



Data Management Plan

Authors: **Mariano Iossa (FiBL Europe)**

Deliverable Number	D8.5
Work Package	WP8
Deliverable type	Report
Dissemination level	PU - Public
Deliverable Lead partner	FiBL Europe
Due date	31 March 2023
Submission date	31 March 2023
Version	V1.3
Reviewers	Monika Messmer (FiBL CH), Ana Marija Špicnagel Ćurko (IPS)
Contact	Mariano.iossa@fibl.org

History of changes

Version	Date	Author	Comments
V1.0	6 March 2023	Mariano Iossa	Draft version
V1.1	21 March 2023	Ana Marija Špicnagel Ćurko	Revision of v1.0 version
V1.2	30 March 2023	Monika Messmer	Revision of v1.1 version
V1.3	31 March 2023	Mariano Iossa	Final version of 1 st DMP

LiveSeeding - Organic seed and plant breeding to accelerate sustainable and diverse food systems in Europe is a 4-year Innovation Action funded by the European Union, the Swiss State Secretariat for Education, Research and Innovation (SERI) and UK Research and Innovation (UKRI). The project started in October 2022 and brings together 37 organisations operating in 16 European countries. LiveSeeding provides science-based evidence and best practice solutions to help achieve 100 % organic seed.

LiveSeeding contributes to the transition towards environmentally-friendly, climate-neutral, healthy and fair food systems through a **PUSH-PULL-ENABLE strategy** to

- enhance the availability and adequacy of organic seeds of cultivars appropriate to organic farming (PUSH),
- increase and stabilise the market demand for organic seeds of cultivars appropriate to organic farming (PULL),
- foster an enabling policy and regulatory environment where both demand and supply can harmoniously and productively negotiate without irrelevant constraints due to legal restrictions and/or regulatory fragmentation (ENABLE).



LiveSeeding addresses the topics in a **holistic multi-actor, multi-stakeholder, participatory approach** involving stakeholders along the value chain in 17 local **Living Labs** (LLs) and 3 established networks of organic breeders (**ECO-PB**), seed savers (**ECLLD**) and Milan Urban Food Policy Pact (**MUFPP**). 15 European countries cover the different pedoclimatic zones and socio-economic contexts, including countries with a low level of development in organic seed and breeding in East and South Europe.

Table of Contents

SUMMARY.....	5
1. DATA SUMMARY.....	7
1.1 DATA TYPE AND FORMATS.....	7
1.2 DATA PURPOSE, SIZE AND SOURCES.....	13
1.3 POTENTIAL USE OF DATA OUTSIDE OF THE PROJECT	15
2. FAIR DATA.....	16
2.1 MAKING DATA FINDABLE	16
2.2 MAKING DATA ACCESSIBLE.....	18
2.2.1 Repositories	18
2.2.2 Data accessibility	20
2.2.3 Metadata:	21
2.3 MAKING DATA INTEROPERABLE	24
2.4 INCREASE DATA RE-USE.....	24
2.4.1 Publication and Creative Commons licenses.....	24
2.4.2 Data Quality assurance processes	25
2. OTHER RESEARCH OUTPUTS.....	26
3. DATA MANAGEMENT AND ALLOCATION OF RESOURCES.....	26
4. DATA SECURITY	27
5. ETHICS.....	28
6. OTHER ISSUES	28
ANNEX 1 OPEN SCIENCE POLICY UNDER HORIZON EUROPE.....	29
ANNEX 2 CREATIVE COMMON LICENSES	30
ANNEX 3 (META)DATA FILE NAMING GUIDELINES (NOMENCLATURE).....	31
ANNEX 4 GENERAL TEMPLATE FOR METADATA OF DATA AND REPORTS OBTAINED IN LIVESEEDING	32
ANNEX 5 TEMPLATE FOR METADATA OF DATA OBTAINED FROM FIELD TRIALS	34
ANNEX 6 EXAMPLES OF PARAMETERS FOR FIELD TRIALS' TEMPLATES	36

ANNEX 7 METADATA PARAMETERS FOR SURVEYS - EXAMPLES FOR TEMPLATES	38
ANNEX 8 EXPECTED DATA GENERATED FROM LIVESEEDING WPS	2
ANNEX 9 PROTOCOL/GUIDELINES TO UPLOAD PROJECT PUBLICATIONS AND DOCUMENTS IN REPOSITORIES	2
ANNEX 10 ONTOLOGY FOR INTEROPERABILITY	5

List of abbreviations

APC	Article Processing Charges
API	Application Programming Interface
BrAPI	Breeding Application Programming Interface
CA	Consortium Agreement
CC	Creative Commons
CPVO	Community Plant Variety Office
DMP	Data Management Plan
DoA	Description of the Action
ECO-PB	European Consortium of Organic Plant Breeding
ERD	European Router Database
EU	European Union
ExCom	Executive Committee
FAIR	Findable, Accessible, Interoperable, and Reusable
H2020	Horizon 2020
IP	Intellectual Property
IPRB	Intellectual Property Right Board
OA	Open Access
PC	Project Coordinator
REST	Representational State Transfer
SCT	Scientific Coordination Team
YY.MM	Year in 2 digits, Month in 2 digits
WP	Work Package

Summary

The LiveSeeding Data Management Plan (DMP) has been developed to illustrate how the project will comply with the grant rules on FAIR Data and Open Science, and to provide project partners with guidance on how to operationalise it. This is the 1st version of the LiveSeeding DMP, which will be updated and expanded by month 36 and month 46 as the project is rolled out and data management needs expand.

The objective of this deliverable on Data Management Plan (DMP) is:

- To ensure good data management throughout the lifecycle of the project but also future sharing and reusing of research data beyond the project and consortium;
- To present the procedures and repositories used for the project in line with FAIR principles and new requirements of Horizon Europe programme;
- To provide clear guidance and practical tools to project partners for the management, storage and sharing of data produced in the scope of the project.

To that aim, this deliverable includes 6 chapters and 10 annexes that cover:

- A presentation of the type of (meta)data, size, purpose of its collection, methodology and standards / controls used within LiveSeeding (**chapter 1**);
- An explanation on how the project will comply with FAIR (Findable, Accessible, Interoperable, and Reusable) principles (**chapter 2**);
- An introduction to procedures and tools used in the project to ensure data security, ethics, privacy, and resources deployed (**chapters 3-6**) and reference to other project deliverables focused on those issues;
- A set of easy-to-use (meta)data management tools (templates, info sheets, etc) for partners (**Annexes 1-10**).

The structure of the deliverable is based on the new [template](#) developed by the European Commission to guide grantees. It reflects changes in the rules of Horizon Europe in respect to the previous funding programme (Horizon 2020). Some important changes are highlighted in the info box below:

Data management - Main changes of rules in Horizon Europe

- **Scientific publications:** No embargo period, immediate OA
- **Re-use:** Author(s) need to retain rights (no copyright transfer) as for a CCby license to assert ownership; acknowledgement for re-use required (see **Annex 2**)
- **Costs of APC:** recoverable only if journal is a full open access one (not hybrid)
- **Metadata:** must be FAIR and under CC0 license (see **Annex 2**); deposit as soon as possible after production/generation or after processing control

- **Metadata:** partners should not forget to add project acronym
- **Data** can be confidential if necessary but grantees need to justify why (default approach being FAIR). For example, commercially valuable data, risk to endanger trade secret (soft IP) or make IP more difficult

All data will be treated as open source unless it is defined beforehand as confidential data to follow the principle ***“as open as possible, as closed as necessary”***, i.e. to allow for circulation of scientific knowledge while protecting commercial interests and confidentiality. Specific restrictions to the openness / dissemination of background data and results as indicated by partners in the Consortium Agreement will be respected (the exact list of restriction is not annexed due to confidential nature of this information and the fact that this deliverable is public).

Each LiveSeeding partner is responsible for careful data recording, validation of data quality, and clear documentation of the data they have generated , in compliance with data protection legislation as for guidelines provided under deliverable [D8.4 on Protection of personal data & recruitment of stakeholders](#). Task and WP leaders are responsible to provide data templates and monitor that data protocols are followed.

The DMP is a living document that will be updated over the course of the project by M36 (version 2) and M46 (version 3), as per Grant Agreement. At this stage, some protocols are still under development and more detailed description of datasets and templates need to be defined by WP and task leaders. The Project Coordinator, Scientific Coordination Team and Executive Committee stir this process.

1. Data Summary

The DMP ensures good data management throughout the lifecycle of the project guaranteeing all information and data collected and produced respects the FAIR principles, i.e. all research data should be Findable, Accessible, Interoperable and Reusable (FAIR). The DMP describes the data (incl. metadata) and software standards, availability, curation and preservation methods and covers the data documents and solutions used and created in the project. This chapter specifically addresses the questions proposed in the deliverable [template](#) proposed by the EC on the data type and format (chapter 1.1); the purpose of its collection, the size and sources, i.e. the origin of the data generated by the project or re-used from other sources/projects (chapter 1.2); the potential use of data outside the project, both during and after the project life span (chapter 1.3).

1.1 Data type and formats

LiveSeeding will generate a diverse set of statistical, agronomic, socio-economic and lab data as well as software codes in order to achieve the overall goal to strengthen the organic seed and plant breeding sector for sustainable and diverse food systems in Europe. Based on the expected results of the individual WPs presented in Table 1, each WP leader has identify the expected data and research outputs including their status in terms of type and/or origin (e.g. reused, collected, generated) listed in Annex 8. These list will be amended and further elaborated during the project. LIVESEEDING also uses existing data and software from previous EU funded projects and other sources (see chapter 1.2) that will be integrated with new data and findings as mentioned in the background data of the consortium agreement as well as existing software that will be adjusted for LiveSeeding purposes. Table 2 lists the ICT tools that will be adapted or newly developed for the use of a broader audience. Personal data collected in the scope of LiveSeeding will be anonymised in compliance with the European Data Protection Regulation (GDPR, Regulation (EU) 2016/679) as well as applicable national legislation on data protection as described in detail in deliverable D8.4.

Table 1: Project results organised by Work Package	
WP	Project Results
WP1	Sound breeding strategies and novel tools suitable for organic farming
	New organic varieties (OV), organic heterogeneous material (OHM) and improved landraces for numerous crops
WP2	Pilots of viable multi-actor organizational models for decentralised, participatory cultivar evaluation

	New trial designs and analysis for on-farm cultivar testing
WP3	Scientific insight on interaction of seed microbiome and seed vigour
	Holistic seed health strategy
WP4	Stepwise approach in the form of national roadmap (tailored for 10 EU countries) to phase out derogations and increase availability of organic seed
	Data on supply and demand of organic seed
WP5	Business models and plans for organic breeding, cultivar testing and seed production initiatives
WP6	Quantification of added value of organic breeding
	Cities food policies including the aspect of farmed diversity from local production
	Marketing strategies for food products derived from OV and OHM
WP1-6	LiveSeeding will have a strong focus on training related to the innovations tested in each WP and specifically targeted to the user target group of each WP outcomes. Trainings modules including different type of material (hand-outs, videos, practical exercises, etc.) will cover all major topics of the project

A short description of the **4 software** that will be developed/adapted under LiveSeeding project to collect and process data is provided in Table 2:

Table 2: Softwares developed/adapted under LiveSeeding project
SHiNeMaS to support participatory breeding will be adapted (WP1): <i>Short description:</i> SHiNeMaS is a digital tool developed by INRAE in former projects and dedicated to seed lots history, phenotyping and cultural practices. SHiNeMaS aims: (1) to study on-farm management of crop diversity, (2) to develop population-varieties adapted to organic and low-inputs agriculture, (3) to co-develop tools and methods adapted to on-farm experiments. <i>Data type:</i> in SHiNeMaS, project partners will collect: - Seed lot description and their genealogy (Reproduction, crossing, mixtures, selection, diffusion (main data collected); - On farm evaluation data of germplasm (plant height, awn, ear weight, thousand kernel weight etc.); - Images illustrating field trials ; - Stock and storage information of seeds. <i>Adaptations:</i> the main developments about SHiNeMaS under/for LiveSeeding will be the implementation of web services to ensure interoperability with other systems

involved in the project. Also, some minor evolutions will take place to make it fit for WP1 requirements, although SHiNeMaS has already been developed to fit participatory breeding purpose. Some elements will be further clarified by the next DMP version as a result of discussion on interoperability with OHMTrack in particular.

Software type/openness: Open source (code available on Renater, French repository for research organisations)

SeedLinked to support on-farm cultivar testing networks will be adapted (WP2):

Short description: is a user-friendly, interactive platform developed by the organisation SeedLinked (subcontractor) that enables simple, decentralized cultivar evaluation designs and inclusion of real-time multi-actor feedback (from farmer to consumer) and allows for data inputs by smartphone. Using crowdsourced data generated by the networks, it provides recommendations to inform the supply chain on what variety does best and where.

Data type:

- Phenotypic data such as disease resistance, earliness, yield, and flavour/taste, physiological stage dates such as planting, transplanting, flowering, and harvest dates; growers' comments and pictures;
- Crowdsourced varietal information is localized and connected to pedo-climatic information (e.g. soil maps or bio-climatic variables such as temperatures and rainfall);
- The combination of the two or genotype by environmental interaction allows varietal recommendation for a given place for the all supply chain from Breeder to the consumer.

Adaptations: SeedLinked will be adapted for the needs of European on-farm cultivar testing networks with localisation and translation of the user interface in 6 EU languages. It will be made available to researchers, organic breeders and seed savers associations for trials with farmers, gardeners and consumers. Also, SeedLinked will add crop-specific ontology and new trial designs, and enable the integration of quantitative data collection and pedo-climatic data (e.g. soil maps or bio-climatic variables such as temperatures and rainfall).

Software type/openness: Coding is owned by SeedLinked and will remain so as for Grant Agreement and subsequent services Agreement (subcontract).

OHMtrack, software for OHM traceability will be newly developed by AEDIT (WP3)

Short description: OHMTrack will support seed producers and traders in managing the traceability and identification of OHM, complying with the delegated act on OHM (Commission Delegated Regulation (EU) 2021/1189).

Data type: users will enter data about

- Seed lots (origin, amount, seed owner, genealogy);

- Seed lot transfers (seller, buyer, amount);
- Seed producing data (location, field surface, seed sown, seed harvested).

Data will be stored in a relational database, made available through API and web and an access control system will be used to enforce data access based on user credentials. See **Annex 10** for more details about data ontologies and the approach currently being privileged for repository and findability.

Development: OHMTrack is a new system/software to trace the distribution of officially notified OHM seed lots, but its focus will also be on interoperability with the other project software tools to maximise effectiveness and impact

Software type/openness: open source with codes available in a public repository (still to be defined as it's in its infancy)

European Router Database - ERD software on organic seed will be adapted (WP4):

Short description: is an online software application developed by FiBL-DE in the scope of H2020 project LIVESEED for the availability of organic plant reproductive material (seed & vegetative material), which enables plant reproductive material suppliers to manage their supply in all EU Member States, Switzerland and UK from one account.

Its objective is to overcome main burdens that hinder cross-country use of national produced organic seed listed in national data bases in local language in order to increase the supply of organic plant reproductive material in the EU, Switzerland and the UK. The database is available in all EU Member State languages, offers one single login for up to 29 national database and an internal message system between national competent authorities and supplier for organic plant reproductive material to enable targeted communication.

Data type:

- Data on national crop and crop classifications according to Cat I, Cat II, Cat III (Reg. EC No 2018/848);
- Yearly national derogation reports on non-organic seed and vegetative propagating material.

These data will be collected and published in the router database and the projects website. The collected data will be available either in PDF reports or Excel tables.

All publications and training materials are open access.

Adaptations: under the LiveSeeding project new features and 2 API will be added to the existing European Router database developed under the H2020 funded LIVESEED project ([D1.5 Report on the European Router Database](#)).

Software type/openness: the software used for the router database is owned by FiBL-DE.

It is worth noting that while these softwares are developed under different work packages and for their respective purposes, their interoperability is a key element of the project and will be addressed under T9.3 on Internal and external interoperability of LiveSeeding digital tools (more details on this are available under chapter 2.3).

Scientific publications, conferences and a summer school will be used to present the project's results to the scientific community. Several dissemination and communication tools (including videos, articles, practice abstracts, policy briefs, and booklets) will be used to present and promote the outcomes to a wide spectrum of actors.

Thus, data acquired in the scope of LiveSeeding will include very diverse datasets from different disciplines and in different formats. This diverse type of data can be grouped in the following categories:

Statistical data:

- Production area of selected crops under organic certification in different European countries added to existing data base of FiBL-CH, supplemented by database from EuroStat, national databases;
- Data on quantity of organic seed produced for selected crops based on survey(s).

Socio-economic data:

- Qualitative data on implementation of organic regulation with regard to seed issues in different countries;
- Quantitative data (e.g. estimates of percentage / amount, prices of organic seed) from expert and stakeholder interviews and surveys;
- Economic data obtained from seed companies, farmers, European Seed Association and existing data base on agronomic production costs.

Agronomic data:

- Yield data of landraces, OV's and OHMs under organic farming in selected geographic regions;
- Phenotypic data (production traits, morphological and phenological traits, functional traits such as resistances, food quality and taste traits) of experimental field trials;
- Germination rate and speed of selected crop species;
- Analytic data (e.g. protein, alkaloid content).

Lab data:

- Data from lab experiments (e.g., new lab protocols);
- Data on seed treatments (e.g., new protocols, reduction of fungal infection);
- Plant genotypic data (e.g., SNP markers);

- Microbiome data (e.g., next generation sequencing data of seed microbiome).

Standard data format for data collection on LiveSeeding SharePoint:

- Tabular data: MS Excel XLS/XLSX or CSV files;
- Reports: MS Word DOC/DOCX or PDF files;
- Images: JPG, PNG, or EPS files;
- Video: MPEG-4 or MP4 files;
- Presentations: PPTX files or PDF files.

These formats are used by all partners and compatible to most computer programs. Other data formats might be used by individual partners but will be converted to the standard formats before uploading to allow proper exchange of data on the SharePoint.

For long term storage of final data, the tabular data will also be saved as .CSV, and reports as .RTF and .PDF to ensure high compatibility and long-term usage of the research data by a broader audience (e.g. scientists, breeders, farmer organisations, public authorities, policy makers) beyond LiveSeeding project and its project partners. Longer term curation is further detailed under chapter 2.5.1.

Over the lifespan of the project, different data types will be produced processed and used. Table 3 here below indicates the source and the level of openness as presented in the grant agreement.

Table 3: Data type by WP, Source and Openness		
Type, Description & WP(s)	Source	Openness
Qualitative and quantitative on organic seed and breeding domain [WP2, 3, 5]	Surveys, Interviews	Public [on repository]
Qualitative and quantitative data of agronomic and organoleptic quality data [WP1, 2]	On-farm trials and LLs activities	Public, Data shared [SeedLinked or repository]
Seed microbiome data [WP3]	Experimental trials	Scientific publications [data and papers on repository]
Phenotypic and genotypic data through develop selection tools [WP1]	Field and lab trials	Scientific publications [data and papers on repository]
Phenotypic and genotypic data for selection purposes [WP1]	Breeding programme	Internal use (confidential, IPR of breeder)

OHM descriptions [WP1, 2, 3]	Notification and experimental trials	Public [on repository]
Performance of organic varieties (WP1, WP2)	DUS and VCU testing and experimental and selection trials	Partly confidential upon agreement with breeder
Data on derogation categories per crop species and Member States [WP4]	Official national databases, workshops	Public EU router database, webpage
Statistic data of organic seed use [WP4]	Aggregation of data on organic farming area and expert interviews	Public [on repository]
Business data on organic seed and plant breeding [WP5]	Expert interviews, case studies, LLs, SMEs, Companies	Partly confidential upon agreement with the source (guidelines in DMP)
Data on economic, social and environmental benefits of organic breeding [WP6]	Expert interviews, case studies, LLs, SMEs, Companies	Partly confidential upon agreement with the source (guidelines in DMP)

Due to the diversity of data sets, different standards and protocols will be defined by Task Leaders in agreement with WP leaders, while rolling out the project. A list of data expected from different WPs and Tasks of LiveSeeding are shown in **Annex 8**. This annex will be updated at least twice as part of the overall deliverable update by month 36 and month 46 as for Grant Agreement.

Each LiveSeeding partner is responsible for careful data recording, validation of data quality, and clear documentation of the data they have generated, in compliance with data protection legislation as for guidelines provided under [D8.4 on Protection of personal data & recruitment of stakeholders](#). Task and WP leaders are responsible to provide data templates and monitor that data protocols are followed during internal semester reporting.

1.2 Data purpose, size and sources

The ultimate goal of LiveSeeding is to facilitate the delivery on the market of high-diversity cultivars for a wide range of crops, as basis for quality organic seed production to foster more sustainable food systems.

In order to achieve this, the project relies substantially on data collection and processing carried out through: the conduct of laboratory experiments (e.g. seed vigour test, molecular analysis, microbiology), field trials (on-station and on-farm,

research led and participatory, for breeding, cultivar testing and demonstration), software coding for data collection and management, quantitative and qualitative surveys and questionnaires (e.g., in the context of Living Labs), statistical data collation on organic farming and organic seed demand and supply, and social benefit analysis.

The collected data are the basis to fulfil the objective of the respective tasks within WPs (see **Annex 8** on Expected data generated from LiveSeeding WPs). The data will also be used across WPs: for example, the data collected in the “Survey on the European organic seed and breeding sector” that has been conducted under T5.1 will be used also in T3.4 on tackling challenges in upscaling organic seed production and T5.2 on financing and governance models for organic plant breeding initiatives. The data will be of interest for scientists and students in the respective disciplines, plant breeders, seed producers, farmers, advisors, national authorities and examination offices, policy makers, stakeholders of the organic value chain, and will fill gaps with respect to statistic data on organic seed production and use in Europe.

Regarding the size of data: the size of cleaned data is estimated to be between 1 and 4 GB. Working documents including such data might be between 10 and 20 GB.

The project will re-use and build upon existing available data and software for certain tasks. In particular, the project will draw on:

- already **existing global database managed by FiBL-CH on comprehensive organic farming statistics** that cover the area under organic management, specific information about land use in organic systems, the number of farms and other operator types, and selected market data¹.
- Data reports and software from **prior EU-funded projects** (e.g., [LIVESEED](#) (2017-2021) [DIVERSIFOOD](#) (2015-2019), [BRESOV](#) (2018-2023), and [ECOBREED](#) (2018-2023) all funded under H2020 RIAs, [SOLIBAM](#) (2010-2014) funded under FP7-KBBE, [COBRA](#) (2013-2016) funded under CORE II, [Organic Data Network](#) (2012-2014) funded under FP7-KBBE, and from prior national research projects in which LiveSeeding partners were involved;
- Data, reports and training material derived from workshops, and surveys conducted by the [European Consortium for Organic Plant Breeding](#) (ECO-PB)²;
- Background information provided by individual partners will also be considered as **baseline** to build up new research.

¹ <https://www.fibl.org/en/shop-en/1254-organic-world-2023>

² ECO-PB - European Consortium of Organic Plant Breeding is a non-profit network of organisations fostering organic plant breeding. It brings together the majority of organic plant breeding initiatives across Europe and is currently coordinated by LiveSeeding partner FiBL-CH

For example, data collected under T4.1 on the changes in supply and demand of organic and farm-saved seeds will be compared with latest data collected under the H2020 funded project "LIVESEED".

1.3 Potential use of data outside of the project

Project partners will be proactive and take specific measures to ensure that their results are exploited outside of the project by a broad range of stakeholders. To that aim a deliverable including dissemination and exploitation plan ([D7.1 Plan for the Exploitation, Dissemination and Communication of Results - PEDR](#)) has been prepared by M06 in order to define plans and procedures to maximise dissemination of project activities and exploitation of results.

PEDR describes main project's dissemination, communication, and exploitation tools, channels and methods that will be used in order to maximize the impact of LiveSeeding across the project's entire range of targeted audiences. During the project various materials will be produced, such as media releases, videos, newsletters, articles, leaflets, posters, roll-ups, presentations, practice abstracts, policy briefs, training materials, booklets and (scientific) publications.

These materials will be distributed among the stakeholders and stored on different repositories depending on the type of data and materials. The project publications and output materials will be deposited on 3 main online open access repositories: Zenodo for peer-reviewed scientific publication and data, Organic Eprints (for all kind of publications, data, reports, videos, presentations) and Organic Farm Knowledge platform, section "Seed" (for practice abstracts). LiveSeeding participants are encouraged to share their results via [Open Research Europe](#) (ORE). If no sensitive data involved, LiveSeeding data will be made available under the terms of a Creative Commons Public Domain Dedication (CC0 license). The project communication team will work to create awareness of the newly uploaded LiveSeeding data sets and their availability for further use, e.g. by social media outlets and news on webpage.

Also, [LiveSeeding project website](#) will be a depository of all project results that have an Open Access status. The website will be timely updated by IPS Konzulting and it will serve as a channel for reaching out to stakeholders, citizens and the wider public by providing information on different events, publications, and project deliverables, as well as an inventory of project activities and fast access to results summarized in the form of booklets, videos, newsletters, and policy briefs and recommendations. For long term storage the LiveSeeding website content will be uploaded on the repository Organic Eprints and moved to the website of ECO-PB before the end of the project.

2. FAIR data

FAIR management of data means that data of the project will comply with the FAIR data principles. The data will be findable and accessible through the use of open access repositories. To make data interoperable, defined standards, formats and vocabularies for data and metadata will be used. The re-usability of the data will be available through Creative Commons Public Domain Dedication (CC0 license).

Chapters 2.1 to 2.4 explain what the project will do to make data FAIR but also guide project partners on how to go about it at practical level.

FAIRification Process

The FAIR Data Principles apply to metadata, data, and supporting infrastructure (e.g., search engines). Most of the requirements for findability and accessibility can be achieved at the metadata level. Interoperability and reuse require more efforts at the data level.

For more guidance on the “fairification” process can also be found on the [GoFair website](#).

2.1 Making data findable

LiveSeeding peer reviewed publications and supplementing metadata and data will be assigned, a unique **Digital Object Identifier (DOI)** which allows to find and track data when deposited in **Zenodo** and linked to the [LiveSeeding](#) community.

All LiveSeeding publications and data as well as other materials, such as booklets, Practice Abstracts, PowerPoint presentations, will be uploaded on the repository in **Organic eprints** and will be findable with keywords (including “LiveSeeding” as a keyword for easy search). Instruction for uploading in Zenodo and organic eprints are given in Annex 9.

LiveSeeding research/experiment data files will include "Metadata" in the excel file containing also the "Data". **All Metadata files will include the Project Acronym.** More specifically:

For agronomic data, following metadata should be accompanied to allow correct interpretation:

- General Information of experiment
- Organisation responsible for experiment
- Objectives

- Clearly defined treatments or factors to be tested
- Statistical design, number of replications and randomisation
- Field plan
- Information of location (postal code, GPS, altitude, soil type, organic or conventional management, crop rotation, soil management)
- Crop species (Latin names),
- Breeding lines and cultivar names, breeder and origin of seed
- Year of trial
- Trial management (weed management, fertilization, crop protection, sowing and harvesting time, sowing density, row spacing, temperature and rainfall, irrigation, etc.)
- Time and developmental stage of data assessment and sampling
- Methodology used for sampling and data collection
- Indication of tools used for assessment
- Clear definition of each qualitative and quantitative parameter assessed
- Methods used for data validation and statistical analysis
- Any specific comments or assumptions made.

For microbiome or plant genetic data, different formats will be used based on the respective data. For these data following metadata must be accompanied to allow correct interpretation:

- Exact procedure of sampling for DNA extraction
- DNA extraction methods, purification and storage conditions
- Detailed methodology of DNA analysis
- Laboratory responsible for analysis
- Time of analysis and respective lab protocols, positive and negative standards used
- Number of field and lab replications used
- Method used for data validation and statistical analysis
- Any specific comments or assumptions made.

For socio-economic data in quantitative and qualitative surveys or interviews, the following metadata must be accompanied to allow correct interpretation:

- Questionnaire / interview guide
- Recruitment method for participants
- Demographic data of participants (e.g. age, gender, affinity to organic production, education, residency, country of origin)
- Country of survey
- Information provided before survey
- Tools used for data recording (e.g. online, phone interview, recording device, written, etc.)

- Clear definition of each qualitative and quantitative parameter assessed
- Methods used for data validation and statistical or other analysis
- Any specific comments or assumptions made.

Templates for metadata are provided in **Annex 4 and 5**. Task leaders together with WP leaders are in charge of developing detailed templates for such experiment descriptions and data entering. The involved partners are responsible to provide these metadata for each dataset. The task leaders will monitor the completeness.

Metadata will be used to index searchable keywords to improve findability. Based on the specified keywords in the metadata (e.g. organic breeding, microbiome), the LiveSeeding data can be searched in a similar modus like publications and reports under Organic Eprints and also under the Zenodo's search engine after publication. For both Zenodo and Organic Eprints "LiveSeeding" community will be implemented to allow to extract all LiveSeeding document from the repositories.

In Zenodo repository the metadata is exported in several standard formats such as MARCXML, Dublin Core, and DataCite Metadata Schema (according to the OpenAIRE Guidelines).

It is worth noting that, several journals already require the deposit of the full research data and script used for statistical analysis as supplement information for publication.



Tip for metadata and data stored in separate files

In case the metadata and data are stored in separate files (for example in case you need to use a format other than excel that allows several sheets in one file) it is crucial that the link between these files is obvious. This can be achieved by submitting several supplementing files together with the publication: supplement I contains the raw data and supplement II the metadata. Making use of the unique identifier DOI obtained when depositing data in Zenodo, the metadata can be clearly linked with the main data set.

2.2 Making data accessible

2.2.1 Repositories

Different types of data will be transferred to different repositories (FAIR):

- **Statistical data** on crop specific level (e.g., area of organic production) will be integrated in existing database of FiBL-CH on Organic Agriculture. Data are publicly available under "[Organic World](#)" and summarized in annual reports

[“The World of Organic Agriculture”](#) published by FiBL-CH and IFOAM Organics International and reported at biofach Nürnberg the world largest organic faire;

- **Practice Abstract** will be transferred to [EiP-AGRI website](#), the [Organic Farm Knowledge Platform](#)³, [Organic Eprints](#)⁴ and [LiveSeeding](#) website (www.liveseeding.eu);
- All **Scientific publications** (both peer-reviewed and non-reviewed) will be uploaded on [Organic Eprints](#) to reach the organic community; Organic Eprints has been established to promote open access to research results, is running on Linux and other free, open source software and has had almost 7.5 million downloads since May 2018. More statistics about Organic Eprints traffic are available [here](#);
- **Peer-reviewed publications** and open research data will also be deposited in [Zenodo](#)⁵;
- **Tools, training and demonstration material** will be made available on the [Organic Farm knowledge Platform](#), Section “Seed” the [LiveSeeding](#) website as well as on [ECO-PB](#) website⁶;
- **Genomic data**, where appropriate, shall be integrated into existing general platforms such as the [Plant Genomics and Phenomics Research Data Repository](#) (GPD), the [Genetic and Genomic Information System](#) (GnpIS), or

³ The Organic Farm Knowledge platform (www.organic-farmknowledge.org) provides access to a wide range of tools and resources about organic farming that can help improve production. It also aims to serve as a virtual meeting place for cross-border learning. The platform promotes the exchange of knowledge among farmers, farm advisers, and scientists, with the aim of increasing the productivity, quality and sustainability of organic farming across Europe. The core of the Organic Farm Knowledge platform is the toolbox ranging from practice abstracts, leaflets, videos, website, etc. The toolbox offers the possibility for an arbitrary text search as well as displaying tools by theme, language, organisation, project, keyword, type, country of origin and year.

⁴ Organic Eprints (is an international open access, multilingual online repository of electronic documents related to research in organic food and farming. The archive contains full-text papers together with bibliographic information, abstracts and other metadata. It also offers information on organisations, projects and facilities in the context of organic farming research. To upload papers to Organic Eprints a free of charge user registration is required. A [User guide](#) for organic e-print is available in different languages and in Annex 9.

⁵ Zenodo (www.zenodo.org) is a free of charge, general-purpose open repository developed under the European OpenAIRE program and operated by CERN. Zenodo makes the research results citable via Digital Object Identifier (DOI) and integrates them into existing reporting lines to funding agencies. A LiveSeeding community has been created and guidelines for uploading are described in Annex 9.

⁶ ECO-PB - European Consortium of Organic Plant Breeding (www.eco-pb.org) is a non-profit network of organisations fostering organic plant breeding since 2001. It brings together the majority of organic plant breeding initiatives across Europe. The open access webpage is managed by LiveSeeding partner FiBL-DE.

[COpenPlantOmics](#) (COPO), as well as crop-specific databases such as [Gramene](#) for cereals or [Cool Season Food Legume Genome Database](#) for legumes;

- **Field trials** will be stored as a combined file across years and locations using the same parameters. Data will be prepared in such a way that they can be used in meta-analysis within LiveSeeding project and beyond (see chapter 2.5.1 for long term curation). Non-confidential data (see chapter 2.2.2 for Background and Results in order to ensure that data is “as closed as necessary”) will be deposited together with the publication as supplementing data on [Zenodo](#) and [Organic Eprints](#). Agronomic data, when appropriate for long term storage, shall be integrated in crop specific repositories like the FAO’s [AgroPortal](#) following common reference ontologies (see chapter 2.3 for more information on ontologies and interoperability) [Github](#), [USDA](#) or others;
- **Socioeconomic data** including those gathered through surveys after anonymization (considering GDPR compliance) can be archived in domain-specific repositories like [Council of European Social Science Data Archives](#), [European Research Infrastructure \(CESSDA\)](#), [Eurobarometer](#), [GESIS Data Catalogue](#), or [GESIS Research Data Centers - International Survey Programs \(ISSP\)](#)).

It is worth noting that, several journals already require the deposit of the full research data and script used for statistical analysis as supplement information for publication.

2.2.2 Data accessibility

All raw data collected by partners and uploaded on the LiveSeeding SharePoint will be treated as confidential. Preliminary data will only be accessible to the project partners. Access to the LiveSeeding SharePoint is regulated by a personal login and password system administered by FiBL Europe and accessed/used by project partners as for the guidelines of project deliverable **D8.1 Internal platform for communication and exchange manual**.

Greatest care will be given to protect personal data obtained during surveys or interviews and confidential information in line with the project deliverable **D8.4 Protection of personal data & recruitment of stakeholders** as well as **D8.3 Ethics Issues and Non-European compliance**. All such deliverables are available in [LiveSeeding SharePoint](#). Without clear consent of each participant to be named, all personal data will be anonymized, before any data will be uploaded on the LiveSeeding SharePoint and made publicly available.

As for the terms of the LiveSeeding Consortium Agreement signed by each partner, all information declared as confidential background information will be kept confidential. Access rights to results and Background needed for the performance of the own work of a Party under the LiveSeeding project shall be granted on a royalty-

free basis, unless otherwise agreed for Background in the Consortium and Grant Agreement.

Also foreground information which is disclosed by one Party (the “Disclosing Party”) to any other Party (the “Recipient”) in connection with LiveSeeding implementation and which has been explicitly marked as “confidential” at the time of disclosure (or when disclosed orally has been identified as confidential at the time of disclosure and has been confirmed and designated in writing within 15 calendar days from oral disclosure as confidential information by the Disclosing Party) is treated as “Confidential Information”.

Finalized data will be made available to LiveSeeding partners and linked parties that need these data as input for their tasks. However, certain data and information might be excluded from the open data strategy:

- if there are commercial interests (e.g., one or several partners claim to protect their IP rights for exploitation of results);
- if data are not yet published in peer-reviewed papers. For example, in case data is collected also as part of a PhD research project, making it public before the publication as a doctorate research publication would make it unusable for the Doctorate project as no longer considered novel data.

The ExCom, on the basis of advice from the IPRB -where relevant- will monitor results and exploitation plan and make sure that commercial interests of involved partners are guaranteed. Data that might be considered confidential will not be disclosed to the broader public before approval by ExCom. The ExCom will act as **Data Access Committee** (e.g. to evaluate/approve access requests to personal/sensitive data outside of the indications already provided in the present deliverable).

If there is no request for confidentiality, compiled data and results will be made available to all LiveSeeding partners and beyond. Thus, data or research results that are not protected by IP rights and are no longer confidential (i.e. after publication) will be published on the LiveSeeding website and added to repositories or databases as for indications under chapter 2.5.1.

2.2.3 Metadata:

Metadata can be defined as information that is needed to be able to use and interpret statistics. Metadata describe data by giving definitions of populations, objects, variables, the methodology and quality⁷.

⁷ <https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:Metadata>

Another, simpler, way to define **Metadata** is "data that provides information about other data",⁸ but not the content of the data, such as the text of a message or the image itself.

Metadata allows users to access resources by "allowing resources to be found by relevant criteria, identifying resources, bringing similar resources together, distinguishing dissimilar resources, and giving location information"⁹

While nowadays metadata typically takes a digital form, to exemplify this, we can refer to traditional catalogue cards used in paper archives. Such cards include the name of the author, title, subject, date of publication, kind of literature, edition etc) to help to find relevant information and sources.¹⁰



Figure 1: Catalogue cards, as an example of metadata source (Wikipedia)

For a visual and easy to understand explanation of what metadata is with several examples of metadata for different data formats (a photo, a Word document, a spreadsheet, an email, a website, etc) it is also possible to consult [Dataedo](https://dataedo.com/) website.

Metadata will be made openly available and licenced under a public domain dedication CC0 (see **Annex 2** on Creative commons), as per the Grant Agreement. As a minimum requirement under Horizon Europe grants, Metadata of deposited publications will provide information at least about the following:

- Publication (author(s), title, date of publication, publication venue);
- Horizon Europe funding, grant project name, acronym and number;
- Licensing terms;

⁸ <https://en.wikipedia.org/wiki/Metadata>

⁹ <http://www.niso.org/publications/press/UnderstandingMetadata.pdf>

¹⁰ https://en.wikipedia.org/wiki/Library_catalog

- Persistent identifiers for the publication, the authors involved in the action and, if possible, for their organisations and the grant;
- Where applicable, the metadata must include persistent identifiers for any research output or any other tools and instruments needed to validate the conclusions of the publication.

Rich metadata allow a computer to automatically accomplish routine and tedious sorting and prioritising tasks that currently demand a lot of attention from researchers. The rationale behind this principle is that someone should be able to find data based on the information provided by their metadata, even without the data's identifier.¹¹

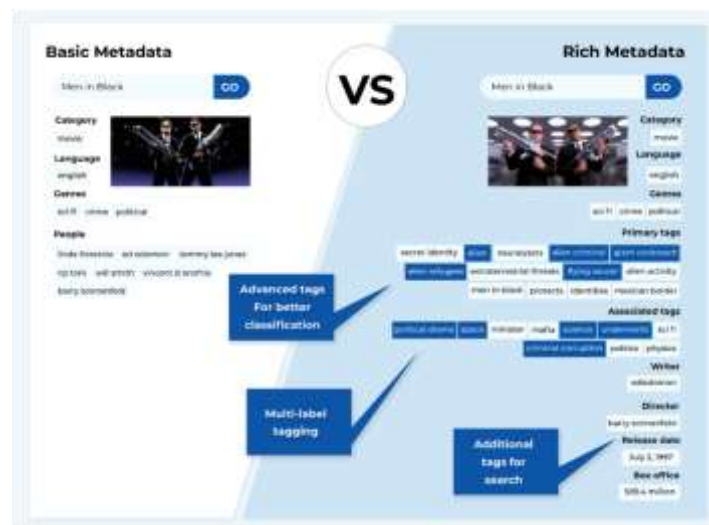


Figure 2: Example of the difference between basic metadata and rich metadata¹²

Since **rich metadata** are essential to ensure that data and other research outputs are reusable, LiveSeeding partners are required to provide more extensive metadata as for guidance/templates under **Annex 4 and 5**.

Metadata will remain available and findable, even when the data is no longer available by depositing the data with the attached metadata in repositories, such as Zenodo and Organic Eprints, as for indications provided under chapter 2.5.1.

Softwares used in the framework of the LiveSeeding project to collect data have been described under chapter 1.1. Once data is collected and extracted, they will be available as .xls, .csv and other formats for which the software itself will no longer be needed and therefore data will continue to be accessible and re-usable.

¹¹ <https://www.go-fair.org/fair-principles/f2-data-described-rich-metadata/>

¹² <https://reconsenselabs.com/blog/rich-metadata-game-changer-for-media-organization>

2.3 Making data interoperable

In order to improve the interoperability of the data produced under the LiveSeeding project, controlled vocabulary will be used to allow the combination of data. In particular, data and metadata should be annotated using vocabularies/ontologies/thesauri that are commonly applied in the field.

LiveSeeding project partners are evaluating the use of Breeding API¹³ an interoperability standard for plant breeding databases¹⁴. BrAPI is a standardized RESTful web service API specification for communicating plant breeding data. This community driven standard is free to be used by anyone interested in plant breeding data management.

CropOntology will be used to manage the ontology of the phenotyping traits. Crop and variety will be managed using the codes and standards from CPVO - Community Plant Variety Office. For identifying the seed lots and germplasm it is planned to test the **Digital Object Identifier** managed by the International Treaty on Plant Genetic Resources for Food and Agriculture.

The vocabularies/ontologies/thesauri that are used to describe the data must also be well documented and easily findable, accessible, and interoperable. A dedicated task will ensure that all the tools developed and used in the project respect the FAIR principles. The main relevant data will be accessible using a set of APIs allowing authorized user to search and access data.

If the data set builds on already existing data or on complementary information stored in a different data set, this needs to be specified. In particular, the scientific link between the data sets needs to be described. Furthermore, all data sets need to be properly cited (i.e. including their persistent identifiers). If applicable, a Uniform Resource Locator (URL) giving reference to a web resource of such additional data (e.g., reference web pages (http), file transfer (ftp), email (mailto), database access (JDBC)) will be provided.

2.4 Increase data re-use

2.4.1 Publication and Creative Commons licenses

All peer-reviewed scientific publications produced under LiveSeeding will be published as open access immediately after being peer-reviewed and validated (see **Annex 1**). In fact, **under the new Horizon Europe rules no embargo period is allowed for scientific publications, and immediate OA is needed.**

Authors of publications produced under the LiveSeeding project will retain enough copyrights to be able to allow others reuse the content of their work, by using Creative Commons licenses, CC-BY or CC0 (see **Annex 2**).

¹⁴ <https://brapi.org/>



Before sending a manuscript out to a publisher

Author(s) can consult [Sherpa Romeo](#), an online resource that aggregates and presents publisher and journal open access policies from around the world in order to comply with the grant requirement.

Also



When sending a manuscript to a potential publisher:

Author(s) can write / clarify: "This work was funded by the European Union under an Horizon Europe grant. As set out in the Grant Agreement with donor, beneficiaries must ensure that at the latest at the time of publication, open access is provided via a trusted repository to the published version or the final peer-reviewed manuscript accepted for publication under the latest available version of the Creative Commons Attribution International Public Licence (**CC BY**) or a licence with equivalent rights. CC BY-NC, CC BY-ND, CC BY-NC-ND or equivalent licenses could be applied to long-text formats."

If the publishing agreement is contrary to the grant agreement obligations, authors should negotiate its terms and, alternatively, look for a different publishing venue/options writing. In other words, if the copyright agreement of the publishing company is in contradiction or doesn't allow you to fulfil the obligations of open access data and publications described above and in **Annex 1**, you need to renegotiate this or choose another journal where to publish the project related publications.

2.4.2 Data Quality assurance processes

All deliverables of LiveSeeding project will be reviewed by a Review Team composed of 4 members of the ExCom (2 Reviewers, a back-up Reviewer and Project Coordinator for format compliance review) as for the procedures and workflows set under the D8.2 Project Management Guidelines.

All scientific publications of LiveSeeding project will be peer-reviewed by an **Editorial committee** composed of the following project partners, selected for their academic profile and/or long experience with scientific research and publications: Dr. Véronique Chable (Institut National De Recherche Pour l'Agriculture, l'Alimentation et l'Environnement, INRAE), Dr. Vladimir Meglic (Agricultural Institute Of Slovenia, Kis),

Prof. Adrian Rodriguez Burruezo (Universitat Politècnica De Valencia, UPV), Prof. Pedro Manuel Mendes-Moreira (Instituto Politécnico de Coimbra, IPC) and Dr. Monika Messmer (Forschungsinstitut Für Biologischen Landbau Stiftung FiBL-CH) plus the Project Coordinator (as Convener) before publication to ensure the quality and coherence of the publications. IP rights of co-authors and potential innovation for exploitation will be screened on request by the IPRs Board and signed off by the ExCom.

2. Other research outputs

OHM Track will be a new software developed under LiveSeeding project to support seed producers and traders in managing the traceability and identification of OHM, complying with the delegated act on OHM (Commission Delegated Regulation (EU) 2021/1189). It is briefly described in Table 2 under chapter 1.1.

European Router Database, **SeedLinked** and **SHiNeMaS** will have new features, adaptations as briefly described in Table 2 under chapter 1.1 (adaptations sections for each software) and APIs as shortly described under chapter 2.3 for better interoperability.

As for the SeedLinked app, which will be localised under T2.4 for use in Europe for trials with farmers, gardeners and consumers, the coding will remain property of SeedLinked company, while data will remain under the ownership of trial managers (project partners) and users. Therefore, data collected from project partners using the SeedLinked app for project purposes will be treated like any other project data and will follow principles and procedures described in previous chapters.

Breeding materials developed under WP1 will be released as OHM or OV and the ExCom, after the advice of the IPRs board, will guide how they will be used and exploited within the framework of the principle of “as open as possible, as close as necessary”.

3. Data Management and allocation of resources

FiBL Europe has established a protected online collaborative workspace using Microsoft SharePoint for collection of data. The collaborative workspace is administered by FiBL Europe (WP 8 - Project coordination Leader) and regulated by a personal login and password system.

Task leaders will be responsible to manage data on the collaborative workspace for their task, under the supervision of WP leaders, as for guidelines of the present deliverable and the deliverable D8.1 Internal platform for communication and exchange manual. To that aim, a set period of time to be dedicated to preparing and collating materials for deposit has been factored and included in the project budget.

A public website of LiveSeeding project has been developed by IPS (WP7 – Communication, Dissemination and Exploitation Leader) and is available [here](#). Specific sections of the website will act as repository for deliverables, scientific and non-scientific publications, practice abstract, and other documents containing project data. Details about the structure and information included in the website is presented in the deliverable D7.2 Project website and social media. The cost of the development and maintenance of the online collaborative workspace and the public website are covered by the project budget. Financial sustainability of long-term curation of data will be ensured at two levels:

- Uploading data and publications in existing database of respective project partners, such as FiBL-CH organic statistical database;
- Uploading data and publications on external, reliable and free of charge repositories such as Zenodo, Organic eprints and others as for specifications under chapter 2.5.1.

As for the scientific publications, it is important to note that **under the new Horizon Europe rules, Article Processing Charges (APC) are not eligible** for reimbursement during the duration of the project, **in case of** traditional subscription-based journal that offers an OA option (so-called **Hybrid journals**). Also, APC costs are eligible if they are included in the project budget.

4. Data security

LiveSeeding SharePoint is the common platform for secure data management. The collaborative workspace has been set up and is administered by FiBL Europe. FiBL Europe is responsible of the setup and maintenance of the LiveSeeding SharePoint and manages access rights to partners in order to guarantee high functionality while protecting confidential data. Instruction on how to use the SharePoint are provided by the deliverable [D8.1 Internal platform for communication and exchange manual](#).

Each WP has a dedicated folder in the project SharePoint for data storage. WP (co)Leaders together with the Task Leaders will be responsible to define templates for (meta)data entry that facilitate exchange and merging of datasets. The WP leaders will ensure that data are correctly and timely deposited on the platform.

Partners will upload metadata and **anonymised data on the project SharePoint**. This means that non-anonymised data files, if applicable, will be stored in their own organisation's computer systems, following data security guidelines contained in [D8.4 Protection of personal data & recruitment of stakeholders](#), while files already anonymised will be uploaded on the common platform.

After the end of the project, FIBL EUROPE will continue to store the data for at least one more year on the LiveSeeding SharePoint and on ECO-PB website for an indefinite period to ensure long term curation and make it available for future analysis by the members of the LiveSeeding consortium (more details under chapter 2.5.1).

5. Ethics

LiveSeeding was assessed as “Ethics ready” and therefore REA required in the DoA the inclusion of a deliverable on Ethics issues and non-European compliance with internal mechanisms only and no need for an external ethics advisor.

Such deliverable [D8.3 on “Ethics Issues and Non-European compliance”](#) has been produced and uploaded on M03 (December 2022) in the EC Portal in line with it and provides the legal and contractual obligations of LiveSeeding project partners in relation to ethical issues and compliance of non-European partners and providers, as well as the guidelines and procedures/ tools in order to ensure such compliance.

Any exchange of Plant Genetic Resources (PGR) for food and agriculture will be done in accordance to the International Treaty for Plant Genetic Resources in Food and Agriculture (ITPGRFA) and the Nagoya protocol and the EU Access and Benefit Sharing Regulation (No 551/2014, Art. 2.2), as well as relevant national regulations. In particular, PGR material obtained by partners of the LiveSeeding consortium will be treated according to multilateral SMTA and, where this is not possible, recurring to bilateral MTA.

As for non-European compliance, all non-EU project partners and providers have confirmed by means of a signed statement that the ethical standards and guidelines of Horizon Europe and relevant legislation will be rigorously applied. Non-EU partners are all based in countries that received an adequacy decision of the European Commission in relation to privacy and data protection. As for the sole non-EU provider, agri-tech company and subcontractor SeedLinked, no ethical issues are envisaged.

6. Other issues

The procedures and compliance level required under the present deliverable are full encompassing. Therefore, no further procedures or measures are required in order to comply with other donors’ requirements (SERI and UKRI).

Annex 1 Open Science Policy under Horizon Europe



Important change under Horizon Europe: No embargo period is allowed for scientific publications, and immediate OA is needed;

All peer-reviewed scientific publications arising from Horizon Europe funding have to be made available in open access. This implies that publications are to be made freely available online, immediately upon publication and with no restrictions on use, by depositing them on a repository

Horizon Europe Open Access Mandate

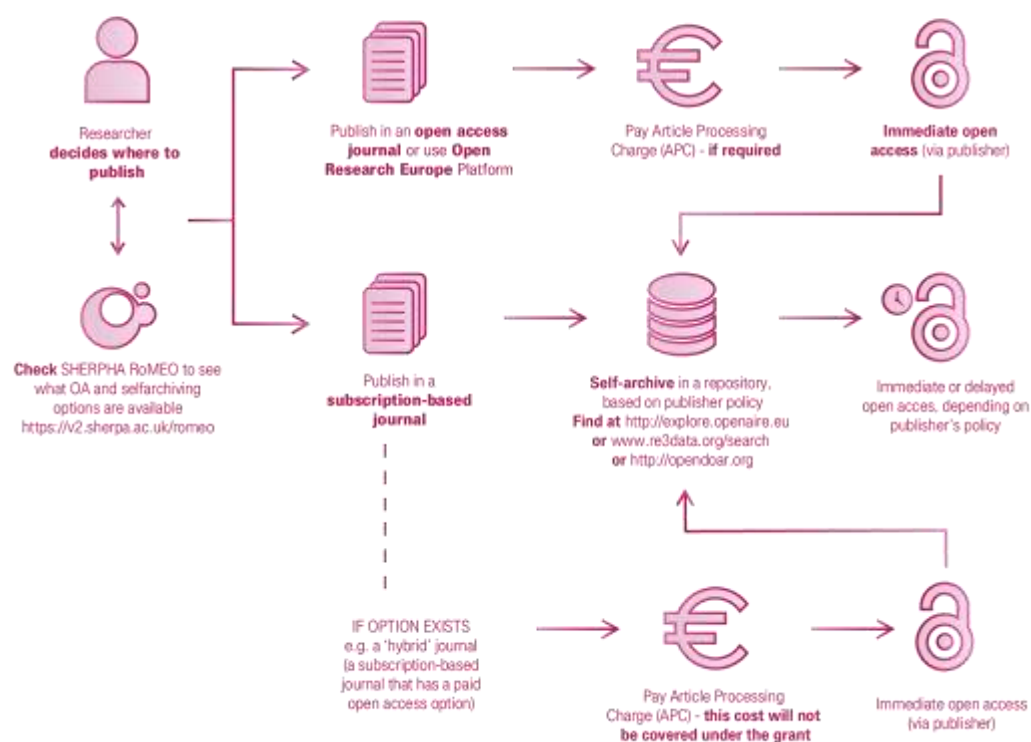


Figure 1 Horizon Europe Open Access Mandate (Source: [OpenAIRE](#), June 2022)

Annex 2 Creative Common Licenses

Creative Commons licenses give everyone from individual creators to large institutions a standardized way to grant the public permission to use their creative work under copyright law. In order to comply with the requirement for immediate open access (no embargo) of all peer-reviewed scientific publications arising from Horizon Europe funding, authors are required to retain sufficient intellectual property rights (IPR) to comply with these open access obligations.



This in practice means that authors should have an Open licence, preferentially CC-BY or CC0 licence

CREATIVE COMMONS LICENSES		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	✓	✓	✗	✗	✗

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

Figure 2: *Creative Commons Licenses*, (Source: Creative Commons, 2015, Wikimedia)

Annex 3 (Meta)Data file naming guidelines (Nomenclature)

The research DATA generated by partners during the project will be stored on the LiveSeeding SharePoint in the respective WP and Tasks folder.

Partners are requested to always name files using the following criteria (nomenclature):

- A. The file name must be unique and informative. The filename should contain involved task using the numbering of the Description of Action, partner organisation (like in GA of LIVESEED), topic, species, year of experiment,
- B. **No spaces** should be made in filename, but underscores can be used
- C. **Version number and status** indicated by following endings:
 - _RAW for preliminary data
 - _FIN for finalized data (validated)
 - _PRO for protected data
 - _OA for open access data
- D. **Language:** language shall be indicated with 2 letters (EN, ES, FR, DE, IT, in capital letters in the file name



LiveSeeding_WPn_Tn.n_DATA_filename_Version_Status_LANGUAGE
Exemple: LiveSeeding_WP3_Task3.3_DATA_FIBL-CH-pottrial-pea-2019_v2_RAW_DE

Final versions of documents from all WPs, used for the dissemination of LiveSeeding project will be stored in the project SharePoint as follows:

Scientific Publications :

- under [Scientific Publications](#)

Other Publications (including practice abstracts, brochures, posters, leaflets)

- Under the [Publications \(PA, brochures, leaflets, factsheets, etc\)](#) section of the WP7 Communication, Dissemination and Exploitation folder

Photos and Videos: under WP7Communication, Dissemination, and Exploitation.

- • Under the [Image database](#)

LiveSeeding SharePoint folders and sections contain also some guidance on what and how to store files.

Annex 4 General template for metadata of data and reports obtained in LiveSeeding

Please always add the project acronym (LiveSeeding) in the (meta)data and use "LiveSeeding" as a key word when uploading materials on organic eprints or other repositories

	
Funding Programme	Horizon Europe
Project Acronym	LiveSeeding,
Project Number	101059872
File name	add Filename
Licensing terms	CC0
Task no.	add Task number related to the file
Task Description	Title of Task or Subtask of LIVESEED
Title/Name	Title or short description of dataset
Involved Organisation(s)	Organisation responsible, who else was involved in data collection (add short name for organisation)
Author(s)	add responsible persons involved in collecting the dataset
Type of data	indicate type of data (agronomic, socioeconomic, molecular data) of qualitative or quantitative nature
Status of data	indicate status and level of protection of research dataset
Date	add year of data collection
Objective of data set	add objective why the experiment/research was performed
Use of existing data	indicate if existing data have been used, if yes, what kind of data were used
Original source(s): Download URL:	If applicable: The URL of the dataset in its original source
Original source(s):Licenses:	If applicable: The license under which the dataset is applicable from its original source
New collected data	indicate if new data have been collected and if yes that kind of data were collected
Size of data	Estimate or give exact size of dataset in kB

References	Publications or reports related to the dataset
Analysis	Indicate software and program of analysis of data
Experimental design	list factors tested, experimental design, number of replications, survey design for socio-economic data
Documentation of experiment	list environmental parameters that are important for interpretation of data (e.g. plot size, soil type, farming management, climatic data)
Standardisation	indicate which standard or template has been used for data collection and if this is in line with international ontology (structures vocabulary)
Parameters used	list of common parameters with clear descriptions
Data repository	indicate suited repository for storage and broader re-use of dataset
Value of dataset	Information about the potential value for exploiting and re-using the dataset. Including references to datasets that can be integrated for added value.
Comments	Special comments for the interpretation of the dataset
Annex a	Template for data collection
Annex b	Report, Presentation, Publications linked to data
Annex c	Demonstration material linked to data

Annex 5 Template for metadata of data obtained from field trials

Please always add the project acronym (LiveSeeding) in the (meta)data and use "LiveSeeding" as a key word when uploading materials on organic eprints or other repositories

	Documentation of Field Trial
File name	Add Filename
Task/Subtask no.	Add Task or Subtask related to the file
Task Description	Title of Task or Subtask of LIVESEED
Title/Name	Title or short description of dataset
involved Organisation(s)	which organisation is responsible, who else was involved in data collection (add short name for organisation)
Name of Site	
Manager of trial	
Farmer involved	
Location (physical address)	
Location (GPS data)	
Year of assessment	
Soil Type	
Soil Analysis	
Average Rainfall per year	
Average Temperature per year	
Year of conversion to organic	
Crop rotation	
Previous Crop	
Previous Intercrop	
CROP Species and Cultivars	(species Latin name, cultivar used, seed source, germination rate)

Experimental design	field	copy of field plan and experimental design
Time and way of soil preparation		
Time, quantity and type of fertilisation (incl. nutrient concentration)		
Time of sowing and type of sowing machine		
Plot size & Sowing density		
Time/frequency and quantity of irrigation		
Time/frequency, quantity type of plant protection (incl. concentration of active compound)		
Time/frequency and quantity of irrigation		
Time/frequency and type of weed control		
Time of harvest and harvesting equipment		
Parameters assessed:		description of measurement, code of parameter, assessment according to Reference or clear description, unit, at which growth stage, when (date), by whom
Average plant height at flowering (Code PH_fl_cm)		average plant height on plot level estimated based on 3 measures per plot from soil till top of plant in cm in 5 cm intervals at growth stage beginning flowering assessed by person xyz on the 21st May 2018
Leaf rust at flowering (Code LR_fl_5)		average severity of leaf rust estimated on plot level rated from 1 = no symptoms till 9= leaves are completely infected = dead tissue according to [give Reference] assessed 5 days after flowering by person xyz on the 26th May 2018

Annex 6 Examples of parameters for field trials' templates

(Source: <https://github.com/craig-willis/icasa/blob/master/icasa-mgmt-info.csv>)

AgMIP Category	ICASA Variable_Name	ICASA Code_Display	Description	Unit_or_type
Silver	harvest_yield_harvest_dw	HWAH	Harvested yield at harvest (dry wt)	kg/ha
Silver	harv_yield_harv_f_wt	HFAH	Harvested yield at harvest (fresh wt)	Mg/ha
Silver	harvest_moisture	HMAH	Moisture content of fresh yield at harvest	kg[water]/kg[harvest]
Silver	anthesis_dap	ADAP	Growth stage as days after planting, anthesis	dap
Silver	anthesis_date	ADAT	Growth stage date, anthesis	date
Gold	tops_dry_weight_maturity	CWAM	Tops dry weight at maturity	kg/ha
Gold	byprod_removed_at_harv	BWAH	By-product removed at harvest as dry wt	kg/ha
Gold	harvest_maturity_dap	R8AP	Growth stage days>planting, harvest mature	dap
Gold	physiologic_maturity_dap	MDAP	Growth stage of physiol. maturity, as days after planting	dap
Gold	physiologic_maturity_date	MDAT	Growth stage of physiol. maturity, date	date
Gold	evapotrans_cumul_mat	ETCM	Evapotranspiration, cumulative to end of season	mm
Gold	rainfall_in_season	PRCM	Rainfall, season	mm
Gold	canopy_height_max	CHTX	Canopy height, max value	m
Platinum	grain_unit_dry_wt_maturity	GWGM	Grain unit dry wt at maturity	mg
Platinum	harvest_no_at_maturity	H#AM	Harvest number per area at maturity	number/m2
Platinum	grain_protein_conc_maturity	GP%M	Grain protein concentration at maturity	%
Platinum	grain_N_at_maturity	GNAM	Grain N at maturity	kg/ha
Gold	date_of_measurement	DATE	Date of measurement	date
Gold	ear_number	E#AD	Ear number	number/m2

Gold	grain_dry_weight	GWAD	Grain dry weight	kg/ha
Gold	tops_dry_weight	CWAD	Tops dry weight	kg/ha
Gold	leaf_area_index	LAID	Leaf area index on a given day	m2/m2
Gold	tops_(ie_canopy)_N	TNAD	Tops (ie canopy) N	kg[N]/ha
Platinum	soil_layer_top_depth	SLDUB	Soil layer depth to upper boundary of sample (top)	cm
Platinum	soil_layer_bot_depth	SLDLB	Soil layer depth to lower boundary of sample (bottom)	cm
Platinum	soil_water_by_layer	SWLD	Soil water measured at a specified depth	cm3/cm3

Annex 7 Metadata parameters for surveys - Examples for templates

(Source: [Guide to Social Science Data Preparation and Archiving](#), 6th edition, Inter-university Consortium for Political and Social Research -ICPSR)

XML

ICPSR recommends using XML to create structured documentation compliant with the Data Documentation Initiative (DDI) metadata specification, an international standard for the content and exchange of documentation. XML stands for eXtensible Markup Language and was developed by the W3C, the governing body for all Web standards. Structured, XML-based metadata are ideal for documenting research data because the structure provides machine-actionability and the potential for metadata reuse.

XML defines structured rules for tagging text in a way that allows the author to express semantic meaning in the markup. Thus, question text — for example, `<question>Do you own your own home?</question>` — can be tagged separately from the answer categories. This type of tagging embeds “intelligence” in the metadata and permits flexibility in rendering the information for display on the Web.

Data Documentation Initiative (DDI)

At the outset of a project, we encourage data producers to generate documentation that is tagged according to the Data Documentation Initiative (DDI) metadata specification, an emerging international standard for the content, presentation, transport, and preservation of documentation (Blank and Rasmussen, 2004). The DDI specification is written in XML, which permits the markup, or tagging, of technical documentation content for retrieval and repurposing across the data life cycle. (See [“Getting Started with the DDI.”](#))

The Data Documentation Initiative (DDI) provides a set of XML rules specifically for describing social, behavioral, and economic data. DDI is designed to encourage the use of a comprehensive set of elements to describe social science datasets, thereby providing the potential data analyst with broader knowledge about a given collection. In addition, DDI supports a life cycle orientation to data that is crucial for thorough understanding of a dataset. DDI enables the documentation of a project from its earliest stages through questionnaire development, data collection, archiving and dissemination, and beyond, with no metadata loss.

DDI authoring options

Several XML authoring tools are available to facilitate the creation of DDI metadata. With a generic XML editor, the user imports the DDI rules (i.e., the DDI XML Schema) into the software and is then able to enter text for specific DDI elements and attributes. The resulting document is a valid DDI instance or file.

There are also DDI-specific tools, such as [Nesstar Publisher](#) and [Colectica](#), which produce DDI-compliant XML markup automatically. For more information on DDI and a list of tools and other XML resources, please consult the [DDI website](#).

Depositing DDI metadata

ICPSR encourages the deposit of DDI metadata with deposits of research data. There are currently two main versions of the DDI specification — DDI Codebook (Version

2.*) and DDI Lifecycle (Version 3.1). Most archives will prefer or at least readily accept documentation submitted in either of the DDI versions. To be in full compliance, a document should have question text integrated into each variable.

It may not be possible for a project to produce documentation that is DDI-conformant. In those situations, using a uniform, structured format with integrated question text is the best alternative, as it will enable the archive to convert the files to XML format easily.

Important metadata elements

Since most standard computer programs will produce frequency distributions that show counts and percents for each value of numeric variables, it may seem logical to use that information as the basis for documentation, but there are several reasons why this is not recommended. First, the output typically does not show the exact form of the question or item. Second, it does not contain other important information such as skip patterns, derivations of constructed variables, etc.

Key Social Science metadata items

A list of the most important items to include in social science metadata is presented below. Note that many of the high-level elements have counterparts in the Dublin Core Metadata Initiative (DCMI) element set. The DCMI is a standard aimed at making it easier to describe and to find resources using the Internet. For more information on the DCMI, please view its [website](#).

Principal investigator(s) [Dublin Core -- Creator]. Principal investigator name(s), and affiliation(s) at time of data collection.

Title [Dublin Core -- Title]. Official title of the data collection.

Funding sources. Names of funders, including grant numbers and related acknowledgments.

Data collector/producer. Persons or organizations responsible for data collection, and the date and location of data production.

Project description [Dublin Core -- Description]. A description of the project and its intellectual goals and indicates how the data articulate with related datasets. Publications providing essential information about the project should be cited. A brief project history detailing any major difficulties faced or decisions made in the course of the project is useful.

Sample and sampling procedures. This section should describe the target population investigated and the methods used to sample it (assuming the entire population is not studied). The discussion of the sampling procedure should indicate whether standard errors based on simple random sampling are appropriate, or if more complex methods are required. If weights were created, they should be described. If available, a copy of the original sampling plan should be included as an appendix. A clear indication of the response rate should be provided, indicating the proportion of those sampled who actually participated in the study. For longitudinal studies, the retention rate across studies should also be noted.

Weighting. If weights are required, information on weight variables, how they were constructed, and how they should be used.

Substantive, temporal, and geographic coverage of the data collection [Dublin Core -- Coverage]. Descriptions of topics covered, time period, and location

Data source(s) [Dublin Core -- Source]. If a dataset draws on resources other than surveys, citations to the original sources or documents from which data were obtained.

Unit(s) of analysis/observation. A description of who or what is being studied.

Variables. For each variable, the following information should be provided:

The exact question wording or the exact meaning of the datum. Sources should be cited for questions drawn from previous surveys or published work.

The text of the question integrated into the variable text. If this is not possible, it is useful to have the item or questionnaire number (e.g., Question 3a), so that the archive can make the necessary linkages.

Universe information, i.e., who was actually asked the question. Documentation should indicate exactly who was asked and was not asked the question. If a filter or skip pattern indicates that data on the variable were not obtained for all respondents, that information should appear together with other documentation for that variable.

Exact meaning of codes. The documentation should show the interpretation of the codes assigned to each variable. For some variables such as occupation or industry, this information might appear in an appendix.

Missing data codes. Codes assigned to represent data that are missing. Such codes typically fall outside of the range of valid values. Different types of missing data should have distinct codes.

Unweighted frequency distribution or summary statistics. These distributions should show both valid and missing cases.

Imputation and editing information. Documentation should identify data that have been estimated or extensively edited.

Details on constructed and weight variables. Datasets often include variables constructed using other variables. Documentation should include "audit trails" for such variables, indicating exactly how they were constructed, what decisions were made about imputations, and the like. Ideally, documentation would include the exact programming statements used to construct such variables. Detailed information on the construction of weights should also be provided.

Location in the data file. For raw data files, documentation should provide the field or column location and the record number (if there is more than one record per case). If a dataset is in a software-specific system format, location is not important, but the order of the variables is. Ordinarily, the order of variables in the documentation will

be the same as in the file; if not, the position of the variable within the file must be indicated.

Variable groupings. Particularly for large datasets, it is useful to categorize variables into conceptual groupings.

Related publications. Citations to publications based on the data, by the principal investigators or others.

Technical information on files. Information on file formats, file linking, and similar information.

Data collection instruments. Copies of the original data collection forms and instruments. Other researchers often want to know the context in which a particular question was asked, and it is helpful to see the survey instrument as a whole. Copyrighted survey questions should be acknowledged with a citation so that users may access and give credit to the original survey and its author.

Flowchart of the data collection instrument. A graphical guide to the data, showing which respondents were asked which questions and how various items link to each other. This is particularly useful for complex questionnaires or when no hardcopy questionnaire is available.

Index or table of contents. A list of variables either in alphabetic order or organized into variable groups with corresponding page numbers or links to the variables in the technical documentation or codebook.

List of abbreviations and other conventions. Variable names and variable labels often contain abbreviations. Ideally, these should be standardized and described.

Interviewer guide. Details on how interviews were administered, including probes, interviewer specifications, use of visual aids such as hand cards, and the like.

Coding instrument. A document that details the rules and definitions used for coding the data. This is particularly useful when open-ended responses are coded into quantitative data and the codes are not provided on the original data collection instrument.

Annex 8 Expected data generated from LiveSeeding WPs

WP and Task	Task lead	Source of Data	Type of data	Form of Data	Ethical/IP issues	Use of existing data	Restriction/confidentiality (require justification)*	Potential Repository
WP1 Increase and manage diversity	Time line							
T1.1 Genetic resources	M1 - M45	AEGILOPS (EL)	field trials	quantitative and qualitative crop data (morphological, agronomic and functional traits); sensory data; pictures	xlsx/csv; photos (jpeg, etc)	data ownership by each testing organisation	no	Seedlinked; PSR data base; UPV seed bank database; other public repository for genetic resources data publication
T1.2 Breeding Strategies	M1 - M48	INRAE (FR)	field/lab trials; updated version of the software SHiNeMaS	crop data for cultivar development; source code of SHiNeMaS	Trials:xlsx/csv SHiNeMaS: input/output data in csv, code in python, SQL, tar.gz	yes (LIVESEED, ReMIX, other past projects in which the breeding programme started); SHiNeMaS is already published, it will be improved during the project.	yes (LIVESEED, ReMIX, other past projects in which the breeding programme started); SHiNeMaS is already published, it will be improved during the project.	SHiNeMaS; other breeding softwares used by involved organisations; SHiNeMaS software: GitLab instance of INRAE https://forgemia.inra.fr/abisoft-qge/shinemas
T1.3 Breeding traits / tools	M1 - M47	KIS (SLO)	field/lab trials, bioinformatic analysis	protocols (phenotyping and screening tools); agronomic data; genomic data	xlsx/csv	Partly confidential	yes (LIVESEED, ReMIX, INVITE, INNOVAR, ECOBREED and other past projects)	Research activities of the task: no (data publication at the moment of scientific manuscript publication) Breeding activities of the task: Breeding data are related to the breeder IP and cannot be disclosed (unless breeder decide to publish)
T1.4 Training in breeding	M1 - M48	UPV (ES)	publicly available training materials including bibliography (books, reviews, journals, etc.); university courses and lessons; newly developed materials by the involved experts; results and methodologies of own LIVESEED and LIVESEEDING breeding activities; interviews	educational materials (video, podcast, booklet, presentations, infographic)	docx, xlsx, ppt, pdf, photos (jpeg, etc), video (mp4, etc)	no	partly, publicly available training materials; materials from past trainings of the involved organisations	no, although in some cases maybe data ownership agreement
WP2 Advancing Organic Cultivar Testing								
T2.1 Market uptake OHM	M1 - M36	DFH (DE)	field trials, surveys, workshops	quantitative and qualitative crop performance data, protocols	Word, Excel/csv	none	Yes, from field trials conducted within national and european research projects (EIP-Agri "Cereal Populations Hesse", LIVESEED), organic breeders (ECO-PB members) and national authorities	no
T2.2 Market access Organic Varieties	M1 - M36	LBI (NL)	surveys, workshops, field trials	field data, protocol recommendations	Word, Excel	none	ECO-PB, LIVESEED	As open as possible; in individual cases, the owner of a candidate OV might object to publish trial data to protect commercial interests
T2.3 on-farm cultivar evaluation	M1 - M48	ITAB (FR)	surveys, field trials	reports, quantitative and qualitative crop performance data, protocols, guidelines	Word, Excel/csv	none	Yes from national official/non-official cultivar testing programmes	no
T2.4 European interactive platform	M1 - M48	FIBL (CH)	On Farm Trials	qualitative and quantitative agronomic and sensory data	CSV files	Yes (data belongs to the testing network)	No	Data ownership agreement
T2.5 Training cultivar testing	M1 - M48	OMKI (HU)	existing, mostly publicly available training materials, interviews, focus groups, partly new project outcomes	text documents, webinars, videos files, infographics	avi, mp4, jpeg, pdf, word, html	none	LiveSeeding website - public	no

WP3 Seed production, seed sourcing, seed quality & health				Type of data	Form of Data	Ethical/IP issues	Use of existing data	Restriction/confidentiality (require justification)*	Potential Repository
T3.1 OHM seed production	M1 - M48	OMKI (HU)	seed producers and breeders, workshops, inetrvies, surveys	Traceability data on OHM breeding activities (pedigree, phenotyping) and OHM seed production (seed lot exchange, farm data) Seed quality data from field, pre-processing OHM maintenance practices	database, json, txt, csv	user should maintain the ownership of data, and fully control who can access their own data; seed quality related data have no ethical/IP issues	partly, some historical data will be acquired, along with crop and variety databases	partly confidential, breeding data are related to the breeder IP and cannot be disclosed (unless breedere decide to publish)	Relational database, reports, OEprints, Zenodo
T3.2 Seed Microbiome & Vigour	M1 - M48	WR (NL)	Experimental lab and field trials	Data on seed vigour, field performance and seed microbiome	Word, Excel	Partly confidential till publication	Partly	Confidential till publication. Scientific journals do not accept already public data	Organic Eprints, Zenodo
T3.3 improving seed health	M1 - M48	ITAB (FR)	Experiments with seed treatments and vigour testing with on farm seed; Internal workshops with partners and Living Labs; Collection of seed sanitation practices among partners and Living Labs; Desk study- bibliography and grain and seed quality / DOK trials	numeric data + pictures from seed testing and experiments; qualitative data and exchange of experience (workshop); protocols + images for seed sanitation	Word, Excel, photos (jpeg, etc)	partly confidential	partly	partly	Organic Eprints, Zenodo
T3.4 scaling up seed production	M1 - M48	FiBL (CH)	Workshops, focus groups, trials data from other projects	Assessments data, data based on stakeholder discussions	Excel, word, Klaxoon data	Partly confidential	partly	partly	Organic Eprints, Zenodo
T3.5 Training seed production	M1 - M48	ITAB (FR)	Produce training devices: workshop, expert interviews, training organization website / Implement training devices: data provided by trainees and trainers on their experience, their expectations	educational materials (videos, booklet, infographic) Produce training systems: list existing training systems (target audience, methods, duration, training organization, contact person, etc.) / Implement training systems: data provided by trainee (Name, First Name, profession, structure, experience, expectations, etc.) and by trainers (CV).	word, excel, pdf, mpeg, Miro data, video	Partly confidential	partly	partly	Organic Eprints, Zenodo
WP4 Improve Organic seed market transparency									
T4.1 Monitoring EU seed sector	M1 - M48	FiBL DE (DE)	expert consultations, survey	statistic data	excell	partly confidentially, Propriarity of FiBL	FiBL organic statistic data	raw data cannot be shared	FiBL database
T4.2 reaching 100% organic seed	M1 - M48	IFOAM OE (SE)	stakeholder interviews and focus group discussion	presentations, minutes, briefings, reports	PPT, pdf	partly confidentially	progress report, workshop report, booklets	Partly (certain data about policy context in specific countries may be confidential)	Organic Eprints
T4.3 EU router database	M1 - M48	FiBL DE (DE)	Software www.seeds4organic.eu; Application Programming interface (API) to www.organicXseeds.com and potentially other national databases	Software	Source code	Propriarity	organicXseeds, national seed databases	GDPR applies to software users; data on availability and derogation is published annually or available upon request	Organic Eprints
T4.4 policy recommendations	M1 - M48	IFOAM OE (SE)	Interview, case studies, focus groups	policy brief	pdf	no	reports of LIVESEED project, ECO-PB and other sources	no	Organic Eprints
T4.5 Train the trainer	M1 - M48	FiBL DE (DE)	external experts, project partners knowledge	report	pdf	no	reports of LIVESEED project and other sources	no	Organic Eprints

WP5 Exploitation: Horizontal proliferation and scaling-up organic seed and breeding initiative		Task lead	Source of Data	Type of data	Form of Data	Ethical/IP issues	Use of existing data	Restriction/confidentiality (require justification)*	Potential Repository
T5.1 Bottlenecks of seed markets	M1 - M24	IPS (HR)	Survey, quantitative and qualitative expert interviews (SMEs, Companies), workshops, review, LiveSeeding partners, SWOT analysis	Statistic data, reports, interviews	Excel, doc, ppt	Anonymization of personall data	LIVESEED survey	partly confidential	LiveSeeding SharePoint, LiveSeeding website, Organic Eprints
T5.2 financing & governance models	M1 - M36	FiBL (CH)	Review, quantitative and qualitative expert interviews (case studies, LLs, SMEs, Companies)	Statistic data, reports	Excel, doc	Anonymization of personall data	LIVESEED	Partly confidential upon agreement with the source (guidelines in DMP)	LIVESEEDING website
T5.3 European Task force for breeding	M1 - M43	IFOAM OE (SE)	workshops, focus group, interviews, survey, case studies	Reports, presentation	Excel, pdf, PPT	no	LIVESEED	partly (if the task force results in financial comitments from funders, the amount or funding source may be confidential)	LIVESEEDING website
T5.4 Training in seed entrepreneurship	M1 - M48	RCxAE (ES)	Survey, quantitative and qualitative expert interviews (case studies, LLs, SMEs, Companies), workshops, review	Statistic data, reports, presentations, photo and video	Excel, doc, ppt, jpeg, zoom, avi	no	LIVESEED	no	LIVESEEDING and ECO-PB website, Organic Eprints
T5.5 business models and plans	M24 - M48	IPS (HR)	LiveSeeding partners, results from tasks T5.1, T5.2, T5.3, T5.4, outputs from LLs, data from SMEs (min 10)	Statistic data, reports, presentations, business models	Excel, pdf, ppt	Anonymization of personal data	LIVESEED survey	confidential as business models are developed to improve competiveness of individual entrepreneurs of partners and Living lab actors.	LiveSeeding SharePoint
WP6 Marketing & Value-chains development, & awareness raising through participation									
T6.1 awareness, Benefits of OV and OHM	M1 - M24	FiBL (CH)	telephone interviews, participatory workshops, literature review	quantitative and qualitative	pdf and excel	Anonymization of personall data	LIVESEED, Engagement.biobreeding	no	Zenodo, Organic Eprints
T6.2 involving citizens and cities	M1 - M48	RCxAE (ES)	City and region responsible interviews, workshops, review	Statistic data, reports, presentations	Excel, pdf, ppt	Anonymization of personall data, unless allowed to mention involved persons	MUFPP position papers	no	LIVESEEDING website, Organic Eprints
T6.3 Cross fertilization living labs	M1 - M48	RSR (IT)	workshops	report	pdf	Anonymization of personall data, unless allowed to mention involved persons	European network of Living Labs, DIVINFOOD, DIVERSIFOOD	no	Organic Farm Knowledge
T6.4 on farm demonstration, knowledge sharing	M1 - M48	INRAE-R (FR)	public events participation	report	pdf	Anonymization of personall data, unless allowed to mention involved persons	templates and reports from LIVESEED	no	Organic Farm Knowledge
T6.5 Exploitation market outreach	M18 - M48	IPS (HR)	LiveSeeding partners, workshops, results from previous tasks T5.2, T6.1, T6.2, T6.3, outputs from LLs	Document, report, Power Point presentations	pdf, PPT	Public	yes	no	LiveSeeding website, Organic Farm Knowledge platform, Organic Eprints

WP7 Communication, Dissemination and training coordination		Task lead	Source of Data	Type of data	Form of Data	Ethical/IP issues	Use of existing data	Restriction/confidentiality (require justification)*	Potential Repository
T7.1 General dissemination PEDR	M1 - M48	IPS (HR)	Grant Agreement, LiveSeeding partners	Document, report	Word doc	Confidential/internal	yes	Yes. This is due to the nature of the PEDR, which represents an internal project's deliverable - plan for dissemination, communication and exploitation activities of the project.	LiveSeeding SharePoint
T7.2 Training task force	M1 - M46	RSR (IT)	Training material provided by partners. Existing and newly produced material	Report, booklet, presentations, Practice Abstract, Video	Pdf, Ppt, Video	Anonymization of personal data, unless allowed to mention involved persons	yes	no	LIVESEEDING website, Organic Eprints, Organic Farm Knowledge
T7.3 Stakeholder workshops	M1 - M46	IFOAM OE (SE)	Stakeholder workshop	Presentation, minutes	PPT, Word	Anonymization of personal data, unless allowed to mention involved persons	Stakeholder database from LIVESEED (stakeholders have been asked to re-confirm interest)	Partially confidential. Reports of workshops are public but contacts of stakeholder database remains confidential for GDPR compliance. Certain workshops with targeted audience will not be public.	LiveSeeding SharePoint, LiveSeeding website, Organic Eprints
T7.4 outreach to other countries	M1 - M48	IPS (HR)	LiveSeeding partners, external stakeholders (existing networks, collaborations, associations of the partners involved in the project)	Presentation	PowerPoint presentation (PPT)	Anonymization of personal data, unless allowed to mention involved persons	yes	no	LiveSeeding website, Organic Eprints
T7.5 Conferences	M36 - M48	FIBL EU	scientific conference and stakeholder/policy events	Presentations, abstract book, briefings	PPT, Pdf	no	yes	no	LiveSeeding website, Organic Eprints, Eucarpia, Zenodo
WP8 Project management									
T8.1 intraconsortium governance	M1 - M48	FIBL EU	PMG guidelines, ExCom meetings, General Assembly, Grant Agreement, Consortium Agreement	reports, minutes	pdf	yes, potentially (IP)	yes	confidential	LiveSeeding Sharepoint
T8.2 project administration & reporting	M1 - M48	FIBL EU	LiveSeeding partners	periodic technical and financial reports	pdf	no	no	confidential	EC portal
T8.3 data management	M1 - M48	FIBL EU	LiveSeeding partners, DMP	report	pdf	no	templates	no (because of its public nature confidential information from partner have not been included)	EC portal
WP9 Innovation Management									
T9.1 innovation steering	M1 - M48	FIBL CH	ExCom meetings, Advisory board, other projects, workshops, interviews	report, policy brief, minutes	pdf	yes, potentially (IP)	LIVESEED and other projects (EU and national funded)	Partly confidential (IP related)	LiveSeeding SharePoint, Organic Eprints
T9.2 IPR	M1 - M48	IPS (HR)	LiveSeeding partners, Grant Agreement	Guidelines	pdf	IP issues	no	Yes. This is due to the nature of the IPR guidelines that will be prepared for consortium use as they represent an internal project's document.	LiveSeeding SharePoint
T9.3 interoperability	M1 - M48	AEDIT (IT)	existing and newly developed ICT tools	database, vocabulary of crops and cultivars	json, txt, csv, XML	yes, potentially (IP)	yes	some data may be confidential due to IP	database, GIT

Annex 9 Protocol/Guidelines to upload project publications and documents in repositories

Dear LiveSeeding partners

We kindly ask you to upload all publications of LIVESEEDING, including scientific publications, presentations, posters, thesis, datasets, flyers, practice abstracts, policy briefs, training/teaching, conferences and workshops.

Please read carefully following instructions on best practice:

Excel table on the SharePoint

1. Fill in your planned scientific publications and conferences in the excel table on the SharePoint (LIVESEEDING_Scientific Publications Conferences Trainings_new 20.03.2019, sheet; planned scientific publications and conferences).
Approval of the Executive Committee (ExCom)
2. Send all publications to the ExCom for approval before submission.
Repositories (Organic eprints and Zenodo)
3. After the approval, you need to upload all publications on the repository organic eprints (Instruction 1) and the peer reviewed publications in addition on the repository Zenodo (Instruction II).
Excel table on the SharePoint
4. Fill in the details of all publications in the excel table on the SharePoint (LIVESEEDING_Scientific Publications Conferences Trainings_new 20.03.2019, sheet; Scientific publications, Training students young breeders or Scientific conferences Workshops).
EU-portal
5. Upload the scientific peer reviewed publications on the EU-portal (see Instruction III).

Please add on every publication following information:

For detailed information read the Data Management Plan

Step by step instructions I-III

I. Organic eprints (compulsory)

1. Go to <http://orgprints.org/>
2. Create an account (you need to register in order to be able to deposit items in the repository)
3. Click on "Manage deposits"
4. Click on "New Item"
5. Choose the type for your deposit
6. Click on "Next" or on "Details"

7. Fill in the details
8. Click on "Next" or on "Subjects"
9. Select at least one main subject category (under Production systems you can find "Breeding, genetics and propagation")
10. Click on "Next" or on "Affiliation"
11. Click on "European Union" and select "LiveSeeding" by clicking on the respective "Add"
12. Select additional names or codes that were involved in creating the deposit
13. Click on "Next" or on "Upload"
14. Upload File
15. Click on "Show options" and choose under "Visible to" the level of access restriction of this document
16. If you have data to upload, click on "Upload" and add a new file
17. Deposit the item
18. Click on "Manage deposits" to make changes or check the Item Status

II. Zenodo (compulsory for peer-reviewed publications)

1. Go to <https://zenodo.org>
2. Create an account (you need to register in order to be able to deposit items in the repository)
3. Click on "Upload"
4. Click on the green "New upload" button
5. Import the file by dragging it into the appropriate field or by clicking "choose files"
6. Click "start upload"
7. In the "community" box, add the LIVESEEDING community (there are also communities for FiBL and Eucarpia)
8. In the "License" box, enter the creative common code: CC BY-NC-ND 4.0 and choose the appearing: Creative Commons Attribution Non Commercial No Derivatives 4.0 International
9. In the "License" box, you can choose the access right
10. In the "Funding" box, find European Commission (EU) and enter the grant number: 727230.
11. In the "Basic information" box, add the DOI as Digital Object Identifier if your publisher already assigned one
12. Fill in all the other relevant information
13. Click "Publish"
14. The curator will have to approve the file in order for it to be added to the community

Note: once a file has been published, it cannot be removed or modified, but the information about the publication can be modified by going to the file and clicking the orange "edit" button. A new version of the file can be added by going to the file and clicking on "new version". Via ZENODO the publications are automatically stored on OpenAIRE, the European Union's electronic gateway for peer-reviewed articles and other important publications (www.openaire.eu).

By storing and publishing (after the embargo period) the scientific publications and other output on ZENODO, each partner of LIVESEEDING ensures that bibliographic metadata are included such as the funding body, the name of the action, acronym and grant number (already predefined in ZENODO), the publication date, and a persistent identifier.

III. EU-Portal (compulsory for peer-reviewed publications)

1. Go to [here](#)
2. Sign up now
3. In the row of the LIVESEEDING project click on "Actions" and choose "Manage Project"
4. Click on "Continuous Reporting"
5. In the new window, click on "Publications"
6. Click on "Manually add publications"
7. Fill in the details
8. In the "Repository Link" box, you should provide the direct link to the publication
9. After you have filled all information, click on "Add publication"

If you have any questions, please contact Mariano Iossa (Mariano.Iossa@fibl.org) - LiveSeeding Project Coordinator, AND Monika Messmer (Monika.Messmer@fibl.org) LiveSeeding Scientific Project Coordinator in CC

Annex 10 Ontology for interoperability

Species Use the UPOV code to ensure agreement on the species among the tools. https://www.upov.int/genie/species.shtml e.g., es. Durum wheat -> TRITI_TUR_DUR	Cultivar For crop variety or OHM use the official name of CPVO (Variety Finder): https://vf.plantvarieties.eu/varieties . Check for misspelling and synonyms. Concatenate the code to the UPOV code: e.g., Senatore Cappelli -> "TRITI_TUR_DUR - Cappelli".
Trait For phenotyping monitoring code the variable measured using cropontology (https://cropontology.org/). e.g., wheat flowering -> https://cropontology.org/rdf/CO_321:0000121	Seed lot + Seed transfer Associate to each seed lot a unique code: Uniform Resource Identifier (URI), to be assigned also to shinemas seed lots and OHMTrack. URI can be: • the accession code of a seed bank e.g., https://db.rsr.bio/?accession_code=RSR03012 • a DOI (FAO) e.g., https://doi.org/10.18730/Z0TYN Keep track of the codes during seed transfers

Standard for Seed Lot - URI

- DOI can be the optimal solutions; but may have some difficulties (owner should have a VIEW code, an xml should be prepared)
- An alternative may be the use of a Unique Resource Identification (URI)
- URI is a string similar to a web address; with a part that identifies an organisation (the url of the org) and a part that is unique for each seed lot of the organisation (usually the accession code)
- example
 - https://db.rsr.bio/accession_code/RSR20001
 - <https://www.organicxseeds.ch/search/listProductDetail?offer=9ec3829f-619a-4542-8457-9adcdf4b571>
- the URI is a code, the organisation does not need to have on the website a responding page
- the organisation needs only to ensure that that code (e.g. RSR20001) is not reused for something else