

# Sharing FAIR soil data

## Repositories, licenses and metadata

**Maria Fantappiè, coordinating WP6 of EJPSOIL, CREA - Council for Agricultural Research and Economics**  
[maria.fantappie@crea.gov.it](mailto:maria.fantappie@crea.gov.it)

**Fenny van Egmond and Paul van Genuchten,**  
**of the coordinating team of WP6 of EJPSOIL, Wageningen University & Research**

EJP SOIL webinar on open science 03/11/2023



**EJP SOIL**  
European Joint Programme

EJP SOIL has received  
funding from the European  
Union's Horizon 2020  
research and innovation  
programme: Grant  
agreement No 862695



# DATA MANAGEMENT PLAN EJPSOIL

The data produced under EJPSOIL with H2020 fundings follows FAIR principles

## Findable

Metadata and data should be findable for both humans and computers

## Interoperable

Data needs to work with applications or workflows for analysis, storage and processing

F

A

I

R

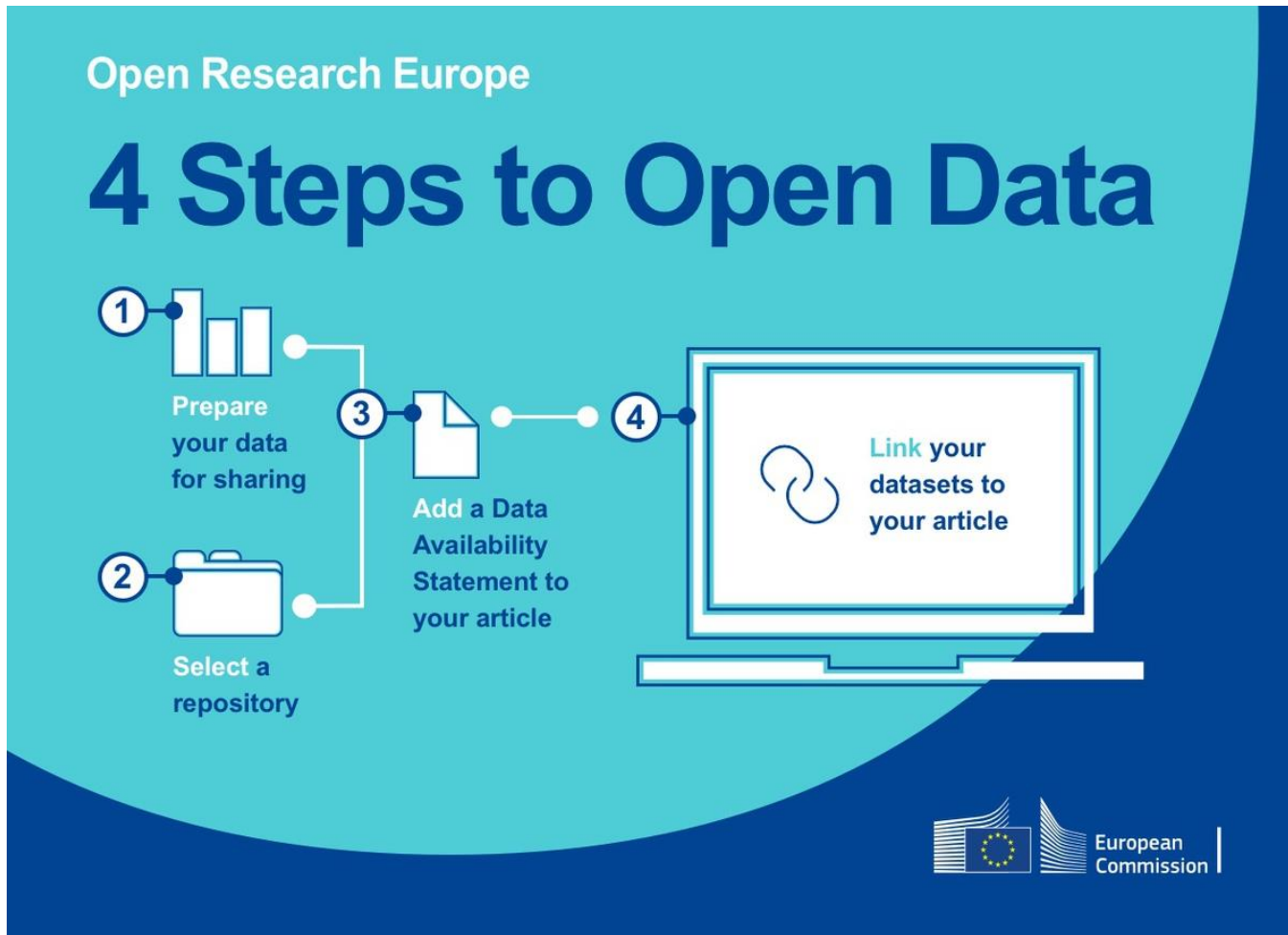
## Accessible

Once found, users need to know how the data can be accessed

## Reusable

The goal of FAIR is to optimise data reuse via comprehensive well-described metadata

# Open Access requirements for data underlying publications

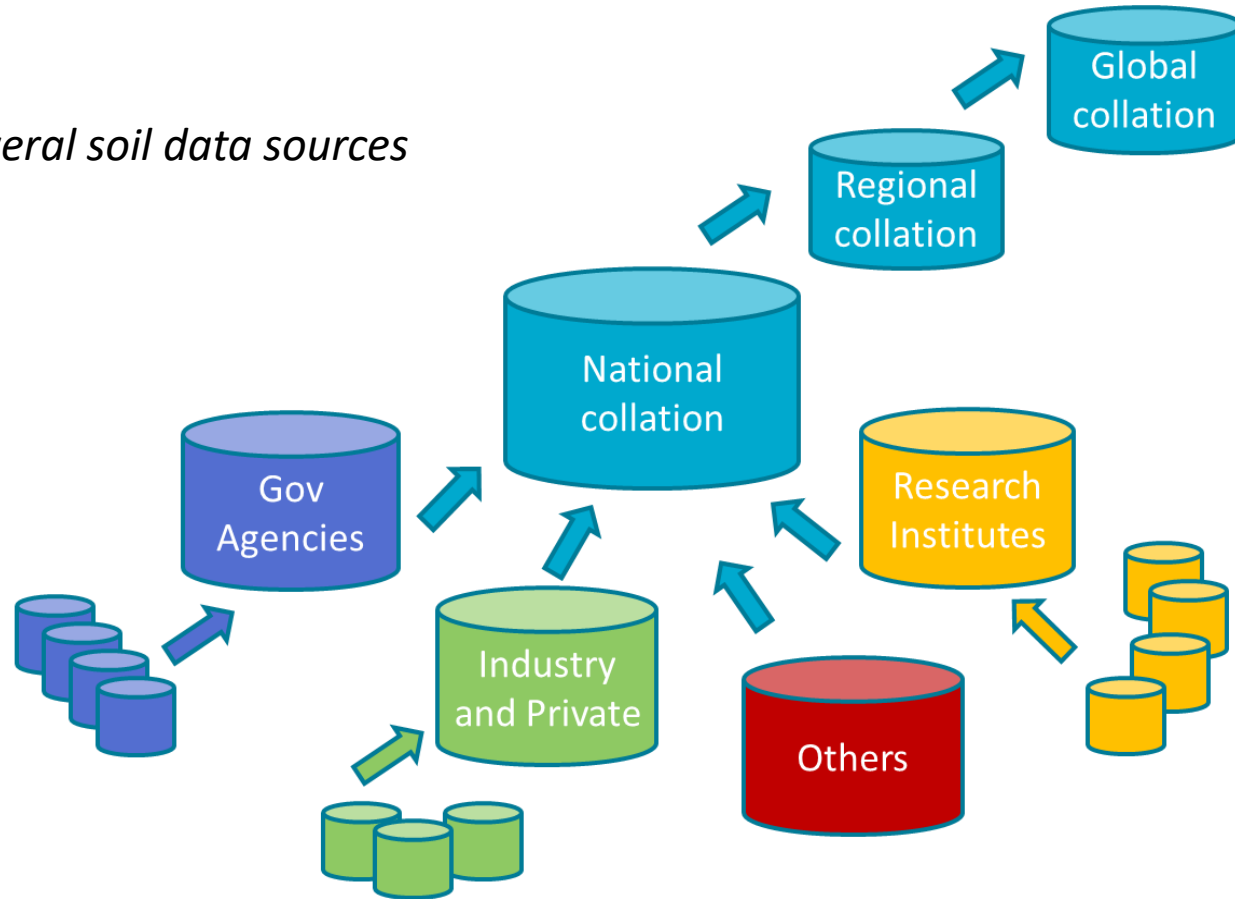


Open Access as soon as possible

- Deposit data (and metadata) in an eligible permanent repository (e.g. Zenodo)
- Provide DOI -> dataset becomes a publication
- CC BY license (open with recognition of authorship)
- Link dataset to article, link article to dataset
- NOT as supplementary material with publisher (it is not permanent repository)
- Be aware of predatory publishers: [checklist](#) to identify a trusted journal

# PRODUCING RESEARCH BY COLLECTING SEVERAL DATA SOURCES

*Several soil data sources*



As researchers you may be using also datasets that you were collecting from other soil data owners/producers.

# DATA MANAGEMENT LEGAL ASPECTS

## Background data

Data produced **OUTSIDE THE EJP SOIL programme** by project partners

The data owners define the sharing rules.

Specific agreements to be produced and signed by the data owners.

**In the D6.2 of EJP SOIL:**  
a draft template of agreement with a list of possible sharing rules to facilitate the sharing.

## EJP SOIL data

Data produced **INSIDE THE EJP SOIL programme** by ALL THE EJP SOIL WPs and by ALL THE EJP SOIL PROJECTS

Sharing rules defined in the Grant Agreement and Consortium Agreement of EJP SOIL.

Open access at the end of the project, respecting an embargo period (to get results published).  
Intellectual properties rights respected.

Permanent repositories.

**NOTE THAT:** The sharing rules for site **coordinates** are in all cases respected, following the **national legislations**.

# 1) MAKE DATA INTEROPERABLE =>> Soil data provisioning workflow: Guidance

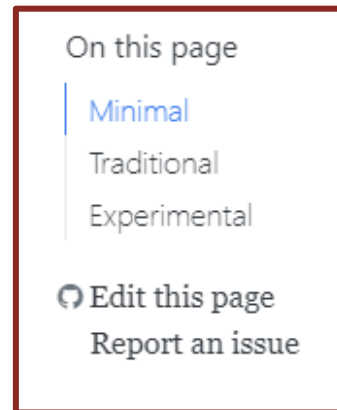
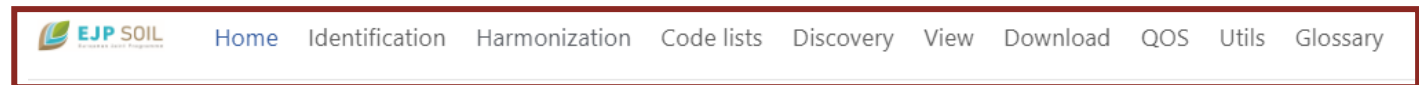


Wiki with guidance on soil data provisioning options:

<https://ejpsoil.github.io/soildata-assimilation-guidance/>

including recordings of two 3 day courses on Soil Data Assimilation

EJPSOIL WP6 has produced/published and is producing several other tools to help making soil data INSPIRE compliant, we will present during the EUSO stakeholders forum online in 15-17 November 2023.



## Introduction

AUTHOR  
Paul van Genuchten

PUBLISHED  
November 10, 2022

This WIKI is a collaborative effort to collect and describe hands-on good practices on data assimilation and dissemination in the [Soil](#) domain, with a focus on Europe. The [INSPIRE](#) directive has been and is an important effort for standardisation in the environmental data domain, therefore this WIKI has a lot of links to INSPIRE sources. Because INSPIRE adopts industry standards, this WIKI does reference common standards from ISO, Open Geospatial Consortium, Global Soils Partnership, IANA and W3C, giving it a global relevance.

## 2) MAKE DATA ACCESSIBLE =>> Upload in repositories (permanent)



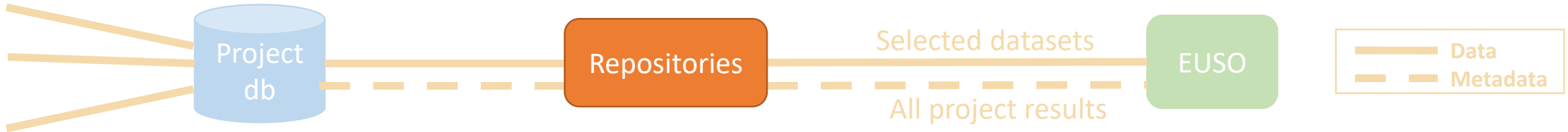
### • Minimum requirements repositories:

- Persistent (>20 yr guaranteed) repositories (e.g. ZENODO for research, national repositories)
- You get a DOI that you can use to cite the dataset





### 3) MAKE DATA REUSABLE =>> Use standard and open licenses



#### • Include the data license in the metadata:

- **OPEN LICENSE** like **CC-BY** must be used **for datasets produced under EJPSOIL (H2020)**
- for datasets produced under EJPSOIL (H2020), **it has to be declared if any further specification is needed** (e.g. for the anonymization of soil coordinates, in the name of which national legislation)
- When you are using **background datasets** with different sharing rules, first make a sharing agreement with the data owner, than publish the sharing agreement, and mention in the metadata the citation of the agreement that you have made. You can the sharing agreement template produced and published by EJPSOIL as annexed to the [D6.2](#)





# Copyright licenses

**CREATIVE COMMONS LICENSES**

**COPY & PUBLISH**   **ATTRIBUTION REQUIRED**   **COMMERCIAL USE**   **MODIFY & ADAPT**   **CHANGE LICENSE**

License	Copy & Publish	Attribution Required	Commercial Use	Modify & Adapt	Change License
PUBLIC DOMAIN	✓	✗	✓	✓	✓
CC BY	✓	✓	✓	✓	✓
CC BY-SA	✓	✓	✓	✓	✗
CC BY-ND	✓	✓	✓	✗	✓
CC BY-NC	✓	✓	✗	✓	✓
CC BY-NC-SA	✓	✓	✗	✓	✗
CC BY-NC-ND	✓	✓	✗	✗	✓

**Legend:**

- You can redistribute (copy, publish, display, communicate, etc.)
- You have to attribute the original work
- You can use the work commercially
- You can modify and adapt the original work
- You can choose license type for your adaptations of the work.

Tool to help you choose the right license:

<https://creativecommons.org/choose/>

## 4) MAKE DATA FINDABLE =>> Use metadata and keywords



### • Minimum requirements for metadata:

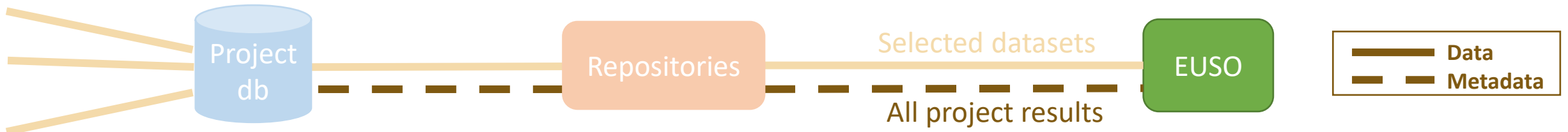
- Include metadata required by repositories to fulfill ISO19139/ DataCite/ DCAT standard

EJPSOIL – WP6 and WP1 are producing an EJPSOIL metadata template (xlsx and git)

- Use keywords in the metadata!

e.g. EJPSOIL, project acronym, soil property, region, country, etc.

## 4) MAKE DATA FINDABLE =>> upload in online metadata catalogues



## • EJPSOIL has now its own soil metadata catalogue online!

With the soil metadata template that you will have compiled the soil datasets will be uploaded in the catalogue. Other data has been and will be retrieved/harvested by other soil metadata catalogues.

[EJPSoil catalogue - Home](#)

The screenshot shows the homepage of the EJPSOIL catalogue. The browser address bar displays 'https://catalogue.ejpsoil.eu'. The page features the EJPSOIL logo (European Joint Programme) and navigation links for 'Home', 'JSON', 'About', and 'Contact'. The main heading is 'EJPSoil catalogue', followed by a brief description: 'These pages present a set of datasets collected in the scope of the EJP Soil project.' Below this, there is a paragraph about the programme's goals: 'EJP SOIL is a European Joint Programme Cofund on Agricultural Soil Management contributing to key societal challenges including climate change, water and future food security.' A longer paragraph describes the objectives: 'The objectives are to develop knowledge, tools and an integrated research community to foster climate-smart sustainable agricultural soil management that: Allows sustainable food production, Sustains soil biodiversity, Sustains soil functions that preserves ecosystem services. EJP Soil is supported by the European Commission through the Horizon 2020 European Union funding for Research & Innovation.' At the bottom, there is a green search bar with the text 'Search the catalogue', a search input field, and a 'Submit' button. Below the search bar, the text 'Recent changes' is visible.

# Discovery of soil data in Latvia

https://catalogue.ejpsoil.eu/collecti https://geolattija.lv/geo/p/319



Home / Collections / EJPSoil catalog

EJP  
"These  
EJP S  
The ot  
produc  
Europ

Title

[Soil agric](#)  
[State Inf](#)  
[Agricultu](#)

[Digital s](#)

[Databas](#)  
[in soils c](#)

[Digital s](#)

[Universi](#)

Prev

Se  
soil l

Recsw 3.0.dev0

St  
Deutschland

Soil substances in  
Germany. In addition  
to substances that  
serve as nutrients fo  
the plants, there are  
also pollutants

#CSW 3.0.dev0

in digital form (parcel-  
specific,no).

tne E  
Polyc

GEO Latvija.lv

Kartes pārļūks

Geoprodukti

Karšu galerija

Teritorijas attīstības plānošana

Metadatu katalogs

LV EN Mana darba vieta

## Digital soil da

Latvia EJP Country survey Dataset

profile sampling, samples from top

### Contacts

Ministry of Agriculture of the Re

Role: pointOfContact

country: Latvia

### Temporal

Created: 1960-1991

Updated: 2023-06-14

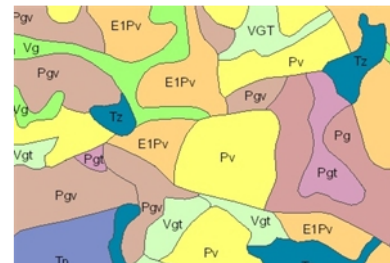
Temporal extent: 1960-1991

### External identification

- <https://geolattija.lv/geo/p/319>

### Links

[55b9911957f49ed6ce00](#)



## Augsnes dziļrakumi (lejupielādes datne)

Zemkopības ministrija

Geoprodukta tips: Lejupielādes datne

Ģeotelpiskā informācija par lauksaimniecībā izmantojamo zemju augsnēm Latvijā, kas kartētas laika periodā no 1960. līdz 1991. g. vairākās kartēs dažādos laika posmos, izmantojot dažādas augšņu klasifikācijas mērogā 1:10 000. Informācija iegūta, digitalizējot Latvijā pieejamās vēsturiskās augšņu kartes. Atbildīgais par karšu uzturēšanu: Zemkopības ministrija.

Kartes mērogs

1:10000

### Licencēšanas noteikumi

[Atvērto datu licence \(CC 4.0\)](#)

Lauksaimniecībā izmantojamo zemju augsnēs Latvijā ir kartētas laika periodā no 1960. līdz 1991. g. vairākās kartēs, dažādos laika posmos, izmantojot dažādas augšņu klasifikācijas mērogā 1:10 000, kā rezultātā iespējamas atšķirības starp rajoniem.

Eiropas ekonomiskās zonas projekta „Nacionālās sistēmas pilnveidošana siltumnīcefekta gāzu inventarizācijai un ziņošanai par politikām, pasākumiem un prognozēm” (Nr. 4.3-23/EEZ/INP-002) zinātniskā pētījuma projekta “Ilgspējīga zemes resursu pārvaldības veicināšana, izveidojot digitālu augšņu datubāzi” rezultātā laika periodā no 2014. gada septembra līdz 2016. gada martam tika veikta esošo ģeodēzisko koordinātu (LKS 92) piesaiste un datubāzes izveide Valsts zemes dienesta Centrālā arhīva materiālos esošajiem 746 augšnes dziļrakumiem.

Lejupielādes datne "Augsnes dziļrakumi" ietver informāciju par dziļrakuma numuru, gadu, integrēto augšnes granulometrisko sastāvu, virskārtas augšnes granulometrisko sastāvu, apakškārtas augšnes granulometrisko sastāvu, brīvo kalcija karbonātu sastopamības dziļumu.

→ ĢEOTELPISKIE DATI

→ METADATI

↓ DATŅU SARAKSTS



Filtrēt pēc datņu nosaukuma

Atcēlt izvēli

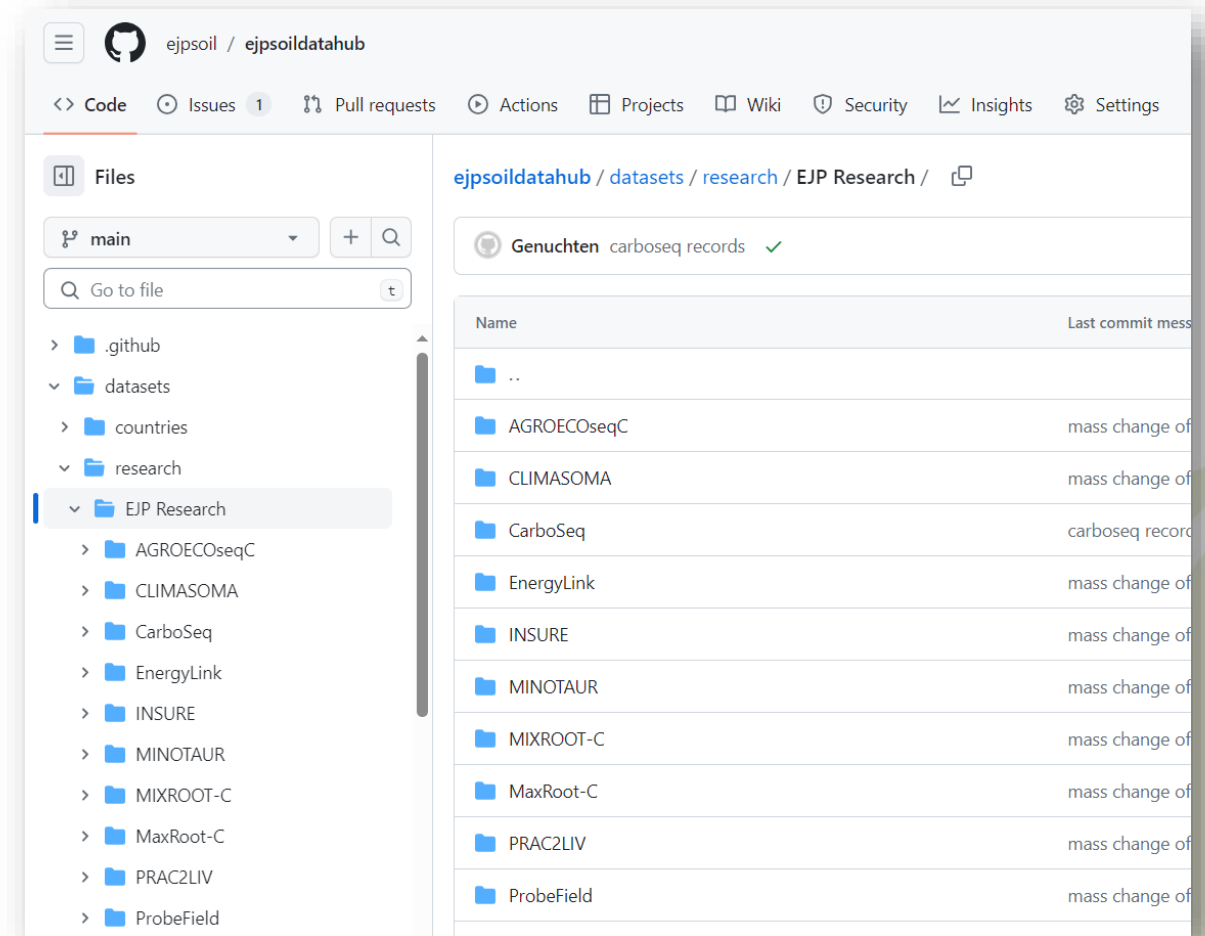
Dzirakumu\_profili\_AP

70,34 KB

Lejupielādēt

# Catalogue co-creation via Github

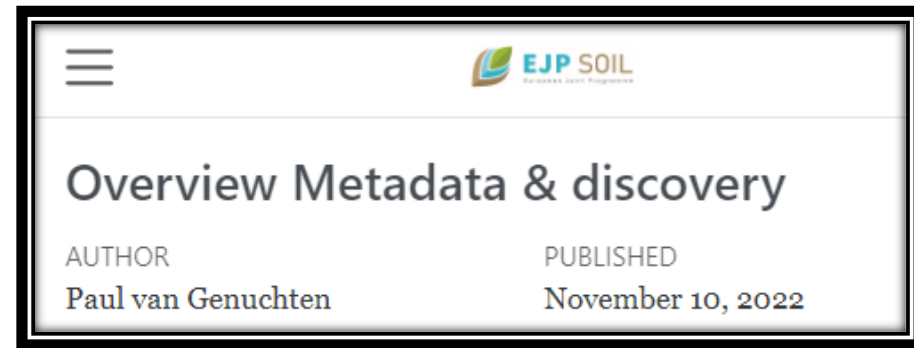
- Records are maintained in Github
- Suggest changes via issues or pull requests
- A folder has been prepared for every project



# Metadata creation options

- **Option 1: Maintain metadata at the source**

- Embed metadata in the resource
- Maintain a separate metadata file for every datafile, service or application, with the same name in that folder



=> link

- **Option 2: Metadata creation as part of upload to a repository (e.g. <https://zenodo.org/>, <https://recherche.data.gouv.fr/en> (former INRAE DataVerse)**

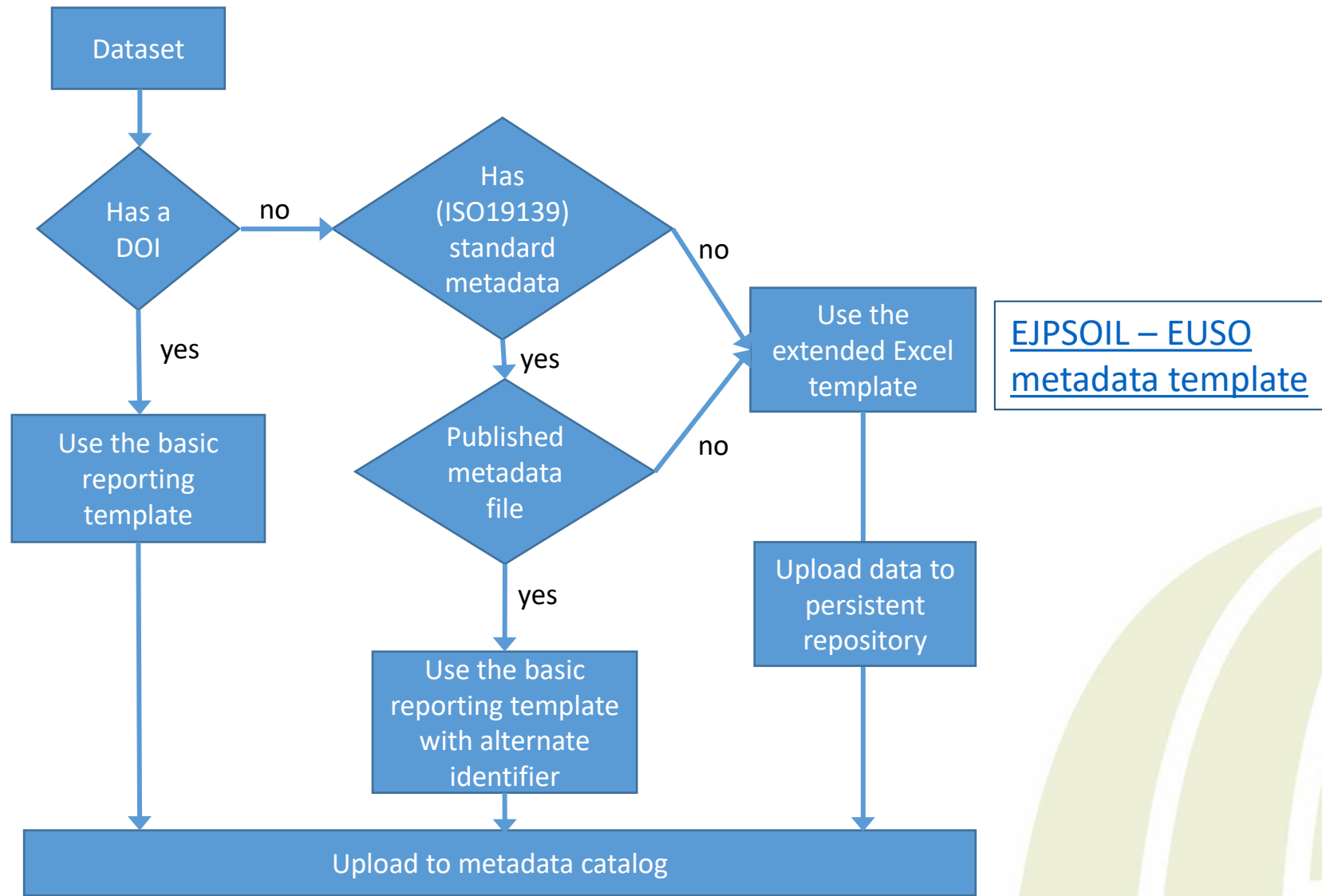
- Guidance document on how to use available metadata properties in Zenodo to create a metadata record which matches the EJP Soil Metadata profile: <https://ejpsoil.github.io/soildata-assimilation-guidance/cookbook/zenodo.html>

- **Option 3: Use the EJP Soil metadata excel template**

[EJPSOIL – EUSO metadata template](#)

# Metadata upload workflow

EJP SOIL Template RP3 - Continuous reporting





## Metadata standards to use

- Use an existing metadata standard

*or*

- Use the EJP SOIL WP6 template which is an extension of existing standards (ISO 19115:2013 and DataCite)

*and*

- Use EJPSOIL and relevant keywords from relevant thesauri (Agrovoc, Gemet, Cordis) and

EJPSOIL, project acronym, soil property, region, country, etc.

A COMPARISON OF METADATA STANDARDS HAVE BEEN PRODUCED UNDER EJPSOIL and is available at <https://ejpsoil.github.io/soildata-assimilation-guidance/cookbook/md-schema-comparison.html>

# Metadata standards

Community	Metadata format	Metadata Tools	Catalogues
<i>Academia</i>	DataCite (DOI)	Dataverse	<a href="https://zenodo.org">zenodo.org</a> , <a href="https://search.dataone.org">search.dataone.org</a>
<i>Open Data / Semantic web</i>	DCAT	CKAN, BRegDCAT	<a href="https://data.europa.eu">https://data.europa.eu</a>
<i>GeoSpatial / INSPIRE</i>	ISO19115:2003	GeoNetwork, ArcGIS, pyCSW	<a href="https://geoportal.org">https://geoportal.org</a> , <a href="https://inspire-geoportal.ec.europa.eu">https://inspire-geoportal.ec.europa.eu</a>
<i>Earth Observation</i>	STAC	STAC Browser	<a href="https://explorer.digitalearth.africa/stac">explorer.digitalearth.africa/stac</a>
<i>Search engines</i>	Schema.org	Rich results test	<a href="https://datasetsearch.research.google.com">https://datasetsearch.research.google.com</a>

# Metadata template excel file

	A	B	C	D	E	F	G	H	I	J	K
nr	Identification	EUSO Data WG subgroup	Context	Title	Abstract	Format	Extent (geographic)	Reference period - Start	Reference period - End	Access constraints	
	<i>Unique identification of the dataset (A UUID, URN, or URI, such as DOI)</i>	<i>The EUSO subgroups which contributed to this record</i>	<i>Context: (e.g. EU-Project SOILCARE, EJP-Soil, Literature, ESDAC, etc.)</i>	<i>Short meaningful title</i>	<i>Short description or abstract (1/2 page), can include (multiple) scientific/technical references</i>	<i>File Format in which the data is maintained or published</i>	<i>Geographical coverage (e.g. EU, EU &amp; Balkan, ...)</i>	<i>Reference period for the data - Start</i>	<i>Reference period - End; empty if ongoing</i>	<i>Indicates if the data is publicly accessible or the reason to apply access constraints</i>	
1											
2											
3											
4											
5											
6											
7											
8											
9											
10											
11											
12											
13											
14											

## EJP/EUSO Metadata Profile for Soil Datasets

Label	Cardinality	Codelist	DataCite	Description
Identification	1-1		yes	Unique identification of the dataset (A UUID, URN, ...)
EUSO Data WG subgroup	0-n	yes		The EUSO subgroups which contributed to this record
Context	0-n	yes		Context: (e.g. EU-Project SOILCARE, EJP-Soil, Literature, ...)
Title	1-1		yes	Short meaningful title
Abstract	1-1		yes	Short description or abstract (1/2 page), can include (multiple) scientific/technical references
Format	0-1	yes	yes	File Format in which the data is maintained or published
Extent (geographic)	0-1		yes	Geographical coverage (e.g. EU, EU & Balkan, ...)
Reference period - Start	0-1			Reference period for the data - Start
Reference period - End	0-1			Reference period - End; empty if ongoing
Access constraints	1-1	yes		Indicates if the data is publicly accessible or the reason to apply access constraints
Usage constraints	1-1	yes	yes	Indicates if there are legal usage constraints (licensing, etc.)
Keywords	0-1		yes	Keywords; separated by ';' and space
Contact	1-n		yes	One Contact per line; name; organisation; email; phone; address
Source	0-n			Source is a reference to another dataset which is used in the processing of the data
Lineage	1-1			Statement on the origin and processing of the data

[EJPSOIL – EUSO metadata template](#)

# Thanks for your attention

[maria.fantappie@crea.gov.it](mailto:maria.fantappie@crea.gov.it)  
[paul.vangenuchten@wur.nl](mailto:paul.vangenuchten@wur.nl)  
[fenny.vanegmond@wur.nl](mailto:fenny.vanegmond@wur.nl)



**EJP SOIL**  
European Joint Programme

EJP SOIL has received  
funding from the European  
Union's Horizon 2020  
research and innovation  
programme: Grant  
agreement No 862695



EJP SOIL internal projects' information and discussion meeting (October 19,  
2023)