AGINFRA

D7.3 - FOOD SECURITY COMMUNITY-CENTRED ASSESSMENT PLAN

Co-funded by the Horizon 2020
Framework Programme of the European Union

DELIVERABLE NUMBER	D7.3
DELIVERABLE TITLE	Food Security Community-Centred Assessment Plan
RESPONSIBLE AUTHOR	Esther Dzalé Yeumo (INRA), Pascal Neveu (INRA)

GRANT AGREEMENT N.	731001
PROJECT ACRONYM	AGINFRA PLUS
PROJECT FULL NAME	Accelerating user-driven e-infrastructure innovation in Food & Agriculture
STARTING DATE (DUR.)	01/01/2017 (36 months)
ENDING DATE	31/12/2019
PROJECT WEBSITE	plus.aginfra.eu
COORDINATOR	Nikos Manouselis
ADDRESS	110 Pentelis Str., Marousi GR15126, Greece
REPLY TO	nikosm@agroknow.com
PHONE	+30 210 6897 905
EU PROJECT OFFICER	Mrs. Georgia Tzenou

WORKPACKAGE N. | TITLE	WP7 | Food Security Community
WORKPACKAGE LEADER	Institut National de la Recherche Agronomique (INRA)
DELIVERABLE N. | TITLE	D7.3 | Food Security Community-centred Assessment Plan
RESPONSIBLE AUTHOR	Pascal Neveu (INRA)
REPLY TO	Pascal.Neveu@inra.fr
DOCUMENT URL	http://www.plus.aginfra.eu/sites/default/files/deliverables/D7.3.pdf
DATE OF DELIVERY (CONTRACTUAL)	30 June 2017 (M6), 30 June 2018 (M18, 1st Updated version), 19 October 2018 (M22, 2nd Updated version)
DATE OF DELIVERY (SUBMITTED)	26 October 2017 (M10), 2 July 2018 (M18, 1st Updated version), 19 October 2018 (M22, 2nd Updated version)
VERSION | STATUS	V3.0 | Final
NATURE	R (Report)
DISSEMINATION LEVEL	PU (Public)
AUTHORS (PARTNER)	Pascal Neveu (INRA), Esther Dzalé Yeumo (INRA), Alice Boizet (INRA)
CONTRIBUTORS	Pythagoras Karampiperis (Agroknow), Panagiotis Zervas (Agroknow)
REVIEWERS	Rob Lokers (Wageningen Environmental Research), Rob Knapen (Wageningen Environmental Research)

VERSION	MODIFICATION(S)	DATE	AUTHOR(S)
0.1	Initial Version	25/09/2017	Esther Dzalé Yeumo (INRA)
0.2	Indicator, organization, communities' identification	06/10/2017	Pascal Neveu (INRA), Esther Dzalé Yeumo (INRA)
0.7	Specification and draft elaboration	16/10/2017	Esther Dzalé Yeumo (INRA)
0.8	Draft for review	20/10/2017	Pascal Neveu (INRA), Pythagoras Karampiperis (Agroknow)
0.9	Review	24/10/2017	Rob Lokers (Wageningen Environmental Research), Rob Knapen (Wageningen Environmental Research)
1.0	Final version	26/10/2017	Panagiotis Zervas (Agroknow)
2.0	1st Updated version	02/07/2018	Alice Boizet (INRA)
3.0	2nd Updated version based on EC reviewers' comments	19/10/2018	Alice Boizet (INRA)

PARTICIPANTS		CONTACT
Agro-Know IKE (Agroknow, Greece)		Nikos Manouselis Email: nikosm@agroknow.com
Stichting Wageningen Research (DLO, The Netherlands)		Rob Lokers Email: rob.lokers@wur.nl
Institut National de la Recherche Agronomique (INRA, France)		Pascal Neveu Email: pascal.neveu@inra.fr
Bundesinstitut für Risikobewertung (BfR, Germany)		Matthias Filter Email: matthias.filter@bfr.bund.de
Consiglio Nazionale Delle Richerche (CNR, Italy)		Leonardo Candela Email: leonardo.candela@isti.cnr.it
University of Athens (UoA, Greece)		George Kakaletris Email: gkakas@di.uoa.gr
Stichting EGI (EGI.eu, The Netherlands)		Tiziana Ferrari Email: tiziana.ferrari@egi.eu
Pensoft Publishers Ltd (PENSOFT, Bulgaria)		Lyubomir Penev Email: penev@pensoft.net

ACRONYMS LIST

API	Application Programming Interface
ESFRI	European Strategy Forum for Research Infrastructure
FAIR	Findable, Accessible, Interoperable, Reusable, as set of principles acting as an international guideline for high quality data stewardship
FAO	Food and Agriculture Organisation of the United Nations
NGOs	Non-Governmental Organisations
RDF	Resource Description Framework
SQL	Structured Query Language
VRE	Virtual Research Environment

EXECUTIVE SUMMARY

This Community-centered Assessment Plan describes a detailed plan regarding the procedures to be carried out for assessing the effectiveness of the AGINFRA PLUS paradigm for research in Food Security communities. It defines the objectives of the pilot trials and their assessment. It also defines the different actions defined in the piloting scheme including the organization of restricted or wider demonstrations and hands-on events, activation of networks, expert sessions, etc. Furthermore, the task will carry out the pilot execution and evaluation activities relative to Food Security foreseen in Task 7.1 and following the established plan and evaluation methodology.

TABLE OF CONTENTS

LIST OF TABLES

1 INTRODUCTION

The "Food Security" use cases to be implemented in the AGINFRA PLUS VRE should leverage the Big Data opportunities in order to sustainable maximize crop performance. That requires to be able to determine, for instance, which plant species and which varieties are most adapted to global changes and natural resource preservation. The Food Security community requires high throughput phenotyping, which is at the heart of plant selection process and produces huge sets of complex data. This plan describes a user-centered assessment plan for the AGINFRA PLUS VRE for the Food Security community. It describes the levels of assessment that will be performed to evaluate a heterogeneous set of quantitative and qualitative indicators.

The objectives of the pilots and the associated pilot assessments deal with the different steps of data processing and analysis, including:

- Discover and access resources
- Exchange plant datasets at different scales (gene, cell, plant, canopy, crop), at different steps (phenology, food processing), environmental datasets (soil, water, life cycle and sustainability), nutrition and biomass production datasets etc.
- Combine, integrate and (pre-) process these variable and huge datasets
- Explore, analyze, model, visualize
- Record the data processing chain in order make it reproducibility

These pilots should also assess the usefulness of the Food Security VRE for the Phenotyping community to work collaboratively on standards definition.

The goal of this document is to describe the procedures to be carried out for assessing the effectiveness of the AGINFRA PLUS VRE in dealing with all these steps. During the assessment phase, we will focus on a set of evaluation indicators for biological questions in environmental stresses such as water stress in the case of wheat species. The assessment plan will include various activities and events such as surveys, trainings, and workshops.

2 ASSESSMENT METHODOLOGY

This part defines how the project intends to expose the use case, how it interacts with the Food Security community in the various phases of the project through targeted sessions and events and how the evaluators, involved in each phase, match the objectives of every phase. Moreover, it defines the setup of the trials and the associated evaluation procedures and metrics used to measure the effectiveness of the AGINFRA PLUS paradigm for research in the targeted research community. These cases will be more focused on high-throughput phenotyping.

2.1 INTERACTION WITH COMMUNITIES

To assess the effectiveness of the AGINFRA PLUS paradigm for the Food Security community, an iterative approach of assessment, according to the proposed three-phase piloting scheme, will be performed. Following the concept of gradual extension of functionality, intended audience and assessment of this scheme, the pilot will interact with the community and the pilot evaluators accordingly.

First Pilot (part of intermediate phase, M13-18)

In order to gather and to analyze needs and requirements, events (such as VRE training) will be organized with partners of WP 2, 3 and 4 (core e-infrastructures), different stakeholders, and various representative users. To start with, we will select around 5 complementary representative persons (a training group) of the phenotyping community to form the testing group. Next, we will organize a small workshop with these representatives during which we will introduce the concept of the VRE to the training group. Then, we will collect the expectations, requirements and fears of the training group regarding the future Food Security VRE.

Available components will also be evaluated by a couple of potential users who don't know how to use the VRE. They will have to follow a user scenario and evaluate every available feature.

Second Pilot (part of summative phase, M27-M30)

We will organize a workshop with people from INRA belonging to the phenotyping community. We will teach them how to use the VRE, then they will follow a user scenario and evaluate the VRE by filling in a questionnaire. Finally, we will have an open discussion to gather more precise feedbacks.

Third Pilot (part of summative phase, M30-M33)

For the third pilot, we aim to involve a much larger community than the 2nd pilot. First, we will invite people of the EMPHASIS community to use the VRE as a forum tool to exchange on phenotyping concepts formalization. Then, we will set up a bigger workshop opened to every member of plant community. This workshop could be jointly organized with ALTERRA and so be opened to the agro-climatic data community. This will help us to involve more people in our respective VREs.

2.2 PILOTS TRIALS

First Pilot (part of intermediate phase, M13-18)

The first pilot trials will focus on the assessment of the usability of the Food Security VRE available features, especially on data and analytics scripts sharing. This initial assessment should help us to identify if we are missing necessary features or, conversely, some features seem useless for the purposes of phenotype community. The first pilot will also be a good opportunity to communicate and introduce the VRE concept to this community.

This first pilot will focus on asking 2 or 3 new users to test the VRE by following a very simple user scenario. Every available feature will be scored and any users' feedback will be collected.

Second Pilot (part of summative phase, M27-M30)

The second pilot will focus on gathering feedback from users following a real user story, from access to data to the exchange of analysis results. They will be asked to fill in a questionnaire to evaluate the VRE. The table below describes the second pilot.

User story	Meta-analysis of phenotyping data
Target users (approximate number)	Geneticists, agronomists, statisticians from phenotyping community (10 users)
Story description	- Access to phenotyping experiments data with the same genotype observed in different contexts. - Treat data (curve fitting to estimate missing points) - Visualize data - Analyse data - Design a workflow (for visualization or analytics) - Report data analysis results - Share the results to the community and outside the community
Tested features	- Data miner algorithms to access the phenotype platforms data and to do data treatment - Visualization tool - Developing environments (Rstudio, jupyter) - Galaxy workflow management - AGINFRA+ catalogue

Third Pilot (part of summative phase, M30-M33)

The third pilot will first focus on semantic features (ontology management and ontology mapping). This use case will also help us to realise if the VRE can be the key instrument for collaborative work between different kinds of user. Two user stories will be tested. The table below describe them.

User story	Concepts formalization & semantics	IT development for phenotyping
Target users (approximate number)	Phenotype managers, agronomists from the Emphasis community (>10 users)	Computer scientists and developers working on the Breeding API (>5 users)

Story description	Defining and formalizing: - Metadata - Phenotype concepts - representation models	Working collaboratively on the specification of new calls for the Breeding API
Useful features	Collaborative features: message posting, workspace Scalability of semantics features	Collaborative features : message posting, workspace

2.3 PILOTS EVALUATION

The evaluation of the Food Security pilot VRE will be performed by INRA along a set of assessment indicators and evaluation questions that are targeted to the specifics of that use case. Due to the difference in the scale, setup and number and type of involved evaluators in each of the piloting phases (as described in the previous section), only the indicators fitting the specific phase will be evaluated. Moreover, the way in which indicators and metrics will be measured will differ over the three evaluations. This is inevitable, because of the type and complexity of the components and workflows to be evaluated as well as because of the different profiles of the evaluators that are involved in every single phase.

The assessment of the effectiveness of a VRE for high-throughput phenotyping has a broad range of objectives, which again leads to a broad and heterogeneous set of potential assessment indicators. Evaluations will vary from the assessment of expert functions required for developing and deploying components and setting up workflows, to the user experience of the VRE's "end-users". While for some indicators it is expected to be feasible to perform quantitative evaluations, the character of some of the other indicators, and of the group of evaluators involved, will result in qualitative evaluations.

To support the overall AGINFRA+ evaluation process, we will apply the following harmonized evaluation methods:

- *Quantitative benchmarking:* for a limited set of indicators, quantitative measurements can be performed, for example for the indicator "performance" by measuring "computation time" in a "conventional" modelling environment compared to a VRE environment.
- *Qualitative scoring:* Quite some indicators will not be suitable for quantitative evaluation. In these cases, at least a qualitative evaluation, for example though ordinal scoring, will be developed. In such cases we also add the option to provide an additional explanation in the form of a short expert review. This will allow to get better insight in the motivation of the individual reviewer.
- *Expert review:* Both in the first and second phase of the piloting scheme, where more experienced evaluators will be involved, a set of more complex, technical and usability indicators will be included in the evaluations. For those indicators that cannot be easily benchmarked or where quantitative assessment alone does not provide sufficient insight, an expert review procedure will be developed. The expert review will allow experts to (1) explain the motivation for their quantitative scoring of indicators, (2) assess indicators, using their past experiences with non-VRE environments and (3) allow evaluators to give their opinion on VRE for modelling through an interview with open questions after an evaluation session.

- *End-user survey:* The broader trials in the third phase of the piloting scheme, the associated character and heterogeneity of the group of evaluators, the end-user-oriented setup of the trial and the limited time available in events will require a targeted approach. The indicators that are specifically evaluated by this group will be assessed through an end-user survey that can be completed in a relatively short time, also by less experienced evaluators. The format will be an on-line survey.

The following table summarizes the indicators from which a selection will be assessed for the WP7 pilot VRE. It also indicates the method of assessment, the pilot VRE versions and the specific questions used to assess each indicator.

Table 1: Evaluation Indicators for Food Security pilot VRE

Indicator	Rationale and evaluation questions	Assessment method	VRE pilot
Usefulness of the VRE pilot for the use case	The pilot VRE should support researchers to perform their work "in the cloud" in a collaborative environment that allows co-development by remote researcher and research teams and that offers performance and scalability in performing Questions: *How do you assess the usefulness of the pilot VRE with regard to the following aspects?* - *Collaboration and communication* - *Data access* - *Data analytics* - *Data visualization* *Do you think that a VRE like D4Science can help you to perform your work "in the cloud" and support co-development by remote researcher and research teams?*	Expert review Qualitative scoring End-user survey	1, 2 1,2 3
Learning Curve	How much time is needed to learn the pilot VRE concepts, new pilot functionalities and to master underlying technology and functionality, before the pilot VRE can be used? Questions: *How do you estimate the effort to make an existing analytical procedure or model run on the pilot VRE?* *How do you estimate the effort to configure and run a model and analyze its outputs?*	Expert review Qualitative scoring End-user survey	1, 2 1,2 3
Performance, Scalability	A VRE and VRE applications should support collaborative research in such a way that performance is (at least) comparable with "traditional environments". They should also offer scalability, meaning that the resources	Quantitative benchmarking Qualitative scoring	2, 3 2,3

	available (for computing, storage etc.) should be able to adapt to the amount and size of the research jobs, the amount of concurrent users etc.

Questions:

How do you rate the performance of the pilot VRE compared to the platform(s) you are used to work with?

How do you rate the scalability of the pilot VRE? | End-user survey | 3 |
| **Openness and FAIRness** | Openness is an important asset of any environment related to the European Open Science Cloud. This means that is should be easy to add new data and analytics to the pilot VRE and to publish and share resulting workflows, components, and data with other researchers. This relates to how the pilot VREs upports in making research data and algorithms FAIR (Findable, Accessible, Interoperable, Reusable).

Questions:

How open do you think the pilot and its tools are from the following perspectives?
- *Ease of finding and accessing and reusing datasets*
- *Ease of finding and assessing and reusing analytics and models*
- *Ease of registering and sharing datasets*
- *Ease of registering and sharing analytics and models*

Are the data and algorithms available through the pilot VRE well documented?

Do you think that the pilot VRE offers the functions to document data and algorithms in such a way that fellow researchers would be able to reuse them? | Qualitative scoring

Expert review | 1, 2, 3

1,2,3 |
| **Uptake potential** | This indicator will give us feedback on how likely it is that the person assessing the pilot VRE would | Qualitative scoring | 2, 3 |

also be willing to use it after the end of the project

Questions:
What is the likelihood that you would use a system as the one demonstrated in the pilot VRE in the future for your work?

In general, the evaluation questions will be defined in such a way that the evaluator will at least provide a qualitative or quantitative rating. Besides, in relevant cases, additional comments will be requested to explain the scoring of the evaluation questions and the indicator in general.

A typical format would be as in the following example evaluation question:

Indicator - <Indicator name>

< short explanation / rationale of the indicator, see also table above >

<Evaluation question 1>

1 - insufficient	2	3 - moderate	4	5 – very good	Not applicable

Further explanation and comments on question 1:

…………………………………………………………………..

<Evaluation question 2>

1 - insufficient	2	3 - moderate	4	5 – very good	Not applicable

Further explanation and comments on question 2:

…………………………………………………………………...

Further explanation and comments regarding <Indicator> in general:

3 USE CASE ASSESSMENT

The use case "Food Security" is defined in deliverable D7.1 (User-driven Requirements & Use Cases – Food Security) as a use case on Food Security with a focus on high throughput plant phenotyping. The main challenge is to design and implement scientific workflows on massive amounts of complex data produced at different scales (population, plant, organ, cells, etc.), different stages (sowing, phenology, harvest) in interaction with various environment components (soil, climate, agricultural practices, biodiversity, etc.). In order to achieve this aim, we must supply analysis combining various heterogeneous data sources and run scientific workflows on a set of datasets, potentially huge and potentially from different data sources. The given workflow may include some steps of data combination and wrangling.

3.1 TARGET USERS

Typical users of this use case will be (see deliverable 7.1):
- Researchers
- Breeders
- Stakeholders (states, farmers' representatives, governmental experts, private companies)
- Engineers
- "Data professionals" (Data scientists, Data analysts, Data managers)

3.2 ASSESSMENT OBJECTIVES

The principal objective of this use case is to characterize, in a background of global changes, an environment and see in return which species and varieties are likely to adapt to this specific environment. To that aim, the user should be able to use the Food Security VRE to import, build, and update workflows to:
- Search for, and access, environmental and experimental data which might be large. The main data sources are the different instances of the EMPHASIS information system (currently under development) which will be accessible through web services.
- Choose the variables that characterize a given environment.
- Identify the varieties or species that can be adapted to the characterized environment.
- Share the results with decision makers through a user-friendly interface so they can decide which species or varieties they want to put in a given environment.

The VRE should allow access control to the data sources and the shared objects (workflows, models and results of workflows). The global assessment will include but not be limited to develop and improve the following:
- Data access: access and integrate data from various data sources of phenomics platforms with a focus on semantic issues.
- Data exploration and visualization: to provide interactive visualization.
- User interface: the VRE must provide coherent "look and feel" and support for a visual component for the design of workflows.
- Machine learning: access to machine-learning approaches and also include support for modern machine-learning approaches like ensemble techniques (boosting, bagging and random forests) and deep learning.
- Flexibility, extensibility and openness: Integration of open-source libraries into the VRE. Work with notebooks should be allowed.
- Delivery: Ability to create APIs or containers (such as code, Predictive Model Markup Language [PMML] and packaged apps) that can be reused.

- Resource management: management of data and tools. Provide reuse and version management of resources, auditing lineage and reproducibility).
- Collaboration: Make users with different skills work together.
- Coherence: provide a seamless end-to-end experience, to make the users more productive across the whole data and analytics pipeline.

3.3 ASSESSMENT INDICATORS

According to the other use cases, the development of the evaluation procedures for the use case will, aligned with the iterative approach of the pilot trials, be developed in a phased way. The relevant indicators from the set as described in the previous section, and the methods of assessment in the different phases, will then be further specified. The level of complexity of workflows can be completed on a given set of datasets. The size of the datasets that the VRE enables to deal with is another quantitative indicator.

4 TENTATIVE PILOT TRIALS SCHEDULE

The execution of pilot trials will start after the completion of the development of required components through WP2, WP3 and WP4.
The overall schedule for the Food Security pilot is the following:

Table 2: The overall schedule for the Food Security pilot

Phase	Time Frame	Activity
Formative	M1 – M12	No pilot assessment Development of first pilot VRE version
Intermediate	M12-M18	First workshop of the testing group Objective: Introduce the concept of VRE 1st assessment pilot
Summative	M25 - M27	Testing new available components in the 2nd pilot VRE version
Summative	M27 - M30	2nd pilot assessment
Summative	M30 - M33	3rd pilot assessment
Summative	M33 - M36	Potential VRE adaptation according to users' feedbacks