SELINA
SCIENCE FOR EVIDENCE-BASED AND SUSTAINABLE DECISIONS ABOUT NATURAL CAPITAL

D7.2 Data management plan (version 1)

31/01/2023

Lead Author/s: Anna Sapundzhieva, Stefania Antoniadou, Pavel Stoev
Contributing Author/s: Joana Seguin

Funded by
the European Union

SELINA receives funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101060415. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Commission. Neither the EU nor the EC can be held responsible for them.

Prepared under contract from the European Commission
Grant agreement No. 101060415
EU Horizon 2020 Research and Innovation Actions

Project acronym:	**SELINA**
Project full title:	**Science for Evidence-based and sustainabLe decIsions about NAtural capital**
Project duration:	01.07.2022 – 30.06.2027 (60 months)
Project coordinator:	Prof Dr Benjamin Burkhard, Gottfried Wilhelm Leibniz University Hannover
Call:	HORIZON-CL6-2021-BIODIV-01
Deliverable title:	Data management plan (version 1)
Deliverable n°:	D7.2
WP responsible:	WP7
Nature of the deliverable:	Other
Dissemination level:	Public
Lead beneficiary:	PENSOFT
Citation:	Sapundzhieva, A., Antoniadou, S., Stoev, P. (2023). SELINA D7.2. Data Management Plan (version 1).
Due date of deliverable:	Month n°6
Actual submission date:	Month n°7

Deliverable status:

Version	Status	Date	Author(s)	Reviewer
1.0	Draft	31 Jan 2023	Anna Sapundzhieva, Stefania Antoniadou, Pavel Stoev PENSOFT	Joana Seguin LUH

Table of Contents

Preface

As a research project funded under the Horizon Europe programme of the European Commission, SELINA will follow all rules on Open Science, Open Research Data and Research. When research data are well organised, well stored and accessible, their validity can be monitored at all times and the result is high-quality, efficient research and cost savings. The SELINA consortium aims to contribute to open science policy and practices by guaranteeing that research data is made findable, accessible, interoperable and reusable (FAIR). This will be achieved through a project-specific Data Management Plan (DMP), which will ensure that research data used and generated within SELINA will be managed in a resilient and transparent way.

This document builds upon the initial version of the data management outline presented in the SELINA proposal and provides further information on the datasets and data-sharing practices that will be used and generated within SELINA. Since this deliverable is submitted at an early stage of the project (M7), it answers data management questions in a matter appropriate to this project stage. Due to the fact that some data management information might not be listed in sufficient detail, the DMP is considered a living document, which will be updated (at least) twice during the SELINA implementation. In M36, D7.9 Data Management Plan (version 2) will give a finer level of granularity in describing the datasets used and generated within the project. In M60, a final version of the DMP will be submitted (D7.8 Data Management Plan (final)), which will capture all changes and updates occurred during the project.

The DMP has been developed in close collaboration with project partners who provided information on their planned usage and generation of research data, along with their institutional policies on data management. This information was provided via a questionnaire distributed in M6 of the project, carried out in the open source online statistical survey platform LimeSurvey. The questionnaire is available as Annex 1 to this document.

Summary

The data management plan of SELINA aims to regulate data throughout the entire project lifetime and beyond by setting rules and recommendations on how data should be made available in a FAIR manner. The data management relates to both the generation of data within the research network of SELINA, as well as the acquisition and processing of primary and secondary data from outside the SELINA internal research network (e.g. use of external datasets).

In accordance with that, this DMP follows a clear structure, providing a Data summary, including the datasets that SELINA plans to use and generate, their purpose, format, origin, expected size, utility and curation (Chapter 1). The Data summary is followed by individual sections on how the project plans to make data findable, accessible, interoperable and

reusable (Chapter 2). The DMP will also describe the allocation of resources for FAIR data management within the consortium during and beyond the project lifetime (Chapter 3), as well as the data security practices which guarantee that the necessary provisions are in place to preserve and curate research data (Chapter 4). Chapter 5 considers the ethical aspects of data sharing, including GDPR-compliance when personal data is concerned.

The DMP ends with concluding remarks on the data management strategy adopted by the project, and it outlines future updates and additions to the plan which are going to be presented at a later stage of the project's development.

The main implications from the data management plan are highlighted in green boxes. These will be consolidated in a Data Management Recommendations one-pager and will be shared with partners to ensure maximum uptake.

List of abbreviations

APC	Article processing charges
DMP	Data Management Plan
DPO	Data Protection Officer
EML	Ecological Metadata Language
EU	European Union
FAIR	Findable, Accessible, Interoperable, Reusable
GDPR	General Data Protection Regulation

1 Data summary

SELINA will use pre-existing and widely available data on different ecosystem types for comprehensive assessments in EU member states. SELINA will also generate datasets in relation to WPs 2-6, 8 and 9. The tables below present an overview of the datasets that partners plan to generate (Table 1. Summary of data planned to be generated within SELINA) and to use (Table 2. Summary of data planned to be used within SELINA) according to partners' responses to the DMP questionnaire carried out in Months 6 and 7 of SELINA.

Table 1. Summary of data planned to be generated within SELINA.

N	Name of dataset	Name of generator	Relevant task	Generated via	Size	Format	Type of data	Sensitive data	Personal data	Delivery date	Utility/Users	Open Access
1	Roadside greenery services in Poznań/Poland	Damian Łowicki	T3.1, 3.2, 3.3, 3.4, 4.1, 4.2, 4.3, 6.1, 6.2, 6.3, 6.4, 8.5	modelling, data processing, literature review, policy review	n/a	shp, GeoTIFF	quantitative data; digital data; GIS data	no	no	2023	City Council of Poznań, Municipal Urban Planning Studio in Poznań, Document entitled: Tree protection standards for the City of Poznań	yes
2	Cooling capacity of green infrastructure in Poznań	Piotr Lupa, Iwona Zwierzchowska, Damian Łowicki	T3.1, 3.2, 3.4, 4.1, 4.2, 4.3, 6.1, 6.2, 6.3, 6.4	modelling, data processing, remote sensing	up to 100 MB	GeoTIFF, xlsx, pdf	quantitative data; digital data; GIS data	no	no	n/a	Decision makers and city planners (City Council of Poznań, Municipal Urban Planning Studio in Poznań, Poznań Metropolis Association), academia (researchers & students)	yes
3	Green infrastructure distribution and social perception of its Ecosystem Services	Małgorzata Stępniewska, Iwona Zwierzchowska	T3.1, 4.2, 4.3, 6.1, 6.2, 6.3, 6.4	data processing	n/a	pdf, shp, xlsx	analogue and qualitative data; GIS data,	no	no (anonymised dataset)	2023	City Council of Poznań, Municipal Urban Planning	yes

							quantitative data, digital data				Studio in Poznań	
4	Food production, resource consumption and social aspects of urban agriculture (Gorzów Wielkopolski city)	Lidia Poniży	T3.1, 3.2, 3.3., 4.1, 4.2, 4.3, 6.1, 6.2, 6.3,	data processing	n/a	xlsx, csv	qualitative data, quantitative data;	no	no (anonymised dataset)	2023	Local authorities, decision makers, urban planners, academia, NGOs	yes
5	Climate regulation service in city resident's perception	Katarzyna Fagiewicz, Piotr Lupa	T3.1, 4.1, 4.2, 4.3, 6.1, 6.2, 6.3, 6.4	data processing	up to 20 MB	shp, xlsx, docx	qualitative data, quantitative data; GIS data	no	no	2023	Local authorities, decision makers, urban planners, academia, NGOs	yes
6	ES supply aggregated from different indicators	BEF-LV	T8.2	data processing	n/a	shp	GIS data	no	no	by end of 2023	marine planners, nature conservation experts	yes
7	Recreational potential of coastal area	BEF-LV	T8.2	data processing	n/a	xlsl; shp	semi-quantitative data; GIS data	no	no	by end of 2024	local and regional authorities, planners	yes
8	Coastal landscape qualities	BEF-LV	T8.2	field work, data processing	n/a	xlsl;.shp	semi-quantitative data; GIS data	no	no	by end of 2024	local and regional authorities, planners	yes
9	Potential areas for marine aquaculture	BEF-LV/MoEPRD	T8.3	data processing	n/a	shp	GIS data	no	no	by end of 2023	marine planners, aquaculture companies	yes

10	Private Sector Needs Analysis	Justine Saunders	WP9	Questionnaire/interview	n/a	docx	qualitative data; semiquantitative data	no	no	Jun 2023	SELINA consortium partners	no, personal data
11	Guidance Materials	Justine Saunders	D9.2	Literature review, Interview, Surveys	n/a	pdf	qualitative data; semiquantitative data	no	no	Dec 2023	Private sector	yes
12	ES evidence generation and uptake	Martine van Weelden	D9.3	Field work at demonstration projects/Survey/interview/Literature review	n/a	pdf	qualitative data; semiquantitative data	no	potentially	Jun 2026	Private sector	yes
13	Scalable ES solutions	Martine van Weelden	D9.4	Data processing	n/a	pdf	qualitative data;	no	potentially	Jun 2027	Private sector	yes
14	Stakeholder database for Hungary	Ildiko Arany	WP2	policy review, interview	small	xlsx	qualitative data	no	yes	Jan 2023	all involved partners	no, interview data
15	Systematic review under WP4	Ildiko Arany, Eszter Tanács, Ulla Mörtberg and Berit Balfors	WP3, WP4, D4.1	literature review	small	xlsx	qualitative data; semi-quantitative data	no	no	Sept 2023	all involved partners, researchers, policy makers	yes
16	Burned area map	CIBIO	T5.1, 5.2	remote sensing	n/a	GIS file	GIS data	No	No	S1 2024	Researchers; Decision-makers; Land managers; General public	yes
17	Fire severity map	CIBIO	T5.1, 5.2	remote sensing	n/a	GIS file	GIS data	No	No	S1 2024	Researchers; Decision-makers; Land	yes

											managers; General public	
18	Maps of fire damage on multiple ecosystem services	CIBIO	T5.1, 5.3	modelling, data processing, remote sensing	n/a	GIS file	GIS data	No	No	S1 2024	Researchers; Decision-makers; Land managers; General public	yes
19	Ecosystem accounts	CIBIO	T5.1, 5.3	modelling, data processing, remote sensing	n/a	GIS file, pdf	semi-quantitative data; quantitative data; GIS data	Yes	No	S2 2024	Researchers; Decision-makers; Land managers; General public	yes
20	PPGIS Zagreb	Martina Kičić	Potential test site	PPGIS questionnaire	n/a	GIS	spatial, quality	no	yes	n/a	scientists	on demand
21	Forest management Istria	Hrvoje Marjanović	Potential test site	Croatian Forests Company	n/a	xls, txt	number, spatial	no	no	n/a	forest managers, scientists	on demand
22	Results from Coca-Cola demonstration projects	denkstatt	WP8	Modeling, data processing, interview	100 Mb	pdf, xlsx	quantitative, semi-quantitative, qualitative	yes - commercial data/IP	no	throughout the project	Coca-Cola	Restricted - commercial
23	NBT.Malta1	Martha Arámbula	WP9, D9.4	Data processing	10 GB	docx; xlsx; shp	GIS data	No	No	1 year	General Public	yes
24	ecosystem condition indicators	Isabel Thomas	T3.2	literature review	1 MB	xlsx	semi-quantitative data	no	no	Jun 2027	Team members and institutions responsible for EU implementation of SEEA	yes

											EA	
25	Carbon sequestration of forest cover	Filipe Bernardo	WP5.2	Remote sensing	<10 GB	Shapefile; Raster	GIS data	No	No	2024	Government entities; Research institutes	n/a
26	Distribution of invasive vegetation species	Filipe Bernardo	WP5.2	Remote sensing	<10 GB	Shapefile; Raster	GIS data	No	No	2024	Government entities; Research institutes	n/a
27	Social media data	Laura Costadone	T9.3	n/a	n/a	n/a	n/a	n/a	yes	n/a	n/a	n/a
28	Policy data	Anu Lähteenmäki-Uutela	T10.2	n/a	n/a	n/a	n/a	n/a	no	n/a	n/a	n/a
29	Ecosystem coverage data	Samuli Korpinen	T3.1	n/a	n/a	n/a	n/a	n/a	no	n/a	n/a	n/a
30	Monitoring data ecosystem condition	Samuli Korpinen	T3.1	n/a	n/a	n/a	n/a	n/a	no	n/a	n/a	n/a
31	Forest Biomass Map of S. Miguel island	Artur Gil	WP5	Remote Sensing	n/a	tiff; shp	GIS data	No	No	n/a	Scientific Community; Public Administration; Forest Companies	yes
32	Updated (ESMERALDA) Methods Database	Sabine Lange, Joana Seguin, et al.	T6.6	literature review	few MBs	xlsx, csv	qualitative, semi-qualitative	no	no	update regularly until project ending	everyone working with the ES concept	yes
33	ESVD	Luke Brander et al.	T6.3, 6.6	literature review	few MBs	xlsx	quantitative	no	no	update regularly until project ending	Researchers and analysts of ES economic values	yes

34	Systematic Review	Sabine Lange, Joana Seguin, et al.	T6.1	literature review	few MBs	csv, xlsx, doc, R	quantitative	no	no	Dec 2023	researchers in the domain of ES	yes
35	Test Site Pollination ES / Ecosystem Condition	LUH (Malte Hinsch, Grazia Zulian, Benjamin Burkhard)	WP3 (3.2-3.4), WP4 (4.1, 4.2), WP6	modelling	several GBs	shp, tiff, py/R	GIS data, quantitative	no	no	ca. 2024	researchers, practitioners in agriculture or planning	yes
36	Surface thermal map of selected parts of the city of Sofia, Bulgaria	Stelian Dimitrov/Martin Iliev/Bilyana Borisova/Lidya Semerdjieva	T9.2	UAV/UAS mapping, processed through Pix4D and ArcGIS	n/a	Geotiff	GRID	No	No	n/a	NA	yes
37	Local climate zones(LCZ) map of Sofia	Stelian Dimitrov/Martin Iliev/Bilyana Borisova/Lidya Semerdjieva	T9.2	ArcGIS	n/a	gdb	vector	No	No	n/a	NA	yes
38	Cooling capacity assessment- Sofia	Stelian Dimitrov/Martin Iliev/Bilyana Borisova/Lidya Semerdjieva	T9.2	ArcGIS	n/a	gdb	vector/raster	No	No	n/a	NA	yes
39	Marine aquaculture development areas	MoEPRD	T8.3	data processing	n/a	shp; gdb	GIS	no	no	within 12 months	regional, local authorities, state institutions, scientists	yes
40	Coastal tourism development areas	MoEPRD	T8.3	data processing	n/a	shp; gdb	GIS	no	no	within 12 months	regional, local authorities,	yes

											state institutions, scientists	
41	Valuable coastal landscape areas	MoEPRD	T8.3	data processing	n/a	shp; gdb	GIS	no	no	within 12 months	regional, local authorities, state institutions, scientists	yes
42	List of stakeholders	Katažyna Bogdzevič	WP8	interviews	n/a	docx	n/a	no	yes (names, surnames, emails), represented organization)	Mar 2023	n/a	no, personal data
43	Satellite images	Sentinel	T4.1; 4.2; 4.3; 5.1; 5.2; 5.3; 8.1; 8.2;	Remote sensing	n/a	raster	GIS data	no	no	n/a	All	yes
44	Ecosystem extent property portfolio	Zander Venter NINA	DP12	remote sensing & data processing	n/a	raster	raster GIS	no	no	2024	Storebrand. TNFD members	yes
45	Habitat map - Salzachauen	Thomas Strasser	T3.1	remote sensing based modelling	20 MB	shp	GIS data	no	no	n/a	decision makers	no
46	Ecosystem Typologies SELINA Partners	Paula Rendon/ Fernando Santos	T3.1	Surveys	n/a	xlsx	semi-quantitative data	no	no	Dec 2023	SELINA Partners	yes
47	Crosswalk matrix of ecosystem typologies	Paula Rendon/ Fernando Santos	T3.1	Surveys and lit review	n/a	xlsx	qualitative data	no	no	Dec 2023	SELINA Partners	yes
48	Forest condition	Paula Rendon/	WP3	Data processing	n/a	csv	quantitative data, GIS	no	no	Jun 2027	SELINA Partners	yes

		Fernando Santos					data					
49	Spanish DP - tbd	Paula Rendon/ Fernando Santos	WP8	Data processing	n/a	csv	quantitative data, GIS data	no	no	Jun 2027	SELINA Partners, others	yes (possibly)
50	Deforestation and degradation map of Sao Miguel	SarVision	5.2	RS & data processing	27 Mb per date; 30 dates per year = 810 Mb, 3 years: 2.43 Gb	tiff	RS	No	No	2024-2026	Selina partners, local/national forest service, ecosystem account users	yes
51	Carbon map of Sao Miguel	SarVision	5.2	RS & data processing & field work	27 Mb per date; 4 dates per year = 300 Mb, 3 years: 0.9 Gb	tiff	RS	No	No	2025-2027	Selina partners, local/national forest service, ecosystem account users	yes
52	Deforestation and degradation map of coastal zone Peloponnesos	SarVision	5.2	RS & data processing	75 Mb per date; 30 dates per year =2 25 Gb, 3 years: 6.75 Gb	tiff	RS	No	No	2024-2026	SELINA partners, local/national forest service, ecosystem account users	yes
53	carbon map of coastal zone Peloponnesos	SarVision	5.2	RS & data processing & field work	27 Mb per date; 4 dates per year = 300 Mb, 3 years: 0.9 Gb	tiff	RS	No	No	2025-2027	SELINA partners, local/national forest service, ecosystem account users	yes

54	WP 3.1 international ecosystem typology database	space4environment	3.1, 5.2	n/a	n/a	gdb	tabular & spatial	no	no	n/a	public	yes
55	Literature Database	Steffen Reichel	T6.6?	n/a	<1 GB	PostgreSQL	List of literature items + Extracted data (used methods etc) from those items	no	Yes, author's emails + names (are public in the paper already normally)	End of T6.6	n/a	yes
56	Floristic diversity	UPATRAS Team	3.1, 3.2, 3.3, 5.1	n/a	n/a	csv, shp	qualitative, GIS data	Yes, include locations of endagered and/or protected species	no	July 2023	Descision and policy makers, Statistical Agengy, Forest Service, Ministry of Environment and Energy	yes, aggregated
57	DP 03 dataset	Jarumi Kato Huerta, Davide Geneletti	8.2	modelling, data processing	1 to 10 GB	csv, shp, tiff, gpkg	GIS data	no	no	2 years	Researchers, policy-makers, NGOs, general public	yes
58	DP 03 qualitative dataset	Jarumi Kato Huerta, Davide Geneletti	8.2	Interviews, surveys	1 to 10 GB	docx, xml	qualitative and semi-qualitative data	no	yes (location data)	2 years	Researchers, policy-makers, NGOs, general public	yes
59	Stakeholder maps	Inge Liekens/partners	2.1	survey	n/a	xlsx	qualitative data	no	yes: name, surname, email	Feb 2023	SELINA partners	no, rules of personal data
60	Seeds of change	Inge Liekens/stakeholders	2.4	survey	n/a	csv	qualitative data	no	yes: name, surname, email	Jan 2024	SELINA partners and DP	yes

61	Analysis seeds of change	Inge Liekens/partners	2.5	interview	n/a	MIRO; cvs	qualitative data	no	no	Jun 2025	DP, SELINA and policy makers	yes
62	Country fact sheets	Inge Liekens/partners	2.4	data processing	n/a	docx; pdf	qualitative data	no	no	n/a	all stakeholders	yes
63	Soil subsidence map of the Dutch peatland areas	WU	5.1	Modelling, remote sensing	n/a	Geodata	GIS data	no	no	Year 3	Policy makers, farming organisations, environmental organisations, researchers	yes
64	CO2 emissions map of the Dutch peatland areas	WU	5.1	Modelling	n/a	Geodata	GIS data	no	no	Year 3	Policy makers, farming organisations, environmental organisations, researchers	yes
65	Natural Capital Indicators	Stijn Schep	Task 3	modelling, data processing, remote sensing, literature review	<1 GB	xlsx, shp	quantitative	Company data	n/a	Second half of 2023	Private sector, public sector	no, commercial data
66	Land use map on Karst	Daniela Ribeiro	mapping land use on pilot area on Karst	data processing from Land use database at Ministry of	n/a	shp	GIS	no	no	April 2023	experts, owners, researchers	yes

				agriculture, forest and food								
67	Spatial data on ecosystem condition and services	Miguel Villoslada	n/a	n/a	n/a	tif, grd, shp, etc.	GIS spatial data, remote sensing data	no	no	n/a	Public agencies, research community, general public	yes
68	Reports on data generation	Miguel Villoslada	n/a	n/a	n/a	pdf	reporting data	no	no	n/a	n/a	n/a
69		Alon Lotan	T3.3, 3.4	data processing	n/a	xlsx; csv; shp; shx; dbf	quantitative and semi-quanti tative	no	no	Jan 2025	researchers	yes

Table 2. Summary of data planned to be used within SELINA.

N	Name of dataset	Relevant task	Size	Format	Sensitive data	Personal data	Open Access	Origin	Ownership	Licence	Reuse requirements
1	Database of Topographic Objects 10K (BDOT10K) for Poland	WP5	n/a	LAS and/or LAZ	no	no	no	Azorean Regional Government	Azorean Regional Government	Depends on the terms defined by the Azorean Regional Government	Yes, with GIS/LIDAR software
2	Spatial Information System (SIP) of the Poznań city	WP5	n/a	tiff	no	no	yes	ESA	EU Copernicus Program	n/a	Yes, with GIS/Remote Sensing software

3	Copernicus Sentinel-2 Data of S. Miguel Island	WP5	n/a	tiff	no	no	yes	ESA	EU Copernicus Program		Yes, with GIS/Remote Sensing software
4	MAREA project modelling results of ES supply	T2.4	n/a	n/a	no	no	yes	ESMERALDA	n/a	n/a	n/a
5	Assessment & Valuation Mapping, Natural Capital Protocol	ES measure	n/a	shp	no	no	no	Trento Municipality	Trento Municipality	Internal use, upon request only	n/a
6	Review of Mapping and Assessment Tools	n/a	n/a	shp	no	no	yes	Trento Municipality	Trento Municipality	CC0	n/a
7	Urban-public census of trees	ES measure	n/a	shp	no	no	no	Trento Municipality	Trento Municipality	Internal use, upon request only	n/a
8	Cities green plans and regulations	guidelines	n/a	pdf	no	no	yes	various Italian municipalities	various Italian municipalities	CC0	n/a
9	Urban-public brownfields	n/a	n/a	shp	no	no	no	Trento Municipality	Trento Municipality	Internal use, upon request only	n/a
10	Time series data on fire characteristics	T9.2	n/a	docx; xlsx; pdf	no	no	no	n/a (depends on materials chosen for review)	n/a(depends on materials chosen for review)	n/a	n/a
11	Ecosystem accounts and ecosystem services models	T9.2 (In collaboration with WP2/3/4/10)	n/a	docx; xlsx; pdf	no	no	no	not yet known (depends on materials chosen for review)	not yet known (depends on materials chosen for review)	n/a	n/a

12	PPGIS Zagreb	WP9	1 GB	shp	no	no	restrictions apply	Malta	Planning Authority	Attribution-NonCommercial-NoDerivatives 4.0 International	GIS software, R, Python
13	Input data from Coca-Cola demonstration projects	T8.2	n/a	docx; xlsx; shp; gdb	no	no	no	BIOR (Institute of Food Safety, Animal Health and Environment	BIOR	n/a	available upon request
14	Data on population (dynamic, density, employment etc.)	T8.2	n/a	docx; xlsx; shp; gdb	no	no	yes	CSB (Central Statistical Bureau)	CSB	CCO 1.0	available upon request
15	Data on seabed sediments	T8.2	n/a	shp; gdb	no	no	no	LHEI (Latvian Institute of Aquatic Ecology)	LHEI	n/a	available upon request
16	Local Councils Units and Boundaries	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
17	Land use Cover	T8.2	30.6 MB	shp	no	no	yes	University of Tartu	University of Tartu; BEF-LV	n/a	n/a
18	Natural Park forest cover	T8.2	1.39 MB	shp	no	no	yes	Field surveys, Latvian Institute of Aquatic Ecology	Latvian Institute of Aquatic Ecology	n/a	n/a
19	Marine sediment map	T8.2	1.39 MB	shp	no	no	yes	Field surveys starting from 1965	Latvian Institute of Aquatic Ecology	n/a	n/a
20	historical data on benthic species distribution and biomass	T8.2	8.99 MB	xlsx	no	no	yes	Field surveys results from 1965-1992	University of Tartu; BEF-LV	n/a	n/a

								stored in the archive of Latvian Environment, Geology and Meteorology Centre			
21	Forest policy interview data	T3.1, 3.2, 3.3, 3.4, 4.1, 4.2, 4.3, 6.1, 6.2, 6.3, 6.4, 8.5	n/a	shp	no	no	yes	Head Office of Geodesy and Cartography "GUGIK", Poland	Head Office of Geodesy and Cartography, Poland	open licence	GIS Software
22	TeRRIFICA Crowdmapping Tool Dataset	T3.1, 4.1, 4.2, 4.3, 6.1, 6.2, 6.3, 6.4	up to 20 MB	xls	no	no	yes	Crowdmapping data generated by the users (volunteers) within the project TeRRIFICA	TeRRIFICA Consortium	open licence	GIS Software, spreadsheet software, e.g. MS Excel
23	Satellite images Landsat-8 retrieved during radiation weather conditions in late spring and summer in Poznań in years 2020-2023	T3.1, 3.2, 3.4, 4.1, 4.2, 4.3, 6.1, 6.2, 6.3, 6.4	up to 2 GB	GeoTIFF	no	no	yes	NASA / U.S. Geological Survey	NASA	open licence	GIS Software
24	FEWMETER v2	T3.1, 3.2, 3.3, 4.1, 4.2, 4.3, 6.1, 6.2, 6.3,	n/a	xlsx, csv	no	no	restricted to June 2023	Data collected within FEW-meter project as part of completting Gardener's diary	Lidia Poniży, participant of the FEW-meter project Consortium	open licence	database software, e.g. MS Excel, Airtable
25	Ecosystem services valuation database (ESVD)	T5.1	n/a	Geodata	no	no	n/a	WU	WU	n/a	n/a

26	LIDAR survey of S. Miguel Island	n/a	n/a	GIS	no	yes	on demand	CFRI	n/a	n/a	available
27	Copernicus Sentinel-1 Data of S. Miguel Island	n/a	n/a	xls	no	no	on demand	Croatian forests company	n/a	n/a	available
28	ESMERALDA Database	T6.6	1.2 MB	xlsx	no	no	yes	https://www.maes-explorer.eu/page/100	ESMERALDA Project	n/a	n/a
29	Nature 2000 network	n/a	several MP	GIS	no	no	yes	https://natura2000.eea.europa.eu/	European Commission	n/a	GIS Software
30	Landuse types	T3.2-T3.4	4.7 GB	shp	no	no	yes	https://gdz.bkg.bund.de/index.php/default/digitale-geodaten/digitale-landschaftsmodelle/corine-land-cover-5-ha-stand-2018-clc5-2018.html	Federal Agency for Cartography and Geodesy (BKG)	Data licence Germany – attribution – version 2.0	n/a
31	data on fish catch by species	denkstatt	100 Mb	pdf, xlsx	yes	no	no	Coca-Cola-funded projects, denkstatt work funded by Coca-Cola	Coca-Cola	restricted-commercial	n/a
32	European Soil Data Centre (ESDAC)	T5.1 and 5.2	n/a	GIS	no	no	yes	Remote sensing analysis of time series satellite information (Copernicus and MODIS)	Public data generated by the CIBIO team		n/a

33	INVEKOS Österreich	T5.2	4.7 Gb per scene, 30 scenes per year, 3 years: 423 Gb	singular complex data (raw)	no	no	yes	Copernicus	Copernicus	n/a	n/a
34	n/a	T5.2	large	tif	no	no	yes	Copernicus	Copernicus	n/a	n/a
35	Gedi	T5.2	large	tif	yes	yes	yes	NASA	NASA	n/a	n/a
36	Nisar (from 2024 if operational)	T5.2	large	singular complex data (raw)	no	no	yes	NASA	NASA	n/a	n/a
37	Sentinel-1	T3.1	2 TB	gpkg	no	no	yes	reporting of planted crops and field location from farmers	Agrarmarkt Austria	CC-BY-AT 4.0	GIS software
38	Land Cover	T8.2	1 to 10 GB	tiff, gpkg	no	no	yes	Copernicus EU	European Union	Copernicus Trade Marks	GIS software (QGIS, ESRI, etc)
39	Soil map	T8.2	1 to 10 GB	csv, xlsx	no	no	yes	Census data from the Italian National Institute of Statistics	European Union/Italy	Creative Commons attributes version 3.0	Microsoft Office
40	Municipality of Trento spatial data	T8.2	1 to 10 GB	tiff, gpkg, shp	no	no	yes	Municipality of Trento (Italy)	Municipality of Trento	General Public Licence	GIS software (QGIS, ESRI, etc)
41	Urban-public green areas	T3.1	239 MB	spatial data	no	no	yes	European Environment Agency	European Environment Agency	EEA standard re-use policy	n/a
42	Habitat type map of Greece	T9.2	n/a	vector	yes	no	yes	Sofia Municipality	Sofia Municipality	n/a	n/a

43	Data on subsidence in individual points, as measured with Sentinel-1	T4.1; 4.2; 4.3; 5.1; 5.2; 5.3; 8.1; 8.2;	n/a	raster/vector	no	no	n/a	https://esdac.jrc.ec.europa.eu/	European Commission	Creative Commons	n/a
44	MAES Estonia database	n/a	n/a	tif, shp	no	no	n/a	Public repositories	Estonian Environmental Agency	n/a	n/a
45	Sentinel-derived products	n/a	n/a	tif, grd	no	no	n/a	Copernicus public repositories	n/a	n/a	n/a

2 FAIR data

2.2 Making data findable

Metadata are standardised and structured dataset characteristics that explain the origin, purpose, time, geographic location, creator, terms of access, and terms of use of a data collection, to name a few. Metadata is commonly used to locate resources and to provide searchable information that helps users easily find existing data, and also as a bibliographic citation record. SELINA will guarantee data findability through the usage of both descriptive and structural metadata. Widely supported descriptional metadata standards have the clear advantage of being easily findable. In generating metadata, the project would aim to follow the EML (Ecological Metadata Language) unified metadata description standard.

An exemplary structure of the minimum characteristics of metadata is proposed below:

- Author(s)
- Year
- Dataset Title
- Data Repository or Archive
- Global Persistent Identifier
- Version, or Subset, and/or Access Date
- Language
- Metadata language
- Licence of use
- Date of metadata creation
- Hierarchy level
- Character encoding
- Format version
- Keywords (if possible)

The metadata generated within SELINA should be produced in an unified form for all project partners. Applying a metadata standard such as the EML - Ecological Metadata Language is recommended.

Naming documents in a standardised, logical, and intuitive way enables team members and collaborators to discover and manage project datasets when needed. SELINA supports the sharing of information consortium-wide and therefore suggests a uniform naming convention for all project-generated datasets.

Datasets processed within SELINA should follow the uniform naming convention: [SELINA_datatset.name_version_creation.date], whereby data format should be DDMMYYYY, numbering style should be 01, 02, 03.
Example: SELINA_Urban-public green areas_v.01_21102023.dbf

2.3 Making data openly accessible

SELINA members will aim to ensure open access to peer-reviewed scientific publications and all underlying datasets related to the project results and funded or co-funded by the project. Both the research paper and the underlying data have to be made available in open access as soon as possible after the paper has been published and no later than the end of the reporting period during which the paper was published. Partners will comply with the rule to deposit a machine-readable electronic copy of the published version or final peer-reviewed manuscript accepted for publication in a trusted open access repository.

SELINA partners will choose between two ways of making data openly available.

1. Uploading the underlying data to an open access research data repository under the Open Data Commons Attribution Licence (ODC-By)(e.g. a generic repository like Zenodo or a thematic repository such as Environmental Information Data Centre).
2. Publishing the datasets as open access data papers in an academic journal, then depositing data in an open access repository.

As seen in Table 1., the majority of partners indicated that they will publish datasets in open access. In the few cases where open access will not be provided, partners have provided an explanation why this is not possible. However, when the datasets are complete, partners will check to see whether datasets can be made available through, for example, aggregation of personal data from interviews or similar measures.

2.3 Making data interoperable

Except for the ESMERALDA Methods Database, which will be updated and made available through a new user interface, SELINA will not produce large databases that need to be made available through Application Programming Interfaces (API). We will rather have a set of small but well curated datasets, so there is no need to adopt a database management that is interoperable with other databases. Instead, SELINA will ensure interoperability of its research data by using common file formats specified in the data summary (e.g. docx, xlsx, pdf, .shp). This facilitates recombination of the data with other datasets from different origins. By using flat text data files (e.g. csv) linked to machine-readable metadata (e.g. EML) and hosted in repositories that provide programmatic access (e.g. DRYAD), we ensure that they can be queried and read by any programming language, and without use of proprietary

software.

These are some types of formats for long-term preservation of research data that we recommend to be used in SELINA.

It is recommended for SELINA partners to use the data formats derived from open source software, i.e. open file formats. The following is advised:

1) data science and data analysis scripts, e.g. .R, .py

2) structured text or mark-up file containing metadata information, e.g. DDI XML file, GeoRSS, Rmarkdown

3) quantitative tabular data, e.g. comma-separated values (CSV) file (.csv) tab-delimited file (.tab)

4) vector and raster data (essential: .shp, .shx, .dbf; optional: .prj, .sbx, .sbn); geo-referenced TIFF (.tif, .tiff, .tfw, .ddf) or CAD data (.dwg); tabular GIS attribute data

Another measure increasing interdisciplinary interoperability of data is the usage of standardised vocabulary for ecosystem service mapping and assessment terminology as, for instance, that developed within the project ESMERALDA, D1.4 Glossary for Ecosystem Service mapping and assessment terminology, (Potschin-Young et al., 2018).

SELINA should use standardised vocabulary as per the Glossary for Ecosystem Service mapping and assessment terminology, developed within ESMERALDA.

2.4 Increasing data re-use

Data produced within SELINA will be licensed using the Open Data Commons Attribution License ODC-By. If they want to publish data associated with a journal article under a license that is different from the Open Data Commons Attribution License (ODC-By), authors should explicitly inform the Project Coordination.

Other open data licenses are Creative Commons CC0 (also cited as “CC-Zero” or “CC-zero”) and the Open Data Commons Public Domain Dedication and License (ODC-PDDL). According to the CC0 license, “the person who associated a work with this deed has dedicated the work to the public domain by waiving all of his or her rights to the work worldwide under copyright law, including all related and neighbouring rights, to the extent allowed by law. You

can copy, modify, distribute and perform the work, even for commercial purposes, all without asking permission." Publication of data under a non-attribution waiver such as CC0 avoids potential problems of "attribution stacking" when data from several sources are aggregated for re-use, particularly if this re-use is undertaken automatically. In such cases, while there is no legal requirement to provide attribution to the data creators, the norms of academic citation best practice for fair use still apply, and those who re-use the data should reference the data source, as they would reference other research articles.

SELINA should use the open Data Commons Attribution License ODC-By.

3 Allocation of resources

Managing data in a FAIR way is associated with various types of costs. They can be grouped into two main categories: 1) article processing charges (APC) for publishing data in open access journals; 2) fees for depositing data in global data repositories. While many data repositories are free of charge to authors to upload their data (e.g. Zenodo, GBIF), some more generic data repositories such as Dryad charge users for publishing their data, unless exemptions apply. SELINA has carefully distributed direct costs among partners so that each partner has an open access publication costs budget (see SELINA Grant Agreement, Figure 3.3: Overview of SELINA budget categories (direct costs)). Each SELINA partner has to use this budget responsibly and prioritise open access publications.

In addition, sufficient resources have been allocated to WP7 leaders responsible for the Data Management Plan, so that updates of the deliverable are performed at least twice within the project duration, as well as upon major changes.

4 Data security

Data security is of high priority in SELINA, which is why partners replied to a series of questions concerning (i) data storage location, (ii) server location, (iii) back up procedures, (iv) data protection practices, (v) data protection responsible persons in the institution.

Partners' replies (a total of 38) can be summarised in the following way.

- The majority of partners store research data on their institutional servers (73%), opposed to web hosting on e.g. Microsoft Teams, Nextcloud or OneDrive.
- The servers of 95% of all partners are located within the EU.
- 86% of all partner institutions have established back up procedures, with a frequency between once a day and once a month.

- Data protection is ensured in all partner institutions either via two-factor authentication method, password protection, local network access or folder encryption.
- 95% of all partner institutions have pointed to a data management responsible person from their institution who is also responsible for the protection of SELINA data.

Overall, data security practices within the consortium are satisfactory. To maximise data security, the following recommendations are in place:

SELINA recommends that back up procedures are performed on a daily basis.

SELINA recommends that data is stored only on servers located in the EU.

SELINA recommends that all partners enquire with full accuracy who the official data protection officer (DPO) of their institution is and obtain their contacts.

5 Ethical aspects

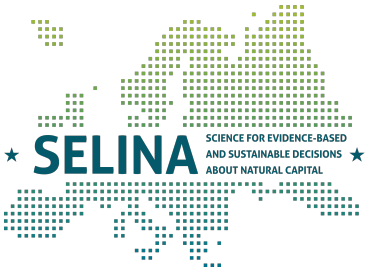
SELINA has not identified any ethical or legal issues related to data management so far. Should issues occur, consortium members will consult the project's Ethics Advisor Dr. Scarlett Sett (consult D11.1 for more information).

6 References

Potschin-Young, M., Burkhard, B., Czúcz, B. and F. Santos Martín (2018). Glossary for Ecosystem Service mapping and assessment terminology. Deliverable D1.4 EU Horizon 2020 ESMERALDA Project, Grant agreement No. 642007, 49 pp.

Annex 1: Data Management Questionnaire

SELINA Data Management Survey



This questionnaire aims to collect information about the various types of data that will be collected, generated or stored by SELINA members. Based on the feedback, a Data Management Plan (DMP) will be created to: 1) document the ownership, licensing and use of the project data; 2) describe the metadata; 3) store safely and enable subsequent use of the research data. The DMP will also define the datasets to be published for open use and the chosen trusted repository. It is a living document that will be updated throughout the project's duration.

There are 13 questions in this survey and it is vital that each partner provides as detailed a response as possible, as this will guide the project's data management practices. We realise you probably don't have the exact answers to some of these questions yet, so feel free to give us your preliminary estimate, which you can then modify or specify during the DMP update in June 2025.

The survey has 13 questions and should take about 20 minutes to complete.

There are 17 questions in this survey.

SELINA Data Management

1. First and last name *

Please write your answer here:

2. Organisation/institution *

Please write your answer here:

Questions 3 and 4 concern the data you will **generate**.

3. Please provide the following provisional information for your **generated data**:

1. **Name of the dataset**

2. **Name of the generator**: name of the person who will generate this data.

3. **Relevant task**

4. **Generated via**: for example, field work, modelling, data processing, remote sensing, literature review, policy review, interview, surveys.

5. **Size**: a rough estimate only if you know.

6. **Format**: for example, .docx; .xlsx; .pdf; .mp4; .xml; .csv.

7. **Type of data**: for example, qualitative data; semi-quantitative data; quantitative data; analogue data; digital data; GIS data.

8. **Sensitive data**: Yes/no. If yes, please specify, for example, racial, political, ethical, health, and more here (https://ec.europa.eu/info/law/law-topic/data-protection/reform/rules-business-and-organisations/legal-grounds-processing-data/sensitive-data/what-personal-data-considered-sensitive_en).

9. **Personal data**: Yes/no. If yes, please specify, for example, name, surname, address, email, IP address, location data.

10. **Delivery**: a rough estimate of a timeline. If there is an embargo period, specify why and how long it will apply.

11. **Metadata**: the metadata accompanying your datasets.

12. **Users**: to whom they might be useful.

13. **Access**: will they be open access? If not, please indicate the reasons, for example, ethical, rules of personal data, intellectual property, commercial, privacy-related, security-related, contract.

14. **Re-use**: potential documents or tools needed to re-use or validate the data.

ⓘ Please fill in at least one answer

	Dataset 1	Dataset 2	Dataset 3	Dataset 4	Dataset 5
Name of the dataset					
Name of the generator					
Relevant task					
Generated via					
Size					
Format					
Type of data					
Sensitive data					
Personal data					
Delivery					
Metadata					
Users					
Access					
Re-use					

This question concerns the data you will **generate**.

4. Would you consider anonymising your **generated datasets** and publishing them in an aggregated form so as to not disclose private information?

If yes, please specify to which dataset(s) that would apply. If not, please explain why. *

ℹ Choose one of the following answers
Please choose **only one** of the following:

- ◯ Yes
- ◯ No
- ◯ Not relevant

Make a comment on your choice here:

This question concerns the data you will **generate**.

Question 5 concerns the data you will obtain from elsewhere and **reuse**.

5. Please provide the following information for the **existing data you will reuse**:

1. **Name of the dataset**
2. **Relevant task**
3. **Size**
4. **Format**: for example, .docx; .xlsx; .pdf; .mp4; .xml; .csv.
5. **Sensitive data**: Yes/no. If yes, please specify, for example, racial, political, ethical, health, and more here (https://ec.europa.eu/info/law/law-topic/data-protection/reform/rules-business-and-organisations/legal-grounds-processing-data/sensitive-data/what-personal-data-considered-sensitive_en).
6. **Personal data**: Yes/no. If yes, please specify, for example, name, surname, address, email, IP address, location data.
7. **Metadata**: the metadata accompanying your datasets.
8. **Access**: open/restricted/closed access.
9. **Origin**: what is the origin of the data?
10. **Ownership**: who owns the data you will reuse?
11. **Licence**: under what licence can you use the data?
12. **Re-use**: potential documents or tools needed to re-use or validate the data.

ⓘ Please fill in at least one answer

	Dataset 1	Dataset 2	Dataset 3	Dataset 4	Dataset 5
Name of the dataset					
Relevant task					
Size					
Format					
Sensitive data					
Personal data					
Metadata					
Access					
Origin					
Ownership					
Licence					
Re-use					

This question concerns the data you will obtain from elsewhere and **reuse**.

Questions 6-13 concern your **data management practices**.

6. Please provide a brief summary of your **institutional data management practices**, specifying:

1. **Data location**: where and how data will be stored, for example, institutional server or web hosting.
2. **Server location**: EU or non-EU; compliant or not-compliant with applicable data protection rules (for example, GDPR).
3. **Backup procedures**: type of backup procedures and their frequency.
4. **Protection**: how data security is ensured, for example, password or two-factor authentication.
5. **Responsible**: name the person from your team who will bear primary responsibility for data management and serve as a contact person if questions arise.

*

Data location	
Server location	
Backup procedures	
Protection	
Responsible	

This question concerns your **institutional data management**.

7. Do you follow a specific naming convention?

*If yes, please specify. **

Choose one of the following answers
Please choose **only one** of the following:

- ◯ Yes
- ◯ No

Make a comment on your choice here:

This question concerns your **data management**.

8. Do you use any standard metadata vocabulary, standards or methodologies when creating your datasets?

*If yes, please specify. **

Choose one of the following answers
Please choose **only one** of the following:

- ◯ Yes
- ◯ No

Make a comment on your choice here:

This question concerns your **data management**.

9. Will you be using standard vocabularies for all data types present in your data set, to allow inter-disciplinary interoperability?

If yes, please specify. *

ℹ Choose one of the following answers
Please choose **only one** of the following:

◯ Yes
◯ No

Make a comment on your choice here:

This question concerns your **data management**.

10. How will you licence your data?

If other, please provide a justification.

*

ℹ Comment only when you choose an answer.
Please choose all that apply and provide a comment:

☐ Creative Commons Attribution International Public License (CC BY) (or equivalent)

☐ Creative Commons Public Domain Dedication (CC 0) (or equivalent)

☐ Other

This question concerns your **data management**.

11. Would you be interested in publishing your data in the form of data papers?

If yes, please give an example of a suitable dataset. If not, please explain why. *

ℹ Choose one of the following answers
Please choose **only one** of the following:

◯ Yes
◯ No

Make a comment on your choice here:

This question concerns your **data management**.

12. Do you have a preference for a trusted repository where to store your research data?

*If yes, please specify. **

Choose one of the following answers
Please choose **only one** of the following:

- ◯ Yes
- ◯ No

Make a comment on your choice here:

This question concerns your **data management**.

13. Can you identify potential obstacles (e.g., technical, social, policies) that would prevent delivering FAIR data during SELINA's lifetime and beyond? Information on FAIR data here (https://www.openaire.eu/how-to-make-your-data-fair).

If yes, please specify.

*

Choose one of the following answers
Please choose **only one** of the following:

- ◯ Yes
- ◯ No

Make a comment on your choice here:

This question concerns your **data management**.

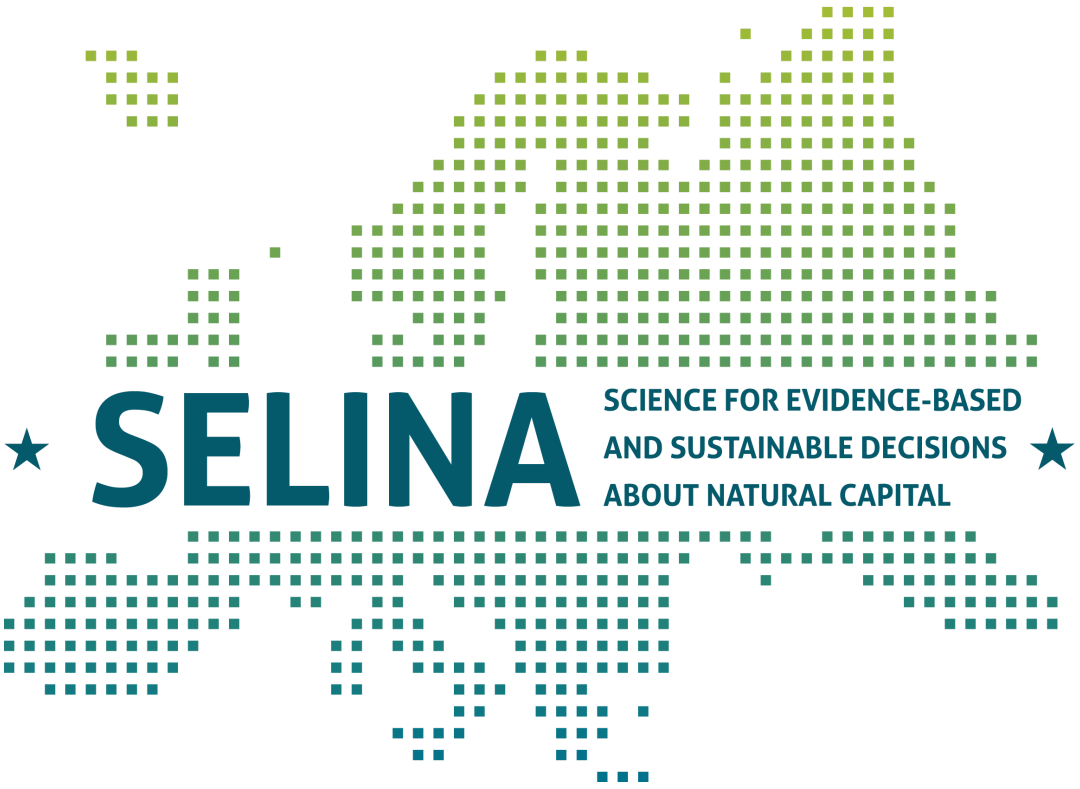
Any additional comments?

Please write your answer here:

Thank you for filling in SELINA's data management questionnaire!

If you wish to receive a PDF copy of your replies, please write to s.antoniadou@pensoft.net.

Submit your survey.
Thank you for completing this survey.



https://project-selina.eu/