D8.3 Data Management Plan

Deliverable Administration

No and name:	D8.3 Data Management Plan
Related task in DoA:	Task 8.4 Development/updating of Data Management Plan
Due date of deliverable:	30.11.2023
Submission date:	30.11.2023
Responsible organization:	GTK
Author name(s):	Marianne Valkama[1] , Timo Tarvainen[1] [1] Geological Survey of Finland (GTK)
Revision number:	1
Status:	Draft
Dissemination level:	SEN - Sensitive

Revision History

Version	Date	Modified by	Comments
0.1	20.10.2023	M. Valkama, T. Tarvainen (GTK)	First draft
1.0	3.11.2023	J.Kaija & M. Valkama	Final version

Disclaimers

- Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or European Research Executive Agency (REA). Neither the European Union nor the granting authority can be held responsible for them.
- UK partner in ISLANDR is supported by UKRI grant.
- Swiss partner in ISLANDR is supported by the Swiss grant.

Funded by the European Union

Table of Contents

Funded by the European Union

Funded by the European Union

List of Acronyms

Abbreviation	Description
CBA	Cost Benefit Analysis
CEC	Contaminants of Emerging Concern
EA	Ethics Advisor
ECTS	ECTS (European Credit Transfer and Accumulation System)
EU	European Union
EUSO	EU Soil Observatory
GA	Grant Agreement
OEI	Other Ethics Issues
SRBLM	Sustainable and Risk-Based Land Management
TENK	The Finnish National Board on Research Integrity
WP	Work package

Funded by the European Union

ISLANDR project in brief

The Information-based Strategies for Land Remediation, in short ISLANDR, is a multidisciplinary project, which is foremost aimed at supporting the execution of the EU mission: *A Soil Deal for Europe*.

More specifically, the ISLANDR research activities are designed to provide tools and methods to support: (1) the delineation of polluted soils across Europe, (2) an evidence-based assessment of the risks posed by polluted soils, (3) the promotion of sustainable and risk-based land management practices, (4) the inclusion of a wider valuation approach in financial and investment cases, and (5) a closer integration of land contamination and spatial planning decision-making. Lessons learnt and experience gained throughout the project duration will be used to (6) deliver key policy-relevant findings related to the Soil Strategy, the proposed Soil Health Law, and other areas of policy where soils are crucial.

To road-test the project's findings, seven test areas across Europe have been identified. To begin with, the ISLANDR Test Areas (ITAs) will provide a real-world context for the planned research activities. More concretely, the ITAs have been selected to cover different land use types, such as urban, peri-urban, rural, agro-forestry, mining, wetlands and coastal areas. Furthermore, the ITAs are characterized by both point and diffuse sources of pollution, as well as by different soil pollution types, such as organic, inorganic, as well as contaminants of emerging concern.

Furthermore, ISLANDR brings a dedicated focus to low input remediation, by including test areas impacted by the consequences of the green transition, such as former mining areas. This will ensure that soil remediation will be facilitated even when the cost of remediation is economically marginal or may even be negative. On the one hand, this necessitates a more thorough understanding of low-input remediation approaches from a technological perspective, yet it also requires a wider value proposition for investment cases and financial planning.

Key actors, stakeholders and end-beneficiaries are at the epicentre of ISLANDR. Through roundtables in the respective ITAs, the foremost assignment of local actors will be to provide feedback and offer insights as to the robustness and effectiveness of the strategies, frameworks and decision-support tools, as well as on the wider valuation approaches and financing mechanisms to be developed over the course of the project's lifetime. Thus, the Roundtables are foreseen to bring an iterative feedback loop to the research process, with a view to ensure the wider uptake of the project's outcomes and achievements.

Last but not least, local communities in the respective ITAs will be invited to participate in a survey organized both during the early stages and towards the end of the project, as a means to document soil literacy among society thereby bringing insight as to whether the

Funded by the European Union

exposure of society to the project's activities on the ground can bring about a strongly desired 'awareness pull' to the benefits to be reaped from healthy soils, thereby leveraging society at large to subscribe to the projects' motto: ISLANDR for Soil Health!

ISLANDR was officially launched on the 1st of May 2023 with a foreseen project duration of 3 years. The project brings together 14 partners from 13 countries, including 10 EU Member States. The project receives €5.8 million in EU funding through the Horizon Europe programme. In addition, the partners based in the UK and Switzerland are foreseen to contribute €1.1 million in funding to the project.

Project partners

	Participant organisation name	Type	Country
1	GEOLOGIAN TUTKIMUSKESKUS (GTK) - COORDINATOR	RES	FI
2	BUREAU DE RECHERCHES GÉOLOGIQUES ET MINIERÈS (BRGM)	RES	FR
3	STICHTING DELTARES (DELTARES)	RES	NL
4	ETHNIKO KENTRO EREVNAS KAI TECHNOLOGIKIS ANAPTYXIS (CERTH)	RES	GR
5	CHALMERS TEKNISKA HOGSKOLA AB (CHALMERS)	UNI	SE
6	UNIVERSIDADE NOVA DE LISBOA (NOVA)	UNI	PT
7	GREENDECISION SRL (GD)	SME	IT
8	INSTYTUT UPRAWY NAWOZENIA I GLEBOZNAWSTWA, PANSTWOWY INSTYTUT BADAWCZY (IUNG)	RES	PL
9	INTEGRATED RESOURCES MANAGEMENT (IRM) COMPANY LIMITED (IRMCo)	SME	MT
1	SUOMEN YMPARISTOKESKUS (SYKE)	RES	FI
1	SANTERRA	SME	BE
1	IRN	SME	XK

Funded by the European Union

1	TEMAS SOLUTIONS GMBH (TEMASOL) (Associate partner)	SME	CH
1	R3 ENVIRONMENTAL TECHNOLOGY LIMITED (R3) (Associate partner)	SME	UK

Deliverable Summary

Deliverable description in the EU Funding & Tenders portal states: "*Initial DMP under the FAIR principles and in compliance with OA policy.*"

The objective of this Data Management Plan (DMP) is to act as guidance for the Consortium Partners in the application of research data management (RDM) practices. The document covers different aspects of RDM, namely, data types, data sharing, data re-use, allocation of resources and data security, in addition to ethical, legal and societal aspects of the data.

The DMP is a living document that will be updated whenever a need arises. Examples of situations where the DMP would be updated include, but are not limited to new datasets, changes to data storage or data sharing practices, and changes in the level of openness of individual datasets. In the ISLANDR project datasets can be fully open, partially open or fully closed. A new version of the DMP will be released and shared each time significant changes take place. The different versions of the DMP will be shared in the CORDIS system of the European Commission as well as in the Research Ideas and Outcomes (RIO) journal.

Funded by the European Union

1 Introduction

1.1 Purpose of the Data Management Plan

According to the Grant Agreement (GA), the initial Data Management Plan (DMP) must be available by month 6. It will be followed by the Updated Data Management Plan (Deliverable D8.4) due in Month 25.

The objective of this Data Management Plan (DMP) is to act as guidance for the Consortium Partners in the application of research data management (RDM) practices. The document covers different aspects of RDM, namely, data types, data sharing, data re-use, allocation of resources and data security, in addition to ethical, legal and societal aspects of the data.

The DMP is a living document that will be updated whenever a need arises. Examples of situations where the DMP would be updated include, but are not limited to new datasets, changes to data storage or data sharing practices, and changes in the level of openness of individual datasets. In the ISLANDR project datasets can be fully open, partially open or fully closed. A new version of the DMP will be released and shared each time significant changes take place. The different versions of the DMP will be shared in the CORDIS system of the European Commission as well as in the Research Ideas and Outcomes (RIO) journal.

1.2 Relation to other project documents and deliverables

The Consortium Agreement and Grant Agreement have elements in them that broadly fall into the domain of a DMP. These elements include confidentiality, intellectual property rights, patents, dissemination of the results and access rights to the results. In the event of discrepancy between the documents, the DMP is overruled by the Consortium Agreement, including its annexes and possible addendums.

Deliverable D9.1 – D9.3 (Ethics Requirements OEI No. 1- No. 3): The ethics advisor must be consulted at least on the following points: Humans, Personal Data and Environment, Health and Safety. This includes the ethics issues related to the involvement of external human participants (e.g. via interviews and questionnaires), A report by the ethics advisor will be submitted as a deliverable at the end of each reporting period.

1.3 Roles and Responsibilities

Following good RDM practices, outlined in this document, is the responsibility of each Consortium Partner. GTK is the Coordinator of the project, but also one of the Consortium Partners, depending on the context.

Funded by the European Union

The first version of the DMP is drafted by the Coordinator in collaboration with the other Consortium Partners (Table 1). After the first version, it is the responsibility of each project partner to notify the Coordinator if updates are needed.

Table 1 The distribution of responsibilities in the ISLANDR project.

Consortium Partner	Coordinator
Ensures that the DMP is in agreement with any relevant institutional policies (e.g. data management handbooks).	**Drafts** the first version of the DMP together with the Consortium Partners.
Informs the Coordinator about any relevant changes to institutional open science and RDM guidelines.	**Updates** the DMP based on feedback from the Consortium Partners. **Shares** the DMP in CORDIS and RIO whenever there are relevant updates.
Informs the Coordinator about any research outputs at least 21 days prior to sharing them.	**Approves** the sharing of datasets and other research outputs together with the other Management Committee members (see the Project Handbook for details). This will be done as soon as possible and no later than within 21 days of the request.
Shares their own research outputs in recommended data repositories.	**Ensures** that the list of datasets and project outputs are always up-to-date.

1.4 Licensing

The licensing policy of the ISLANDR project follows the guidance of the EC, as outlined in the Annotated Grant Agreement and Programme Guide documents. As listed in Table 2, the CC0 license will be used for data and metadata, whereas for publications, the CC BY 4.0-license will be used. However, open access publishing is beyond the scope of this DMP document and will not be discussed further.

As the EC does not impose software licensing obligations, permissive licenses will be used to license software and codes in the project.

Table 2. The licenses used in the ISLANDR project.

License	Explanation	Used for
CC0 1.0 Universal Public Domain Dedication	"No copyright. The person who associated a work with this deed has dedicated the work to the public domain by waiving all his or her rights to the work worldwide under copyright law, including all related and neighboring rights, to the extent allowed by law. You can copy, modify, distribute, and perform the	Data and metadata

Funded by the European Union

	work, even for commercial purposes, all without asking permission".	
CC BY Attribution 4.0 International	"Free to Share — copy and redistribute the material in any medium or format Adapt — remix, transform, and build upon the material for any purpose, even commercially, under the following terms: Attribution — You must give appropriate credit, provide a link to the license, and indicate if changes were made. You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use. No additional restrictions — You may not apply legal terms or technological measures that legally restrict others from doing anything the license permits". (https://creativecommons.org/licenses/by/4.0/)	Publications (open access publishing will be discussed in more detail in the Project Handbook).

1.5 File and folder naming conventions

The project uses a self-explanatory folder structure to safeguard its data. Each work package has its own structure. The exact content of the site folders depends on the data that will be acquired in them. The folder structure is aimed at internal sharing (within the consortium) and storing of the data; hence, it is not used to share data with third parties. The shared project data are stored in the online collaboration platform hosted by Greendecision. After the main project period, the data will be stored in a cloud server for a minimum of 10 years or for as long as the data have scientific value. This switch from one platform to another reflects an IT policy change of the Coordinator.

Within the cloud services, file names follow an agreed consistent format, which comprises the following elements:

- File names should be relatively short but descriptive (<35 characters)
- Special characters and spaces are not allowed.
- Underscores will be used instead of periods, spaces, or slashes.
- Use date format ISO 8601: YYYYMMDD.

This naming is modified from https://libguides.princeton.edu/c.php?g=102546&p=930626 and applied also in the EU MultiMiner project (Grant Agreement No: 101091374).

Funded by the European Union

2 The Data Management Plan

2.1 General framework for data collection

In this section, the data to be collected in each WP of the ISLANDR project will be presented and described in order to define the purpose of the collection as well as to previously define the type, format and origin.

This knowledge will be semi-structured (e.g., formatted WP datasets) or non-structured (documents, graphs, tables, pictures, etc.), and properly indexed in order to be retrievable and exploitable.

Concerning the type of data to be produced, it is equally split between reports and internal databases followed by articles, web service, video clips and photo albums. Most of the data will be in pdf format followed by .docx, xlsx and then jpeg/avi/wmv and similar audio/video formats.

A preliminary table for new data gathering from the test sites has been created. It contains information e.g. on the following: Site location, Type of survey, Names of organizations who need data, Names of organizations who deliver the data and Contact persons.

WP1 – Overview of soil pollution in Europe

WP1 will provide an overview of the state of play of various types of soil pollution in Europe. The work will be strongly data-oriented and consider both diffuse and local soil contamination. WP1 will establish a robust methodology for identifying hotspots with reference to natural or anthropic chemical backgrounds, defined at relevant scales of investigation. After identifying the most important data gaps, WP1 will suggest methodologies to collect EUSO-compatible data on soil pollution. The Metadata catalogue of the European soil pollution data sources will serve as one input to WP2 and WP5.

WP2 – Risk-based contaminated land assessment for the promotion of soil health

WP2 will build upon well-established risk-based assessment methodologies specifically tailored for contaminated soil and land with enhanced insight into the Source-Pathway-Receptor (SPR) approach, in order to:

- Prioritise contaminants of emerging concern (CECs) in soil;
- Upscale existing risk-based approach in order to perform risk assessment for a large set of soil quality data at different spatial scales and
- Develop a risk-based soil health assessment.

Funded by the European Union

WP3 – Sustainable and risk-based remediation

The main objective of WP3 is to develop a systematic and integrated sustainable and risk-based decision-making framework for the restoration and decontamination of both point source and diffuse land contamination.

WP4 – Financial models for decontamination and reuse of land

The overall aim of WP4 is to develop and demonstrate financial models that can be used for prioritising decontamination and reuse of land and promote remediation action on the ground.

WP5 – Spatial planning models, prioritization and approaches for contaminated soil and land reuse

WP5 aims to explore and integrate approaches promoting transition in the thematic of contaminated soil and land reuse.

WP6 – Policy support and dialogue

WP6 aims to ensure the fulfilment of the EU's policy objectives linked to contaminated soil as well as promote and ensure the broad adaptability of the outcomes of ISLANDR.

WP7 – Communication, dissemination and exploitation

Overall objective is to ensure that the promotion of ISLANDR's solutions is not only driven by a 'technological push' effort, but an equal - if not higher level of attention - is given to the 'demand pull' effort.

WP8 – Project coordination

WP7 ensures the achievement of the project's objectives, in terms of quality, timely delivery, and contribution to the expected impact of the project. No specific research data is planned to be collected from WP8, as it is devoted to management. Personal GDPR-compliant data from project partners, advisory boards and interested stakeholders, will be collected. There is no further processing of personal data, and it is not planned to export personal data from the EU to non-EU countries or vice versa.

2.2 Use and Re-Use of the data

The data collected and generated by the consortium will be useful to the development of further activities related to other WPs within the project, to specific end-users and to some EU regulators.

Some existing datasets are proposed to be used, such as GEMAS data, data gathered during other EU-projects (e.g., H2020 project MIREU – Mining and Metallurgical Regions of EU) and the background data provided by the ISLANDR partners. The project will also

Funded by the European Union

register the institutional publication and data repositories of each partner, in order to provide information about the current systems that the project partners maintain.

Furthermore, ISLANDR will benefit from developments for other past and ongoing EU projects in order to maximize the efficiency of EC funding.

As part of WP6, data will be re-usable only within the consortium, and for ISLANDR-related activities, such as newsletter dissemination, poster presentations, invitations to events or specific news on the website.

2.3 Data summary

Data summary
Will you re-use any existing data and what will you re-use it for? State the reasons if re-use of any existing data has been considered but discarded.
The project will use re-existing data as detailed in Table 3.
What types and formats of data will the project generate or re-use?
Listed in Table 3.
What is the purpose of the data generation or re-use and its relation to the objectives of the project?
In WP1, the project will use existing data on contaminants in topsoil in the modelling and produce estimated areas of diffuse soil contamination as a GIS layer.
What is the expected size of the data that you intend to generate or re-use?
Listed in Table 3.
What is the origin/provenance of the data, either generated or re-used?
Listed in Table 3.
To whom might your data be useful ('data utility'), outside your project?
The data might be useful to the different stakeholders of the project, namely the landowners, consulting companies, geological surveys, governmental organizations, non-governmental organizations, companies working with remediation of contaminated sites, the scientific and R&D communities, policy / decision makers and local authorities.

2.4 FAIR Data

2.4.1 Making data findable, including provisions for metadata

Will data be identified by a persistent identifier?

The recommended data repository of the project is Zenodo (https://zenodo.org) which provides DOIs for every published record.

Funded by the European Union

Will rich metadata be provided to allow discovery? What metadata will be created? What disciplinary or general standards will be followed? In case metadata standards do not exist in your discipline, please outline what type of metadata will be created and how.

The ISLANDR project will follow ISO 19(115) standards and INSPIRE profile for metadata. At present no specific soil contamination extension is foreseen. ISLANDR will use Geonetwork architecture.

Will search keywords be provided in the metadata to optimize the possibility for discovery and then potential re-use?

Yes.

Will metadata be offered in such a way that it can be harvested and indexed?

Yes.

2.4.2 Making data accessible

Will the data be deposited in a trusted repository?

Yes.

Have you explored appropriate arrangements with the identified repository where your data will be deposited?

The metadata catalogue of ISLANDR will be on GreenDecisions server after the project in a static manner.

Does the repository ensure that the data is assigned an identifier? Will the repository resolve the identifier to a digital object?

The data users for the public or Open Access data will be able to locate the original data sources, as links to these web services will be included in deliverables and the project website. Scientific publications will include Digital Object Identifiers (DOI) that lead to the associated open data sets.

Will all data be made openly available? If certain datasets cannot be shared (or need to be shared under restricted access conditions), explain why, clearly separating legal and contractual reasons from intentional restrictions. Note that in multi-beneficiary projects it is also possible for specific beneficiaries to keep their data closed if opening their data goes against their legitimate interests or other constraints as per the Grant Agreement.

Project partners may acquire restricted data from ISLANDR Test Areas that will be used

Funded by the European Union

by the contact partner only. The restricted data may contain confidential information that cannot be released to third parties.

A consortium agreement signed at the start of the project, will inter alia specify the terms and conditions pertaining to ownership, access rights, exploitation of background and results and dissemination of results, in compliance with the grant agreement and Regulation n°1290/2013 of December 11th, 2013. Information exchanged during closed consultation with external participants (e.g. Stakeholder workshops, 1:1 interviews), will be subject to a Non-Disclosure Agreement. The classification level of the information to be shared will be defined by the Management Committee (public or restricted) and monitored by the Project Coordinator and the Project Manager.

Concerning the type of data to be produced, it is equally split between reports and internal databases followed by articles, web service, video clips and photo albums. Most of the data will be in pdf format followed by .docx, xlsx and then jpeg/avi/ wmv and similar audio/video formats.

If an embargo is applied to give time to publish or seek protection of intellectual property (e.g. patents), specify why and how long this will apply, bearing in mind that research data should be made available as soon as possible.

During the Project and for a period of 1 year after the end of the Project, the dissemination of own Results by one or several Parties including but not restricted to publications and presentations, shall be governed by the procedure of Article 17.4 of the Grant Agreement and its Annex 5, Section Dissemination, subject to the following provisions.

Prior notice of any planned publication, including a draft of the proposed publication, shall be given to the Management Committee and the Parties concerned at least 45 calendar days before the submission of the manuscript for publication. Any objection to the planned publication shall be made in accordance with the Grant Agreement by written notice to the Coordinator and to the Party or Parties proposing the dissemination within 30 calendar days after receipt of the notice. If no objection is made within the time limit stated above, publication is permitted. Where necessary, the Management Committee may propose and arrange an expedited procedure for obtaining the Parties' consent.

Will the data be accessible through a free and standardized access protocol?
Yes.

If there are restrictions on use, how will access be provided to the data, both during and after the end of the project?

The background data provided by the partners is confidential. Geological, geochemical, spatial planning and financial data gathered from the ISLANDR Test Areas may also be

Funded by the European Union

subject to confidentiality rules and in some cases cannot be released or given to a third-party use without permission from the relevant test site owner and the Management Committee.

How will the identity of the person accessing the data be ascertained?

According to the GDPR, the data controller should use all reasonable measures to verify the identity of a data subject who requests access, particularly in the context of online services and online identifiers. The following methods can be used to identify the person accessing the data: Knowledge-based authentication and Multi-factor authentication.

Is there a need for a data access committee (e.g. to evaluate/approve access requests to personal/sensitive data)?

There is no need for a data access committee.

Will metadata be made openly available and licensed under a public domain dedication CC0, as per the Grant Agreement? If not, please clarify why. Will metadata contain information to enable the user to access the data?

Yes. Metadata will be made openly available.

How long will the data remain available and findable? Will metadata be guaranteed to remain available after data is no longer available?

The metadata catalogue will be provided to be used as a part of the Soilwise project repository and finally as a part of EUSO.

Will documentation or reference about any software be needed to access or read the data be included? Will it be possible to include the relevant software (e.g., in open source code)?

No for both questions.

2.4.3 Making data interoperable

What data and metadata vocabularies, standards, formats or methodologies will you follow to make your data interoperable to allow data exchange and re-use within and across disciplines? Will you follow community-endorsed interoperability best practices? Which ones?

ISLANDR project will follow ISO 19(115) standards and INSPIRE profile for metadata. At present no specific soil contamination extension is foreseen. ISLANDR will use Geonetwork architecture.

In case it is unavoidable that you use uncommon or generate project specific ontologies or vocabularies, will you provide mappings to more commonly used ontologies? Will you openly publish the generated ontologies or vocabularies to allow reusing, refining or extending them?

Funded by the European Union

There is a need for Ontology in the project. This has already started in WP1, WP2 and WP5. In fact, one of the deliverables of WP1 is D1.3 – Ontology. Mappings to more commonly used ontologies will be provided. The ontology document will have 3 hierarchy levels:

- 1st level of hierarchy: soil and land and its use
- 2nd level of hierarchy: soil management
- 3rd level of hierarchy: achieving the first six Soil Mission objectives

Will your data include qualified references to other data (e.g., other data from your project, or datasets from previous research)?

Yes, it will include. See details in Table 3 and chapter Use and Re-Use of the data.

2.5 Increase data re-use

How will you provide documentation needed to validate data analysis and facilitate data re-use (e.g. readme files with information on methodology, codebooks, data cleaning, analyses, variable definitions, units of measurement, etc.)?

Data validation is the process of checking the accuracy, quality, and completeness of data before using it for analysis or reporting. Data validation will be documented, and it will be done in the WPs in question. Furthermore, the ISLANDR Management Committee is responsible for reviewing the planned publications before they are sent for publication.

Will your data be made freely available in the public domain to permit the widest re-use possible? Will your data be licensed using standard reuse licenses, in line with the obligations set out in the Grant Agreement?

The metadata catalogue will be provided to be used as a part of Soilwise project repository and finally as a part of EUSO.

Will the data produced in the project be usable by third parties, in particular after the end of the project?

Yes. The ISLANDR website serving the data and deliverables will be in operation also after the end of the project. All deliverables will be uploaded to the EU´s Funding and Tenders portal, and after the Project Manager has approved them, the public deliverables can be accessed through the EU´s CORDIS portal.

Will the provenance of the data be thoroughly documented using the appropriate standards?

Yes.

Describe all relevant data quality assurance processes.

Funded by the European Union

- Determining and screening anomalies by means of data profiling,
- Removing obsolete information, removing outliers, missing data interpolation
- Data integrity checks
- Calibration procedures
- Data capture resolution and repetitions
- Checks and corrections of transcripts.

2.6 Other research outputs

In addition to the management of data, beneficiaries should also consider and plan for the management of other research outputs that may be generated or re-used throughout their projects. Such outputs can be either digital (e.g. software, workflows, protocols, models, etc.) or physical (e.g. new materials, antibodies, reagents, samples, etc.).

Beneficiaries should consider which of the questions pertaining to FAIR data above, can apply to the management of other research outputs, and should strive to provide sufficient detail on how their research outputs will be managed and shared, or made available for re-use, in line with the FAIR principles.

2.7 Allocation of resources

What will the costs be for making data or other research outputs FAIR in your project (e.g. direct and indirect costs related to storage, archiving, re-use, security, etc.)? How will these be covered? Note that costs related to research data/output management are eligible as part of the Horizon Europe grant (if compliant with the Grant Agreement conditions)

The ISLANDR project has been designed so that the data is FAIR (findable, accessible, interoperable and re-usable), but it is not yet possible to express in numbers what is the cost of that.

Who will be responsible for data management in your project?

In ISLANDR, GTK as the Coordinator has the main responsibility of data management, however, the decisions will be made by the Management Committee.

How will long term preservation be ensured? Discuss the necessary resources to accomplish this (costs and potential value, who decides and how, what data will be kept and for how long)?

The ISLANDR website will be hosted on IRMCO´s server during the lifetime of the project. A lasting solution for hosting the ISLANDR data beyond the lifetime of the project will have to be sought. Most likely part of the data produced by the ISLANDR will be stored in EUSO´s servers.

2.8 Data security

What provisions are or will be in place for data security (including data recovery as well as

Funded by the European Union

secure storage/archiving and transfer of sensitive data)? Will the data be safely stored in trusted repositories for long term preservation and curation?

The servers serving the data are secured by standard means such as firewalls and automatic backing up. The collected personal information will be stored on password-protected devices, and it will not be used or made available, for any other reason, without additional permissions (avoidance of mission creep). Data that can lead to individuals will be deleted after the end of the project. However, this data can be stored for other research purposes with the informant's consent.

2.9 Ethical, legal and societal aspects

Are there, or could there be, any ethics or legal issues that can have an impact on data sharing? These can also be discussed in the context of the ethics review. If relevant, include references to ethics deliverables and ethics chapter in the Description of the Action (DoA).

ISLANDR has nominated an experienced Ethics Advisor for the project who will assist with any questions related to ethics.
Will informed consent for data sharing and long-term preservation be included in questionnaires dealing with personal data?

Yes. When necessary personal data is collected, an individual's free fully informed consent needs to be collected. The values, rights and interests of all research participants needs to be protected by minimizing the collected personal data, processing and storing it carefully and making sure that identification of a natural person is not possible in the project outcomes.

2.10 Other issues

Do you, or will you, make use of other national/funder/sectorial/departmental procedures for data management? If yes, which ones (please list and briefly describe them)?

Not applicable.

3 The Research Output of the Project

In this section, all research outputs of the ISLANDR project will be collected. Research outputs include, but are not limited to, publications, data, software, codes and models. This section will be updated as the project progresses.

3.1 Data

Funded by the European Union

To be updated as the project progresses**. It is not expected that ISLANDR will produce any publicly available raw data.**

3.2 Software and algorithms

To be updated as the project progresses. **It is expected that algorithms will be developed under WP1 and WP2.**

Funded by the European Union

Table 3 The ISLANDR datasets, their origin, formats, sizes and level of confidentiality. TBD=to be defined.

	Dataset/type of data	Origin	Format	Size	Confidential	Rationale for confidentiality	Comments
1	GEMAS – geochemistry of agricultural and pastureland topsoils in Europe	Acquired from EuroGeoSurveys	Excel	TBD	No	N/A	
2	MIREU data on mining regions	MIREU project website and EU Cordis website	Pdf	N/A	No	N/A	GTK coordinated MIREU project and knows what these data contain.
3	REMIX data on mining regions	REMIX INTERREG Europe project website	Pdf	N/A	No	N/A	
4							
5							
6							
7							
8							