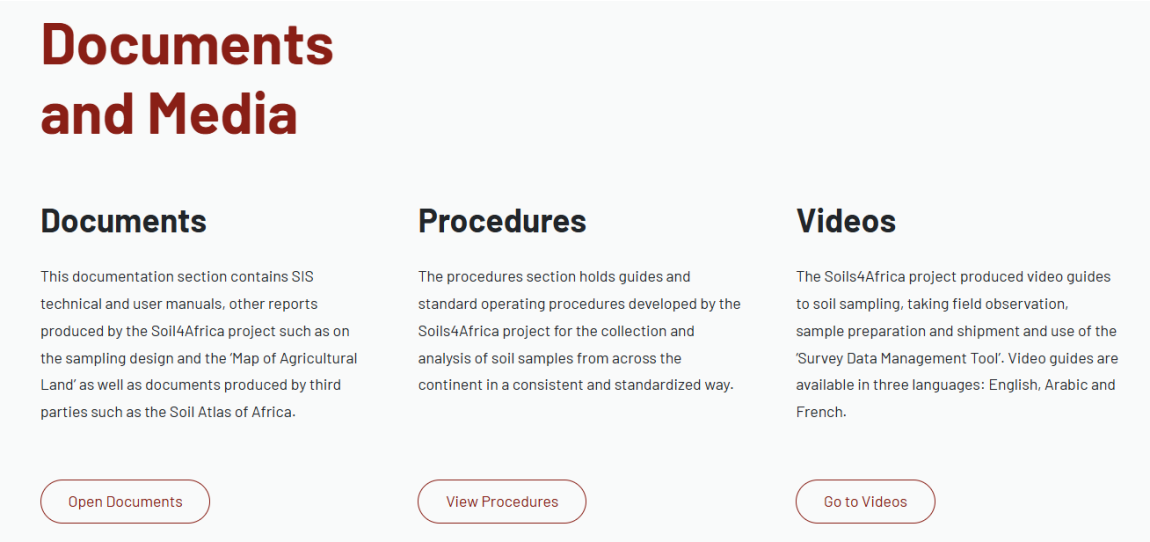





This project has received funding from the
EU Horizon 2020 Research and Innovation Programme
under Grant Agreement 862900.

## Documents and Media

The **Documents and Media** section hosts **documents, procedures and videos** developed by the Soils4Africa project.




**Documents:**

- This section includes technical manuals and Soils4Africa project reports.
- The **button** labelled **'Open Documents'** opens a subpage that allows users to access or download these materials.
- The subpage features a list of documents presented as expandable boxes. Clicking the '+' icon expands a box to reveal a brief abstract, the citation, and one or more download buttons for the document.



This project has received funding from the
EU Horizon 2020 Research and Innovation Programme
under Grant Agreement 862900.

# Documents

This page provides access the technical manuals developed under the Soils4Africa project, as well as additional project reports that support the development of this SIS, and several other relevant soil resources.

## Technical Manuals

- SIS user manual
- Survey Data Management Tool user manual
- Soils4Africa data model documentation

## Soils4Africa Project Reports

- Soils4Africa sampling design
- Map of Agricultural Land of Continental Africa
- Inventory of soil data for Africa represented in the ISRIC-WDC holdings
- User requirements for the IT infrastructure of the Soil Information System
- Technical design of the Soil Information System
- A set of use cases and supporting soil quality indicators
- Methods for deriving selected soil quality indicators

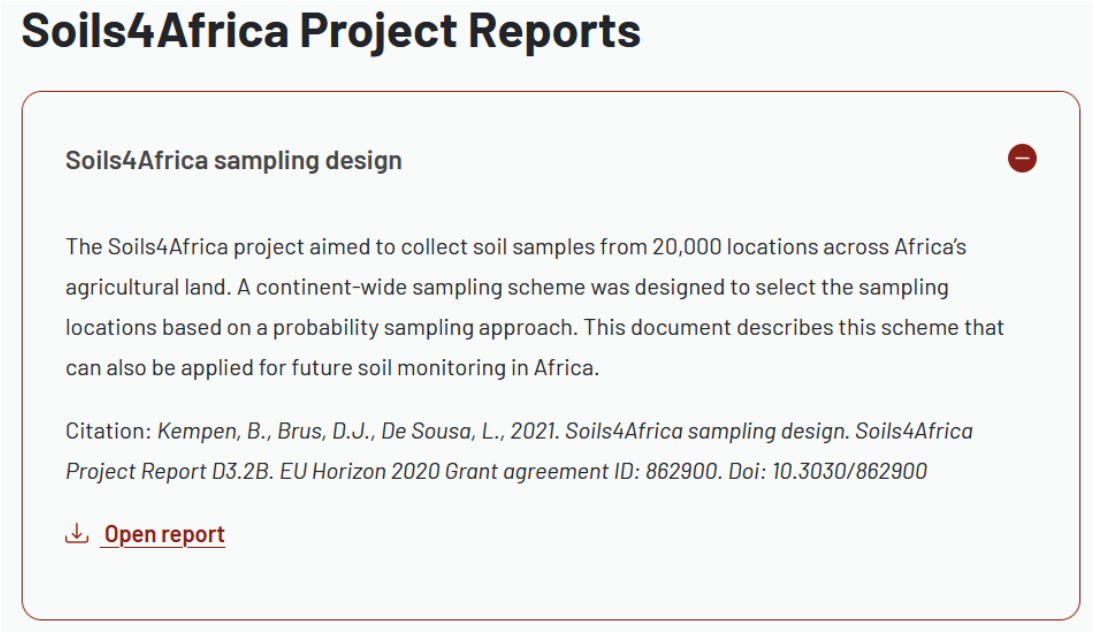
## Other



European Commission | Horizon 2020 European Union funding for Research & Innovation

This project has received funding from the EU Horizon 2020 Research and Innovation Programme under Grant Agreement 862900.

SOILS4AFRICA




**Procedures:**

- This section includes guides and standard operating procedures (SOPs) developed for the collection and analysis of soil samples as part of the Soils4Africa project, as well as several other manuals produced for continental-level soil assessment.
- The **button** labelled **'View Procedures'** opens a subpage enables users to explore and download the procedural guides.
- The subpage features a list of documents presented as expandable boxes. Clicking the '+' icon expands a box to reveal a brief abstract, the citation, and one or more download buttons for the document.



This project has received funding from the
EU Horizon 2020 Research and Innovation Programme
under Grant Agreement 862900.

# Procedures and Protocols

This page list the standard operating procedures, guidelines and protocols developed and used by Soils4Africa project for continental soil monitoring and assessment, as well as several other relevant resources for continental-level soil survey in Africa

- Guidance for Fieldwork
- Guidance for Laboratory Analysis
- Protocol for Field Survey
- Protocol for Field Survey Management
- Standard Operating Procedure: Soil Sample Collection and Observation
- Standard Operating Procedure: Sample Preparation and Shipment
- Reference Site Soil Sampling Protocol
- AfSIS Technical Specifications
- Land Degradation Surveillance Framework (LSDF): field guide



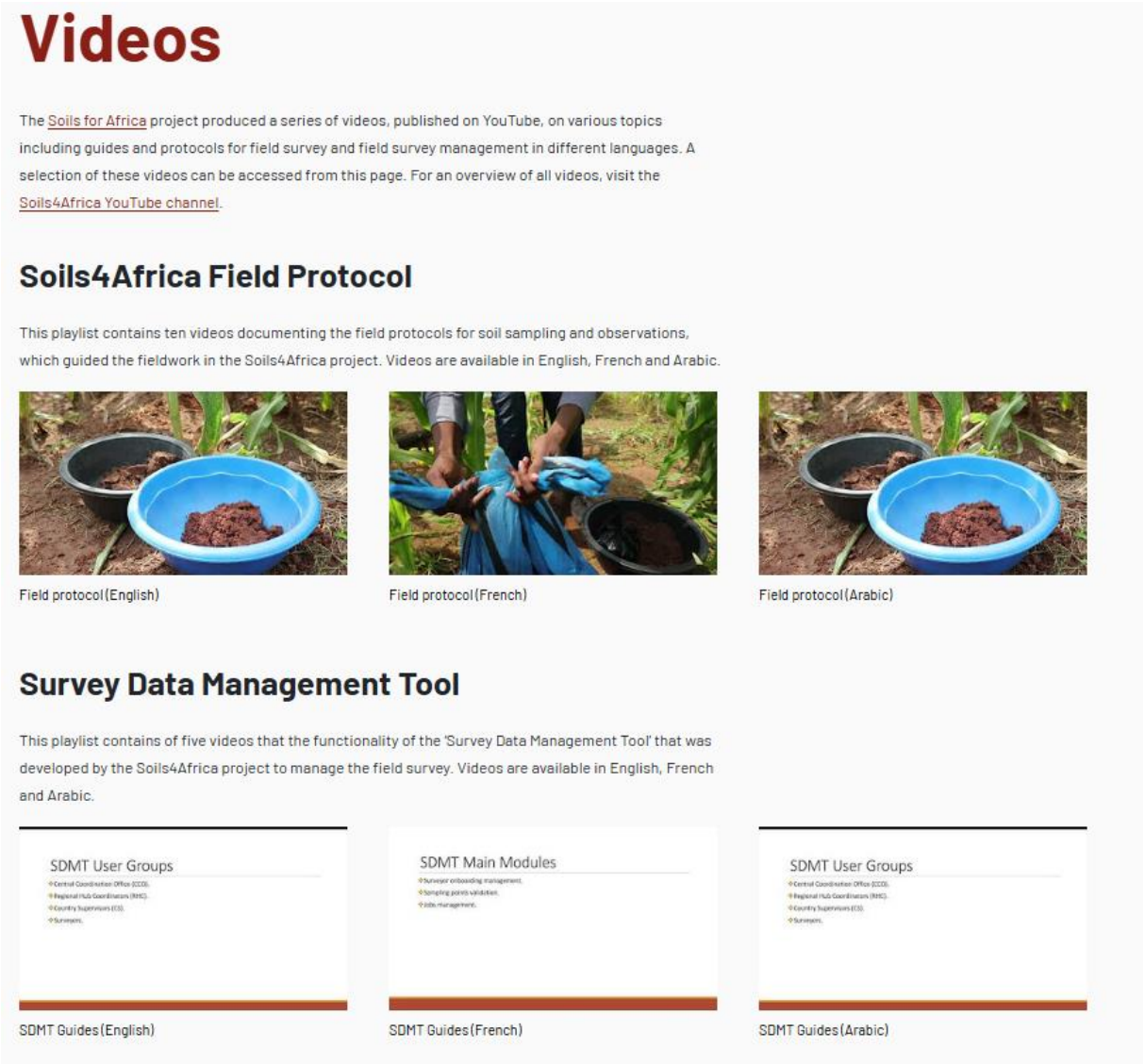
European Commission | Horizon 2020 European Union funding for Research & Innovation

This project has received funding from the EU Horizon 2020 Research and Innovation Programme under Grant Agreement 862900.

SOILS4AFRICA

**Videos:**

- Provides video guides created by the Soils4Africa project. These guides cover topics such as:
    - Soil sampling techniques.
    - Field observation procedures.
    - Sample preparation and shipment processes.
    - The use of the 'Survey Data Management Tool'.
- The videos are multilingual and available in English, Arabic, and French.

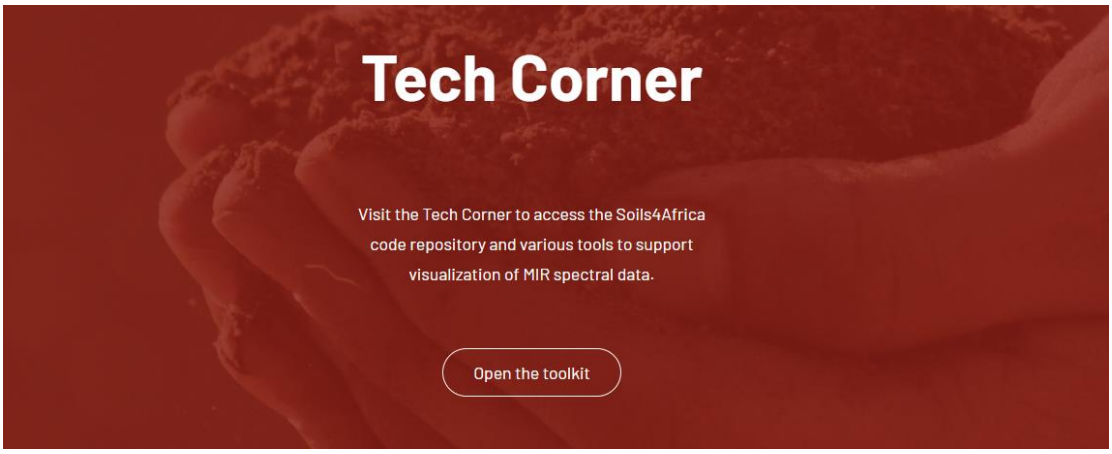





This project has received funding from the
EU Horizon 2020 Research and Innovation Programme
under Grant Agreement 862900.

## Tech Corner

The *Tech Corner* offers scripts for selecting sample locations based on a statistical sampling design, developing soil maps, and calibrating spectral models, available from the toolkit section.

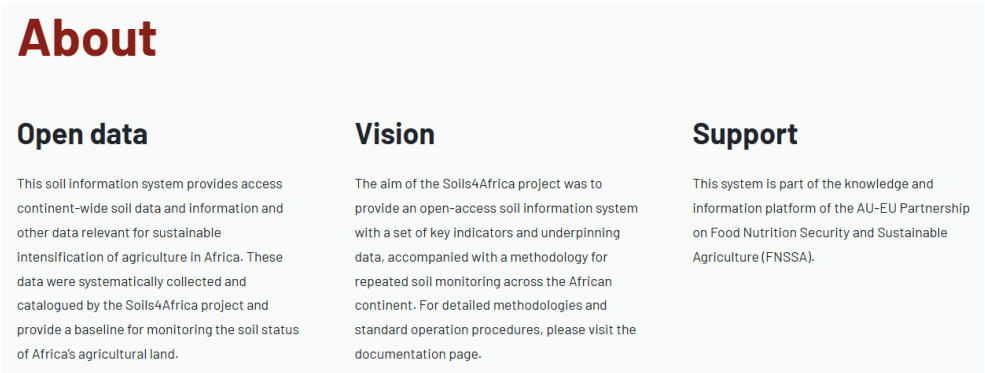



A **click** on the **Open the toolkit** takes the user to the Tech Corner.

**Important Note**: The development of the *Tech Corner* section was still being finalized at the time of writing this manual. As such, the final version of *Tech Corner* may differ from what is presented here.

## About

The 'About' page provides an overview of the **purpose, vision, and support** structure of the Soils4Africa project. The content is organized into three subsections: **Open Data, Vision, and Support.**






European Commission | Horizon 2020 European Union funding for Research & Innovation

This project has received funding from the EU Horizon 2020 Research and Innovation Programme under Grant Agreement 862900.

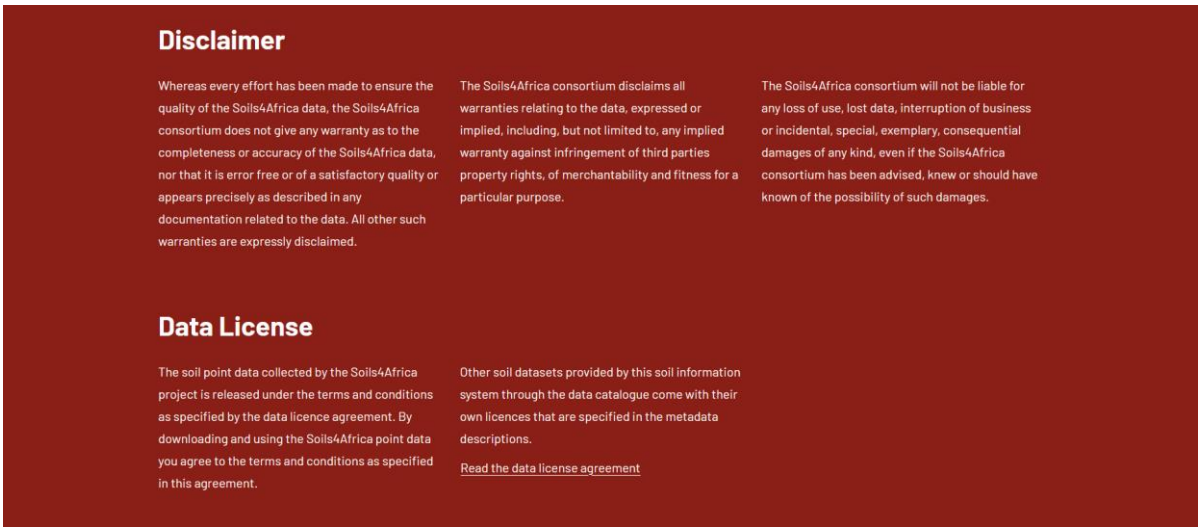
SOILS4AFRICA

## Partners

Lists the 'Partners' who collaborated to develop this soil information system under the Soils4Africa project.


## Disclaimer and Data Licence

The SIS homepage is concluded with a disclaimer and a link to the Soils4Africa data sharing licence.






European Commission | Horizon 2020 European Union funding for Research & Innovation

This project has received funding from the EU Horizon 2020 Research and Innovation Programme under Grant Agreement 862900.

SOILS4AFRICA

## 3. WebGIS

The WebGIS service is an innovative platform designed to enhance how you interact with geographic data. It offers a user-friendly interface that enables visualization, analysis, and sharing of geographic information more efficiently. This platform stands out by allowing users to upload their datasets or utilize our extensive catalogue of geographic data, ranging from soil properties to ecological datasets. With features such as dynamic 3D visualizations, customizable map settings, and robust sharing options, including downloadable maps and URL links for collaboration, the WebGIS service is an indispensable tool. It supports a wide range of applications, enhancing project planning, research, and presentation capabilities, ultimately fostering a better understanding and management of spatial phenomena. The soil information system developed for the Soils4Africa project uses a TerriaJS implementation as WebGIS. The manual below provides an overview of the functionality, a full manual of the TerriaJS library is available at https://userguide.terria.io.

This chapter provides several examples of how to find, select, and add data layers to the WebGIS. Additional examples can be found in the Annex.

### Accessing the 'Soil Data of Africa' WebGIS platform

1. To access the Soil Data of Africa webGIS platform enter the URL **"https://africasis.isric.org/map/"**

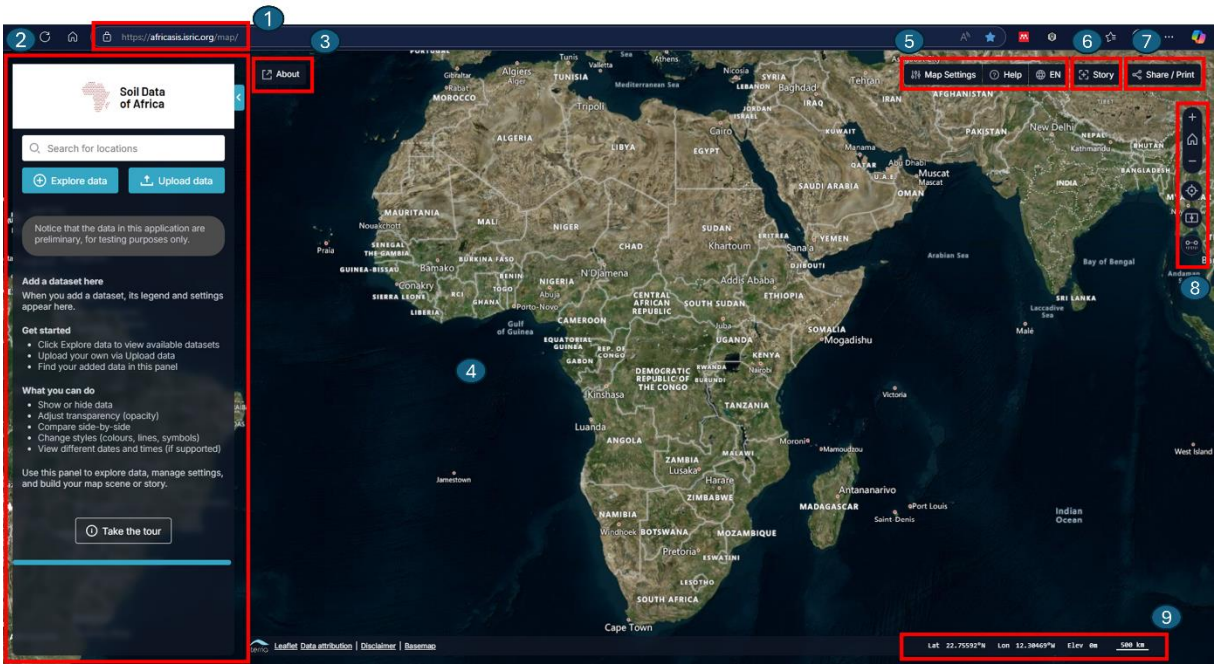



This project has received funding from the
EU Horizon 2020 Research and Innovation Programme
under Grant Agreement 862900.

## WebGIS Interface

The WebGIS presents a user-friendly interface focused on enhancing user interaction with geographical and soil data specific to Africa.

1. The **Browser Tab and URL** shows the URL of the WebGIS site
2. **Navigation and Search Bar** on the left with options to search for locations, explore map data, and upload data. The sidebar also contains Helpful hints and notices regarding the data and functionality of the site.
3. **About**
4. **Main Map Window** is a displays window which displays information.
5. **Map Settings** and **Help** options. Here you can select 2D or 3D map view and the base map.
6. **Story** feature to create and share interactive stories from the maps.
7. **Share / Print** to share a URL that gives access to the map, download the map as PNG or print the map.
8. **Zoom settings** (zoom in, reset zoom, zoom out), **Location, Compare**, **Line measure** features.
9. **Bottom Bar** displays map coordinates (Latitude and Longitude), elevation, and a scale bar indicating distance.






This project has received funding from the
EU Horizon 2020 Research and Innovation Programme
under Grant Agreement 862900.

## Navigating the Workbench

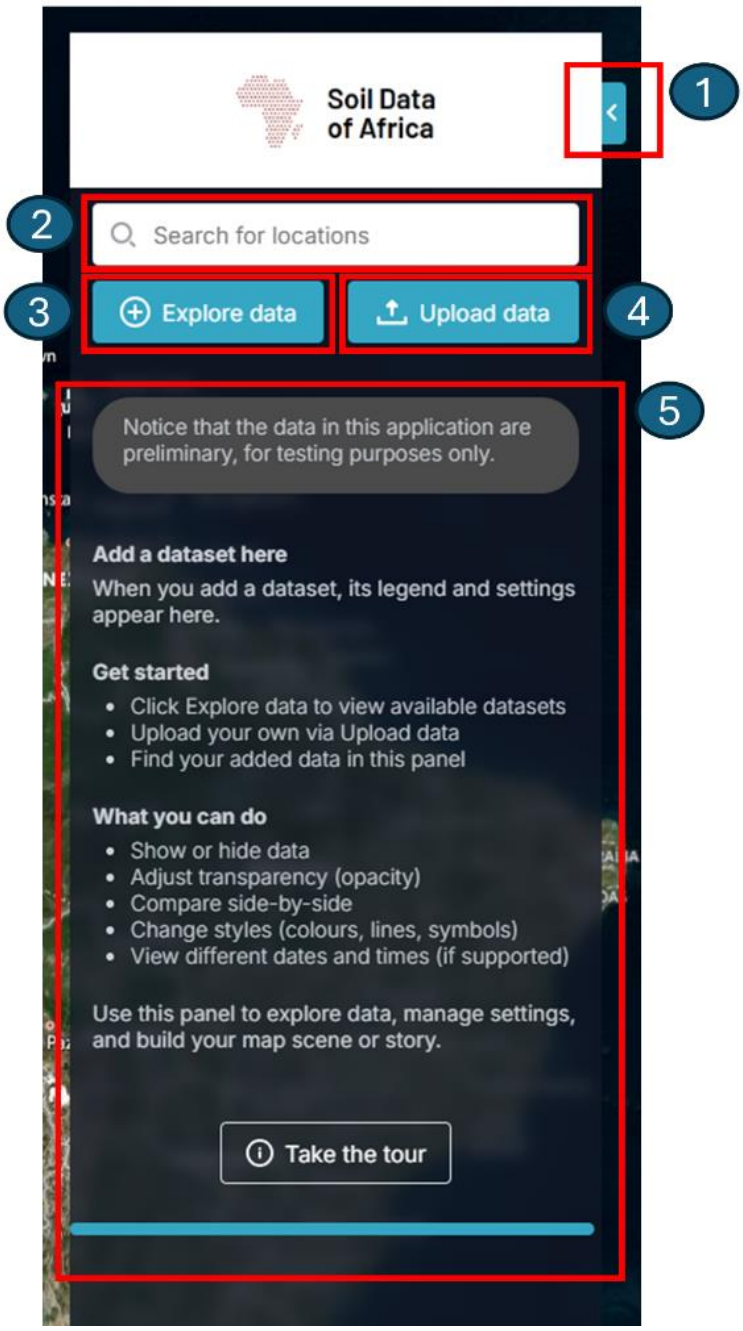
1. **Hide Workbench (<)** located at the top right corner allows users to close the navigation and search bar.
2. **Search Bar** enables users to search for specific locations within the map interface.
3. **Explore Map Data** enables the user to view different types of data layered in the map.
4. **Upload** allows users to upload their own data layers to the map.
5. The information and helpful hints section provides information about how to use the application effectively. It shows where **data legends and settings appear** when added to the map window. How to **get started** and **what you can do** using the platform and **Take the tour** button, when clicked shows a guide through the functions on the platform.



European Commission | Horizon 2020 European Union funding for Research & Innovation

This project has received funding from the
EU Horizon 2020 Research and Innovation Programme
under Grant Agreement 862900.

SOILS 4 AFRICA

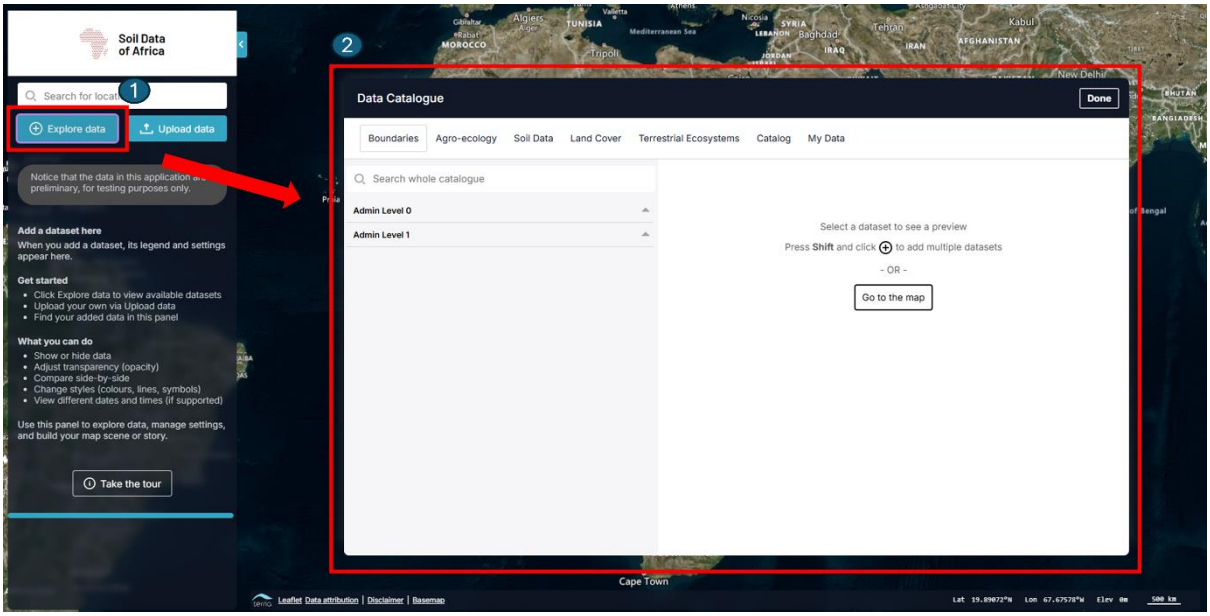





This project has received funding from the
EU Horizon 2020 Research and Innovation Programme
under Grant Agreement 862900.

## Accessing Map Datasets

1. Use the **Explore Map Data** button to explore data specific to Africa.
2. Click on the **Explore map data** to expand and show the various datasets in the WebGIS.
3. The data catalogue is displayed in the main map window displaying the datasets to explore.

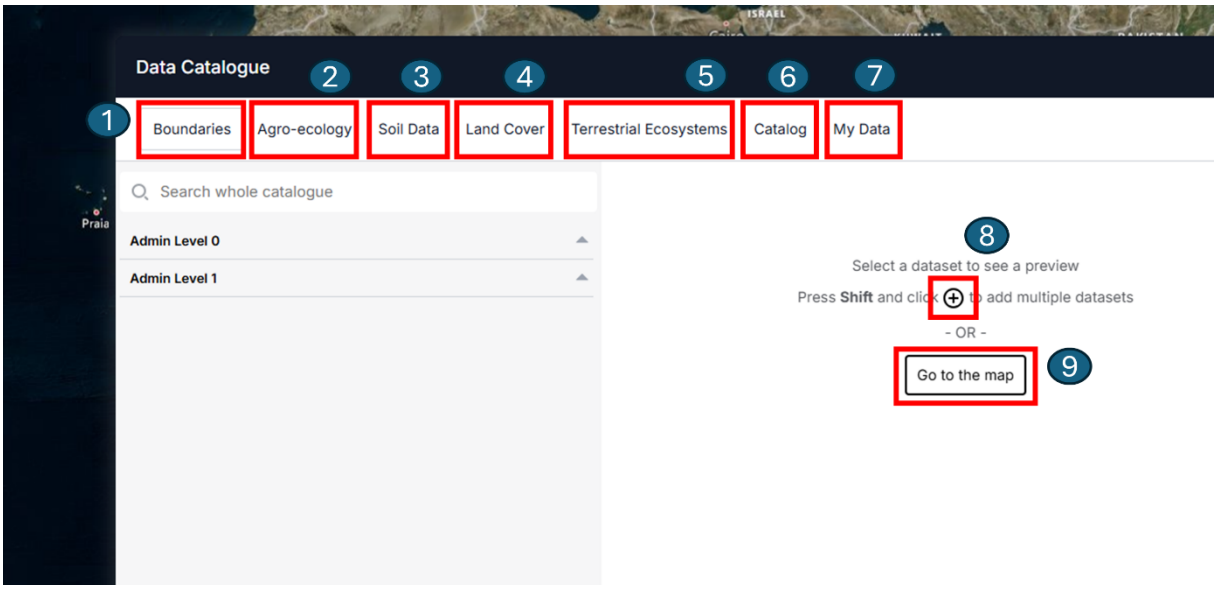



The **Explore map data** button when clicked lists **various datasets** and other **functions**, such as:

1. Boundaries
2. Agro-ecology
3. Soil Data
4. Land Cover
5. Terrestrial Ecosystems
6. Catalogue
7. My Data
8. Add Multiple Datasets
9. Go to the map



This project has received funding from the
EU Horizon 2020 Research and Innovation Programme
under Grant Agreement 862900.




On the next pages we provide three examples of adding layers to the WebGIS. Examples of the layers can be found in the annex.

## Display Soils4Africa field survey datapoints

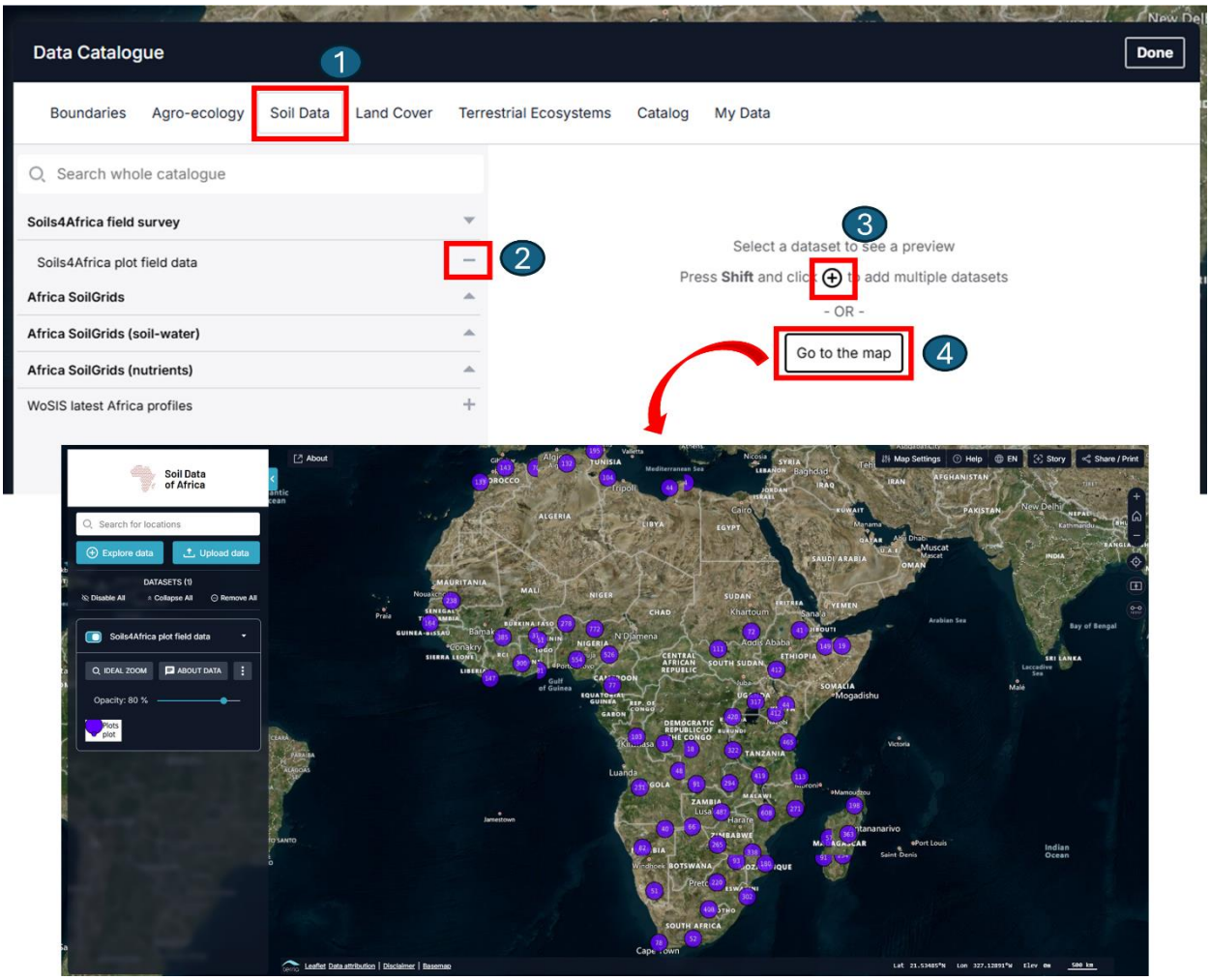
1. In the data catalogue window displaying in the main map window, click on **Soil Data** to display the various soil datasets.
2. Click on **Soils4Africa field survey** to expand to display **"Soils4Africa field survey"** in the data catalogue.
3. Click on the **+** symbol to load the **Soils4Africa field survey** into the map window.
4. The **Go to the map** when clicked also displays the **Soils4Africa field survey** into the map window.



European Commission | Horizon 2020 European Union funding for Research & Innovation

This project has received funding from the EU Horizon 2020 Research and Innovation Programme under Grant Agreement 862900.

SOILS 4 AFRICA

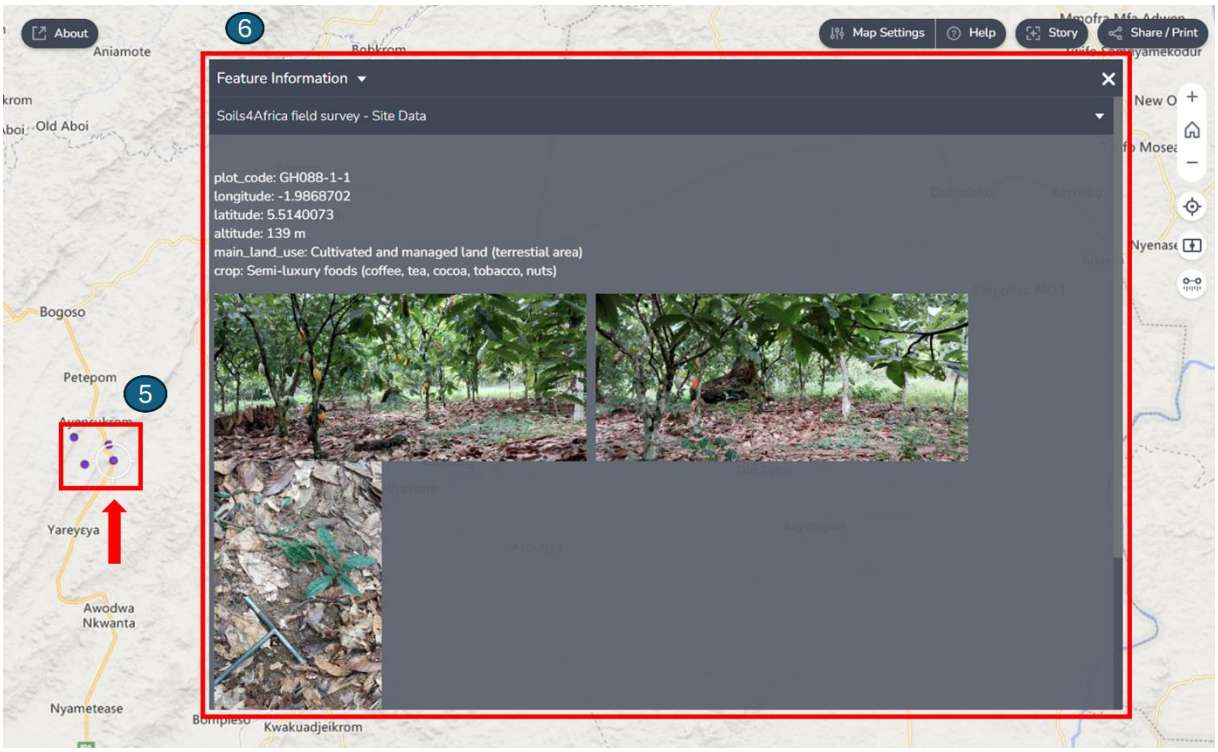



5. **Zoom out and Click on a Data Point** 

to display information available for a particular data point.
6. A display pops up showing the information available to a particular selected data point (e.g. the coordinate, land use, cultivated crop).



This project has received funding from the
EU Horizon 2020 Research and Innovation Programme
under Grant Agreement 862900.





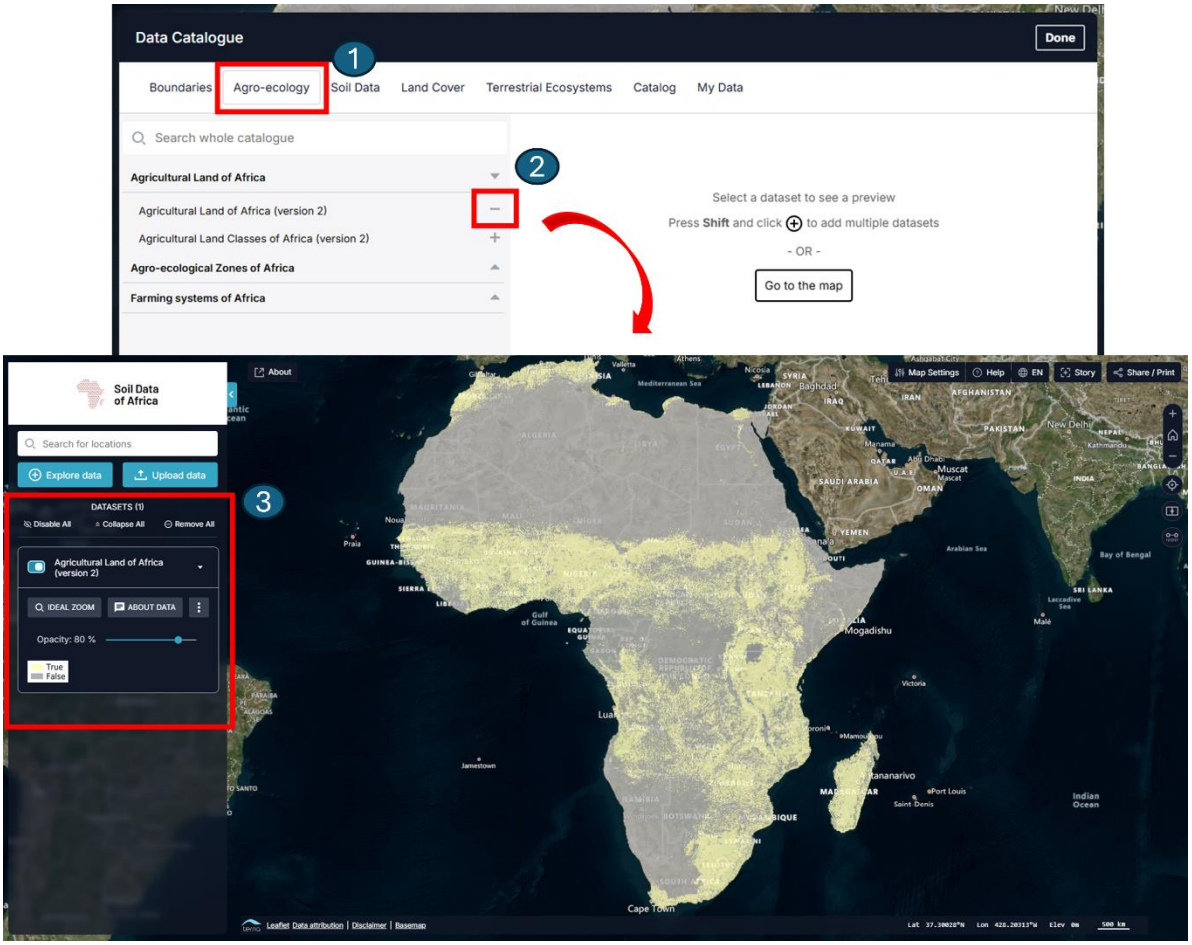

This project has received funding from the
EU Horizon 2020 Research and Innovation Programme
under Grant Agreement 862900.

## Display agricultural lands of Africa (version 2)

1. In the data catalogue window displaying in the main map window, click on **Agro-ecology** to display the various datasets under this sub-domain.

2. Click on **Agricultural Land of Africa** to expand to display;
    - "Agricultural Land of Africa (version 2)" and,
    - "Agricultural Land Classes of Africa (version 2),
    - Click on the **-** symbol to load the **Agricultural Land of Africa (version 2)** into the map window.

3. The left panel of the map window shows the **Active Dataset** that is currently active in the map window. In this case, "Agricultural Land of Africa (version 2)".






This project has received funding from the EU Horizon 2020 Research and Innovation Programme under Grant Agreement 862900.

## Display agricultural lands classes of Africa (version 2)

1. In the data catalogue window displaying in the main map window, click on **Agro-ecology** to display the various datasets under this sub-domain.

   - Click on **Agricultural Land of Africa** to expand to display;
     - "Agricultural Land of Africa (version 2)" and,
     - "Agricultural Land Classes of Africa (version 2),

2. Click on the **-** symbol to load the **Agricultural Land of Africa (version 2)** into the map window.

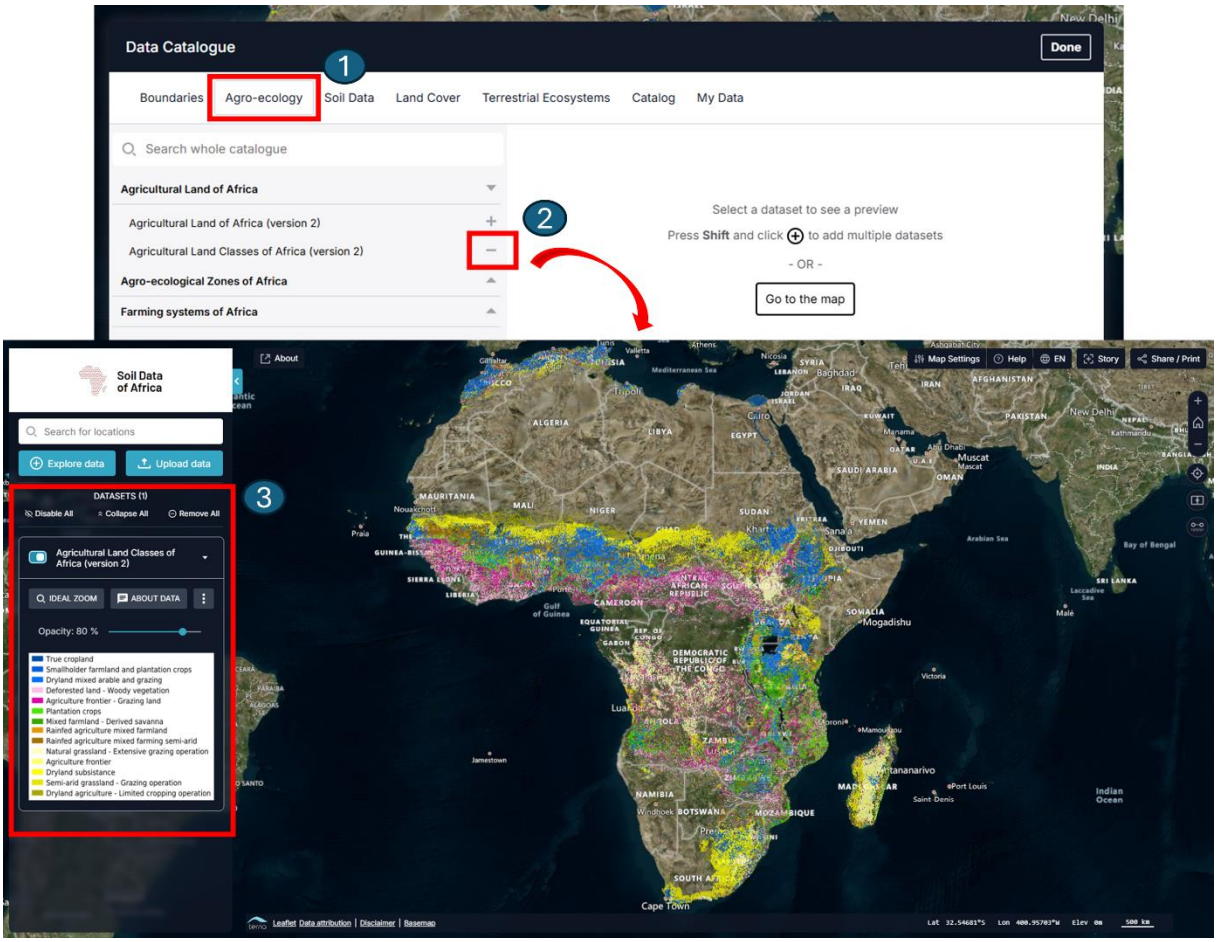
3. The left panel of the map window shows the **Active Datasets** that are currently active on the map.
   a. Users can adjust the **opacity** of the dataset display on the map and select different types of land classifications to visualize, such as true cropland, smallholder farmland, rainfed agriculture, natural grassland, etc.
   b. The map is colour-coded according to the different types of agricultural land use. Each colour corresponds to a different type of land use as listed in the **Active Dataset Options Panel**, illustrating a diverse range of agricultural practices across the continent.



European Commission | Horizon 2020 European Union funding for Research & Innovation

This project has received funding from the
EU Horizon 2020 Research and Innovation Programme
under Grant Agreement 862900.

SOILS 4 AFRICA

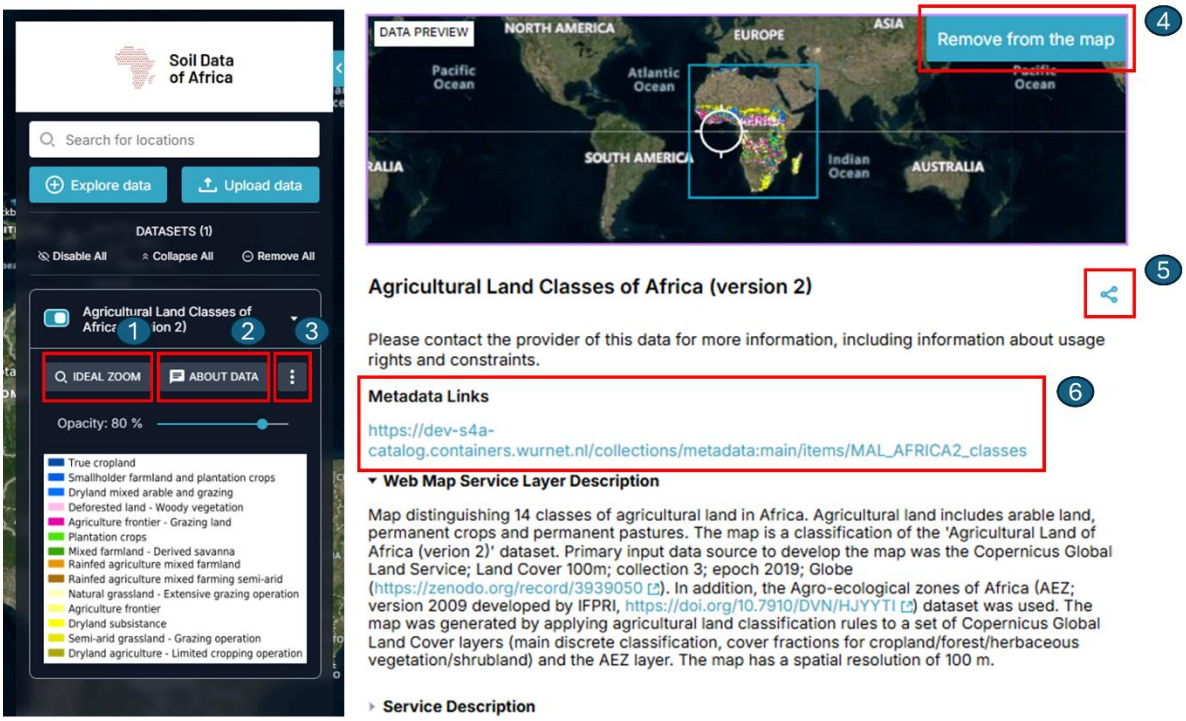



## Additional Workbench and About Data functions

1. Click on the **Ideal Zoom** button to automatically adjust and centre the map window with focus on the African continent.
2. Click on the **About Data** to display Metadata on the data.
3. Click on the **three vertical dots** to display more actions to perform on the data.
    - Click on the **Remove from the map** to remove data already displaying in the map window.
    - Click on the **Share your catalogue with other** to share your catalogue.
    - Display the link to the **Metadata** of the selected data.



This project has received funding from the
EU Horizon 2020 Research and Innovation Programme
under Grant Agreement 862900.

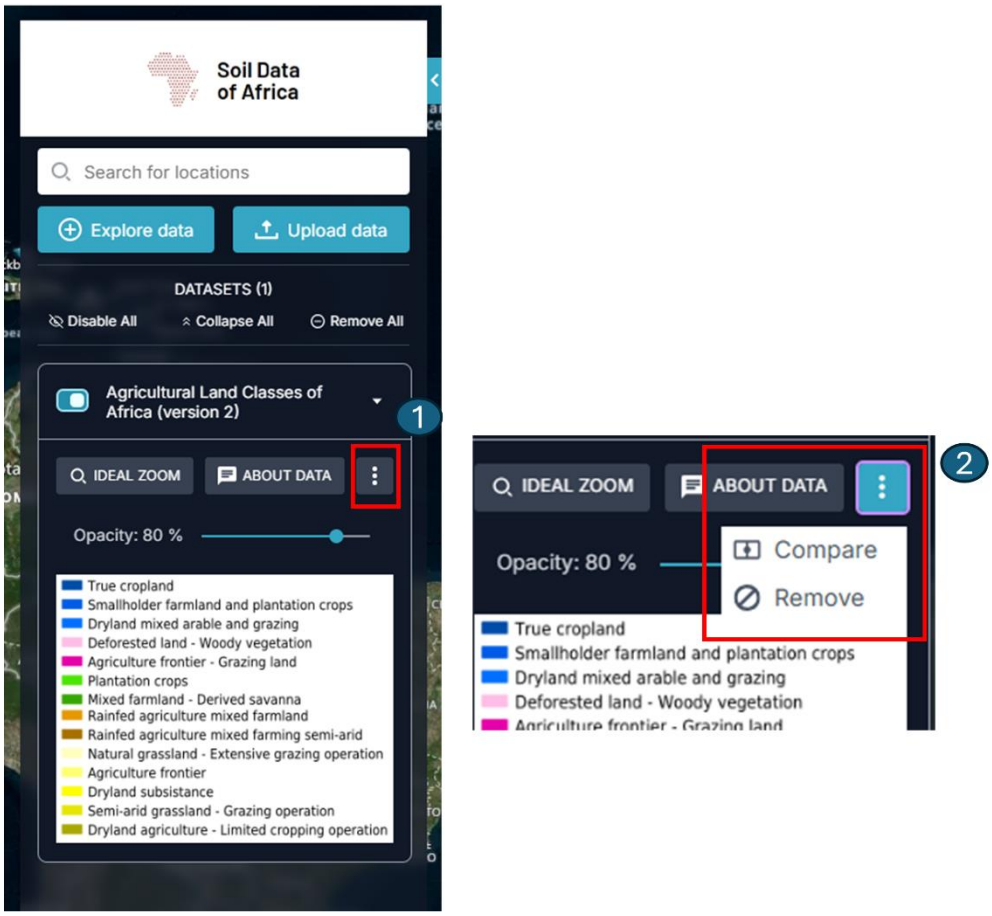



## Additional Workbench function cont.

1. Click on the **three vertical dots** to display more actions to perform on the data.
2. The right side shows the expanded view of the **three vertical dots** icon revealing two additional options:
    - **"Compare"** button may allow the user to compare this dataset with another dataset for analytical purposes.
    - **"Remove"** button allows for the removal of the dataset from the current map view.



This project has received funding from the EU Horizon 2020 Research and Innovation Programme under Grant Agreement 862900.




The **"Compare"** button allows a *"Split-Screen Control"* that enables toggling between a split view and a single view of the map, aiding in comparative analysis of different data layers.

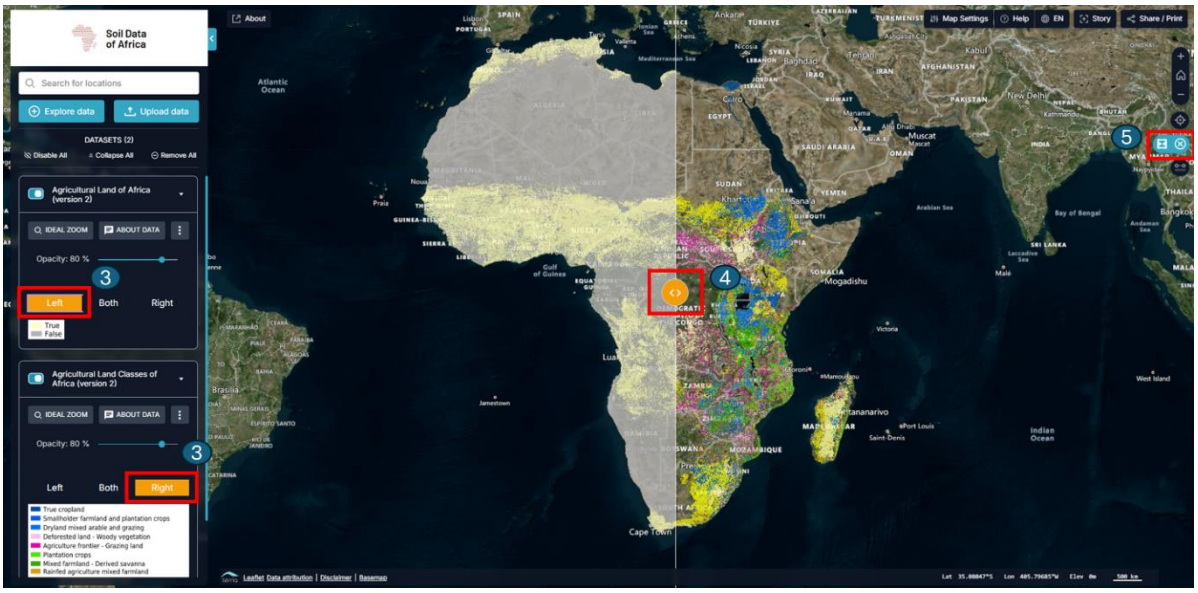
3. The **"Left and Right"** button allows the display of datasets on either side of the Map window.
4. The **Drag Left or Right Adjustment** button allows to adjust views to either side of the map window for image comparison.
5. The End Side-by-Side comparison button allows users to close the comparison view.



European Commission | Horizon 2020 European Union funding for Research & Innovation

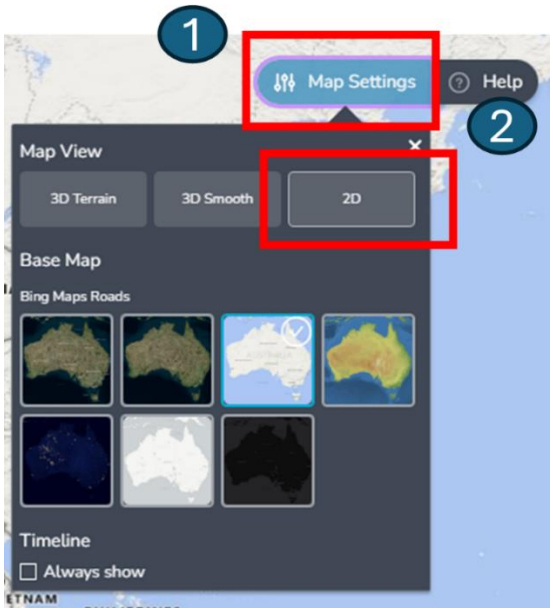
This project has received funding from the EU Horizon 2020 Research and Innovation Programme under Grant Agreement 862900.

SOILS 4 AFRICA




## Map Settings Panel

1. The **Map Settings** panel displays different selection modes;
    - "Map View", "Base Map" and "Timeline" selection options.
    - The Map view selection option displays different viewing modes for the map which are available in:
        - "3D Terrain" and "3D Smooth" options, which allows the user to view the map in a three-dimensional format providing either a textured or smoothed representation of terrain.
2. **"2D" option**, highlighted, displays viewing the map in a two-dimensional, "flat view".






European Commission
Horizon 2020
European Union funding for Research & Innovation

This project has received funding from the
EU Horizon 2020 Research and Innovation Programme
under Grant Agreement 862900.

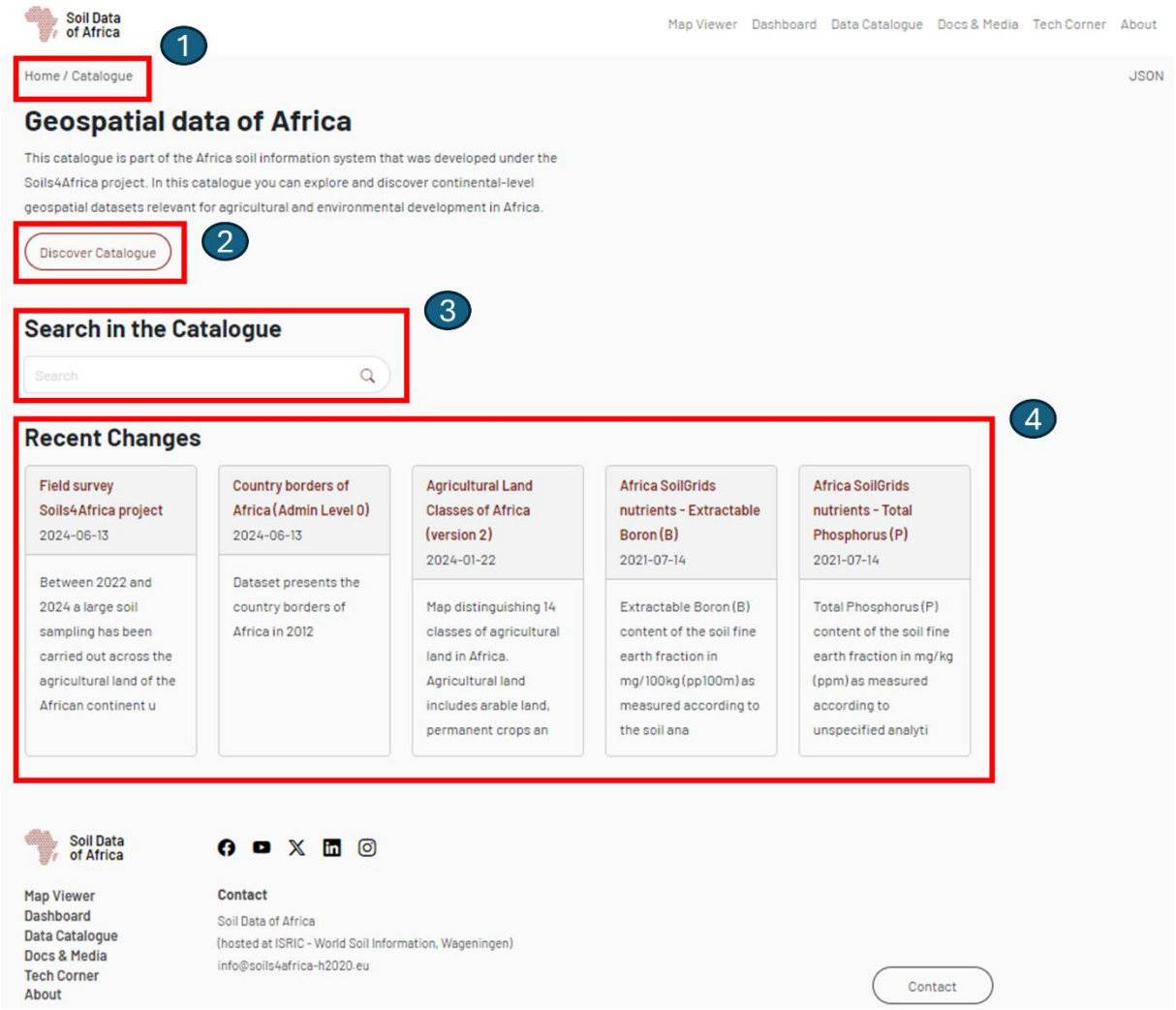
SOILS4AFRICA

## 4. Data catalogue

The catalogue is accessible via this URL https://africasis.isric.org/cat/. It has three components **(i) Catalogue homepage, (ii) Content List, and (iii) Metadata Records.**

The **catalogue home page** displays key features:

1. Home/Catalogue Button: Provides easy navigation to the home page.
2. Discover Catalogue Button: Allows users to explore the data catalogue.
3. Search in the Catalogue: Allows users to search for specific topics within the catalogue.
4. Recent Changes section: Provides a display of recently or added datasets.

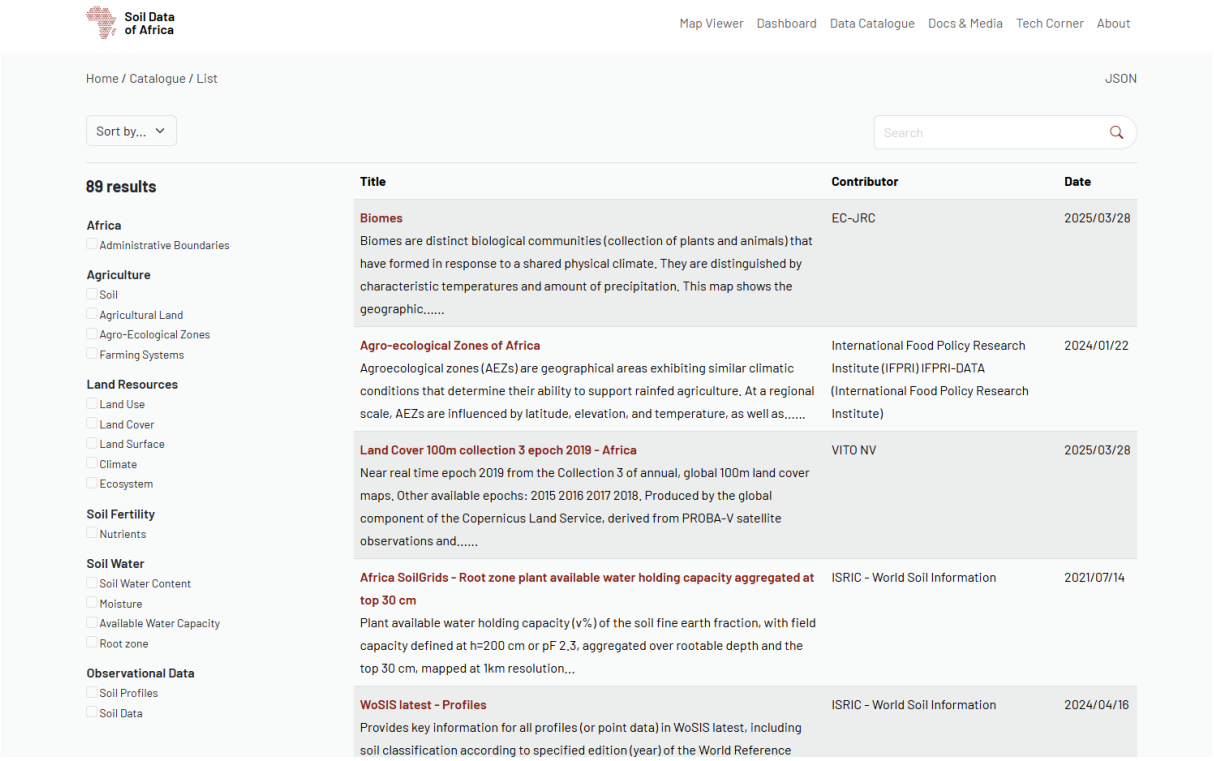





This project has received funding from the
EU Horizon 2020 Research and Innovation Programme
under Grant Agreement 862900.

## The content list

This section displays a searchable and sortable catalogue of available datasets.




## Metadata records

This section provides detailed metadata about a specific dataset selected from the content list. Example shown below:



This project has received funding from the
EU Horizon 2020 Research and Innovation Programme
under Grant Agreement 862900.

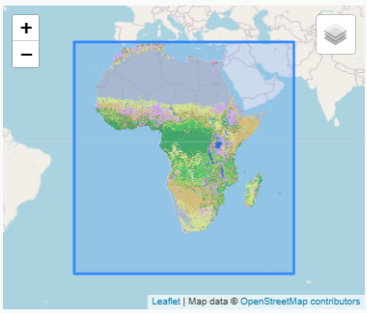
Soil Data of Africa

Map Viewer Dashboard Data Catalogue Docs & Media Tech Corner About

Home / Catalogue / List / Land Cover 100m collection 3 epoch 2019 - Africa JSON | XML

## Land Cover 100m collection 3 epoch 2019 - Africa

Near real time epoch 2019 from the Collection 3 of annual, global 100m land cover maps. Other available epochs: 2015 2016 2017 2018. Produced by the global component of the Copernicus Land Service, derived from PROBA-V satellite observations and ancillary datasets. The The map shows a main discrete classification with 23 classes aligned with UN-FAOs Land Cover Classification System




Open Land Cover 100m coll in map

land cover | land use | copernicus | Land Resources | earth | PROBA-V | satellite | remote sensing

### Contacts

**Marcel Buchhorn (VITO NV), Bruno Smets (VITO NV), Luc Bertels (VITO NV), Bert De Roo (VITO NV), Myroslava Lesiv (IIASA), Nandin-Erdene Tsendbazar (Wageningen University), Martin Herold (Wageningen University), Steffen Fritz (IIASA) VITO NV**

Role: author

### Temporal

Updated: 2025-03-28

Temporal extent: 2025-03-28

### Formats

- image/tiff
- OGC:WMS
- OGC:WCS
- canonical

### Links

- https://files.africasis.isric.org/cgls_2018_2019/PROBAV_LC100_global_v3.0.1_2019-nrt_Discrete-Classification-map_EPSG...
- https://africasis.isric.org/ows/cgls_2018_2019?service=WMS&version=1.3.0&request=GetCapabilities
- https://africasis.isric.org/ows/cgls_2018_2019?service=WCS&version=1.3.0&request=GetCapabilities



European Commission | Horizon 2020 European Union funding for Research & Innovation

This project has received funding from the EU Horizon 2020 Research and Innovation Programme under Grant Agreement 862900.

SOILS4AFRICA

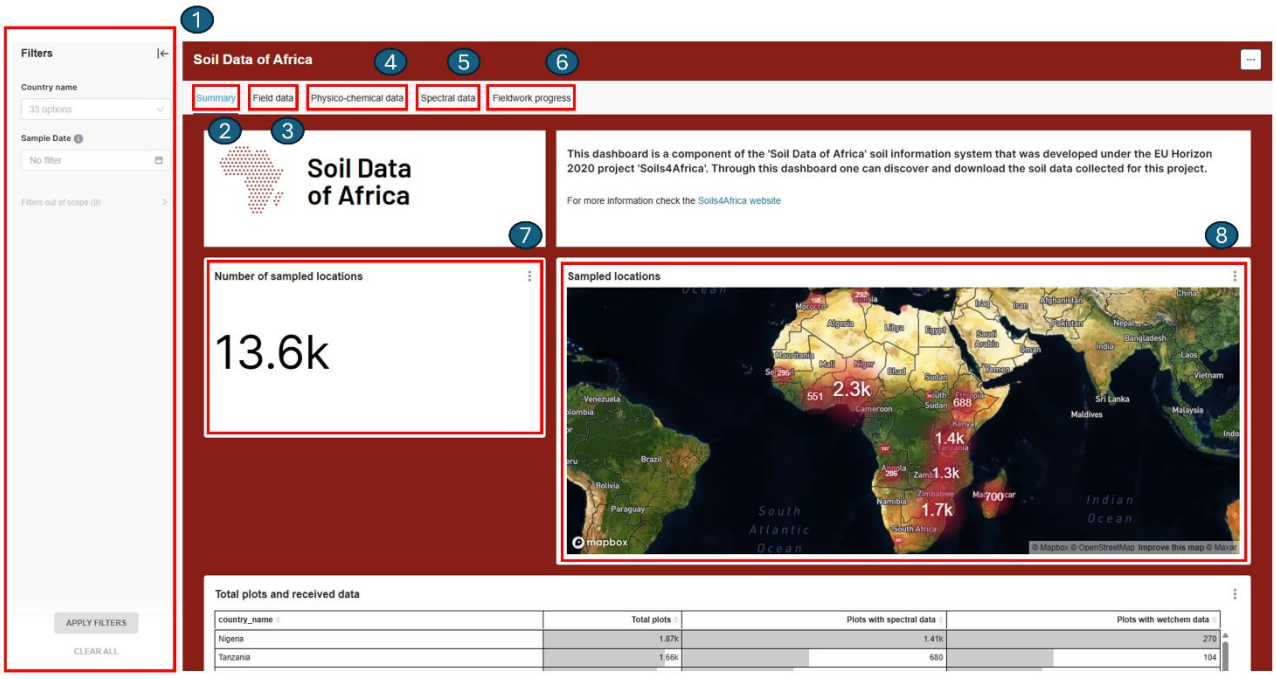
## 5. Data dashboard

The Soils4Africa dashboard provides an overview of soil data collected in the framework of the EU Horizon 2020 Soils4Africa project as well as links to download the data.

Note that the design of the dashboard and display of data summaries and statistics is fully customizable and is likely to change and get richer over time as the continental data assets evolve.

Currently, the homepage of the dashboard has the following functionalities:

1. **Filter panel** allowing users to refine the data displayed on the dashboard, using **Country name** and **Sample date.**
2. Data tab to display the data based on **Summary** of the datasets.
3. Data tab to display the data based on the **Field data** collected.
4. Data tab to display the data based on **Physico-chemical data**.
5. Data tab to display the data based on **Spectral data**.
6. Data tab to display **Fieldwork Progress**.
7. Displays the total number of sampled locations.
8. Display of data density from various countries or regions.

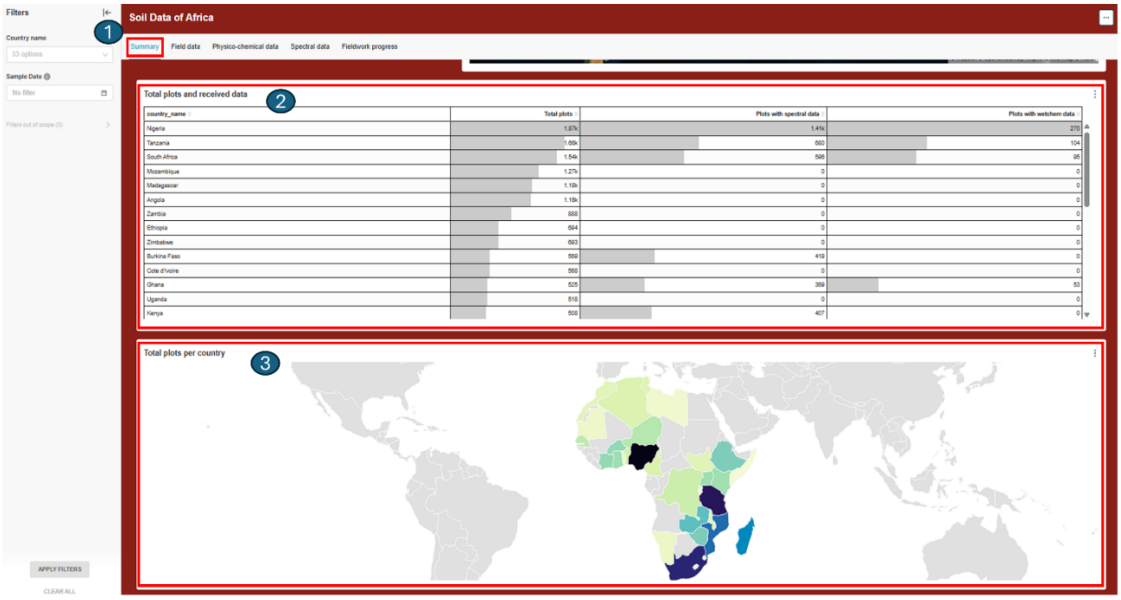





This project has received funding from the
EU Horizon 2020 Research and Innovation Programme
under Grant Agreement 862900.

**The Summary Tab**

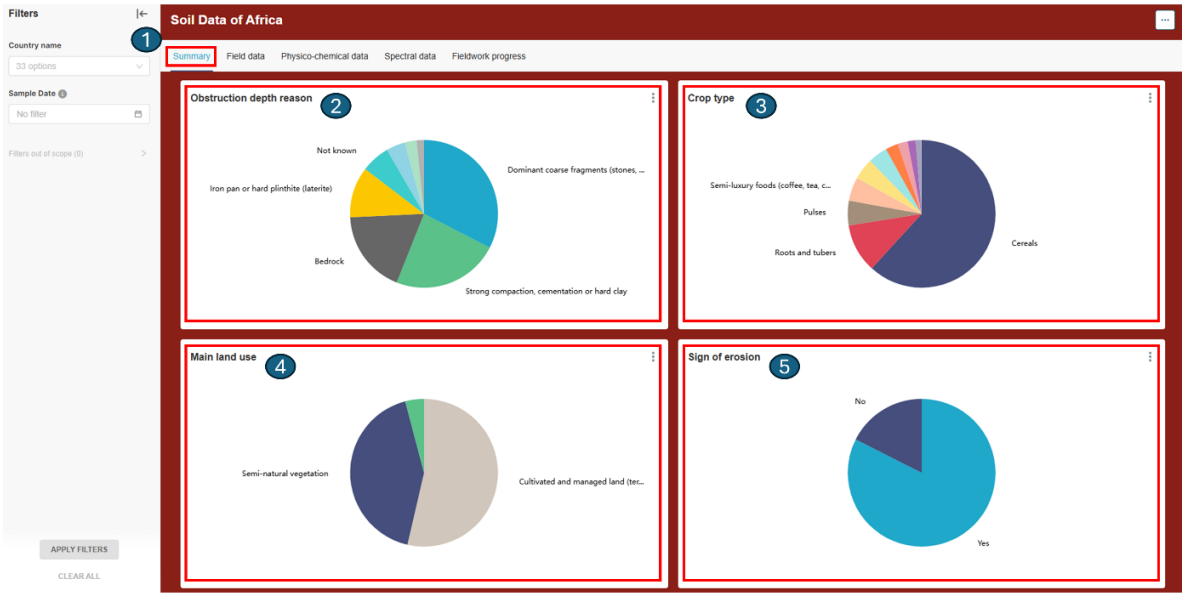
1. A click on the **Summary** tab displays summaries of data based on the total plots and data received per country.
2. Highlights the summary of the total plots and data received from each country.
3. Highlights the total plots per country on the Africa map.




1. The Summary tab cont.
2. Highlights the reason to obstruction by depth.
3. Highlights the of crop type on the visited sampling locations.
4. Highlights the main land use of the sampling locations.
5. Indicate the ratio of any sign of erosion from the sampling locations.

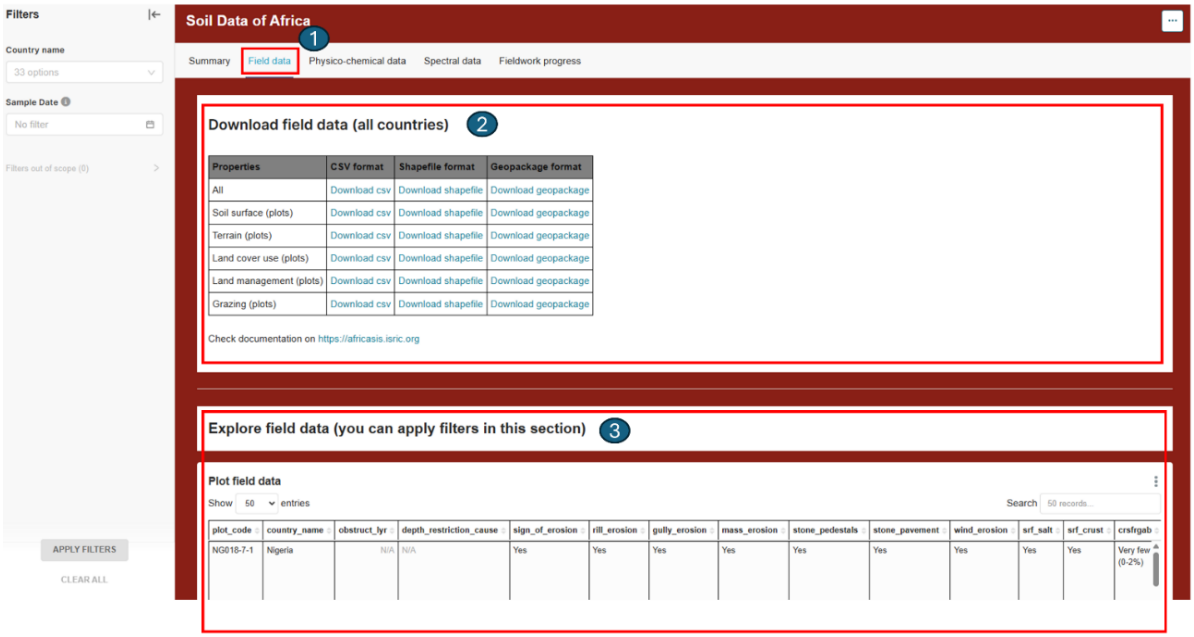


This project has received funding from the
EU Horizon 2020 Research and Innovation Programme
under Grant Agreement 862900.




## The Field DataTab

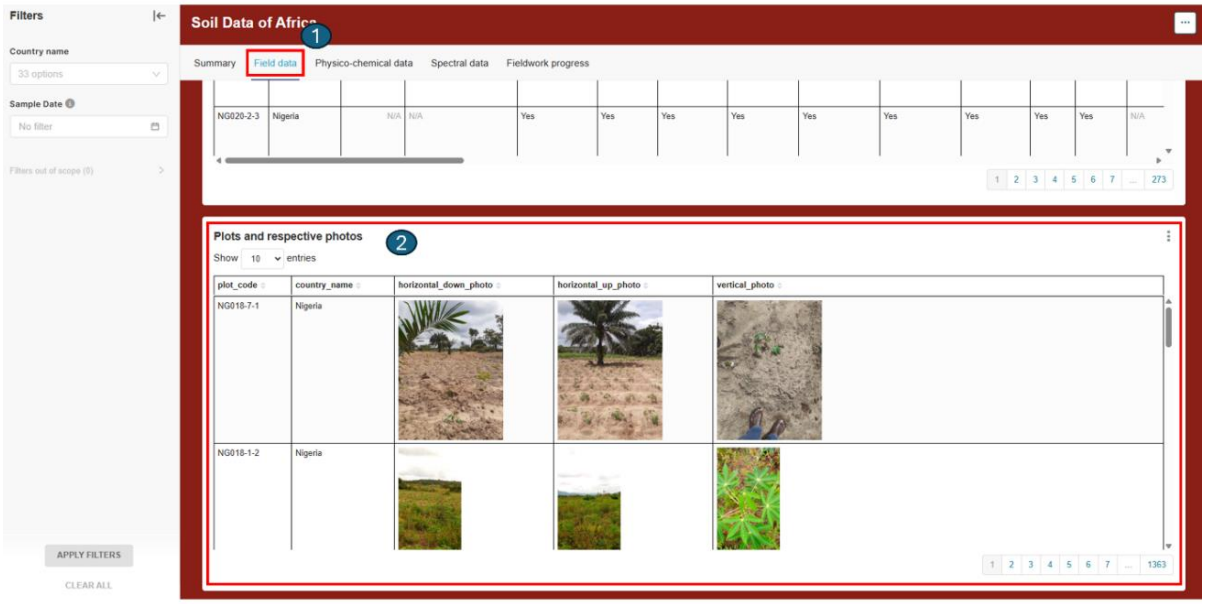
1. A click on the **Field Data** tab highlights field data that can be downloaded and explored.
2. The **Download field data (all countries)** displays the various data properties and file formats that are available for download.
3. The **Explore field data** displays a tabular presentation of the plot data from the various sampling locations.






This project has received funding from the
EU Horizon 2020 Research and Innovation Programme
under Grant Agreement 862900.

1. The **Field data** tab cont.
2. The **Plots and respective photos** section displays the field photos taken at each sampling location.

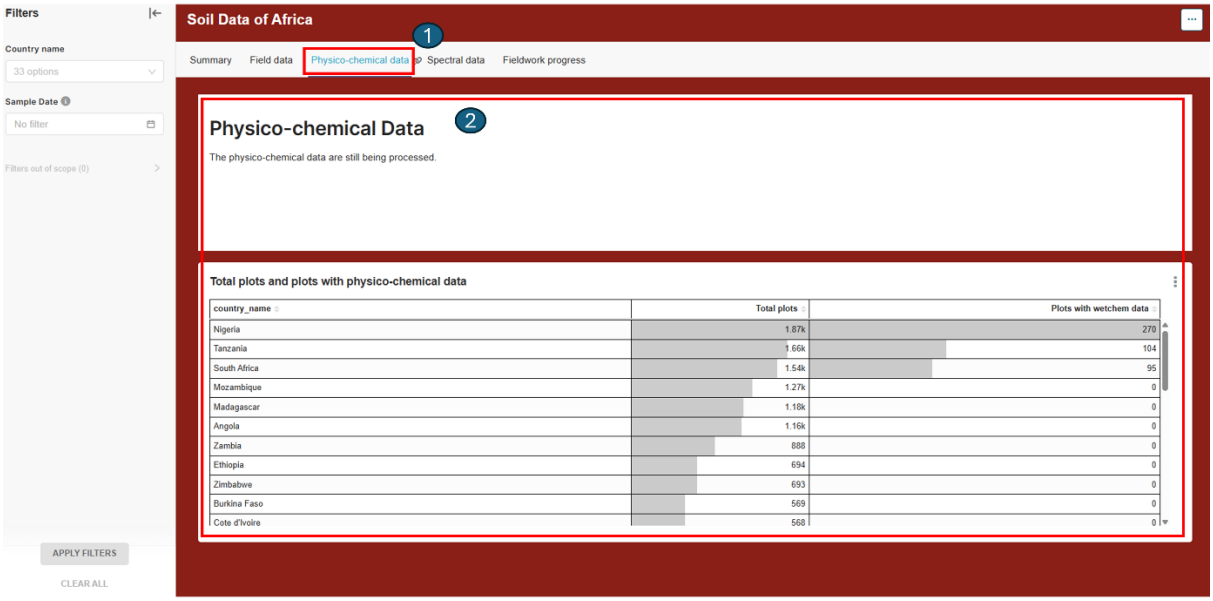



## The Physico-chemical data

1. A click on the **Physico-chemical data** highlights physico-chemical data that can be downloaded and explored.
2. The **Total plots and plots with physico-chemical data** displays the tabular distribution of data from various countries in Africa. Download of these data is expected by June 1, 2026.

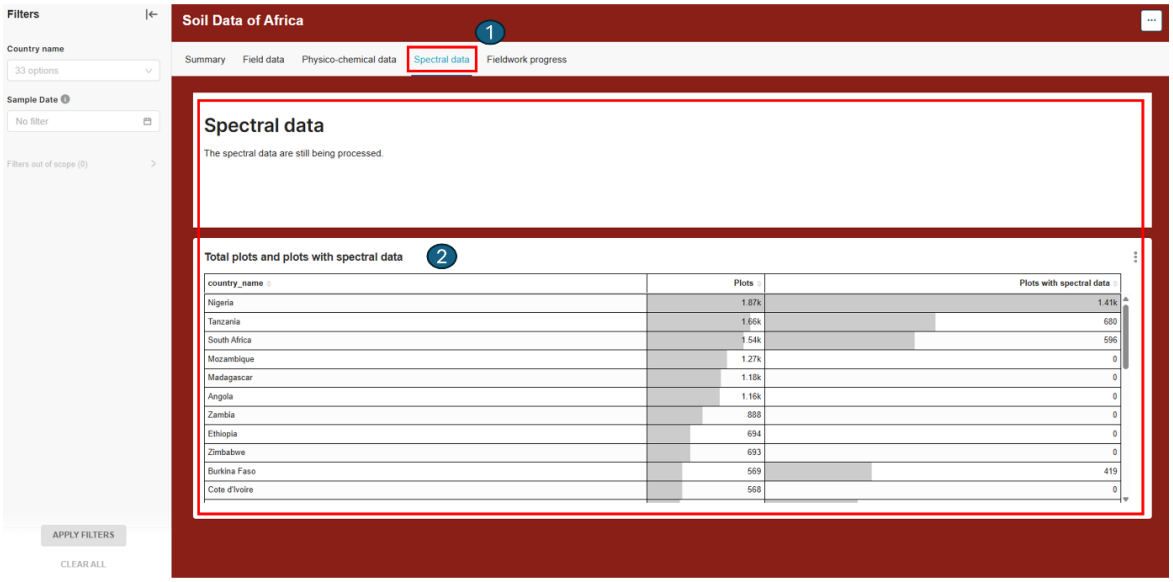


This project has received funding from the
EU Horizon 2020 Research and Innovation Programme
under Grant Agreement 862900.




## The Spectral data

1. A click on the **Spectral data** tab highlights spectral data that can be downloaded and explored.
2. The **Total plots and plots with spectral data** displays the tabular distribution of data from various countries in Africa. Download of these data is expected by June 1, 2026.



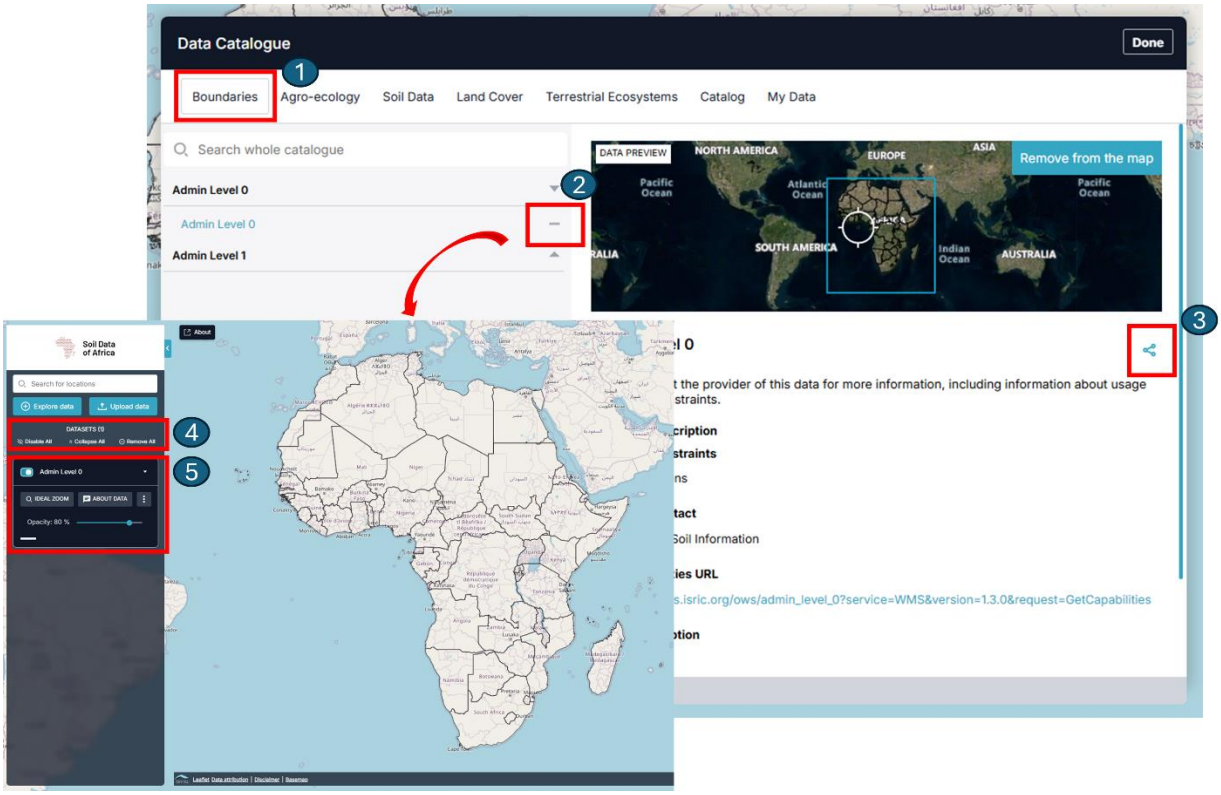



This project has received funding from the
EU Horizon 2020 Research and Innovation Programme
under Grant Agreement 862900.

## ANNEX

### Display administrative boundaries of countries

1. In the data catalogue window displaying in the main map window, Click on **Boundaries** to expand to display "Admin Level 0" in the Data Catalogue.
2. Click on the **+** symbol to ADD the administrative boundaries into the map window.
3. The **Share** button allows users to generate a URL for sharing the current map configuration.
4. Shows the **number of active datasets** in the map window, **and other actions** you can perform to the data, such as **Remove All** or **Collapse All** data.
5. **Shows Opacity and Zoom Controls:** The **IDEAL ZOOM** for Adjusting zoom settings and **OPACITY Slider** for Adjusting the transparency of data overlaps on the map.

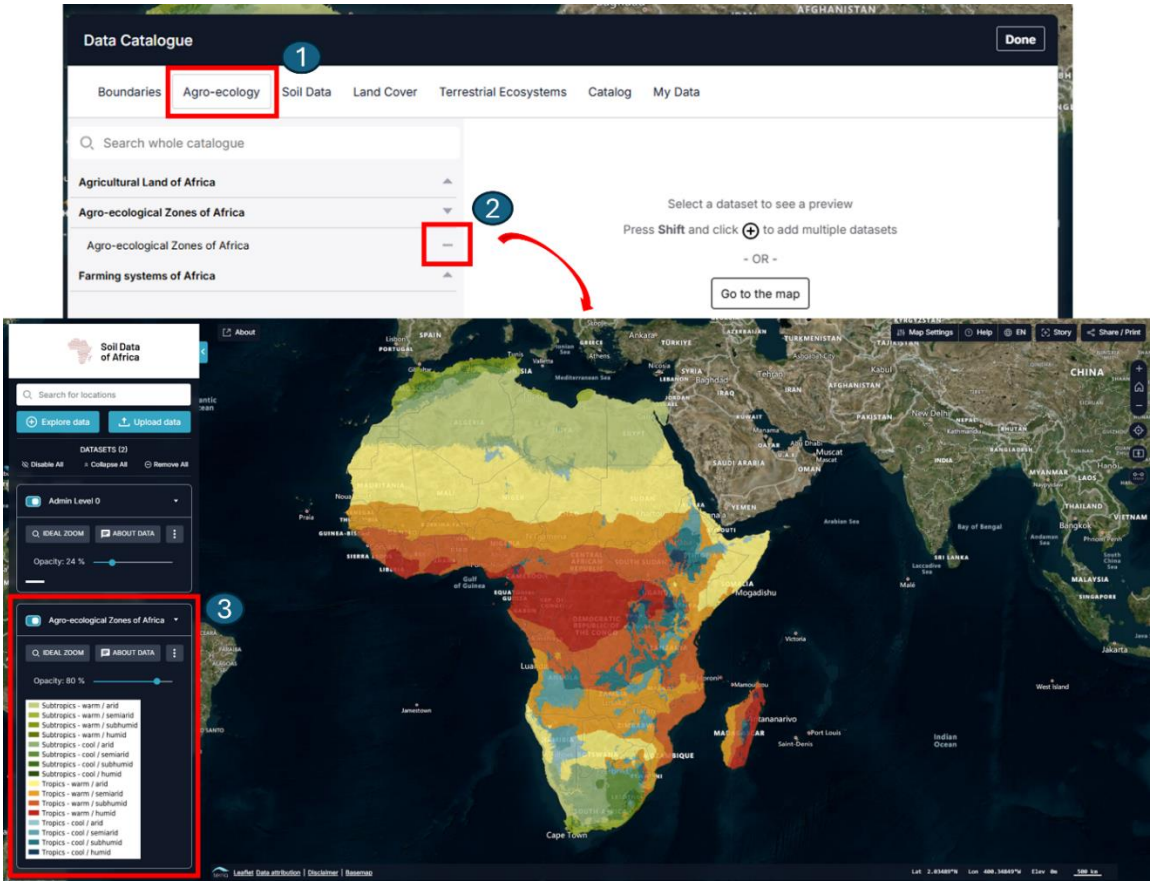





This project has received funding from the
EU Horizon 2020 Research and Innovation Programme
under Grant Agreement 862900.

## Display Agro-ecological zones of Africa

1. Click on **Agro-ecology** to expand and display "Agro-ecological Zones of Africa".
2. Click on the **-** symbol to load the **Agro-ecological Zones of Africa** into the main map window.
3. The left panel of the map window shows the **Active Datasets** that are currently active on the map.
    a. This panel is adjacent to the sidebar and displays the selected dataset "A**gro-ecological Zones of Africa**". It includes an **"Ideal Zoom"** button for optimal map viewing and an **"About Data"** button for detailed information about the dataset.
    b. The **opacity slider** allows users to adjust the transparency of the dataset overlay on the map.
    c. Below these controls, a list of different ecological zones is presented, each labeled with specific climatic characteristics such as "Subtropics warm/arid," "Tropics warm/humid," etc., indicating the diversity of ecological zones within Africa.






This project has received funding from the
EU Horizon 2020 Research and Innovation Programme
under Grant Agreement 862900.

## Display Farming systems of Africa

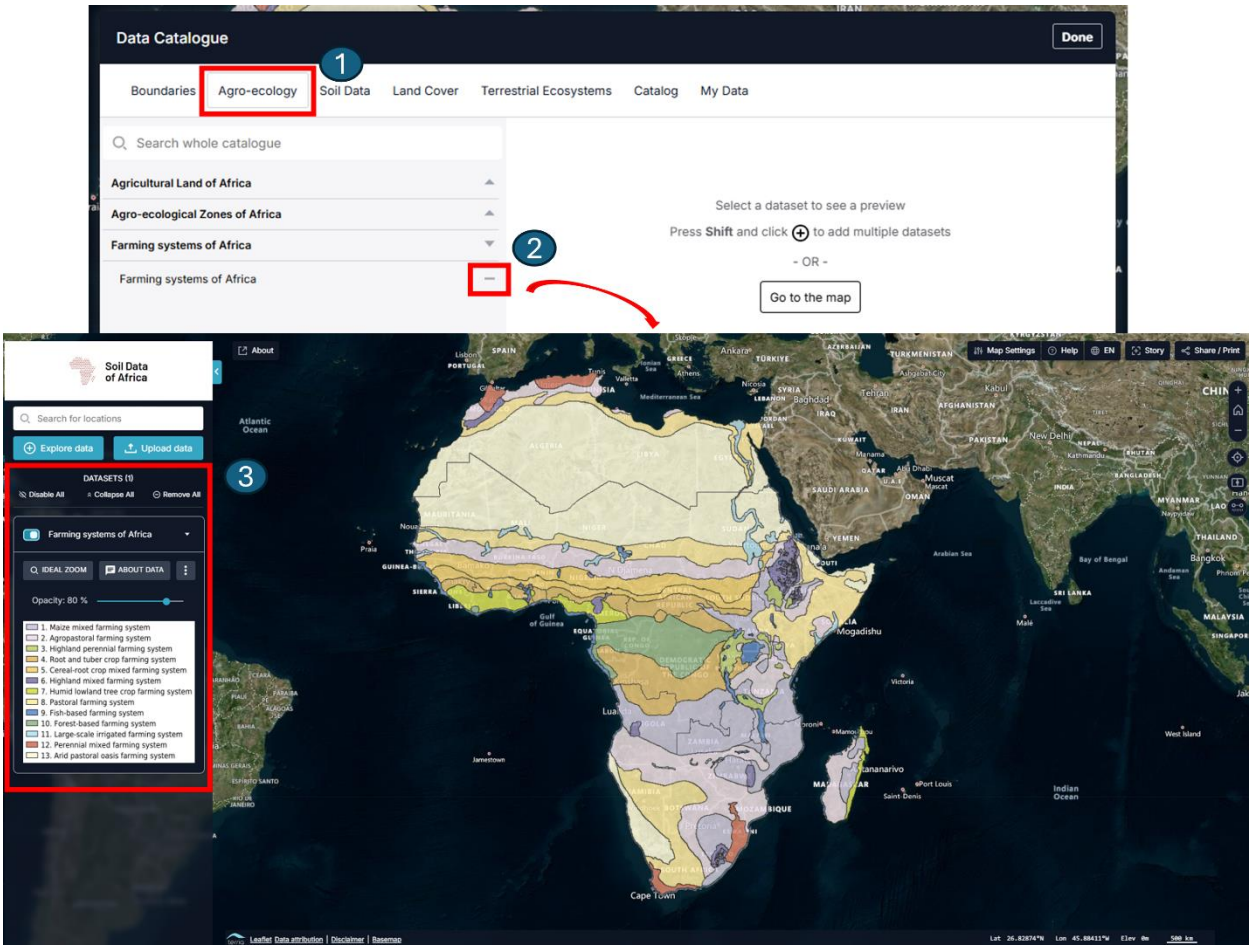
1. Click on the **Agro-ecology** to display the list of datasets under this category.
    - Click on **Farming systems of Africa** to expand and display "Farming systems of Africa".
2. Click on the **-** symbol to load the **Farming systems of Africa** into the main map window.
3. The left panel of the map window shows the **Active Datasets** that are currently active on the map.
    a. This panel is adjacent to the sidebar and displays the selected dataset "A **Farming systems of Africa**". It includes an **"Ideal Zoom"** button for optimal map viewing and an **"About Data"** button for detailed information about the dataset.
    b. The **opacity slider** allows users to adjust the transparency of the dataset overlay on the map.
    c. Below these controls, there is a list of various farming systems categorized and numbered, such as "1. Maize mixed farming system", "2. Agropastoral millet/sorghum system", etc. Each system is represented by different colours or patterns on the map.



European Commission | Horizon 2020 European Union funding for Research & Innovation

This project has received funding from the
EU Horizon 2020 Research and Innovation Programme
under Grant Agreement 862900.

SOILS 4 AFRICA



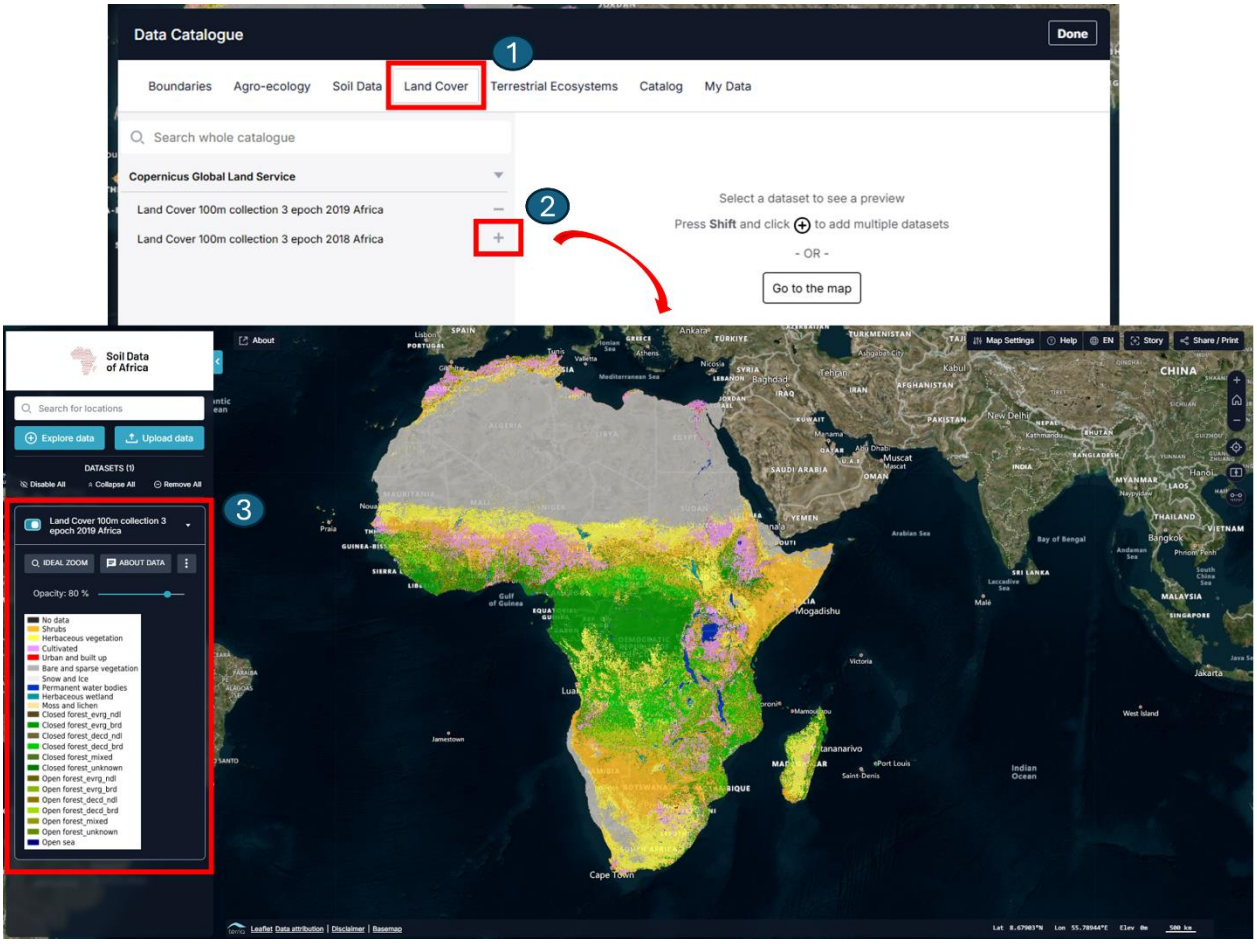

## Display Copernicus Global Land Service data

1. Click on the **Land Cover** to display the list of datasets under this category.
    - Click on **Copernicus Global Land Service** to expand to display;
        - "Land Cover 100m collection 3: epoch 2019: Africa" and,
        - "Land Cover 100m collection 3: epoch 2018: Africa".
2. Click on the **-** symbol to load the **Land Cover 100m collection 3: epoch 2019: Africa** into the main map window.
3. The left panel of the map window shows the **Active Datasets** that are currently active on the map.
    a. This panel is adjacent to the sidebar and displays the selected dataset "**A Land Cover 100m collection 3: epoch 2019: Africa**". It includes an **"Ideal Zoom"** button for optimal map viewing and an **"About Data"** button for detailed information about the dataset.



This project has received funding from the
EU Horizon 2020 Research and Innovation Programme
under Grant Agreement 862900.

b. The **opacity slider** allows users to adjust the transparency of the dataset overlay on the map.
c. Below these controls, there's a detailed list of land cover types.



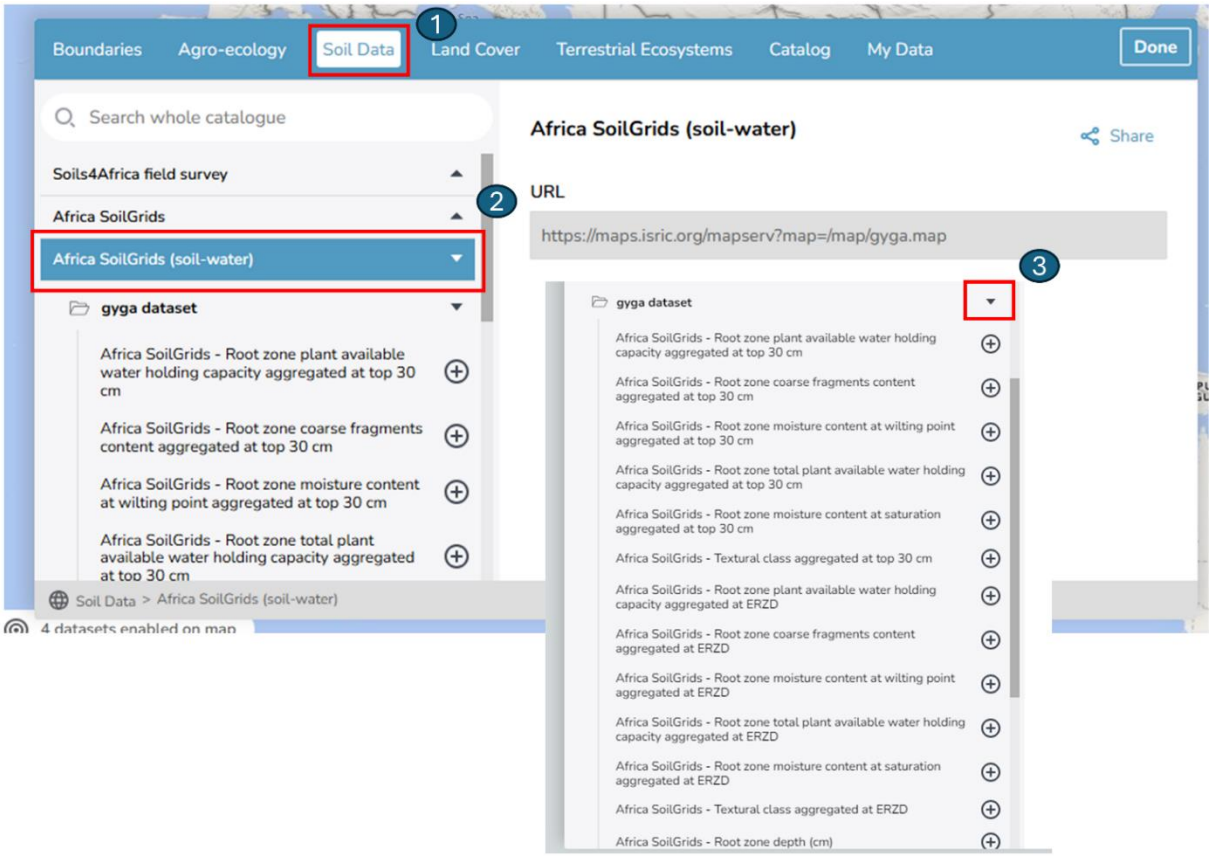

## Display Africa SoilGrids (soil-water) data

1. In the data catalogue window, click on **Soil Data** to display the list of datasets under this category.
2. Click on **Africa SoilGrids (soil-water)** to expand to display;
    - "gyga dataset"
3. Click on the **Δ** symbol to expand the **gyga dataset**; the expanded view shows various subcategories or specific data types available under the "Africa SoilGrids (soil-water)" category. Each item represents a different aspect of soil-water data, such as:
    - Root zone plant available water holding capacity aggregated at top 30 cm
    - Root zone coarse fragments content aggregated at top 30 cm



This project has received funding from the
EU Horizon 2020 Research and Innovation Programme
under Grant Agreement 862900.

- Root zone moisture content at wilting point aggregated at top 30 cm



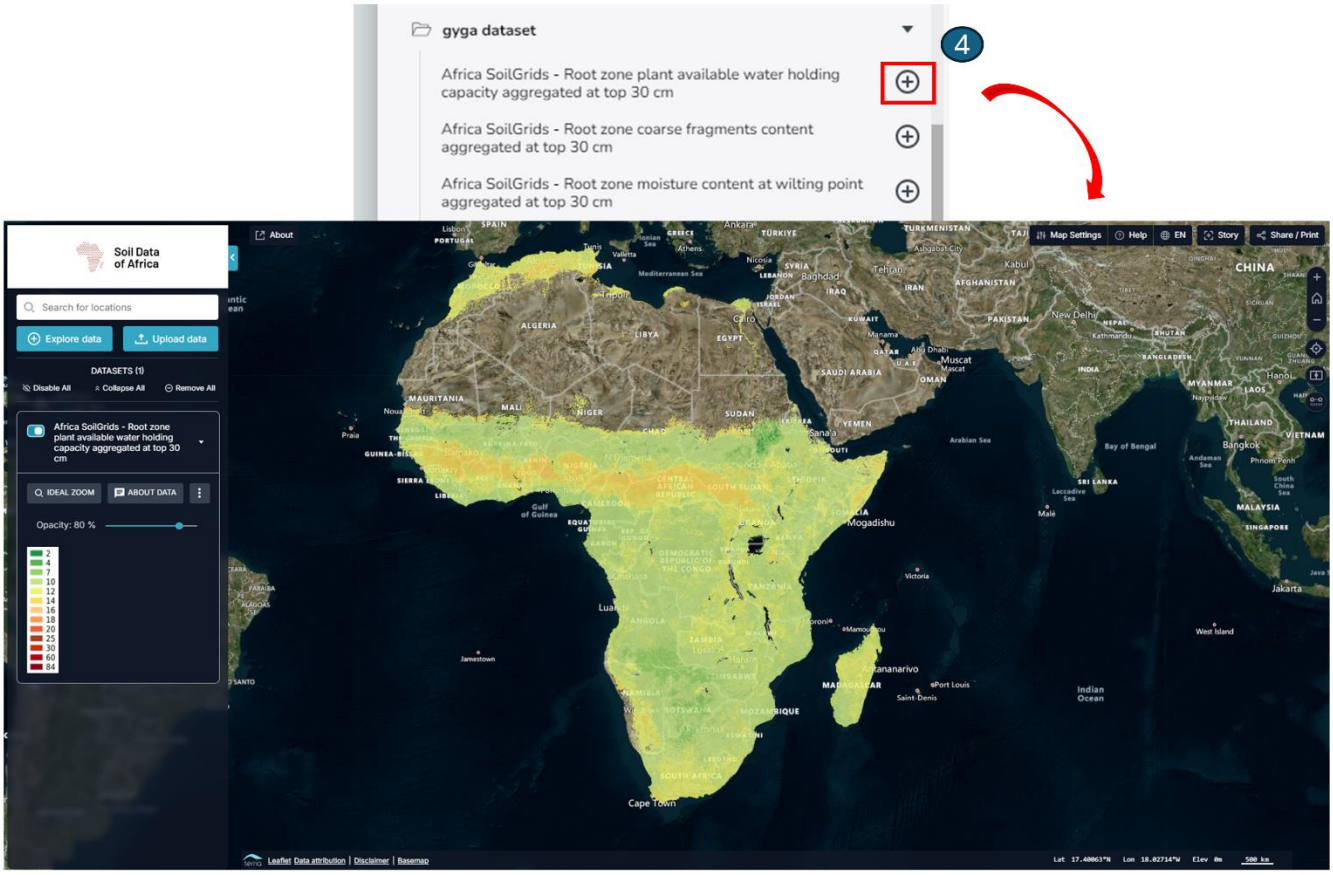

4. Click the button **(+)** to add the **"Africa SoilGrids – Root zone plant available water holding capacity aggregated at 30 cm"** layers data into the main map window



European Commission | Horizon 2020 European Union funding for Research & Innovation

This project has received funding from the EU Horizon 2020 Research and Innovation Programme under Grant Agreement 862900.

SOILS4AFRICA

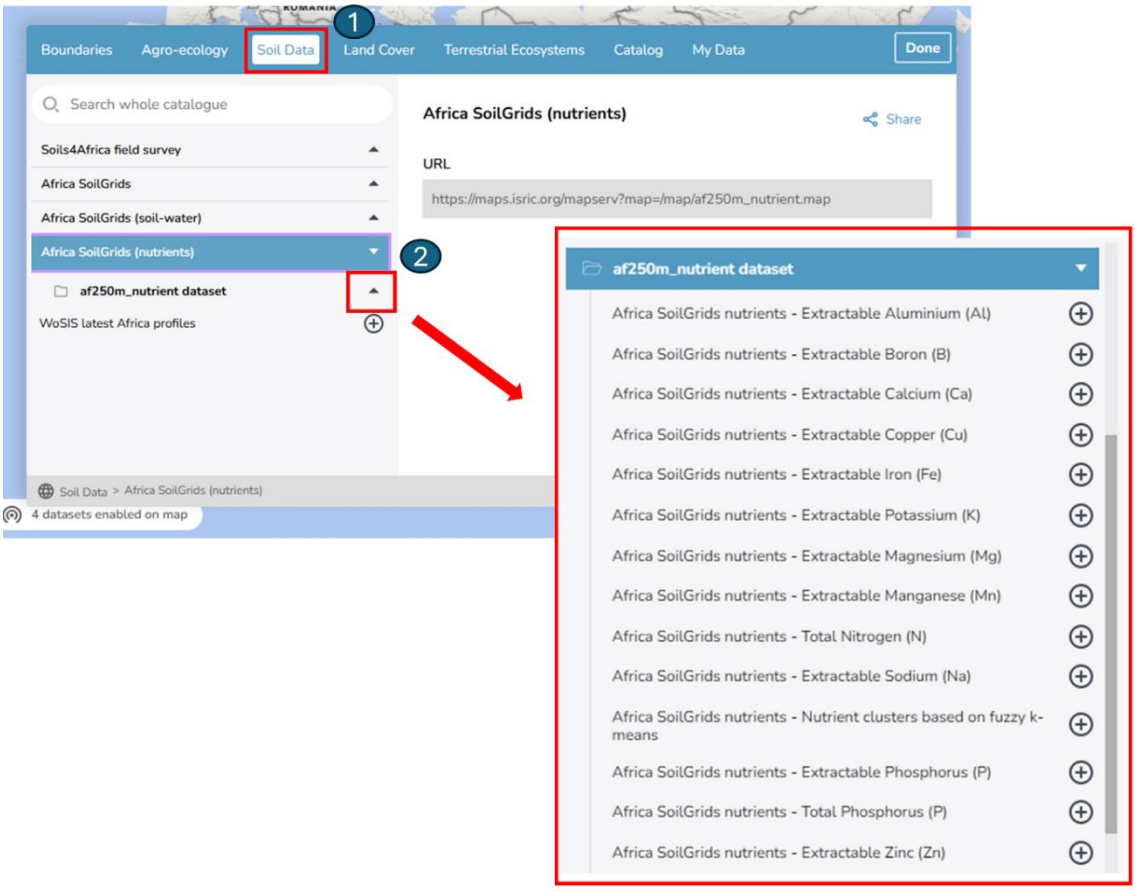





This project has received funding from the
EU Horizon 2020 Research and Innovation Programme
under Grant Agreement 862900.

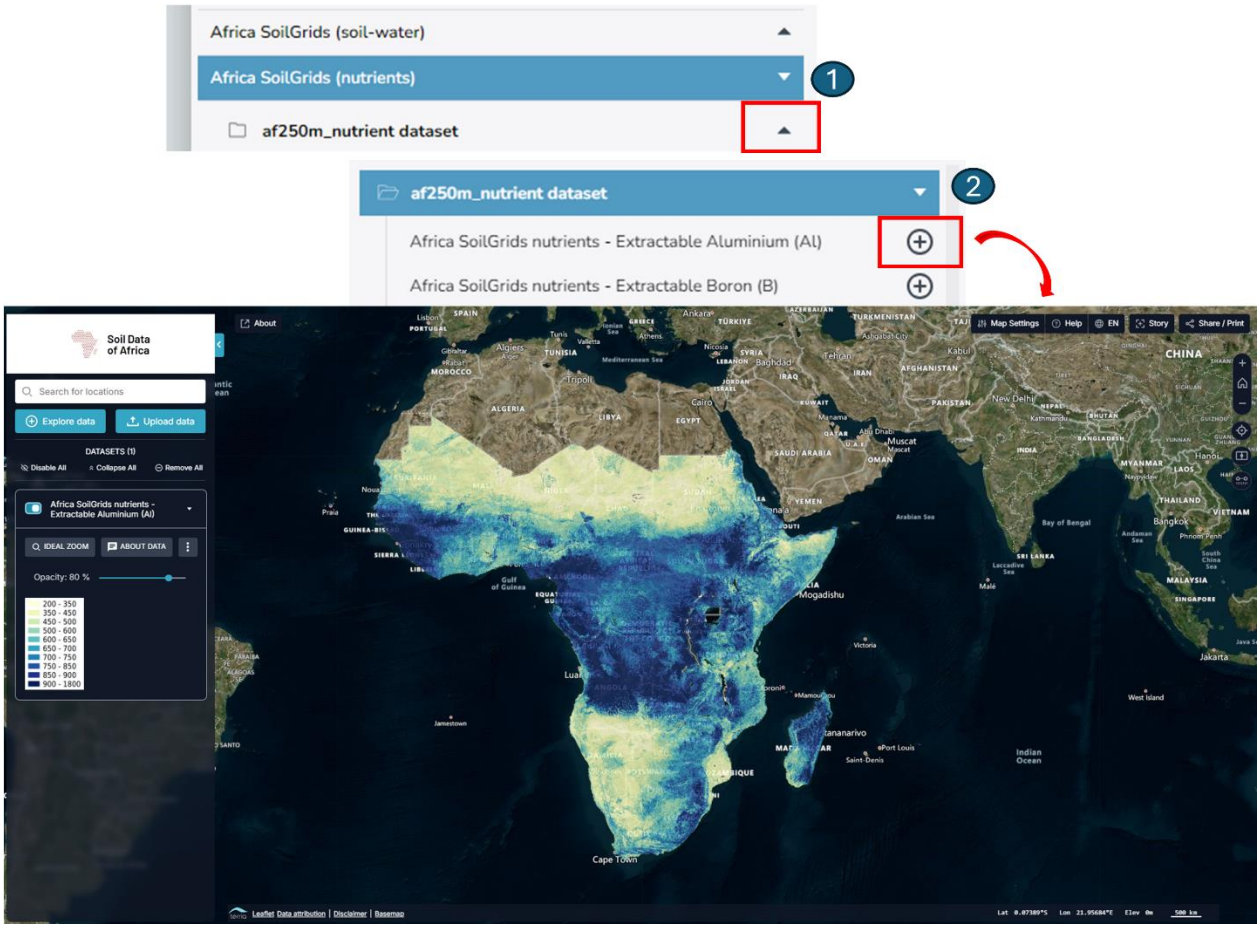
## Display Africa SoilGrids (nutrients) data

1. In the data catalogue window, click on **Soil Data** to display the list of datasets under this category.
2. Click on **Africa SoilGrids (nutrients)** to expand to display;
    - "af250m_nutrient dataset"
    - Click on the **Δ** symbol to expand the **af250m_nutrient dataset**.
    - The expanded view shows various subcategories or specific data types available. under the "Africa SoilGrids (nutrients)" category. Each item represents a different aspect of nutrient data at a resolution of 250 meters.
    - Click the button **(+)** to add these specific data layers to the main map window.






This project has received funding from the
EU Horizon 2020 Research and Innovation Programme
under Grant Agreement 862900.

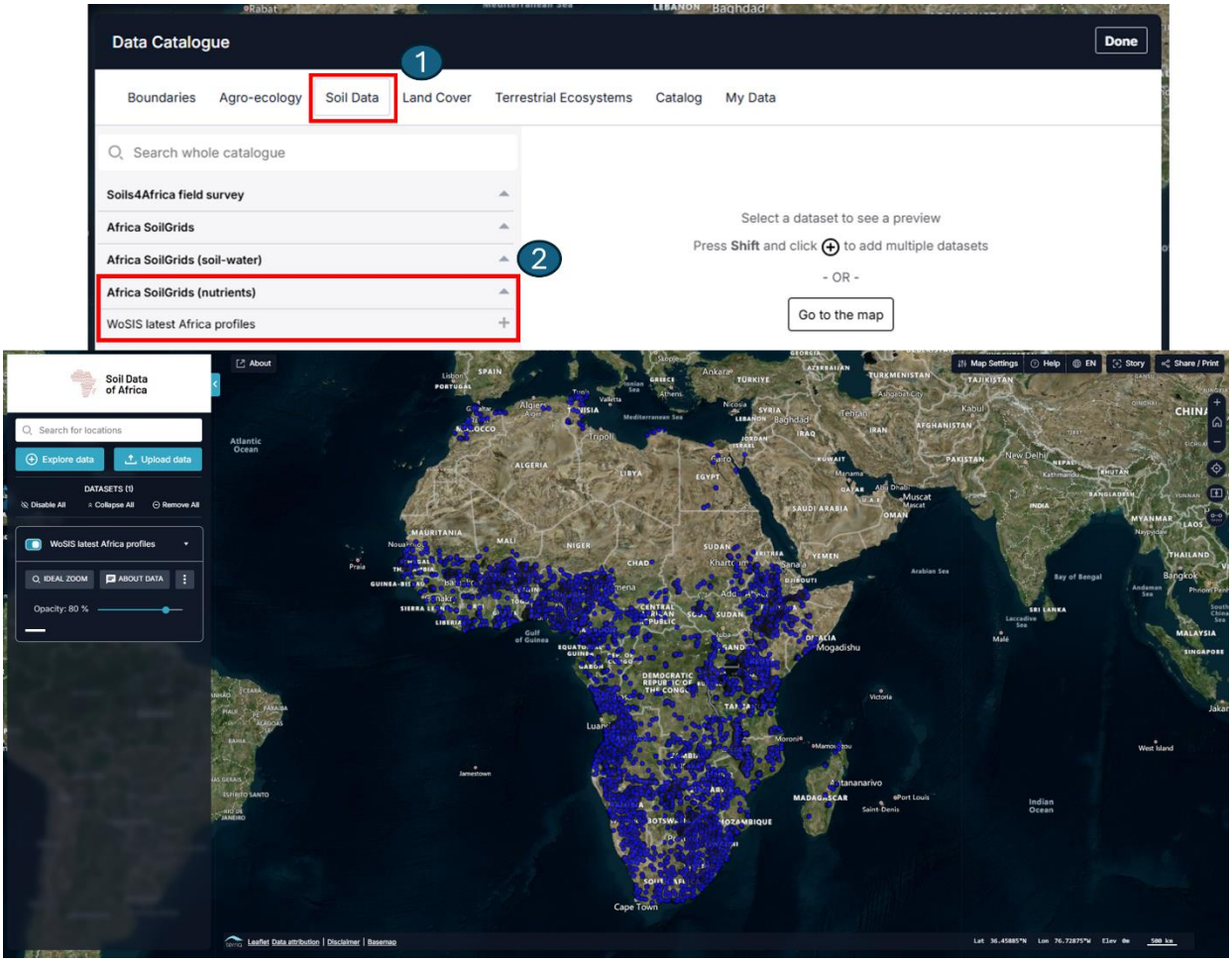





This project has received funding from the
EU Horizon 2020 Research and Innovation Programme
under Grant Agreement 862900.

## Display WoSIS latest Africa profiles data

1. In the data catalogue window, click on **Soil Data** to display the list of datasets under this category.
2. Click on the button **(+)** WoSIS latest Africa profiles to add the data to the main map window.





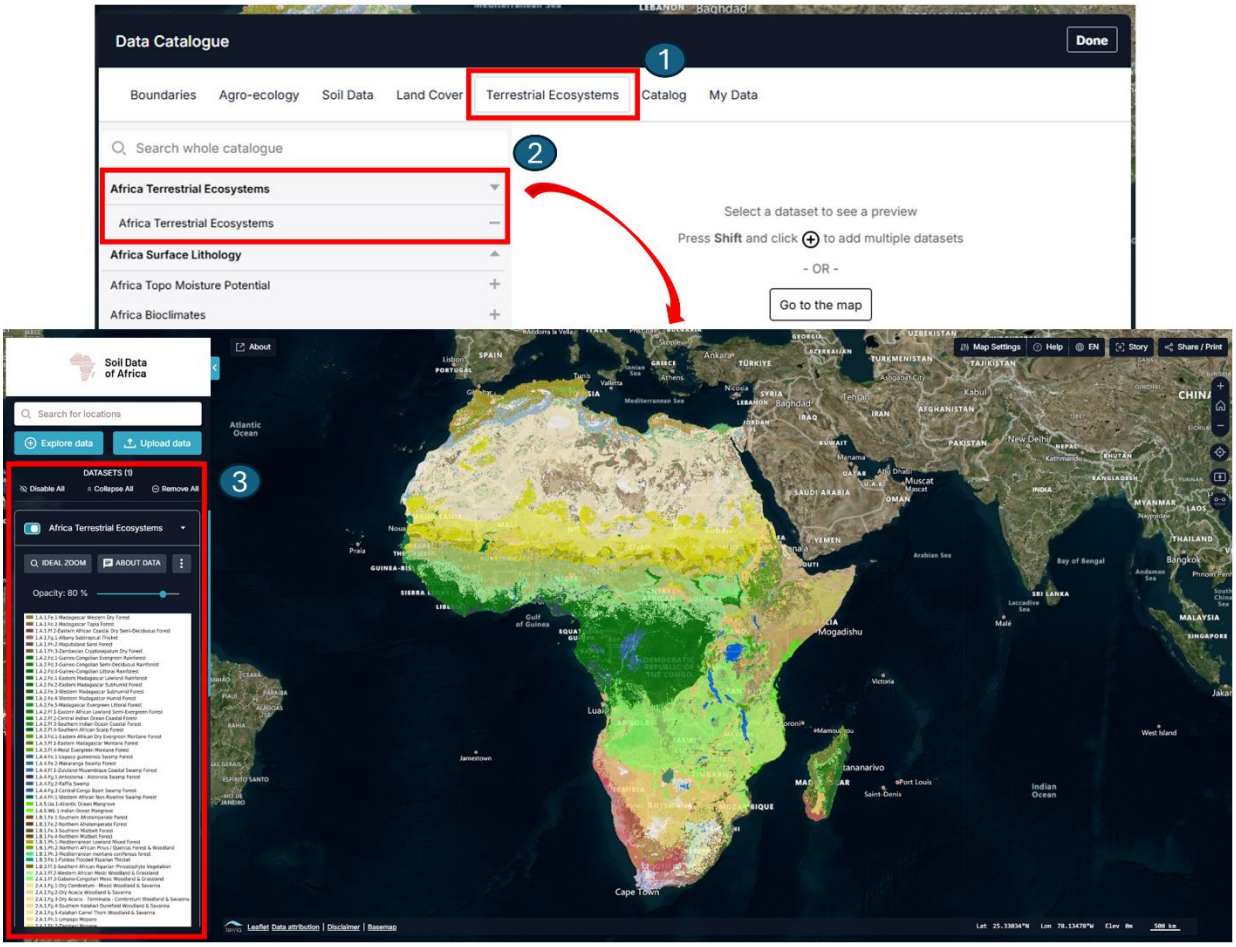

This project has received funding from the
EU Horizon 2020 Research and Innovation Programme
under Grant Agreement 862900.

## Display Africa Terrestrial Ecosystems data

1. In the data catalogue window, click on **Terrestrial Ecosystems** to display the list of datasets under this category.
2. Click on **Africa Terrestrial Ecosystems** to expand to display;
    - "Africa Terrestrial Ecosystems"
    - Click on the **+** symbol to expand the **Africa Terrestrial Ecosystems** and display on the main map window.
3. The left panel of the map window shows the **Active Datasets** that are currently active on the map.
    a. This panel is adjacent to the sidebar and displays the selected dataset "**Africa Terrestrial Ecosystems"**. It includes an **"Ideal Zoom"** button for optimal map viewing and an **"About Data"** button for detailed information about the dataset.
    b. The **opacity slider** allows users to adjust the transparency of the dataset overlay on the map.
    c. Below these controls, there's a detailed list of various ecosystems across Africa displayed.



This project has received funding from the
EU Horizon 2020 Research and Innovation Programme
under Grant Agreement 862900.






This project has received funding from the
EU Horizon 2020 Research and Innovation Programme
under Grant Agreement 862900.

## Display Africa Biomes data

1. In the data catalogue window, click on **Terrestrial Ecosystems** to display the list of datasets under this category.

2. Click on **Africa Biomes** to expand to display;

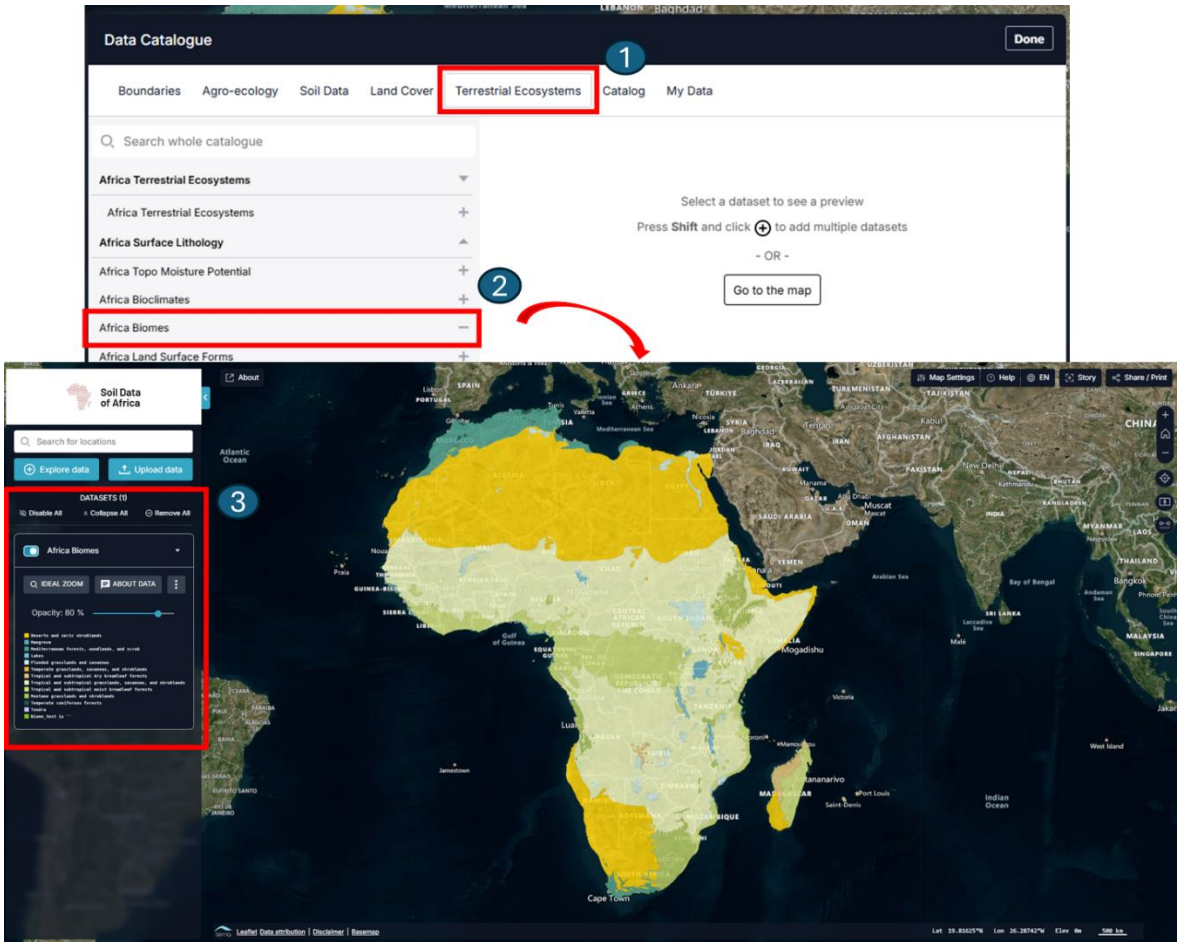
   - Click on the **+** symbol to expand the **Africa Biomes data** and display it in the main map window.

3. The left panel of the map window shows the **Active Datasets** that are currently active on the map.

   a. This panel is adjacent to the sidebar and displays the selected dataset "**Africa Biomes**". It includes an **"Ideal Zoom"** button for optimal map viewing and an **"About Data"** button for detailed information about the dataset.

   b. The **opacity slider** allows users to adjust the transparency of the dataset overlay on the map.

   c. Below these controls, there's a detailed list of various Africa biomes displayed in the legend.






European Commission | Horizon 2020 European Union funding for Research & Innovation

This project has received funding from the EU Horizon 2020 Research and Innovation Programme under Grant Agreement 862900.

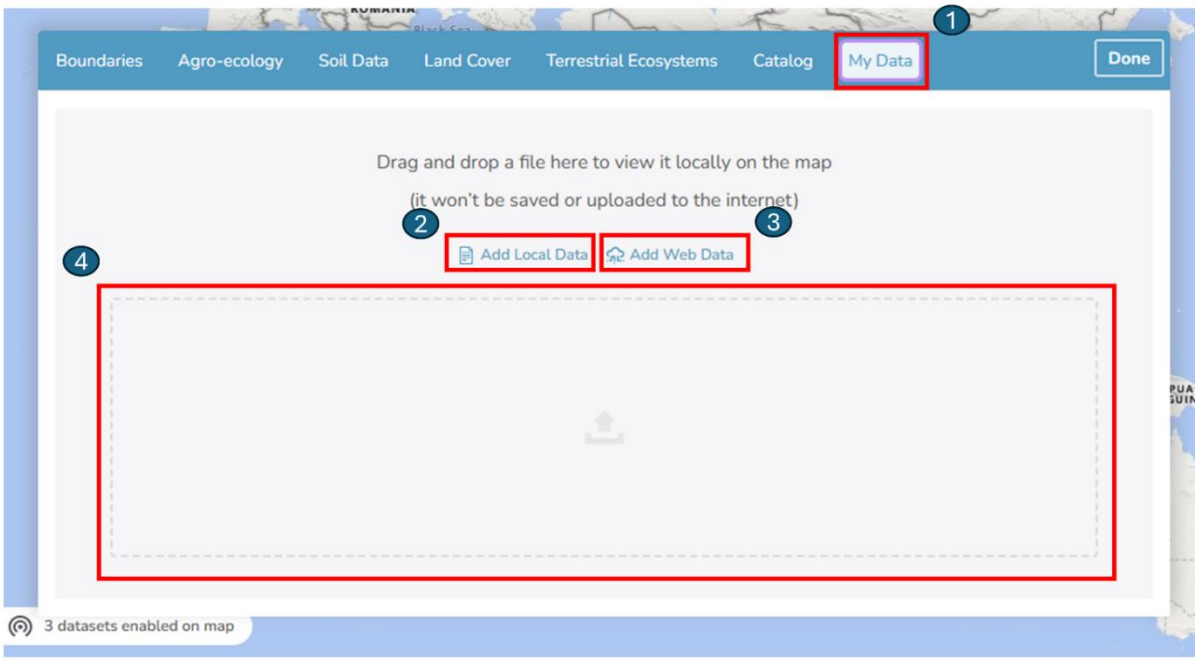
SOILS 4 AFRICA

## My Data

1. The **"My Data"** section of this platform is designed for users to upload and manage custom datasets.

**Data Upload Options** have been categorised under;

- **Add Local Data (2):** Allows users to upload files stored on their local device.
- **Add Web Data (3):** Enables users to input or link datasets from an online source.
- **Upload Area (4):** Where files can be dragged and dropped to be loaded into the map window.



# PART B

# Soil Information System Installation Manual

**Project No. 869200**

31 May 2025

**Deliverable: Soils4Africa_D6.4b**

Version 1.0



European Commission | Horizon 2020 European Union funding for Research & Innovation

This project has received funding from the
EU Horizon 2020 Research and Innovation Programme
under Grant Agreement 862900.

SOILS4AFRICA

## Contact details

**Director of Coordinating Institute – ISRIC:** Rik van den Bosch

**Project Coordinator:** Mary Steverink-Mosugu

**Address:** Droevendaalsesteeg 3, 6708 PB Wageningen (Building 101), The Netherlands

**Postal:** PO Box 353, 6700 AJ Wageningen, The Netherlands

**Phone:** +31 317 48 7634

**Email:** mary.steverink-mosugu@isric.org

## Project details

| | |
|---|---|
| Project number | 862900 |
| Project acronym | Soils4Africa |
| Project name | Soil Information System for Africa |
| Starting date | 01/06/2020 |
| Duration In months | 60 |
| Call (part) identifier | H2020-SFS-2019-2 |
| Topic | SFS-35-2019-2020 Sustainable Intensification in Africa |

## Document details

| | |
|---|---|
| Work Package | WP6 |
| Deliverable number | D6.4b |
| Version | V1.0 |
| Filename | Soils4Africa_D6.4b_SIS_installation_manual.pdf |
| Type of deliverable | Report |
| Dissemination level | Public |
| Lead partners | ISRIC |
| Contributing partners | ISRIC |
| Author | Ulan Turdukulov, Paul van Genuchten, Luis Calisto, Islambek Urazov, Bas Kempen |
| Contributors | - |
| Due date | 30/11/2024 |
| Submission date | 31/05/2025 |





European Commission | Horizon 2020 European Union funding for Research & Innovation

This project has received funding from the EU Horizon 2020 Research and Innovation Programme under Grant Agreement 862900.

SOILS 4 AFRICA

# Table of Contents





European Commission | Horizon 2020 European Union funding for Research & Innovation

This project has received funding from the
EU Horizon 2020 Research and Innovation Programme
under Grant Agreement 862900.

SOILS4AFRICA

This project has received funding from the
EU Horizon 2020 Research and Innovation Programme
under Grant Agreement 862900.

# Abbreviations

| | |
|---|---|
| API | Application Programming Interface |
| ARC | Agricultural Research Council, South Africa |
| CSW | Catalogue Service for the Web |
| CI/CD | Continuous Integration/Continuous Development |
| DBMS | Database Management System |
| EC | European Commission |
| ETL | Extract-Transform and Load workflow |
| EU | European Union |
| GSP | Global Soil Partnership |
| IBEC | Interbalkan Environment Center |
| ICRAF | International Centre for Research in Agroforesty |
| IITA | International Institute of Tropical Agriculture |
| IaC | Infrastructure-as-Code(IaC) process |
| ISRIC | International Soil Reference and Information Centre |
| LTR | Long-Term Release software |
| NAS | Network-Attached Storage |
| NFS | Network File System |
| OGC | Open Geospatial Consortium |
| PV | Persistent Volume |
| SIS | Soil Information System |
| WUR | Wageningen University & Research |
| WP | Work Package |



This project has received funding from the
EU Horizon 2020 Research and Innovation Programme
under Grant Agreement 862900.

# Introduction

The Soils4Africa[1] project (2020-2025) aimed to provide an open-access **Soil Information System** (**SIS**) hosting a set of key soil quality indicators on agricultural land use across Africa. These indicators are based on field data collected from sampling sites across the continent, according to a methodology suited for repeated soil monitoring. If the monitoring is repeated and similar procedures are used that will make the results comparable in the future.

This aim has been achieved and currently, all the soil data collected and generated during the project time will be made available to the users via the SIS portal (https://africasis.isric.org/). The SIS requirements were translated into a technical design that identified the software and hardware components needed to make the SIS operational. The SIS will eventually be hosted and maintained by an African institution to be identified during the course of the project, so they can benefit from the data, procedures and tools developed in the project.

The goal of this report is to provide step by step guide on how to setup the SIS required for handover of the SIS to an African institution that will be hosting it in the future.

The report is organised as follows: the next chapter outlines the key requirements for deploying the continental SIS developed under the Soils4Africa project. It starts with a description of three interlinked systems that must be configured before deploying the software components. The first chapter also explores how these systems can be set up using either cloud hosting or on-premise installation, with open-source software provided for each approach. The subsequent chapters examine each of the three systems in detail, covering: software components and their functionality, deployment requirements and offering installation options through use of containerization tools. The final chapter provides an overview of supporting services, including monitoring and APIs, followed by a summary of the report's key findings.

[1] https://www.soils4africa-h2020.eu/the-project



This project has received funding from the
EU Horizon 2020 Research and Innovation Programme
under Grant Agreement 862900.

# Prerequisites

The architecture of the continental Soil Information System was outlined in the project deliverable 6.1 (Turdukulov et al., 2021). That report states that SIS deployment and web services prioritize the use of open-source software solutions and Long-Term Release (LTR) versions of all software. Additionally, the report identifies the database as an important infrastructure component, serving as the central hub for storing and processing soil data collected from field observations and laboratory measurements. Given the continental scale of the field campaign and the diverse range of project partners, the database model had to be integrated with a version control system, such as Git, to ensure reliability and support for the automated processing and data ingestion workflows.

The guiding principle for deployment of the S4A Soil Information System was that the SIS components will be containerized and managed through an Infrastructure-as-Code (IaC) process. IaC involves managing infrastructure through code, which can be versioned using any version control system like Git, ensuring reliability and better continuous delivery practices.

Therefore main prerequisites for the deployment of S4A are the availability and linkages of the following three main systems:

1. Database Management System,
2. Code versioning control system,
3. And a container orchestration system for automating software deployment, scaling, and management

Code versioning control system is an essential element of the infrastructure that seamlessly integrated with both a database and a container orchestration platform. Together these three systems enable management of data processing workflows, computer infrastructure using code rather than manually configuring hardware or software. By leveraging these three components, SIS achieves greater efficiency, reliability and flexibility, making it an ideal choice for the IaC based practice on which the SIS is based.

## On-cloud versus on-premise installation

Before offering the installation instruction, we need to distinguish between two types of hosting models of the SIS: cloud hosting and on-premise installation.

Cloud hosting is the process of outsourcing an organization's computing, hosting and storage resources to a service provider that offers its infrastructure services in a utility model. The cloud provider administers the setup, cloud infrastructure, security and maintenance, while allowing clients to customize hardware and applications and scale servers online. All major cloud providers (Google, Azure, Amazon) have the ability to configure the above mentioned three systems on-cloud.

The traditional on-premise installation model requires organizations to invest into hardware, build and manage their own systems and data centres. In the traditional model, servers and storage, with dedicated hardware and virtual resources, reside on premises in the SIS hosting institute.

Various factors will affect the choice of on-cloud or on-premise infrastructure. One of the cost benefits of moving to the cloud is that one does not need to make sizeable initial hardware investments and to keep costs of maintaining IT infrastructure. On other hand, costs of on-cloud services, if not appropriately managed, may add up. Additionally, a SIS hosting institute may have regulations that do not allow cloud hosting of national/continental scale soil data. Alternatively, they may be part of bigger institutes/national infrastructures that already run their own data centres.

Table 1 below gives an overview of on-cloud and on-premise for the three main prerequisites of continental SIS.



European Commission | Horizon 2020 European Union funding for Research & Innovation

This project has received funding from the
EU Horizon 2020 Research and Innovation Programme
under Grant Agreement 862900.

SOILS4AFRICA

| | On-cloud alternative | On-premise open source alternative |
|---|---|---|
| *Database* | All major cloud providers offer database hosting service | PostgreSQL (version 13 or higher) |
| *Code versioning system* | GitHub, BitBucket | GitLab (version 17.8 or higher) |
| *Container orchestration system* | All major cloud providers offer container orchestration system hosted on their platform (i.e, Google and Azure Kubernetes Services) | Kubernetes, OpenShift, Docker Swarm |

*Table 1 On-cloud and on-premise alternatives for three main systems*

Since the cloud hosting is easy to configure, we will proceed with the assumption that the setup is based on on-premise installation. We presume that these systems will be installed and managed by the SIS hosting institute.

## On-premise installation and hardware requirements

Below, we provide general instructions and hardware requirements for setting up the three specified systems: a database management system, a code versioning control system, and a container orchestration system.

### Database Management System

As part of the current infrastructure we use PostgreSQL (version 13 or higher) for storing soil datasets. PostgreSQL is a free and open-source Relational Database Management System (RDBMS) emphasizing extensibility and SQL compliance. PostgreSQL is designed to handle a range of workloads, from single machines to data warehouses or Web services with many concurrent users. It has automatically updatable views, materialized views, triggers, foreign keys, and stored procedures (The PostgreSQL Global Development Group, 2021). With PostGIS extension, PostgreSQL supports a wide variety of geographic information system (GIS) data types. Support for geospatial data types is especially important in S4A project since soil data are inherently spatial. GIS data type support are essential for ETL and in other parts of DSM workflows.

Hardware requirements for dedicated PostgreSQL Server (version 13 or higher) as follows:

- Processor: Minimum 2 GHz.
- RAM: At least 8 GB.
- Disk Space: 1GB for installation, with additional space required for data (250 GB).

These are currently used requirements and that may vary based on the future use. For more detailed installation information, refer to the PostgreSQL documentation.

The Soil Profile Database is hosted within the IT infrastructure of Wageningen University & Research (WUR). It consists of multiple SQL scripts managed using the Graphile Migrate tool, available at GitHub. This approach allows the S4A data model to be exported and redeployed on any compatible PostgreSQL database server.

To streamline deployment, a Docker image for PostgreSQL with the PostGIS extension is available at Docker Hub. Databases are automatically populated by applications such as Apache Superset and pycsw, or through configuration scripts, which can be accessed at WUR GitLab. These scripts are built using Graphile Migrate, an open-source, cross-platform JavaScript runtime environment. General steps for setting up and configuring database are:

1. Copy .env.template into .env and make sure you edit the correct parameters according to your needs: DATABASE_URL ROOT_DATABASE_URL



European Commission | Horizon 2020 European Union funding for Research & Innovation

This project has received funding from the EU Horizon 2020 Research and Innovation Programme under Grant Agreement 862900.

SOILS 4 AFRICA

2. Make sure you perform URI Encode for **special characters** you may have in .env specially in your connection strings user password etc.
3. Make sure you have Node.js installed on your system and run corepack plugin enable from the terminal.
4. From inside the terminal, go to this repo root folder yarn install
5. After the previous step run: `yarn run graphile-migrate reset –erase`
6. You should now have a new database ready to go.

### Code versioning control system

A code versioning control system plays a central role in the SIS infrastructure. It is a linkage that seamlessly integrated with both a database (PostgreSQL) and a container orchestration platform (Kubernetes, see the next section). It is used to store code fragments, configuration scripts, datasets and trigger workflow and pipelines which run on changes of repositories.

Soils4Africa uses GitLab version 17.8 installed and managed by Wageningen University and Research. GitLab is an Open Source code repository and collaborative software development platform for large DevOps projects, such as the SIS. In GitLab, the SIS project is built around a number of GitLab repositories. The hardware requirements for installing GitLab depend on the expected workload and the number of users. Here are the general guidelines:

GitLab 18 server requires:

- Processor: 8 CPU for up to 1,000 users or 20 requests per second.
- RAM: Minimum 8 GB, recommended 16 GB.
- Disk Space: 2.5 GB for installation, with additional space depending on repository size.

These requirements may vary based on the specific setting of the SIS hosting institute and possible other use cases. For more detailed information, please refer the GitLab documentation. For SIS infrastructure, we recommend implementing two environments, a staging environment and a production environment. For every repository in GitLab, a "**`main`**" branch includes code targeting the production environment and a "**`dev`**" or "**`test`**" branches targeting the staging environment. The current infrastructure also maintains several repositories on a GitHub code hosting platform at https://github.com/soil-data-of-africa. GitHub allows sharing code repositories, and collaboration with the wider soil research community.

### Container orchestration platform

As mentioned earlier, one of the requirements is to use a container orchestration system to automate deployment, scaling, and management of containerized (Docker) applications. This ensures efficiency, reliability, and consistency across different environments. Kubernetes is a portable, extensible, open source platform for managing containerized workloads and services, that facilitates both declarative configuration and automation (The Kubernetes Authors, 2025). There are several alternatives to Kubernetes for container orchestration and management. Here are some notable ones:

- Docker Swarm: An open-source platform that allows you to manage a cluster of Docker engines as a single virtual system.
- Google and Azure Kubernetes Services: A managed container orchestration service based on Kubernetes, which simplifies deploying a managed Kubernetes cluster in Google and Azure respectively.
- OpenShift: A Kubernetes distribution from Red Hat that provides developer and operational tools on top of Kubernetes.



European Commission | Horizon 2020 European Union funding for Research & Innovation

This project has received funding from the EU Horizon 2020 Research and Innovation Programme under Grant Agreement 862900.

SOILS 4 AFRICA

The hardware requirements for installing Kubernetes can vary depending on the size and purpose of the cluster. Here are some general guidelines:

- Kubernetes cluster with at least 5 nodes, at least 8GB RAM for the master node and 4GB+ per worker node is optimal (24 GB minimum). A 120GB storage is sufficient for a small Kubernetes test cluster. A multi-core CPU, and Gigabit Ethernet for smooth networking is desirable.
- Storage provisioner for Kubernetes persistent volume (PV) storage management. Kubernetes supports multiple storage types, including local storage, network-attached storage (NAS) or Network File System (NFS) storage, and cloud storage platforms, allowing flexible storage options.

A well-configured Kubernetes installation involves multiple physical or virtual machines and necessitates proper networking and node setup to ensure seamless communication between all containers within the cluster. **Setting up such an environment requires advanced expertise from the IT and networking professionals of the SIS hosting institute**. As this falls beyond the scope of this document, we recommend that interested readers refer to the official Kubernetes installation documentation for further details. Note, we propose setting up such container orchestration platform only if the expertise to setup the system is available or if the system is already in place in the SIS hoisting institute. However, in cases where no local expertise or system is available, we recommend opting for the cloud hosting solution. We do not advise establishing a cloud hosting platform as part of this project.

Just like with GitLab, we recommend implementing two environments for each deployment in Kubernetes: a staging environment and a production environment. Changes can first be tested on a staging before being deployed on a production environment.

The Wageningen University IT infrastructure team managed the installation of Kubernetes version 1.30. At the heart of the SIS are Docker container images, which house the various SIS components and are deployed within the Kubernetes environment. In the following section, we will take a detailed look at how these components are deployed.



This project has received funding from the
EU Horizon 2020 Research and Innovation Programme
under Grant Agreement 862900.

# System overview

*Figure 1* shows the graphical summary of three interconnected systems via which most of the SIS components are deployed:

- **`PostgreSQL database`** of soil monitoring results: a single place to store data from the field observations and laboratory and spectral measurements. Once data are analysed and merged on a Git repository, it triggers the Extract-Transform and Load (ETL) workflow that inserts results in the database and alerts of possible mismatches.
- **`GitLab`** repositories for storing codebases and initiating workflows. An update of the repository will trigger a workflow, which deploys a new version in the database and/or deployment in the Kubernetes platform. Gitlab also contains a container registry facility, in which images produced by the workflows are stored before being deployed.
- **`Kubernetes`** platform where all deployments take place upon the triggers from the GitLab repository. All of the described SIS components that are presently deployed in Kubernetes clusters will be delivered as containers upon the completion of the project.

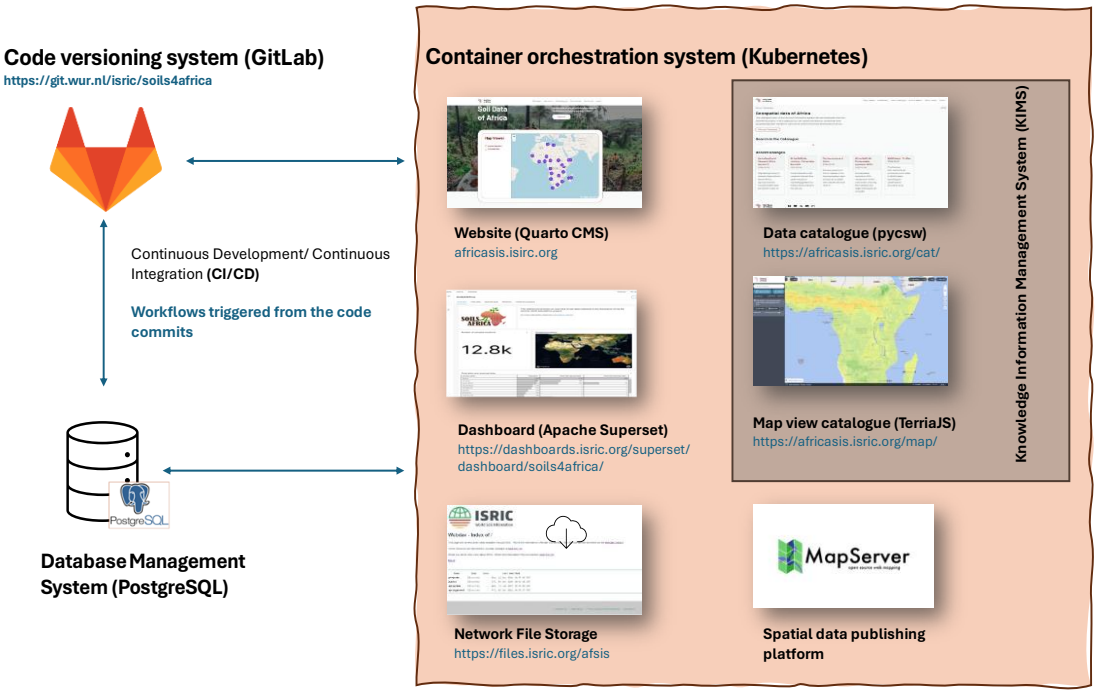



*Figure 1 Three interconnected systems that serve as the foundation of SIS infrastructure.*

The key software components of the SIS (i.e., Docker container images) in Kubernetes platform are:

- Container with the project website as a simple Content Management System (CMS) that serves as entry point to the catalogue, mapviewer, dashboards. S4A uses Git based open source Quarto framework as CMS.
- Knowledge Information Management System (KIMS) that consists of:
  - A containerised map based metadata viewer based on open source TerriaJS. TerriaJS is a client side library providing a GIS functionality in a web browser published by Mapserver geospatial services.



This project has received funding from the
EU Horizon 2020 Research and Innovation Programme
under Grant Agreement 862900.

    - A containerised catalogue based metadata viewers based on open source **pycsw** server that allows view of the uploaded soil data and based on these datasets (including derived raster maps).
- Containerized Mapserver application that enables geo webservices and APIs that serves the map based view of soil data (TerriaJS) and the map data in the SIS.
- Containerised dashboarding mechanism based on open source Apache Superset for creating interactive visualisations used for both internal workflows and on the dedicated use cases websites.
- Raster covariates and soil maps are stored in data volumes over the network, known as Network File System (NFS). NFS is a networking protocol for distributed file sharing data volumes that is later used by deployments in Kubernetes cluster. Generic Web Distributed Authoring and Versioning (WebDAV) interface was setup to the NFS volumes so that users can upload/download data.

In the subsequent chapters, each of these software components will be described in depth.



European Commission | Horizon 2020 European Union funding for Research & Innovation

This project has received funding from the
EU Horizon 2020 Research and Innovation Programme
under Grant Agreement 862900.

SOILS4AFRICA

# Soil Observations Database

This component of the infrastructure is responsible for some basic quality control and storing soil and land observations as well as wet chemistry and spectral measurements results coming from the laboratory. Related field observations and laboratory data is matched via a common sample identifier and further quality-assessed and processed in an Extract, Transform and Load (ETL) workflow. Field and lab data have been processed following the specifications from both WP4 (field campaign) and WP5 (lab analyses). All results are stored in a relational database model directed at the storage of soil property observations.

## Data model

It emerged that S4A project required a database model to store soil properties observations. Therefore, various options were considered and agreed that S4A will develop a more general data model following the state-of-the-art in soil ontology. Several models have been proposed as basis for the exchange of soil data that offer a starting point for a data model. A data model for the S4A database is based on the existing data models for storing soil data. ISO 28258 model (de Sousa et al., 2023; https://iso28258.isric.org/) is the only truly international standard directed at soil data exchange, with the alternatives being primarily regional in nature.

The data model for ISO-28258 builds on and is further specialized for the Soils 4 Africa project resulting in S4A specialised data model (de Sousa et al., 2024) with an extension to "spectral data". In this extension, spectral observations thus start with the introduction of the sensor entity, harbouring the minimal description of the hardware used. **`procedure_spectral`** provides the details on the modes of operation of each sensor, specifying the range of the electromagnetic spectrum observed and the respective spectral resolution. **`observation_spectral`** brings both together, further allowing for a unit of measure, plus minimum and maximum admissible values. The actual result is stored as a JSONB object in the **`result_spectral`** entity (see *Figure 2* below).

From the spectral observations, derived soil property values are obtained by running a model that predicts soil property values from a spectrum or spectra. 25% of the samples have also been analysed through wet chemistry. These results have been used to build the prediction model. The **`observation_derived`** entity thus refers to a specific model and a physio-chemical observation (the target of the model). The **`result_derived`** entity represents actual derived results, that beyond referring to a derived observation must also refer to the physio-chemical result on the which the model was applied (see Figure 3 below).

A dataset in this model is defined as a series of observations having similar lineage. In practice, it can be a range of laboratory reports within a certain timeframe, on samples from a certain region, sampled by a certain group of surveyors, and/or modelled using a certain model calibration. Lineage of a dataset is captured in a metadata record describing that dataset.



European Commission | Horizon 2020 European Union funding for Research & Innovation

This project has received funding from the
EU Horizon 2020 Research and Innovation Programme
under Grant Agreement 862900.

SOILS 4 AFRICA

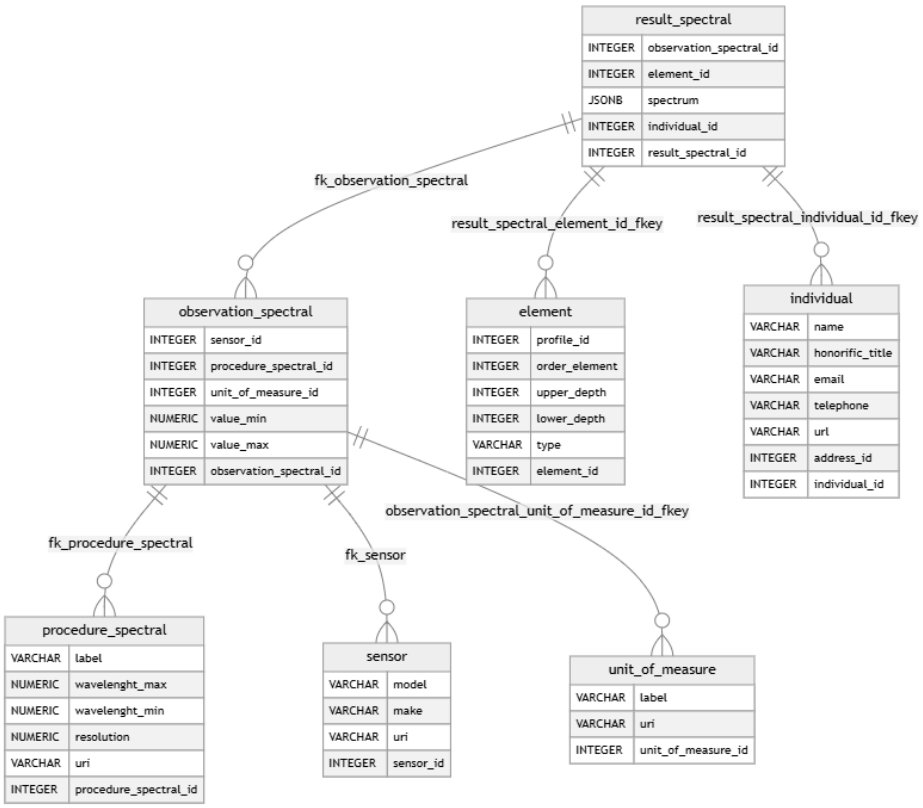



*Figure 2. A simplified representation of the S4A data model describing spectral observations associated with a soil profile.*

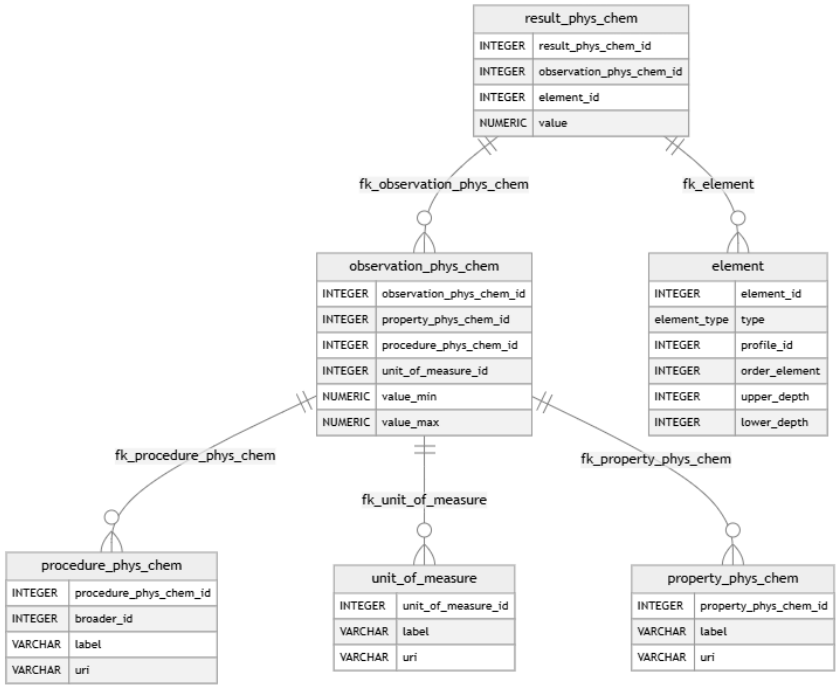



*Figure 3. A simplified representation of the S4A data model describing physio-chemical observations associated with a soil profile.*



This project has received funding from the
EU Horizon 2020 Research and Innovation Programme
under Grant Agreement 862900.

## Data upload

Workflow for submissions of soil physical and wet chemistry results and spectral measurements by project partners were organized as follows:

- Project partners had access to a shared GitLab repository where they can drop **.csv** files. The GitLab repository contains four subfolders: wet chemistry, spectral measurements, modelled data and metadata.
- At every upload a data processing pipeline Extract-Transform-Load (ETL) workflow will start automatically which validates and harmonises the dataset. See the next section for more information on ETL workflow.
- During the ETL workflows and checks, wet chemistry results and spectral measurements are loaded into S4A database.
- Modelled data folder will hold predicted from the spectral observations soil properties in the future.
- The metadata files contain for each uploaded result-table the definition of the columns of the table.

Regarding spectral measurements, the current practice to share spectral calibration models is to prepare one table (for instance a **.csv** table) with both the spectra and the wet chemistry calibration data, linked through a sample identifier. For instance, ICRAF develops its models in the **R** language and bundles these models with the datasets which that model uses (de Sousa et al., 2021). However, no code marking mechanisms are used presently (e.g., tags, branches) and models are just stamped with a date and hash to identify a particular state of the repository (de Sousa et al, 2021).

## ETL workflow

The ETL workflow is the procedure of importing data into the standardized S4A data model. The database includes a set of rules and procedures that perform some level of validation during the ingest process. The ETL validates the correspondence of land and soil profile observations collected in the field campaign, with the wet chemistry and spectral measurements of the samples collected at those locations. During the ingest process plausibility checks of soil property values (and units of measurement) and, where possible, harmonises using consistency procedures.

The ETL workflow is implemented as a Python program. The ETL is set up to run as either CI/CD workflow pipelines when data is uploaded to the GitLab or to run at the scheduled intervals.

There are four main flows of information incoming via the ETL workflow.

- From IITA managed an online platform to collect field survey description information, including: where are samples collected, landscape and landcover observations, etc.
- From the ARC Physio-chemical and spectral laboratory analyses of soil chemical parameters.
- From the ARC BIO; laboratory analyses of soil biological parameters.
- Analysis results of contaminations samples performed by Wageningen University laboratory.

The ETL program uses the following approach:

- https://git.wur.nl/isric/soils4africa/database/-/tree/main/etl?ref_type=heads is GitLab repository that contains the scripts to build a container image in which the scripts run.



This project has received funding from the
EU Horizon 2020 Research and Innovation Programme
under Grant Agreement 862900.

- https://git.wur.nl/isric/soils4africa/s4a-data-shop is the GitLab repository where the raw data is uploaded by the laboratory. After lab users data CI/CD pipelines run by fetching the ETL image created before. The repository is configured with a scheduled pipeline to run the ETL every weekend. Data from the SDMT is processed during this scheduled pipeline



This project has received funding from the
EU Horizon 2020 Research and Innovation Programme
under Grant Agreement 862900.

# GitLab code versioning

GitLab repositories play a central role in the SIS platform. The GitLab repositories in SIS are used to:

- Maintain the website content, catalogue content (metadata), map configurations, deployment configurations.
- Control codebases developed in the S4A project: digital soil mapping codebase, ETL workflows, etc. Make releases and build docker images.
- Manage user feedback via issues and discussions.
- Run triggers and pipelines when changes are made to the content updates, harvesting, deployments, data upload, validations.

A version control system is essential for managing code changes and modifications, as multiple project partners and national institutes collaborate on the same codebase using non-linear workflows.

GitLab serves as both an issue tracker and a code workflow manager. It also enables the creation of wiki pages, ensuring effective communication and documentation of various software components deployed in SIS—an important requirement for the SIS hosting organization.

Beyond documentation and code management, GitLab offers social features such as user profiles, commenting, and rating systems. These features foster community engagement around each repository, allowing users to flag issues, suggest improvements, and contribute new content.

One last useful feature of GitLab  is that it can be enabled with a trigger mechanism to generate certain workflows based on the recent code merges (i.e., ingest data, quality control workflow, soil mapping, etc). This feature is known as Continuous Integration/Continuous Development (**CI/CD**) method for automating the code development and can be useful in future extension of SIS. The CI/CD pipelines are configured in the application repository usually in .**`gitlab-ci.yml`** file.



European Commission | Horizon 2020 European Union funding for Research & Innovation

This project has received funding from the
EU Horizon 2020 Research and Innovation Programme
under Grant Agreement 862900.

SOILS4AFRICA

# Kubernetes platform

Kubernetes is an open-source platform designed for automating the deployment, scaling, and management of containerized applications. It was initially developed by Google and later donated to the Cloud Native Computing Foundation. Kubernetes is highly regarded for deployment due to its robust features and benefits, especially when managing microservices. Here are some key advantages of microservices based on Kubernetes (applicable for Kubernetes alternatives as well):

1. **Scalability and Flexibility**: Kubernetes efficiently manages resources, ensuring that microservices use resources optimally, which can lead to cost savings. Each microservice can be developed, deployed, and scaled independently, allowing the use different technologies and tools best suited for each microservice.
2. **Enhanced Resilience**: The self-healing nature of Kubernetes ensures that microservices and deployments remain available even in the face of failures, improving the overall resilience of the application.
3. **Faster Development and Deployment**: Kubernetes simplifies the development and deployment process and automates many of the complex tasks involved in managing microservices, such as service discovery, load balancing, and scaling, making it easier to manage large, distributed systems.

We suggest the ArgoCD tool as a management layer on top of Kubernetes, facilitating the interaction between GitLab and the Kubernetes platform. The configuration of the services running in Kubernetes are defined in configuration scripts stored on Gitlab. ArgoCD facilitates the automated deployment of updated configurations

Below follows the description of a number of containerized applications that are deployed in Kubernetes platform as part of SIS and will be delivered as containers to the hosting institution.

## CMS – project website (ngnix + Quarto)

The content management system (CMS) has been designed with tailored branding to introduce users to various aspects of the platform. It is implemented using Quarto (https://quarto.org/ ), with content stored and served from GitLab. The Quarto build process can be executed either as a script or within a dedicated Docker container.

The CMS container runs on Ubuntu 22.04, equipped with Quarto (version 1.5.57) to convert GitLab pages into .html files, which are then served via an NGNIX server (version 1.21.4). Web page content is maintained as markdown files in GitLab, and each time a file is updated, a fresh HTML version of the website is generated and deployed using a plain NGINX container.

In this setup, the generated HTML files are transferred via a GitLab CI/CD pipeline to a storage volume, specifically a Network File System (NFS) managed through a WebDAV service (for further details, see *Storage volumes (NFS and WebDAV)* section for more information. The NFS is directly mounted on the NGINX container, ensuring seamless hosting of the updated web pages.

As illustrated in Figure 4, the project website's landing page reflects the latest content. Each time an administrator approves repository merges, Git triggers the CI/CD pipelines to rebuild and update the CMS.



This project has received funding from the
EU Horizon 2020 Research and Innovation Programme
under Grant Agreement 862900.

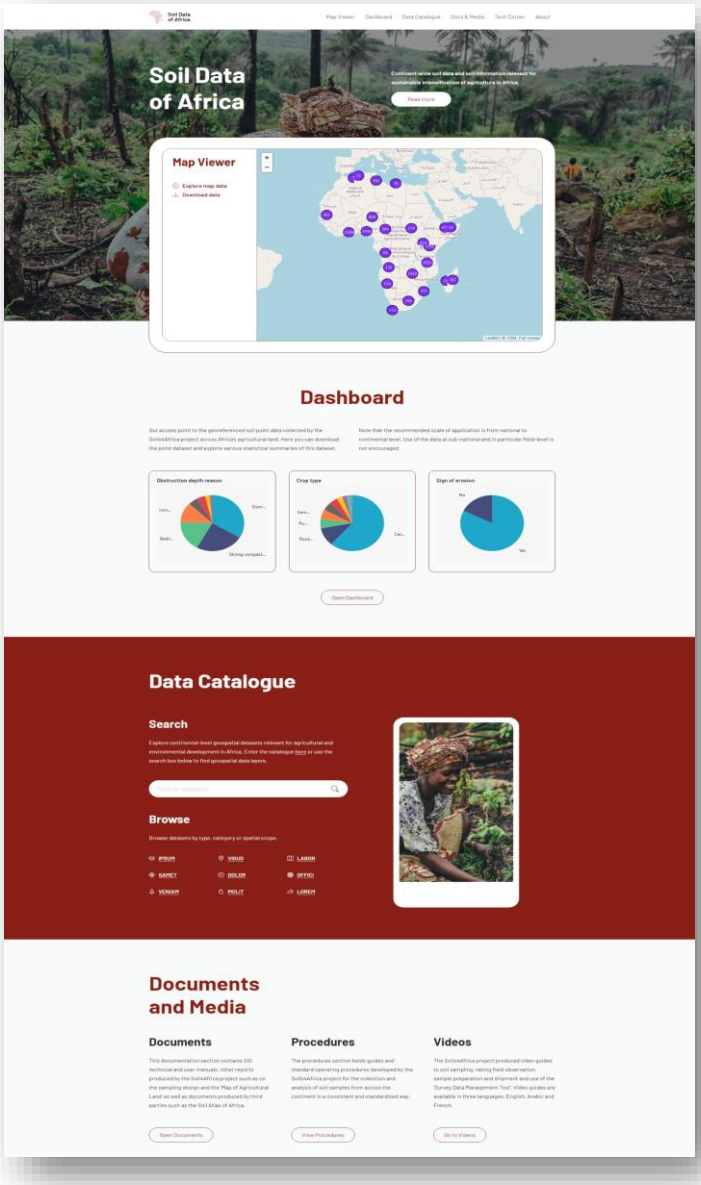



*Figure 4 Landing page of project website based on Quarto (https://africasis.isric.org/).*

## Dashboarding mechanism (Apache Superset)

A dashboarding mechanism is extensively used on the project website to track the progress of the field campaign by developing interactive web visualisations. We use containerised deployment of Apache Superset (https://hub.docker.com/r/apache/superset). It is a modern, enterprise-ready business intelligence web application built on open-source tools. Superset is cloud-native, it was designed to scale out to large, distributed environments and can run analytic workloads against most popular database technologies (Superset, 2021).

Container with Apache Superset is currently deployed in the Kubernetes platform at WUR (https://dashboards.isric.org/superset/dashboard/soils4africa/). A docker image of the superset is available at https://hub.docker.com/r/apache/superset . Superset requires a PostgreSQL database backend, to store dashboards and authorisations.

Figure 5 shows a snapshot of Apache Superset dashboard in the SIS.



This project has received funding from the
EU Horizon 2020 Research and Innovation Programme
under Grant Agreement 862900.

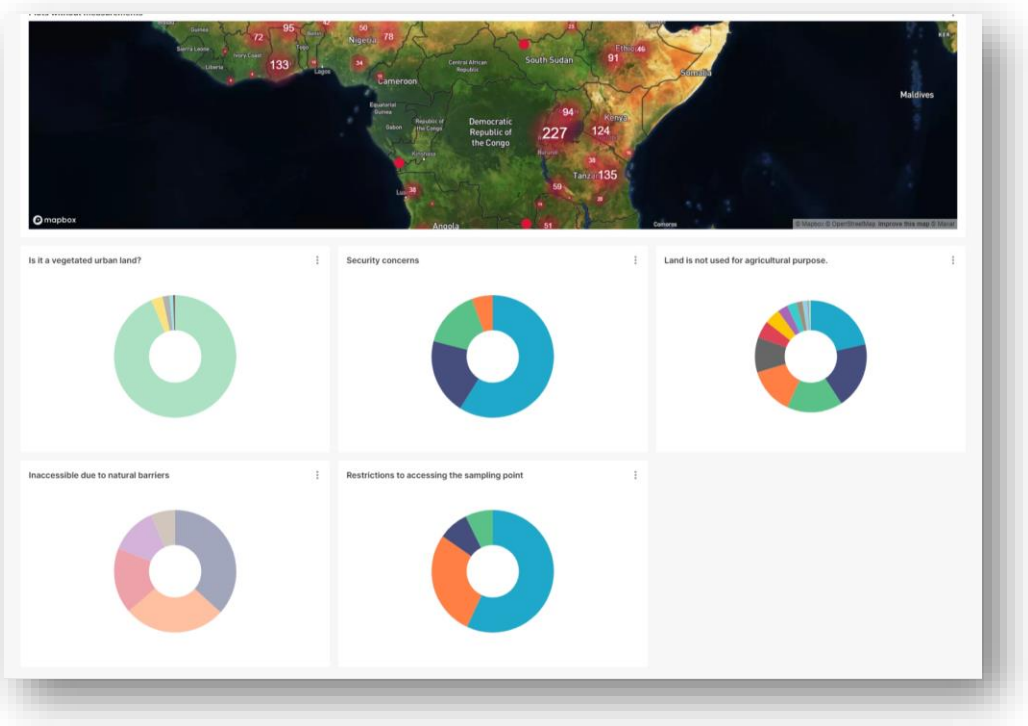



*Figure 5 SIS dashboard implementation based on Apache Superset.*

## Storage volumes (NFS and WebDAV)

Soil properties maps, in cloud optimized GeoTiff format, covariates and other ancillary environmental data necessary to run the DSM workflow are stored in data volumes over the network, known as Network File System (NFS). NFS is a networking protocol for distributed file sharing data volumes.

S4A installed generic and extendable Web Distributed Authoring and Versioning (WeDAV) interface written in Python to the NFS volumes. Initial docker file container was adopted from (https://github.com/mar10/wsgidav/blob/master/Dockerfile) and modified to the needs of the SIS. Deployed WebDAV are available for use by other components of the infrastructure like Mapserver, TerriaJS, pycsw and the ETL workflow. Storage volumes facilitate backup operations and make easy migration of the SIS to future hosting organization.

## KIMS catalogue

The main aim of Soils4Africa project is to make all data available as open data and shared according to FAIR principles through user-friendly web services allowing data to be reviewed and downloaded. A Knowledge Information Management System (KIMS) fully facilitates FAIR data exchange and forms the core of the SIS architecture. It has two interfaces for data exchange:

- map based interface for quick spatial preview of the datasets (TerriaJS) and
- classical web catalogue interface (pycsw) that allows filter and search functionality.

### Map interface for metadata (TerriaJS)

The mapviewer is based on TerriaJS (https://terria.io). TerriaJS is a library for building rich, web-based geospatial data explorers. It can handle catalogues of tens of thousands of layers, with dozens of geospatial file and web service types supported. It is almost entirely JavaScript in the browser, meaning it can even be deployed as a static website, making it simple and cheap to host.



European Commission | Horizon 2020 European Union funding for Research & Innovation

This project has received funding from the EU Horizon 2020 Research and Innovation Programme under Grant Agreement 862900.

SOILS4AFRICA

It runs as a container at ghcr.io/terriajs/terriamap:0.2.3. The configuration of the mapviewer is maintained in GitLab repository at https://git.wur.nl/isric/ict/k8s-deployments/-/blob/master/s4a-deployments/overlays/dev/config/simple.json?ref_type=heads. Configuration is injected into the TerriaJS container while the service is deployed.

*Figure 6* below shows map interface of the KIMS based on TerriaJS.


*Figure 6 Map interface of KIMS that allow view of spatial data based on TerriaJS.*

### Web interface for metadata (pycsw)

**pycsw** is a python based OGC reference implementation of Catalogue Service for the Web and an early adaptor of OGC API Records and STAC Catalogue (Tom Kralidis, 2025). **Pycsw** offers standardised access to metadata records within the catalogue and harvesting records from remote sources through the following specifications:

- Catalogue Service for the Web – the existing OGC CSW standard.
- OGC API – Records. New OGC API family specification that offers the capability to create, modify, and query metadata on the Web.
- The Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH) is a low-barrier mechanism for repository interoperability.



This project has received funding from the
EU Horizon 2020 Research and Innovation Programme
under Grant Agreement 862900.

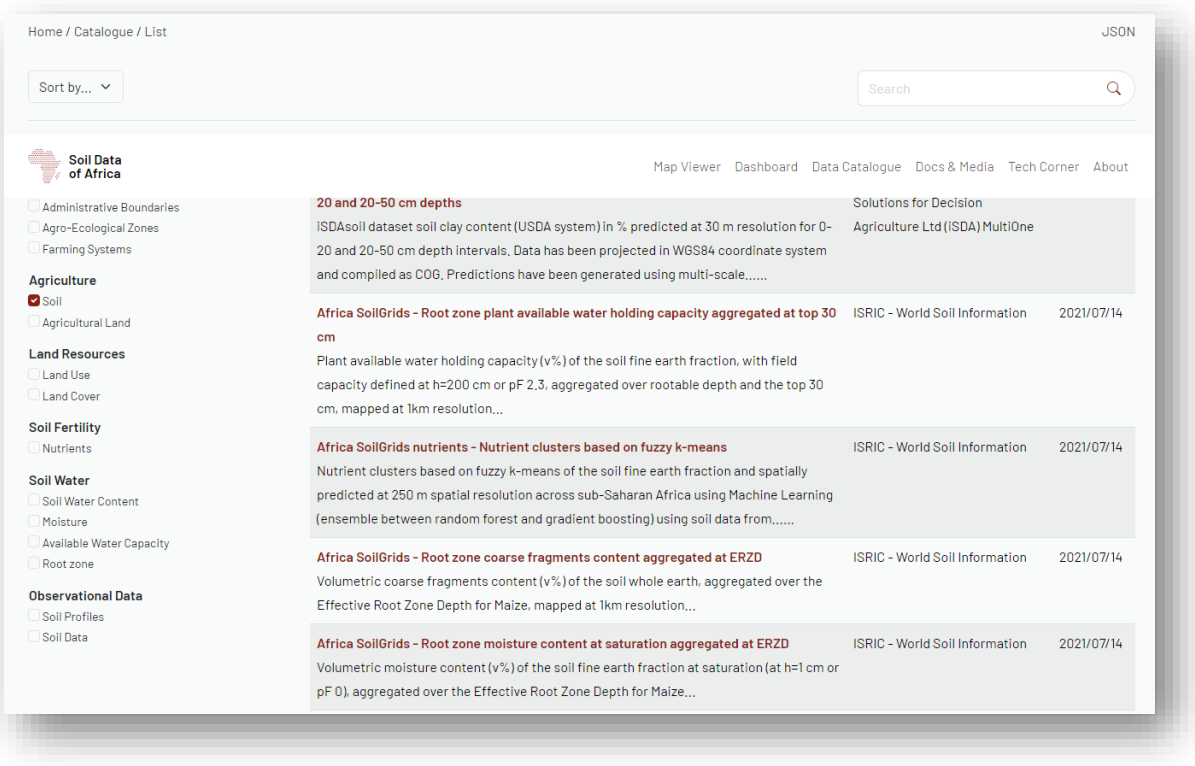



*Figure 7 Web based view of S4A datasets based on **pycsw**.*

**pycsw** uses a PostgreSQL database backend, which is automatically configured once configured. **Pycsw** supports metadata transformation across ISO19139, DCAT, and Dublin Core standards, enabling the SIS to sync with metadata catalogues like GeoNetwork, CKAN, Dataverse. This integration allows the SIS to function as an OGC-compliant node within GLOSIS, the federation of soil information systems.

**Pycsw** enhances searchability, making metadata records discoverable via search engines like Google and Bing. A customized instance provides search and filtering features, utilizing a PostgreSQL database within a CI/CD-managed container.

The **pycsw** Docker image is available at https://hub.docker.com/r/geopython/pycsw . Its catalogue layout is tailored to match the S4A project's branding, with HTML templates injected at deployment. These templates are accessible here .

### Metadata workflow

The GitLab repository containing metadata uses a basic metadata model encoded in YAML format, called Metadata Control File (MCF). MCF is a subset of iso19139:2007. YAML is easy to read by humans and an optimal encoding for content versioning. Once YAML records are updated in the GitLab repository, the content of the repository is pushed to a PostgreSQL database using CI/CD workflow. We use **pycsw** software to publish the metadata records to the PostgreSQL database. The Geodatacrawler tool has the ability to scan a local folder of dataset files and perform basic metadata operations on each of the files in order to load them in the catalogue. A typical usage scenario is that an organisation maintains a repository of data files on a network share. Each data file on a file system is accompanied by a metadata file. Crawler scripts will pick up these metadata files and publish them as iso19139 (or alternative models) on a catalogue. iso19139 is the metadata model currently mandated by INSPIRE and very common in the GeoSpatial domain.



European Commission | Horizon 2020 European Union funding for Research & Innovation

This project has received funding from the EU Horizon 2020 Research and Innovation Programme under Grant Agreement 862900.

SOILS4AFRICA

Other communities tend to use different standards, such as STAC (Earth Observation), DCAT (Open Data) or DataCite (Academia), etc.

### KIMS catalogue user administration and management

Managing user rights is essential for adding or editing metadata in KIMS. In this approach, user management is delegated to GitLab's code versioning system. All registered project partners in GitLab have the ability to flag issues related to metadata, suggest improvements, or add new metadata content. This makes GitLab a valuable tool for metadata usage, management, and quality control.

Metadata is stored as simple text files in a GitLab repository and published in a searchable catalogue whenever changes occur. Users can then update the content or flag issues directly via GitLab.

Additionally, GitLab provides social features such as user profiles, commenting, and rating systems, fostering community engagement around each repository.

## Mapserver:

**MapServer** is an Open Source platform for publishing spatial data and interactive mapping applications to the web. Mapserver implements industry standard OGC protocols such as:

- Web Map Service – WMS (graphical representation of map layers)
- Web Feature Service – WFS / OGC API – Features gives direct API access to vector data
- Web Coverage Service – WCS / OGC API – Coverages gives direct API access to grid data
- Web Processing Service – WPS/ OGC API – Processes provides standardised access to processing API's

We are using a container image developed by Camptocamp Geospatial Solutions which is the most common MapServer image in the open source community (https://github.com/camptocamp). At the time of writing this manual, Mapserver version 8 is used (https://github.com/camptocamp/docker-mapserver). Mapserver is deployed in the Kubernetes platform hosted by Wageningen University. It reads source data and mapfiles (instructions on how to display those data) from NFS where it was uploaded via WebDAV interface. The MapServer configuration files are generated by the metadata workflow.

Note, the OGC API family of standards are a new set of OGC standards being developed to make it easy for anyone to provide geospatial data to the web. These standards build upon the legacy of the OGC Web Service standards (WMS, WFS, WCS, WPS, etc.), but define resource-centric APIs that take advantage of modern web development practices. These standards are being constructed as "building blocks" that can be used to assemble novel APIs for web access to geospatial content (OGC, 2021). Therefore, it is important to ensure future proof design of the SIS by enabling support for recently developed OGC API standards, as these implement OpenAPI Specification allowing them to be used by third party websites, mobile phone applications and data science platforms in the future.

The Annex offers an additional (in depth) installation and configuration steps for the deployment of the SIS software components in Kubernetes cluster.



European Commission | Horizon 2020 European Union funding for Research & Innovation

This project has received funding from the EU Horizon 2020 Research and Innovation Programme under Grant Agreement 862900.

SOILS4AFRICA

# Supporting services

In addition to the core architectural components of the SIS that were described in previous chapters, there are several supporting services implemented in the SIS, which could be beneficial for implementation at the future SIS hosting institute."

- **Backup service**: The contents of the PostgreSQL database needs to be persisted for all deployed software components. A backup service or synchronization service needs to be set up to maintain data persistence over time.
- **Availability/Usage monitoring**: Container orchestration platforms typically include built-in services for availability monitoring, incident recovery, and usage monitoring. The primary goal of these services is to ensure system reliability by minimizing downtime, swiftly recovering from incidents, and tracking usage patterns. Through continuous monitoring, SIS hosting organization can gain valuable insights into which components of the system are functioning efficiently, identify potential bottlenecks, and assess overall performance. This proactive approach helps maintain a stable and optimized infrastructure, supporting seamless application delivery and user experience.
- **Convenience API's**: If needed, S4A soil profile datasets can be made downloadable through one of its interfaces discussed above (Dashboarding, Catalogue or Mapserver). Besides a data dump the platform also includes a number of convenience API's to access subsets of the data and near real-time data.
- **OGC:WMS/WFS webservices**: Using Mapserver container, configuration is at https://git.wur.nl/isric/soils4africa/s4a-registry



European Commission | Horizon 2020 European Union funding for Research & Innovation

This project has received funding from the
EU Horizon 2020 Research and Innovation Programme
under Grant Agreement 862900.

SOILS 4 AFRICA

# Conclusion

This report provides a comprehensive installation guide for the continental SIS developed under the Soils4Africa project, detailing the essential requirements for deploying its software components. The three interconnected systems necessary for installation include: database system, code version control, and a container orchestration system. Various deployment options, such as cloud hosting and on-premise installation, are outlined alongside alternatives, hardware requirements and installation steps.

The core software 77components of the SIS that are presently deployed within Kubernetes orchestration system will be delivered as Docker containers, ensuring compatibility with major container deployment standards. Emphasis has been placed on using open-source tools to avoid vendor lock-in and adhering to open standards, including OGC Web services, OGC API standards, metadata exchange protocols, extensibility options for API endpoints.

Most software components follow a microservices architecture and utilize the Infrastructure as Code approach to enhance portability and scalability. This guide ensures that the SIS supports all necessary workflows, from data ingestion to mapping and publishing soil data, contributing to sustainable agricultural intensification in Africa and laying the foundation for long-term soil monitoring across the continent under the emerging Soil Initiative of Africa.



European Commission | Horizon 2020 European Union funding for Research & Innovation

This project has received funding from the
EU Horizon 2020 Research and Innovation Programme
under Grant Agreement 862900.

SOILS4AFRICA

# Annex

This annex offers an additional (in depth) installation and configuration steps for the deployment of the continental SIS software components in Kubernetes cluster.

## 1. Prerequisites:

- Kubernetes cluster with at least 5 nodes.
- Storage provisioner for PVC management (e.g., NFS).
- Docker images for: CMS (nginx + Quarto), Metadata Catalogue (pycsw), MapServer, TerriaJS, and WebDAV.

## 2. Deployment Steps:

### Step 1. Configure the namespace:

- create a namespace called s4a-dev for isolating resources:

```
kubectl create namespace s4a-dev
```

### Step 2: Deploy WebDAV with PVC

- Create a Persistent Volume (PV) and Persistent Volume Claim (PVC) for storing and serving TIFF files.
- Deploy the WebDAV container with the PVC mounted at **`/files`**.

***Example YAML snippet:***

```
volumes:
- name: webdav-storage
  persistentVolumeClaim:
      claimName: webdav-pvc
  volumeMounts:
      - mountPath: /files
        name: webdav-storage
```

### Step 3: Deploy the Catalogue Metadata Service (pycsw)

- Deploy a pycsw container to provide CSW (Catalogue Service for the Web) support.
- Configure pycsw with the database connection and catalogue configuration.

### Step 4: Deploy the MapServer

- Use a Docker image for MapServer configured to access raster files via WebDAV.
- Mount the **`s4a/data`** path from the WebDAV container or PVC to serve map layers.

***Example YAML snippet:***

```
volumeMounts:
    - name: nfs-s4a-data
      mountPath: /srv/data
      subPath: s4a/data
```



This project has received funding from the
EU Horizon 2020 Research and Innovation Programme
under Grant Agreement 862900.

**Step 5: Deploy TerriaJS**

- Deploy the TerriaJS frontend to visualize the data served by MapServer.

**Step 6: Deploy the CMS using the NGINX container image:**

- Configure it to serve static content and provide backend integration.

## 3. Additional Configuration:

Ingress Configuration:

- Set up an Ingress resource to expose the services (MapServer, TerriaJS, pycsw, CMS) over HTTP/HTTPS with appropriate hostnames.

***Example YAML snippet:***

```
host: mapviewer.s4a.yourdomain.com
backend:
      service:
            name: terria-map
      port:
            number: 80
```

## 4. Deployment Verification:

- Verify that all pods are running in the s4a-dev namespace:

```
kubectl get pods -n s4a-dev
```



This project has received funding from the
EU Horizon 2020 Research and Innovation Programme
under Grant Agreement 862900.

# Bibliography

The Kubernetes Authors, 2025. Kubernetes Documentation – concepts. Retrieved from https://kubernetes.io/docs/concepts/overview/#why-you-need-kubernetes-and-what-can-it-do

de Sousa, L., Turdukulov, U., Kempen, B., 2021. User requirements for the IT infrastructure. Soils4Africa Project Report D3.5. EU Horizon 2020 Grant agreement ID: 862900. Doi: 10.3030/862900

de Sousa, L. M. Calisto, L., van Genuchten, P, Turdukulov, U., Kempen, B., 2023. Data model for the ISO 28258 domain model. ISRIC World Soil Information. Retrieved from https://iso28258.isric.org/

de Sousa, L.M., Calisto, L., van Genuchten, P., Turdukulov, U., Kempen, B., 2024. Data model for the Soils4Africa project (v20240603). Soils4Africa Project Report. EU Horizon 2020 Grant agreement ID: 862900. Doi: 10.3030/862900

OGC, 2021. OGC API Context. Retrieved from https://ogcapi.ogc.org/

Superset, 2021. What is Apache Superset? Retrieved from https://superset.apache.org/docs/intro#what-is-apache-superset

The PostgreSQL Global Development Group, 2021. PostgreSQL: The World's Most Advanced Open Source Relational Database. Retrieved from https://www.postgresql.org/

Turdukulov, U., Kempen, B., Mendes de Jesus, J., Calisto., L., van Genuchten, P., Poggio, L., 2021. Technical design of the Soil Information System. Soils4Africa Project Report D6.1. EU Horizon 2020 Grant agreement ID: 862900. Doi: 10.3030/862900



European Commission | Horizon 2020 European Union funding for Research & Innovation

This project has received funding from the EU Horizon 2020 Research and Innovation Programme under Grant Agreement 862900.

SOILS4AFRICA