



INnovations in plant Varlety Testing in Europe

Deliverable D7.2 Data Management Plan

Technical References

Project Acronym	INVITE
Project Title	INnovations in plant Variety Testing in Europe
Project Coordinator	François Laurens
Project Duration	60 months
Deliverable No.	D7.2
Dissemination level ¹	Public
Work Package	WP 7 - Database management and data interoperability
Task	T 7.1 – Data management plan and operational implementations
Lead beneficiary	24 (ACTA)
Contributing beneficiary(ies)	
Due date of deliverable	31 December 2019
Actual submission date	23 December 2019

¹ PU = Public

PP = Restricted to other programme participants (including the Commission Services)

RE = Restricted to a group specified by the consortium (including the Commission Services)

CO = Confidential, only for members of the consortium (including the Commission Services)

Document history

V	Date	Beneficiary	Author
1	18/12/2019	Internal validation (project coordinators and WP7 members) + Submission to WP leaders	Géraldine Hirschy (WP7 leader)
2	Next step	All project participants for a review and check if there is any specific data constraints (early January)	



Summary

This document presents the Data management Plan (DMP) of the INVITE project. It describes the data management policy adopted within the project for all the datasets used or generated in the project. The plan covers every stage of the data lifecycle and provides guidelines and procedures to use data generated and used in the project. It will be regularly updated as a living document. The plan specifies which type of data repositories partner institutions will use, and which procedures will be put in place for long-term and safe preservation of the data. The document will also describe the data management procedures for the project, including measures to comply with expectations of the Open Research Data Pilot. The main elements of this data management plan include dataset reference and name, dataset description, standards and metadata, data sharing, data archiving and preservation.



Table of content

1	INTRODUCTION	5
2	DATA SUMMARY	6
2.1	DATA USED IN THE WORK PACKAGES ACTIVITIES	6
2.1.1	PRIMARY DATA	7
2.1.2	SECONDARY DATA	7
2.2	PROJECT MANAGEMENT DATA	8
2.3	PROJECT DOCUMENTS	8
2.4	DATA MANAGEMENT TOOLS	9
2.5	DATA ORIGIN	10
2.6	DATA MANAGEMENT RULES	10
3	FAIR DATA	10
3.1	MAKING DATA FINDABLE	10
3.2	MAKING DATA OPENLY ACCESSIBLE	11
3.2.1	ACCESS TO RESEARCH DATA	11
3.2.2	ACCESS TO SCIENTIFIC PUBLICATIONS	11
3.2.3	AGREEMENT ON ACCESSIBILITY TO DATA AND MATERIALS FROM PARTNER EOs AND PROs FOR INVITE	12
3.3	MAKING DATA INTEROPERABLE	12
3.4	INCREASE DATA RE-USE	13
3.4.1	DATA QUALITY ASSURANCE PROCESSES	14
3.5	ALLOCATION OF RESOURCES	14
3.6	DATA SECURITY	14
3.7	ETHICAL REQUIREMENTS	15
	CONCLUSION	16
	ANNEXES	17
	ANNEX 1 – LIST OF INVITE DELIVERABLES	17
	ANNEX 2 – LIST OF VCU AGRO-ENVIRONMENTAL VARIABLES REQUESTED BY WP4	20
	ANNEX 3 – LIST OF VCU VARIABLES REQUESTED BY WP4 FOR WHEAT	24
	ANNEX 4 – LIST OF VCU VARIABLES REQUESTED BY WP4 FOR MAIZE	25
	ANNEX 5 – LIST OF VCU VARIABLES REQUESTED BY WP4 FOR SUNFLOWER	26
	ANNEX 6 – PROPOSED PROCEDURE FOR ACCESS TO DATA AND MATERIAL (DEC 17TH 2019)	27



1 Introduction

A well devised Data Management Plan is essential in designing and producing a high-quality research project. The Data Management Plan covers every stage of the data lifecycle and provides guidelines and procedures to use datasets generated in the project. The plan specifies which type of repositories the partner institutions will use, and which procedures will be put in place for long-term and safe preservation of the data.

Researchers may continue to work on data after funding has ceased, follow-up projects may analyze or add to the data, and data may be re-used by other researchers. The Data Management Plan therefore needs to go beyond the life of the project. The plan is a living document that must react, evolve and develop, to respond to changes in the data and or project.



2 Data summary

The INVITE project aims at improving efficiency of variety testing and availability of information to stakeholders on variety performance under diversified production conditions and on biotic and abiotic stresses.

One expected impact is the introduction of plant traits into the testing protocols that respond to new challenges and demands in the conventional and organic sectors, while also considering the economic return of growers.

INVITE will collect, analyze and disseminate different types of data, in order to (1) simplify the variety description, (2) enable exchanges of varieties performance information between stakeholders, (3) increase the efficiency in variety testing and (4) ease the promotion of varieties that are more adapted to sustainable agricultural practices and more resistant to climate change.

The data managed during this project can be divided in two main categories: the data collected and/or used in the work packages activities (including primary and secondary data), and the data used for the project management.

In addition to these datasets, documents and codes will also be produced during the project. These information materials are different from datasets, as they are not only a set of raw and unorganized facts, but supports to provide information. Information is processed, structured, or presented in a given context to make it meaningful and useful.

2.1 Data used in the work packages activities

Primary data refer to the data originated by the researcher for the first time. It is all the new data that will be produced (collected and generated) by the consortium during INVITE for the specific needs of the project. Primary data in INVITE include:

- **Raw data, that will be collected** through various means;
- **Generated data, that will be computed** based on this collected raw data.

Such primary data may be of use for future research, after INVITE has ended, and for this reason as many data as possible will be put in open access. Any reuse will nevertheless require the permission of the data owner.

Secondary data refer to already existing data collected earlier. It is all the data that had been generated outside INVITE for other purposes. INVITE will make use of existing secondary datasets, from European, international, national and local sources.



2.1.1 Primary data

2.1.1.1 Raw data (collected data)

Raw data are any kind of data that will be directly collected by INVITE partners. It includes:

- Phenotypic data: plant measurements (transpiration efficiency), morphological data (roots, aerial parts), etc.
- Genotypic data: molecular data, genome sequences, etc.
- Agronomic data: disease observations, number of spores in fields, grain yields, phenological stages, etc.
- Environmental data: weather data, soil characteristics, etc.
- Field management data: as-applied nitrogen doses, farming operations dates and types, etc.

These raw data will be stored on the project's data sharing system, with restricted access for the partner who collected them, and the WP participants who will analyze them.

2.1.1.2 Generated data

Generated data include any data that will be generated by the project partners. It includes:

- DUS data (Distinctness, Uniformity and Stability).
- VCU data (Value for Cultivation, Use and Sustainability).
- Crop models outputs.

All generated data will be stored in the project's data sharing system with restricted access to the partner who collected them, and the WP participants who will analyze or use them.

2.1.2 Secondary data

The secondary data (also called in the project "historical data") include all the additional information on varieties traits already available at public or private levels.

Two types of secondary data will or may be used:

- Public data from national, European or international data on varieties, published by public authorities. These data will be stored on the data sharing system if necessary.

Private variety data from partners or stakeholders, coming from other research projects or from private institutions (most of the time Examination Offices). A translated and pseudonymized synthesis of these data will be stored with restricted access to the partner who collected them, and the WP participants who will analyze them (see list of requested traits in Annex 2 – List of VCU agro-environmental variables requested by WP4, Annex 3 – List of VCU variables requested by WP4 for wheat, Annex 4 – List of VCU variables requested by WP4 for maize and Annex 5 – List of VCU



variables requested by WP4 for sunflower). See §3.2.3 of this document for the rules to exchange historical data.

2.2 Project management data

In INVITE, project management data include contact details of persons will be collected and used during the project for management (in WP9- Consortium coordination and project management) and communication (in WP8 - Dissemination and optimization of exploitation of results) purposes.

More specifically, management data are:

- Project's management data collected and processed by INRA Transfert and INRA (for WP9) and any other partners, such as contact details and other personal information of INVITE consortium partners, including personal data collected for organizing consortium meetings;
- Project's communication data collected and processed by ARCADIA and EUROSEEDS (the partners responsible of WP8), for the project's communication and dissemination activities;
- Human resources data for each partner, e.g. when recruiting short-term employees or trainees to work in the project.

This list of data produced will be reviewed and updated periodically to ensure all data sets are included in the data management plan.

2.3 Project documents

During the project and in addition to the datasets that will be used, several documents and deliverables will be produced by the participants:

- Methodological data: trials protocols, etc.
- Project deliverables: all the deliverables and milestones listed in the grant agreement (see Annex 1 – List of INVITE deliverables).
- Code: code of the database prototype.

In addition to the deliverables listed here general activities of the INVITE consortium will generate:

- Press articles in partners' media or other media related to agriculture, publications internet posts through social media;
- Pictures from events, demonstration activities, meetings;
- An email list, used for INVITE's outputs dissemination, which will be confidential;
- PowerPoint presentations or posters from events / fairs / conferences;
- Scientific publications (researchers will be encouraged to publish the data used to generate scientific papers in public repositories such as [OpenAIRE](#)).



2.4 Data management tools

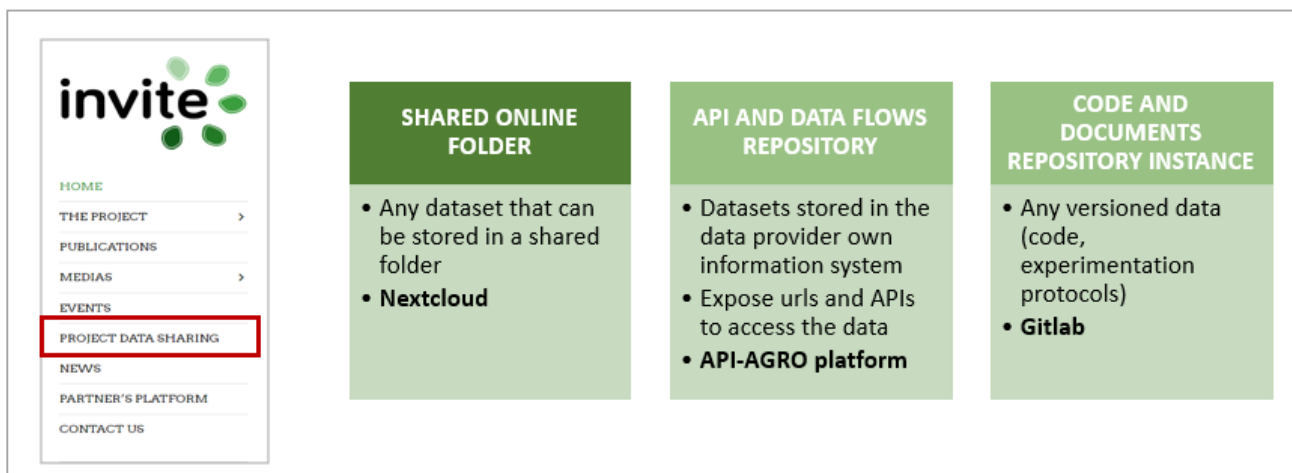
All the datasets (from historical datasets or generated in the tasks) and tools that will be used, generated and updated by the project's participants will be centralized in a single place, the **data sharing system (see description in the Deliverable 7.3)**.

The WP participants will use and combine the datasets stored in this system in order to produce new results analysis and to validate the applicability of the developed tools.

The data sharing system will be a web portal (hosted in the INVITE website) including three tools:

- 1) A shared online folder (Nextcloud software, hosted exclusively in the EU) to store any data types and datasets.
- 2) An API and data flows repository (using API-AGRO platform) to expose remote files via URLs and API where some datasets can be accessed. In this case, the data provider is responsible for the data storage and the data access management through the API-AGRO platform.
- 3) A code and documents repository instance (Gitlab software) hosted by Naktuinbouw for the prototype code and versioned datasets. It enables version control on codes and documents.

All three systems accessible via the web portal will be password protected. Content can only be shared within the project according to the access rights defined by the data owners. All WPs will input data to the listed systems by providing different datasets for the different crops. Any relevant data use limitations will be made clearly visible in the project data repository according to the access rights inventory made in Task 7.3.



In order to guide the project partners to choose the right system according to their use cases, a decision tree is provided in this document.



This decision tree will be used in the user interface of the web portal. It is essential that the decision tree remains easy to understand and simple to use, in order to provide clear and actionable information to the data providers and consumers.

2.5 Data origin

Data origin will be reported in each deliverable. The use of existing data or the generation of new data will be clearly identified in each deliverable produced. The ethical clearance of all data both new and existing will be addressed by each work package leader and will be clearly reported upon. For further ethical information see section 8.0 and deliverable reports D10.1 and D10.2.

2.6 Data management rules

Specific rules will be enforced as regards data collection, management, processing, storage and erasure, whether they are primary data, secondary data or management data.

Firstly, when processed data during the project contain personal data, the application of the EU General Data Protection Regulation (GDPR) 2016/679 will be strictly followed (see Del 10.1 and Del10.2). The five principles relating to processing of personal data (i.e. Purpose limitation, Accuracy and data minimization, Storage limitation, Security, integrity and confidentiality, Rights of the data subject) will be accordingly addressed. Moreover, when research data (primary or secondary) data will encompass personal data, only pseudonymous data will be centralised and made available for the use of INVITE partners. Secondly, each INVITE partner will commit to complying with the EU and national legal rules governing the use, transfer, storage and erasure of secondary data. Such data will be obtained from authorities or organizations that are knowledgeable as regards the rules on data protection and that will transfer data to INVITE partners in compliance with these rules.

3 FAIR data

3.1 Making data findable

The data produced and/or used in the project, and stored in the project data sharing system, will be identifiable and locatable by means of a standard identification mechanism.

Keywords and tags will be assigned to each dataset and subsets in order to facilitate their re-use. For more precision on Metadata, see Making data interoperable section.



3.2 Making data openly accessible

INVITE is included in the European Commission's Open Research Data Pilot that "aims to make the research data generated by Horizon 2020 projects accessible with as few restrictions as possible, while at the same time protecting sensitive data from inappropriate access".

3.2.1 Access to research data

As described in section 2, the raw and generated data will be stored in the data sharing system, with restricted access. Each tool of the web portal is confidential and members can access it via login and password.

Results (including intermediary results) will be stored with open access to all the consortium and the registered stakeholders. Consortium members will have a complete access to the results on the project data sharing system, as well as on the collaborative platform (for project deliverables). Registered stakeholders will have access to some results (only the ones that will compromise any future publications – for ex. Charts but not raw data) on the stakeholders' platform that will be accessible on internet through the project's website with a registration/identification procedure.

INVITE will make publicly available as many of its research data as legally possible, except management and personal data. The related data will be made available through open data repository(-ies) (to be defined) and will be given appropriate key-words, along with a Digital Object Identifier (DOI), and be accompanied by appropriate metadata.

Each project partner is accountable for the dissemination of the research data, according to the Grant Agreement and consortium constraints.

3.2.2 Access to scientific publications

In INVITE, all written outputs (deliverables, academic and professional articles) that are not confidential will be published in 'green' open access at minimum, i.e. will be made publicly available through personal or institutional repository as soon as they are delivered to the European Commission via the portal. When appropriate, peer-reviewed academic articles will be published with 'gold' open access. The costs for such open access are eligible costs during the life of the project¹.

¹ Extract from the European Commission's Open Research Data Pilot (http://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-data-management/open-access_en.htm): 'The 2 main routes to open access are:



Each project partner is accountable for the dissemination of its own scientific publications, according to the Grant Agreement and consortium constraints.

The bibliographic metadata must be in a standard format and must include all of the following:

- The terms “European Union (EU)” and “Horizon 2020”;
- The name of the action, acronym and grant number;
- The publication date, and length of embargo period if applicable ;
- A persistent identifier.

3.2.3 Agreement on accessibility to data and materials from partner EOs and PROs for INVITE

To reach the objectives described in the INVITE R&D proposal (H2020-SFS29), access to specific data and materials held by the partner Examination Offices (EOs) and Post-Registration Organizations (PROs) is necessary and will be given under conditions validated by all partners in the future Consortium agreement.

A procedure has been defined by the partners and will be submitted to the Examinations Offices (in January 2020).

The procedure needs to be formalized and validated by all the EOs ; a first version is described in the Annex 6 ; the DMP will be updated with the validated procedure as soon as it is finalized.

3.3 Making data interoperable

INVITE partners will try and make the open data as interoperable as possible by using common data format in the disciplines present in INVITE (and converted into open format whenever possible), and standard vocabularies of the disciplines where data are usually used, and when possible alternative vocabularies so that it can be understood in other disciplines.

- Self-archiving / 'green' open access – the author, or a representative, archives (deposits) the published article or the final peer-reviewed manuscript in an online repository before, at the same time as, or after publication. Some publishers request that open access be granted only after an embargo period has elapsed.

- Open access publishing / 'gold' open access - an article is immediately published in open access mode. In this model, the payment of publication costs is shifted away from subscribing readers. The most common business model is based on one-off payments by authors. These costs, often referred to as Article Processing Charges (APCs) are usually borne by the researcher's university or research institute or the agency funding the research. In other cases, the costs of open access publishing are covered by subsidies or other funding models.'



In order to ensure that the metadata are machine-readable, the metadata will be described according to the standard [Dublin Core Terms](#) vocabulary and provide the following information:

- Contributor
- Coverage (country, actor...)
- Creator
- Date
- Description (summary)
- Format and size
- Identifier
- Language
- Publisher
- Relation (url...)
- Rights
- Source
- Subject (keywords)
- Title
- Type (text, figure, image...)

If this is required by the project participants, additional metadata may be added in this initial scope (after having assessed their added-value and having checked the feasibility in the Data Sharing System).

Also, in order to ease the implementation of common crop models and increase the re-use of the data, the models will follow the AgMIP standard (Agricultural Model Intercomparison and Improvement Project), that ensures the compatibility among models, and thus allows the comparison of observed and simulated data.

3.4 Increase data re-use

Generated data will be made re-usable after the end of the project or after their use in publications by the project members, with the permission of the owner. No time restrictions will apply.

Raw data will only be re-usable by project members, with the permission of the owner.

Standard API is also a way to make the data visible at the end of the project and can increase the data re-use and data interoperability among different information systems.



3.4.1 Data quality assurance processes

There will be a quality control step of the datasets provided by the Examination Offices (historical phenotypic data): a data quality assurance process will be defined prior to any use of the data by the WP4 partners (modelling).

The detailed procedure will be added in this document as soon as it will be validated by all the project partners.

3.5 Allocation of resources

Each INVITE partner is responsible for collecting and storing their own data, whether primary research data, secondary research data or management data.

Once the activities of the project are terminated, long-term preservation of data is preferred, where possible, to data destruction for historical and scientific interests of future generations.

- Data that can be made available in open access will be deposited on a data repository as explained in section “Access to scientific publications”, whose cost is supported by the repository provider and not by INVITE partners.
- **Five years after the project has ended, primary data** (including accompanying documents such as signed consent forms) **will be destroyed or archived** to comply with national laws. The place for archiving will be explored during the project and discussed among INVITE partners.
- **Management and communication data** will not be made publicly available, unless based on an explicit consent, and **will be destroyed 3 years after the end of INVITE or archived** if requested by EU or national laws.

The **costs incurred for data management during the project will be supported by each partner with the INVITE budget** and the costs that were stipulated in the Description of the Action (e.g. Gold Open Access fees) or their own institutional budget (e.g. salary of dedicated people for data management).

3.6 Data security

Data security is needed for personal data, as well as for secondary data governed by specific rules. This is important to establish trust between the data providers and the data (re)users.

INVITE partners will ensure full protection of these data (technical and organizational security measures), including protection against unauthorized or unlawful processing and against accidental loss, destruction or damage. In particular, uploading files and data will be done on electronic or computer-based supports that fully comply with the EU GDPR security requirements.



Prior the sharing for the analysis all data containing sensitive personal information has to be anonymized. Anonymization refers to removing any identifier that can reveal identity of the participants both from data and metadata.

On a daily basis, when INVITE partners will be working on the project data, they will be stored on computers, laptops, intranets or hard-drives accessible through institutional credentials, with passwords periodically modified according to national law provisions for data security and protected by regularly updated antiviruses. None of the project data will be left inadvertently available, by being left on desks or in unlocked rooms. If mobile devices will be used to store data files (e.g. backup files), they will be kept in a safe place accessible only to the researchers involved or will be encrypted with ad-hoc software (e.g. in case of personal/sensitive data). Regarding data recovery, all the research materials stored in computers will be subjected to regular backup in order to safeguard them from accidental losses.

In terms of long-term preservation, as explained above, the place for archiving INVITE data will be decided jointly during the project. The objective is to find an archiving place which respects integrity and authenticity of the data, guarantees the readability of it (e.g. if formats and software change in the future), the understanding of the data in the mid and long-term, and provides access.

3.7 Ethical requirements

INVITE partners will comply with the ethics WP (WP10) that has been created by the European Commission in line with the ethics requirements raised in the Ethics Summary Report.

An ethics mentor has been appointed on Month 3 (Deliverable 10.5) to advise the project participants on ethics issues. A report on the activities of the ethics mentor will be kept on file.

The Ethical Deliverables 10.1, 10.2., 10.3, and 10.4 are submitted to the European Commission at the same time as the Data Management Plan.

Partners appointed Data Protection Officers and these ethical reports explain how the project will comply with GDPR principles for personal data, as well as the specificity for research undertaken in non-EU-countries. See these deliverables for further information.



Conclusion

This document has presented the Data Management Plan of INVITE, including the rules for data protection and storage, as well as strategies for making data findable, accessible, interoperable and reusable (FAIR data).

This is the first version of the DMP and it will be regularly updated along the project; a review of the DMP is planned every 6 months.

The lists of VCU variables added from annexe 2 to annexe 5 are not the final lists, they will be confirmed during January 2020.



Annexes

Annex 1 – List of INVITE deliverables

Deliverable	Short description	Data type
D1.1	Probes specific to stress related genes in tomato designed	Report
D1.2	Correlations between root, yield stability and plant physiological traits as a response to drought stress	Report
D1.3	Bioindicators for crop resilience against biotic stress	Report
D1.4	Bioindicators for wheat tolerance to water stress	Report
D1.5	Bioindicators for tomato adaptation to challenging conditions	Report
D1.6	Epigenetic bioindicators for variety identification and evaluation of plant adaptation to environment	Report
D1.7	Evaluation of the potential extrapolation of WP1 bioindicators to other crops.	Report
D2.1	Cost-benefit analysis for low-cost sensors to decide the best technologies to be validated in WPs 5 and 6	Report
D2.2	“Shopping list”, “Do it yourself custom list”, and measurement procedure to guide EOs and PROs in the implementation of the two non-RGB lowcost phenotyping tools	Report
D2.3	Dataset of DUS and VCU trial with low-cost sensors and standard reference measurements	Other
D2.4	Annotation of the dataset of DUS and VCU trial with low-cost sensors and standard reference measurements	Other
D2.5	Classifier and regression software to automate DUS and VCU trials	Other
D2.6	RGB low-cost phenotyping tools selected to be demonstrated for their applicability in WP5 and 6	Other
D3.1	Genome-wide matrix of polymorphism data of PRG varieties and reduced marker set array	Other
D3.2	Genome-wide matrix of polymorphism data	Other



	and reduced marker set array of wheat, maize and soybean varieties	
D3.3	Gene-specific markers implemented in genotyping array for potato, apple and tomato	Other
D3.4	Marker trait associations for at least 1 DUS characteristic for each of wheat, maize, PRG, and soybean	Other
D3.5	Novel marker-based methods for DUS testing implemented in prototype software	Other
D4.1	Efficiency assessment of current national performance and VCU trials	Report
D4.2	Crop growth model predictions for environmental characterisation	Report
D4.3	Environmental characterisation of current trial networks	Report
D4.4	Enriched mixed model methodology for high precision variety prediction	Report
D4.5	Genotype specific crop growth models for variety performance prediction	Report
D4.6	Environmental characterisation for future trial networks	Report
D4.7	High resolution Decision Support System for variety choice	Report
D4.8	Evaluation of prediction methodology for potato, tomato, lucerne, soybean, and oil seed rape	Report
D5.1	Recommendations for the testing of heterogeneous plant material	Report
D5.2	Proposal of a multi criteria evaluation in variety testing	Report
D5.3	Recommendations for variety testing for organic farming	Report
D5.4	Applicability of phenotyping tools for variety testing	Report
D5.5	Protocols for sustainability assessment in variety testing	Report
D5.6	Applicability of molecular markers for distinctness and management of reference collection	Report
D5.7	Assessment of the added value of phenotyping, genotyping tools and sustainability criteria for variety testing	Report
D6.1	Handbook of harmonised assessment/scoring scales in VCU testing. Guidelines towards more integrated and efficient DUS and VCU testing	Report
D6.2	Criteria for the socio-economic and environmental evaluation of innovations in variety testing	Report
D6.3	Demonstration of optimised METs compared with current practices. Technical advantages of adopting new MET configuration. Guidelines for the implementation of newly defined METs	Report



D6.4	Cost/benefit model for EOs to estimate cost and benefits of practices in their specific testing conditions	Other
D6.5	Full analysis for decision makers of the impacts of modifying official rules in DUS and VCU at both EU and national levels	Report
D7.1	Decision tree for data categorisation	Other
D7.2	Data Management Plan	ORDP: Open Research Data Pilot
D7.3	Operational web portal for storing and sharing project data including a data repository with API, a shared online folder, and a private code repository	Websites, patents filling, etc.
D7.4	A unified ontology for DUS and performance testing data	Other
D7.5	Guidelines and specifications for future databases for storing DUS/VCU and genetic data	Report
D7.6	European database prototype	Other
D7.7	European database userfriendly interface and API	Demonstrator
D8.1	Communication toolkit for electronic communications (logo and graphical chart, leaflet, standard PPT and poster templates, factsheets, website, and social network)	Other
D8.2	Updated project dissemination plan with all targeted events and actions including a mapping of actors to be targeted	Report
D8.3	Layout for the ENewsletters and publication of the first one	Report
D8.4	Minutes of the workshops, stakeholders' workshops, and field days organised during the second year	Report
D8.5	Minutes of the workshops, stakeholders' workshops, and field days organised during the third year	Report
D8.6	Minutes of the workshops, stakeholders' workshops, and field days organised during the fourth year	Report
D8.7	Potential evolution of variety testing regulation as seen by policy makers	Report
D8.8	Compilation of all ENewsletters published during the project	Report
D8.9	Webspace on the project website that will include all innovation factsheets (about 15 recommendation sheets) compiled in WP6 and based on results from WP1 to WP5	Other



D8.10	Minutes of the workshops, stakeholders' workshops, field days, and Final conference in Brussels organised during the final year	Report
D9.1	Kick-off meeting and its minutes	Report
D9.2	Management guidelines	Report
D9.3	Project collaborative workspace	Other
D9.4	First annual meeting and its minutes	Report
D9.5	Second annual meeting and its minutes	Report
D9.6	Third annual meeting and its minutes	Report
D9.7	Fourth annual meeting and its minutes	Report
D9.8	Review of the collaboration with Innovar during the first year	Report
D9.9	Review of the collaboration with Innovar during the second year	Report
D9.10	Review of the collaboration with Innovar during the third year	Report
D9.11	Review of the collaboration with Innovar during the fourth year	Report
D10.1	H - Requirement No. 1	Ethics
D10.2	POPD - Requirement No. 2	Ethics
D10.3	NEC - Requirement No. 3	Ethics
D10.4	OEI - Requirement No. 4	Ethics
D10.5	GEN - Requirement No. 5	Ethics

Annex 2 – List of VCU agro-environmental variables requested by WP4

The lists from Annex 2 to Annex 5 are not final lists, they will be confirmed during January 2020

Description	Requirement level
Names of institution responsible for the experiment or trial	required
Name of country where field is located	required
Third admin level for location of site for soil profile (usually county or municipality)	optional
Name of site for experiment	required
Name of experiment or trial	desired
Experiment (trial) identifier	optional
General description of experiment. Can include noteworthy events, findings, etc.	desired



Field, identifier usually consisting of institution + site + 4 digit number	required
Harvest operations year	required
Treatment name	required
Unique identifier linking from other group	required
Plot layout or experimental design (e.g., "RCBD with four blocks")	desired
Number of blocks or replicates	required
Plot area (gross) per rep	desired
Plot row number for single plot or treatment	optional
Plot harvest row number	optional
Row spacing	optional
Spacing between individual plots, if any	optional
Plot harvest length	optional
Plot width (gross) for single plot or treatment	optional
Plot harvest width	optional
Plot length (gross) for single plot or treatment	optional
Plot harvest length	optional
Plot row number for single plot or treatment	optional
Plot harvest row number	optional
Field latitude	required
Field longitude, E positive, W negative	required
Elevation of field site	optional
Indicates whether or not the field was irrigated (Y or N)	required
Indicates whether or not fertilizer was applied (Y or N)	required
Weather station identifier to link to site information	required
Weather station name	desired
Unique soil identifier	required
Unique identifier linking from other group	required
Previous crop code	optional
Initial conditions measurement date	optional
Initial water content of entire profile, expressed as percent of available water	desired
Initial profile inorganic nitrogen amount	desired
Unique identifier linking from other group	required
Crop (or weed) species identifier	required
Cultivar, line or genotype identifier	required
Genotype level, as listed in Treatments table under GE	desired
Cultivar notes	optional



Planting or sowing date	required
Planting depth	optional
Plant population at planting	required
Unique identifier linking from other group	required
Irrigation date	required
Irrigation amount, depth of water	required
Unique identifier linking from other group	required
Fertilization date	required
Fertilizer material	required
Nitrogen in applied fertilizer	required
Unique soil identifier	required
Soil texture	desired
Maximum depth of root growth	desired
Unique soil identifier	required
Soil layer base depth	desired
Soil water, saturated	required (if not available will be estimated from %silt, %clay, and %sand)
Soil water, drained upper limit	required (if not available will be estimated from %silt, %clay, and %sand)
Soil water, lower limit	required (if not available will be estimated from %silt, %clay, and %sand)
Soil bulk density when moist for layer	desired (required if SLNI not available)
Total soil organic carbon by layer	desired (required if SLNI not available)
Soil texture, clay (<0.002 mm), weight percent of fine earth	required (if not available will be estimated based on SLTX)
Soil texture, silt (0.05 to 0.002 mm), weight percent of fine earth	desired (required if SLDUL and SLLL not available)
Soil texture, sand (0.05 to 2.0 mm), weight percent of fine earth	desired (required if SLDUL and SLLL not available)
Soil texture, coarse fraction (>2 mm), weight percent of fine earth	desired (required if SLDUL and SLLL not available)
Nitrogen, total soil organic	required (if not available estimated from SLOC, SLBDM, SLOM)
Cation exchange capacity	optional
Soil CaCO ₃ content	optional
pH of soil in water, from profile description	optional
Total soil organic matter by layer	desired
Weather station identifier to link to site information	required
Weather station name	desired
Latitude of station, degrees decimal with North as positive value (4 digits precision)	desired
Longitude of station, degrees decimals with East as positive value (4 digits precision)	desired



Elevation of weather station	optional
Reference height for windspeed measurement	desired
Weather station identifier to link to site information	required
Date for daily weather, field observations, etc.	required
Solar radiation	required
Temperature of air, maximum	required
Temperature of air, minimum	required
Rainfall, including moisture in snow, in one day	required
Wind speed, daily mean or "wind run"	desired
Vapor pressure, average daily	desired
Relative humidity, average daily value	desired
Temperature, dewpoint	desired
Evapotranspiration, potential	desired
Date of sampling	desired
Soil layer depth to upper boundary of sample (top)	desired
Soil layer depth to lower boundary of sample (bottom)	desired
Soil water measured at a specified depth	desired
Soil nitrogen in a soil layer	desired
NO ₃ [N], for a given soil layer	desired
NH ₄ [N], for a given soil layer	desired



Annex 3 – List of VCU variables requested by WP4 for wheat

Description	Requirement level
Harvested yield at harvest (fresh wt)	required
Standard deviation of harvested yield at harvest (fresh wt)	required
Moisture content of fresh yield (e.g., grain, fruit, leaves) at harvest	required
Standard deviation of harvest moisture fraction	required
Total above ground dry mass at maturity	desired
Ear height, max value	optional
Ear height, std dev of max value	optional
Plant height (e.g., including spikes above canopy per se) at maturity	desired
Grain protein concentration at maturity	desired
Ear number	desired
Harvest number per unit at maturity (e.g., seed per spike or pod)	optional
Percent or score lodging	desired
Plant population at emergence	desired
Average grain dry mass at maturity	desired
Growth stage date, emergence	desired
Growth stage date, ear emergence (heading)	required
Growth stage of physiol. maturity, as date	desired
Growth stage date, harvest	desired
Vernalization factor, alternativity (e.g. 0 to 1 scale)	desired
Quality class	desired
Frost tolerance	desired
Mildew	optional
Yellow rust	optional
Brown rust	optional
Septoria tritici	optional
Fusarium ear blight	optional



Annex 4 – List of VCU variables requested by WP4 for maize

Description	Requirement level
Harvested yield at harvest (fresh wt)	required
Standard deviation of harvested yield at harvest (fresh wt)	required
Harvested yield at harvest (dry wt)	required
Moisture content of fresh yield (e.g., grain, fruit, leaves) at harvest	required
Standard deviation of harvest moisture fraction	required
Total above ground dry mass at maturity	desired
Ear height, max value	desired
Ear height, std dev of max value	desired
Plant height (e.g., including spikes above canopy per se) at maturity	desired
Ear number	desired
Harvest number per unit at maturity (e.g., seed per spike or pod)	optional
Percent or score lodging	optional
Plant population at emergence	desired
Average grain dry mass at maturity	desired
Growth stage date, emergence	desired
Growth stage date, anthesis	required
Growth stage of physiol. maturity, as date	desired
Growth stage date, harvest	desired
Maturity group	desired
Maturity type	
Plant early vigour	desired
Plant vigour at leaf stage 6 to 10	optional
Frost tolerance	optional
Corn smut, Ustilago Maydis	optional
Fusarium head blight	optional
Helminthosporium	optional
Rust infection	optional
Fusarium stalk rot	optional
Kabatiella infection	optional
Frit fly	optional



Annex 5 – List of VCU variables requested by WP4 for sunflower

Description	Comment2
Harvested yield at harvest (fresh wt)	required
Standard deviation of harvested yield at harvest (fresh wt)	required
Moisture content of fresh yield (e.g., grain, fruit, leaves) at harvest	required
Standard deviation of harvest moisture fraction	required
Total above ground dry mass at maturity	desired
Ear height, max value	optional
Ear height, std dev of max value	optional
Plant height (e.g., including spikes above canopy per se) at maturity	desired
Grain oil concentration at maturity	required
Ear number	desired
Harvest number per unit at maturity (e.g., seed per spike or pod)	desired
Percent or score lodging	desired
Plant population at emergence	desired
Average grain dry mass at maturity	desired
Growth stage date, emergence	desired
Growth stage date, anthesis	required
Growth stage of physiol. maturity, as date	desired
Growth stage date, harvest	required
Precocity at maturity	
Frost tolerance	desired
Botrytis	optional
Phomopsis	
Sclerotinia	optional



Annex 6 – Proposed procedure for access to data and material (Dec 17th 2019)

INVITE



INnovations in plant Varlety Testing in Europe to foster the introduction of new varieties better adapted to varying biotic and abiotic conditions and to more sustainable crop management practices

Access to historical data and
reference materials



INVITE – Proposed procedure for access to data and material, 17 dec 2019

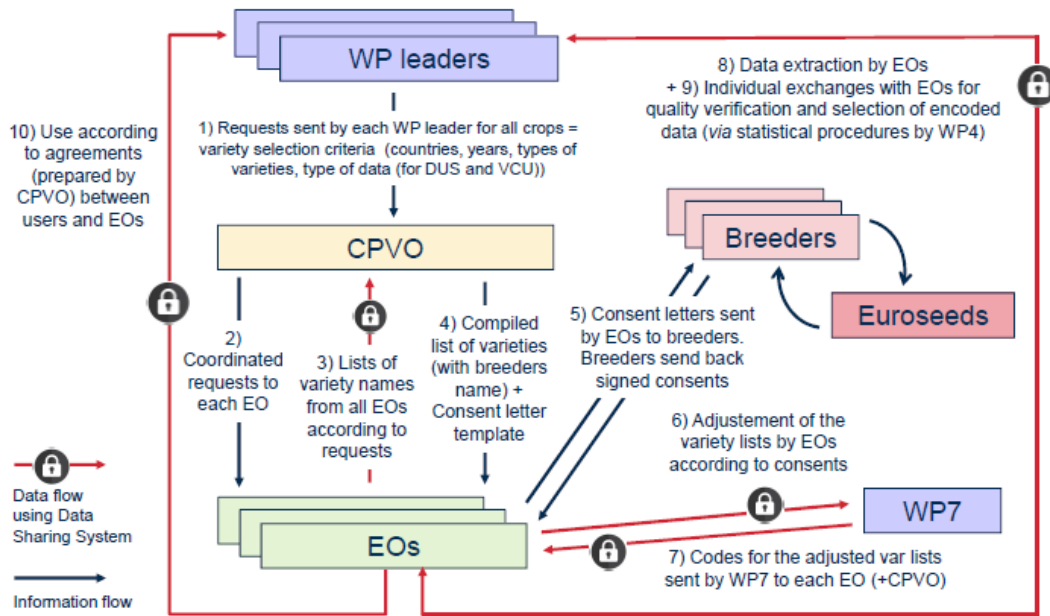


Requests to EOs for access to reference material and historical data by third parties in the framework of INVITE

- DUS historical data : **WP2, WP3, WP6**
 - VCU historical data : **WP4, WP6**
 - Reference materials : **WP1 ?, WP2, WP3**
- **WP leaders have to define selection criteria for each crop (types of varieties, traits, years, countries...) with the relevant experts/scientists and send these criteria to CPVO before 15 Jan 2020.**



Proposed procedure for access to historical data

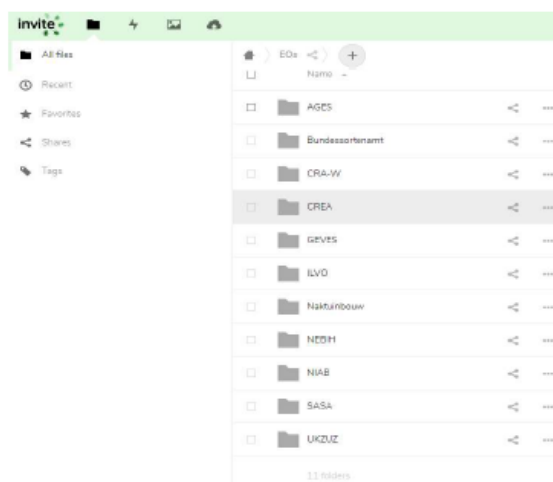


INVITE – Proposed procedure for access to data and material, 17 dec 2019



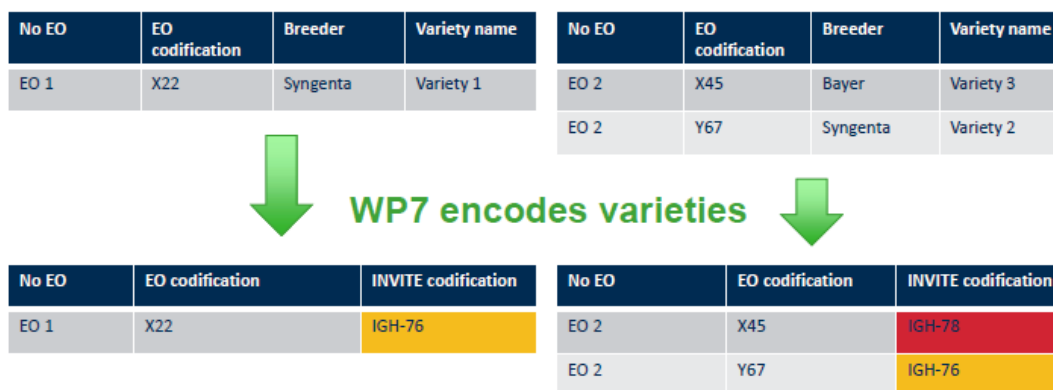
Data sharing system

- No anonymous access → All access are secured (login/password)
- Permission management: the provider chooses who can access to its files
- Each EO have its own folder for their varieties datasets (no mixing between EO data)



Encoding the varieties

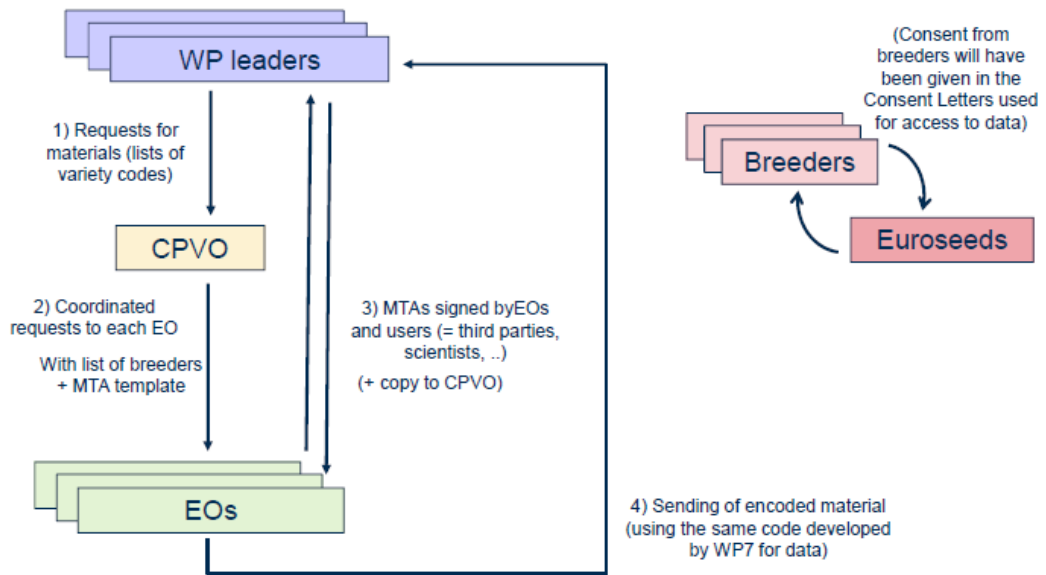
EO codification: unique identifier from each EO
 INVITE codification: unique identifier from INVITE - WP7



INVITE – Proposed procedure for access to data and material, 17 dec 2019



Proposed procedure for access to reference materials



INVITE – Proposed procedure for access to data and material, 17 dec 2019

