

Deliverable N.: D4.14

Title: Data management plan

(ORDP: Open Research Data Pilot)

Funding scheme: Coordination and support action

Project Acronym: LEAP4NSSA

Project Full Title: Support to the implementation of the Long-term EU-AU Research and Innovation Partnership for Food and Nutrition Security and Sustainable Agriculture

Grant Agreement n°: 817663

Project duration: 48 months

Published by the LEAP4FNSSA Consortium

Dissemination Level: Public



This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No. 817663

DOCUMENT INFORMATION

Project number	817663	Acronym	LEAP4FNSSA
Full title	Support to the implementation of the Long-term EU-AU Research and Innovation Partnership for Food and Nutrition Security and Sustainable Agriculture		
Project URL	www.leap4fnssa.eu	Document URL	
EU Project Officer	Soriani Cristina		

Deliverable	Number	D4.14	Title	Data Management Plan
Work Package	Number	WP4 WP5	Title	Dissemination and Communication Coordination

Date of delivery	Contractual	30/04/2019	Actual	03/04/2019
Status	Version 1		Draft	
Nature	<input type="checkbox"/> prototype <input checked="" type="checkbox"/> report <input type="checkbox"/> demonstration <input type="checkbox"/> other			
Dissemination level	<input checked="" type="checkbox"/> public <input type="checkbox"/> restricted			

Authors (Partner)	CIRAD, with contributions from WP4 and all SC partners			
Authors and contributors	Name	Philippe Petithughuenin Laurence Dedieu	Email	Laurence.dedieu@cirad.fr
	Partner	Cirad CIHEAM	Phone	+33 (0) 467 614 965

Abstract (for dissemination)	This document represents the first version of the LEAP4FNSSA Data Management Plan that is due by the end of April 2019. The V1 describes the life cycle of the data to be collected, generated and processed in LEAP4FNSSA. It reflects the current state of the discussions within the consortium about Data management strategy.
Keywords	Open data, data dissemination, data management, personal data protection

LEAP4FNSSA CONSORTIUM

The project consortium is composed of:

AUSTRIA	Universitaet fuer Bodenkultur Wien (BOKU)
BURKINA FASO	Fonds National de la Recherche et de l'Innovation pour le Developpement (FONRID)
CZECH REPUBLIC	Ceska Zemedelska Univerzita V Praze (CULS)
DENMARK	Kobenhavns Universitet (UCPH)
EGYPT	Ministry of Higher Education and Scientific Research (MHESR) Knowledge Economy Foundation (KEF)
ETHIOPIA	African Union – Union Africaine - The African Observatory for Science Technology and Innovation (AOSTI)
FINLAND	Luonnonvarakeskus (LUKE) Jyvaskylan Yliopisto (JYU)
FRANCE	Agence Nationale de la Recherche (ANR) Centre de Cooperation Internationale en Recherche Agronomique pour le Developpement (CIRAD) – COORDINATOR
GERMANY	Bundesanstalt für Landwirtschaft und Ernährung (BLE) Deutsches Zentrum für Luft und Raumfahrt EV (DLR) Go Africa Katharina Kuss (GAKK) Universitaet Hohenheim (UHOH)
GHANA	Council for Scientific and Industrial Research- Science and Technology Policy Research Institute (CSIR-GH) Forum for Agricultural Research in Africa (FARA)
GREECE	Knowledge & Innovation Consultants - Symvouleftiki Monoprosopi Epe (KINNO)
HUNGARY	Szent Istvan University (SIU)
ITALY	Centro Internazionale di Alti Studi Agronomici Mediterranei (CIHEAM-IAMB)
KENYA	Ministry of Education Science and Technology (MEST) Network of African Science Academies (NASAC)
NIGERIA	International Institute of Tropical Agriculture (IITA)
PORTUGAL	Fundacao para a Ciencia e a Tecnologia (FCT)
SENEGAL	Ministere de l'Enseignement Superieur, de la Recherche et de l'Innovation (MESRI-DFRSDT)
SOUTH AFRICA	Agricultural Research Council (ARC) Department of Science and Technology (DST) National Research Foundation (NRF)
SPAIN	Instituto Nacional de Investigacion y Tecnologia Agraria y Alimentacion OA MP (INIA)
SWEDEN	Sveriges Lantbruksuniversitet (SLU)
THE NETHERLANDS	Stichting Wageningen Research (WR)
UGANDA	Regional Universities Forum for Capacity Building in Agriculture Limited (Forum) by Guarantee (RUFORUM) The Registered Trustees of the Association for Strengthening Agricultural Research in Eastern and Central Africa (ASARECA)
UNITED KINGDOM	Association of Commonwealth Universities (ACU) Natural Resources Institute-University of Greenwich (NRI)

Disclaimer

This document reflects only the author's views and the European Union is not responsible for any use that may be made of the information contained therein.

INDEX

DMP version History.....	5
1. Data description	5
1.1. Purpose of the data collection/generation and relation to the objectives of the project	5
1.2.Dataset description	5
1.3.Methodology to collect/generate the data	6
1.4.Data documentation, standards and metadata, format	7
1.5.Specify if existing data is being re-used	8
1.6.Outline the data utility	8
1.7.FAIR Data.....	8
2. Data storage and security.....	8
2.1.Personal data.....	9
3. Ethical aspects	9
4. Data sharing	10
4.1During the project.....	10
4.2.At the end of LEAP4FNSSA.....	10

DMP version History

This document represents the first version of the LEAP4FNSSA DMP that is a deliverable due by the end of April 2019. The DMP V1 describes the life cycle of the data to be collected, generated and processed in LEAP4FNSSA. It reflects the current state of the discussions within the consortium about Data management strategy.

A simplified template has been used for this first version, but the main elements of a Data management policy has been kept to structure the DMP. Further and more elaborated versions will follow the structure based on the FAIR principles as recommended by the [Guidelines on FAIR Data Management in H2020](#) (Version 3.0, July 2016).

All DMP versions will be available to the participants of the project through the project's sharing platform, along with all project documents.

1. Data description

1.1. Purpose of the data collection/generation and relation to the objectives of the project

The objective of LEAP4FNSSA is to establish a sustainable structure, or 'Platform', for the efficient and coherent implementation of the AU-EU Research and Innovation Partnership as described in the FNSSA Roadmap. Under the aegis of the HLPD and its Bureau, and by building upon related projects such as RINEA, CAAST-Net Plus, ProIntensAfrica and LEAP-Agri, LEAP4FNSSA aims to achieve this overall objective through three specific objectives:

- Increase the synergies and coherence between actors, research and innovation projects, initiatives and programmes, through the development of institutional alliances and clusters of projects;
- Develop a learning environment and a large knowledge base, including monitoring and evaluation activities, creating communication and links between different initiatives to improve STI cooperation;
- Establish a long-term and sustainable governance and funding mechanism for the Platform.

WP3 will provide the core information system for the LEAP4FNSSA Platform. The system will be composed of a database on major projects, particularly EU- and AU-funded ones. The database will be used to produce different kinds of relevant knowledge (domains, good practices, achievements, impacts, etc.) through clustering.

The second main component of the system will be a tool designed to explore data beyond the existing database of EU and AU supported projects. This tool will "mine" the information available online on projects, programmes and policies, with advanced instruments of data and text mining and with analyses of big data.

By the end of the project, the components developed in WP3 and the other WPs will form an information system meant to gather, analyse and report on information, institutions, and actors, providing for the Platform a fully operational efficient system of information sharing. As such, it will form an integral part of a Knowledge and Communication System.

1.2. Dataset description

The data collected and generated by the different LEAP4FNSSA partners will have multiple types and formats, not all clearly yet defined, and will vary in size from a few MB's to several GB's. Due to the fact that data collection and creation is an ongoing process over the lifetime of the project, the detailed description of data nature, exact scale, to whom those data may be useful or if these data underpin a scientific publication will be answered in an updated version of the DMP.

Three main types of data will be collected/generated by LEAP4FNSSA:

- Documents: R&D projects, research programs and initiatives, publications, technical papers, strategic documents, international/national/regional FNSSA-relevant research and innovation policies, funding mechanisms, etc.

Most of these documents will be collected from a corpus of relevant public websites and bibliographic databases where they will be freely accessible. Other documents will come from more restricted data sources but will be available specifically for the LEAP4FNSSA project and its future platform. Indeed, all these documents will be compiled, described and organized to create a consistent and well-structured FNSSA knowledge database.

- Datasets generated by text and data mining (TDM) from this FNSSA database to extract structured information and data resulting from specific requests. Examples of datasets: mapping of projects, clusters of projects, identification of research gaps, alignment of policies, etc.
- Qualitative (and quantitative) data collected through stakeholder consultations, workshops, surveys, interviews, including exchanges such as, formal and informal dialogues between actors operating in research, extension and education and between programmes' projects.

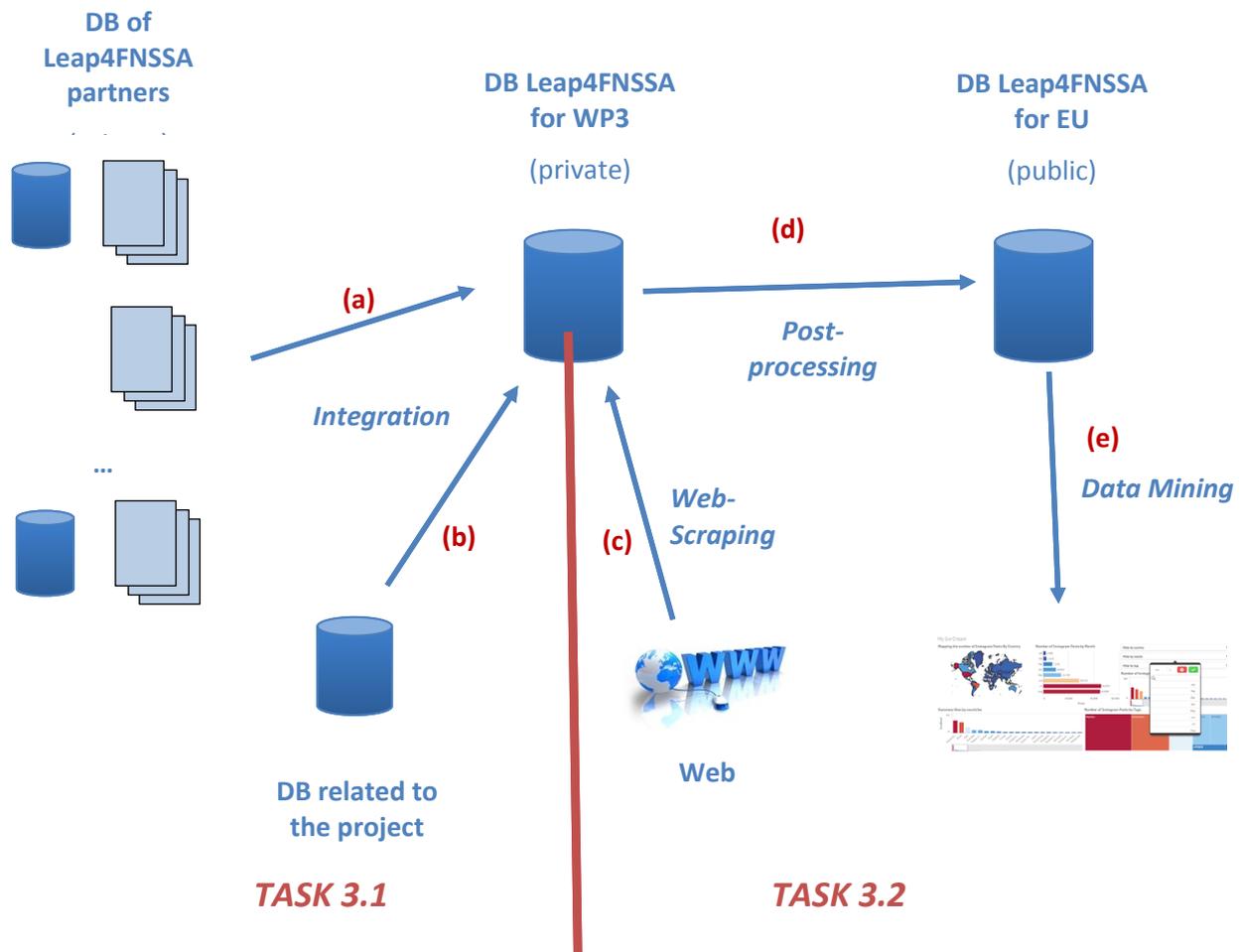
As a principle, LEAP4FNSSA will limit as much as possible the collection of personal data and will rapidly anonymise the data after collection. In addition, all LEAP4FNSSA's partners have approved the ethical procedures detailed in the ethical deliverables as described below and will implement them during the project.

1.3. Methodology to collect/generate the data

A database with information compiled from existing projects (goals, beneficiaries, geographic coverage, development approaches, upscaling theories, Partnership management, etc...) will be created. This database will use information sources from different partners and other organisations such as: the EU and AU, FARA, the Association for Strengthening Agricultural Research in Eastern and Central Africa (ASARECA), the West and Central African Council for Agricultural Research and Development (CORAF/WECARD), CCARDESA, EU and AU member states, and international organisations such as FAO, and CGIAR centers.

The processing of Leap4FNSSA data according the subtasks of WP3 is detailed in the following schema.

First the subtask 3.1 consists of identifying and gathering documents and/or databases of the partners that are relevant for the project. These data will be integrated (**a**) in a database (i.e. "DB Leap4FNSSA for WP3" in the schema) according to standard attributes. Moreover, other structured data relevant for Leap4FNSSA will be integrated in this database (**b**); this integration process will be able to continue at the end of the project. In the context of subtask 3.2, a web scraping approach (**c**) based on 3 steps (i.e. automatic acquisition of data, classification and information extraction) will extract relevant information on the Web. This NLP (Natural Language Processing) approach will be implemented by Cirad. The relevant and non-confidential attributes of the database will be selected (**d**) in order to produce a public database (i.e. "DB Leap4FNSSA for EU" in the schema). Finally, specific data mining methods (**e**) will be proposed by Cirad in order to highlight new knowledge (for instance, associations between specific information).



1.4. Data documentation, standards and metadata, format

To ensure data understanding and interoperability, and to fulfil the FAIR principles (Findable, Accessible, Interoperable and Reusable), recommended best practices will be implemented such as:

- the use of names for the datasets that are based on a comprehensive pattern for naming defined for the project and shared between the data collectors.
- the use of open formats for data files, such as: .csv, ASCII, tab-delimited, etc.
- the use of standards for data description (standard procedures and metadata, and controlled vocabulary) such as the Dublin Core generic standard and the Data Documentation Initiative (DDI) norm for project documentation and qualitative data.
 - for document description: the Dublin Core standard rely on 15 metadata such as: Title, Contributors, Subject, Description, Date (ISO 8601), Source, Time period covered, geographic coverage, Rights, etc.
 - for R&D projects: some supplementary metadata will be added such as: Funders (International, National, Co-funding), Call identifier, Topic of the call, Beneficiaries, Goals, Project duration, Coordinators, etc.
 - for datasets resulting from TDM: the metadata will specifically describe: Nature of the datasets, Sources of the datasets, Methods of dataset production, Property rights associated with the resource, including intellectual property rights, type of licence, etc.

- for qualitative data: key elements highlighted by DDI 3.1 will be included in field notes: Mode of data collection (survey questionnaire, focus group, interview, workshop, etc.), Type of population (governance bodies, decision-makers, funders, policy-makers, position of the interviewee/respondent in the organisation, etc.), Number of persons, Sampling procedure, Workshop Subject or Interview/Survey Topics, Period of reference (ISO 8601), Key informations, Level of anonymisation, etc.

Each dataset will be accompanied by a “Read me” file or “fieldwork report” (.txt) containing the specific protocol that explain how the data were collected, curated, processed, and reported, the description of variables, the questionnaires (.txt or pdf/A) if relevant, etc.

1.5. Specify if existing data is being re-used

LEAP4FNSSA will produce data but will also use existing data. Indeed, existing data will be extracted from different sources from which the selected data will be freely accessible and with a right for re-use (at least for this project, e.g. for data coming from EU or AU archives with restricted access).

In addition, the FNSSA platform will also collect data coming from other FNSSA-related projects. These data will come from other organisations that will be invited to become part of the FNSSA Platform, to feed raw information about the initiatives and programs they support but also to access consolidated, synthetic and elaborated information on the implementation of the FNSSA RoadMap.

1.6. Outline the data utility

The LEAP4FNSSA platform will strengthen the knowledge base on FNSSA and will increase the efficiency of the AU-EU Research and Innovation Partnership on FNSSA. This knowledge sharing platform will indeed provide a well-organized structure with new instruments for data analysis, to make it easy for partners and relevant stakeholders to find, better understand and reuse the various data and documents.

All the stakeholders of the entire value chain from research to innovation, including all relevant actors from the private sector (including farmers), civil society, government and research organisations, will benefit from this FNSSA platform that will enhance impact at local level, generating locally relevant innovation and exchangeable knowledge and know-how.

1.7. FAIR Data

LEAP4FNSSA participates in Open Research Data Pilot, which requires the policy of FAIR data (findable, accessible, interoperable and re-usable research data). The concept of FAIR data is not applicable to most of LEAP4FNSSA data which are not research data. However, to fulfil the FAIR principles, the data will be accompanied with metadata clarifying the meaning of the data and how the data has been collected, and data set naming will clearly describe the content of the data; the data will be stored in a format readable by commonly used data management tools or office software; and clear data access and re-use policy will accompany the datasets, to facilitate re-use, as long as the data source is acknowledged.

2. Data storage and security

During data collection and processing, data will be stored on password-protected computers and laptops, with access limited to the research team. The data and all associated documents will be backed up and secured by each data collector on a regular basis, with clear labelling of versions and dates. As a best practice to secure the data, the 3-2-1 Back-up rule will be applied, i.e.: to keep at least three (3) copies of the data, with regular (1/week) back-ups on at least two (2) different storage media, with one (1) located offsite.

2.1. Personal data

Concerning personal data, the LEAP4FNSSA members will take good care to collect minimal personal data for the work packages. The procedure for data security has been described in the deliverable D6.1: Description of the technical and organizational measures to safeguard the rights and freedoms of the data subjects.

The project will develop two databases with personal data.

At the Consortium level, a personal database will be developed to assist the partners to link with each other and conduct their activities. This database will gather the personal details of all the individual members of the Consortium (name, surname, title, gender, institution, position in the institution, physical address, email address, fix and mobile telephone, skype address, role in the LEAP4FNSSA project). This base will be developed, updated, maintained and stored by the Coordinator, with appropriate rules to protect individual rights. It will also be accessible to all the Consortium partners and to the EC officers through the LEAP4FNSSA intranet portal. Each individual member of the Consortium partners has agreed to contribute to this database, and at all time, he/she will be able to withdraw, partially or totally, his/her personal details. This database will not be maintained at the end of the project.

At the project level, another personal database will be developed in view of the EU-AU partnership platform to be formed at the end of LEAP4FNSSA. It will gather some personal information about the persons identified throughout the activities (name, surname, title, gender, institution, category of institution, position in the institution, physical address, email address, telephone, skype address, expected role in the platform). The database will mainly be composed with the persons invited by WP2 as representatives of important stakeholders of the FNSSA sector (policy makers, private or public funders, research or farmers organizations *etc.*), which intend to play a role in the EU-AU platform under construction. This database will be developed by the Coordinator and fed online, through the LEAP4FNSSA intranet portal, by the SC members as contributors. It will be accessible to the persons themselves, to every Consortium partners, and to the EC officers. Authorization will be requested through a consent form before entering the details of a person into the database. At all time, this person will be given the possibility to withdraw, partially or totally, his/her personal details. This database will be part of the legacy to the EU-AU platform. A one page instruction leaflet will be attached to the database.

3. Ethical aspects

All ethical aspects concerning LEAP4FNSSA are covered by ethics deliverables uploaded in the EU Single Electronic Data Interchange Area (SEDIA). These deliverables are:

- D6.1: Protection of personal data (Potential data processing and privacy issues related to the project's proposed dissemination activities and communication channels)
- D6.2: Involvement of non EU countries (personal data exchange).
- D6.3: Humans (questionnaires and surveys).

All LEAP4FNSSA's partners have approved these principles and measures and will implement them during the project.

All ethics deliverables are available to the partners through the project-sharing platform, along with the DMP versions.

In addition, the LEAP4FNSSA Consortium is in a process to set an Ethics Advisory board to advise on ethics issues involved in the project. The EAB will meet upon request of the SC but at least once every twelve (12) months during the Project.

4. Data sharing

4.1. During the project

The LEAP4FNSSA project will handle two types of data and information. The first type of data and information will be managed through an internal communication system designed with the open-source Liferay software, and accessible only by the Consortium partners. The second type of information will be shared with the HLPD policy makers, and with the general public, through the website of the project.

In each case, there will be a need to constitute transitory databases with personal details, projects description, policy documents or papers, *etc.* These databases will be stored temporarily on a Cirad server and made accessible by the other partners of the Consortium. However, once cleaned of sensitive data, the databases which could be of interest for a wider community of users may be published and shared for further utilization.

The LEAP4FNSSA consortium will define during the project, which, when and how datasets will be shared and disseminated, depending on their strategic nature, their origin, or the potential beneficiaries. Some datasets might, indeed, be declared confidential by the HLPD bureau or by the European Commission or African Union.

4.2. At the end of LEAP4FNSSA

LEAP4FNSSA will develop an open AU-EU FNSSA partnership platform that will gather all the African and European institutions, stakeholders, networks committed to work in a coordinated way in the implementation of the HLPD FNSSA Roadmap. The Platform will be open to the entire research to innovation chain and will address the fragmentation of initiatives on FNSSA by embedding activities in the HLPD policy process.

The institutions will become members of the LEAP4FNSSA platform by signing a Charter. By doing this, they will commit themselves to contribute to the promotion of scientific cooperation between both continents. In return, they will have access to the various findings, concepts algorithms, information and mechanisms developed in the course of the project. FARA will be key in disseminating the results of the project and convincing other African actors to benefit of the achievements in becoming part of the Platform, and signing its Charter.

The members of the LEAP4FNSSA community will also inherit of the external communication website developed by the project, and decide what information to share with the other policy makers and to public and private decision makers (in particular through international Senior Official Meetings and multistakeholder dialogue).

The FNSSA Platform will be a long-term knowledge management, communication and governance mechanism and could take the form of an International Research Consortium.