



LEAP4FNSSA Project



Deliverable N.: D3.6

Title: Report and terms of reference for the query interface: indicators to be produced, initial concepts and information resources list with report on workshops W3-2-1/1 and W3-2-1/2

Funding scheme: Coordination and support action

Project Acronym: LEAP4FNSSA

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

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.

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	D3.6	Title
			Report and term of reference for the query interface: indicators to be produced, initial
Work Package	Number	WP3	Title
			Information Generation and Knowledge Management
Date of delivery	Contractual	31/03/2020	Actual
			V2 08/09/2020 V1 16/04/2020
Status	Version 2.0		Final
Final review	V2 08/09/2020 V1 15/04/2020	Formatting by WP5	16/04/2020
Type of document	<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"/> confidential		
Authors (Partner)	ARC, CIRAD		
Responsible Authors and contributors	Name	Dr Shadrack Moephuli Dr Petronella Chaminuka	Email
			Moephulis@arc.agric.za Chaminukap@arc.agric.za
	Partner	ARC	Phone
	Name	Dr Thierry Helmer Dr Pierre Martin Dr Mathieu Roche	Email
			thierry.helmer@cirad.fr pierre.martin@cirad.fr mathieu.roche@cirad.fr
	Partner	CIRAD	Phone
Abstract (for dissemination)	<p>In support to the long-term EU-AU Research and Innovation Partnership for Food and Nutrition Security and Sustainable Agriculture, the project LEAP4FNSSA is a Coordination and Support Action (CSA), whose main objective is to provide a tool for European and African institutions to engage in a Sustainable Partnership Platform for research and innovation on Food and Nutrition Security, and Sustainable Agriculture (FNSSA). Work Package 3 of the project will provide the core information system for the partnership platform, which will strengthen the decision-making process of the High Level Policy Dialogue (HLPD) and other stakeholders, through producing information on the progress and achievements of the Roadmap and evidence-based information justifying additional interventions. The information system comprises two main components i.e. a database of existing and past projects implemented jointly under the FNSSA partnership, and a knowledge management system (KMS). A series of actions have to date been implemented in the development of the knowledge management system. These include a WP3 technical workshop to identify resources to input into the KMS, to define key concepts and thematic fields in the roadmap, and to identify information needs of the HLPD and other potential KMS users. Another action has been to analyse through text mining the current project database developed in the project to identify semantic information useful for the KMS. There was stakeholder consultation on the development of the KMS which also took place in Addis Ababa in November 2019. This report shows the preliminary indicators and initial concepts defined, and the identified resources list and results of analysis of semantic information to feed the KMS. The report of the workshop held in Pretoria and results of the analysis of semantic information are also shared.</p>		
Keywords	AP4FNSSA, KMS, resources inventory (max 3 words)		

Index

Executive Summary	6
1. Introduction	7
1.1. Purpose, context and scope of this deliverable.....	7
2. Progress towards developing the KMS	7
2.1 Resources inventory workshop (3.2-1) summary.....	7
2.2 Information on list of resources to feed the KMS	8
2.3 Initial concepts and analysis of semantic information.....	8
3. Workshop 3.2-1 report detail	9
3.1 Workshop objectives.....	9
3.2 Participants	9
3.3 Day 1: 30 September 2019	9
3.4 Day 2: KMS Pipeline approaches and tools	11
3.5 Day 3: 02 October 2019	12
3.6 Resolutions of the workshop	14
4. Semantic Information to Integrate in the LEAP4FNSSA KMS	15
4.1 Objective: Evaluation of the semantic aspects of the KMS pipeline	15
4.2 Material and Method.....	15
4.3 Results.....	15
5. Way forward	16
References.....	16

Abbreviations

AU	African Union
CSA	Coordination and Support Action
DMP	Data Management Plan
EU	European Union
EU-AU	European Union-African Union
FARA	Forum on Agricultural Research in Africa
FNSSA	Food and Nutrition Security and Sustainable Agriculture
HLPD	High Level Policy Dialogue
KMS	Knowledge Management System
LEAP4FNSSA	Long-term EU-AU Research and Innovation Partnership for Food and Nutrition Security and Sustainable Agriculture
W	Workshop
WP	Work Package

Executive Summary

In support to the long-term EU-AU Research and Innovation Partnership for Food and Nutrition Security and Sustainable Agriculture, the project LEAP4FNSSA is a Coordination and Support Action (CSA), whose main objective is to provide a tool for European and African institutions to engage in a Sustainable Partnership Platform for research and innovation on Food and Nutrition Security, and Sustainable Agriculture (FNSSA). Work Package 3 of the project aims to provide the core information system for the partnership platform. The information system will help to strengthen the decision-making process of the High Level Policy Dialogue (HLPD) and other stakeholders, through producing information on the progress and achievements of the Roadmap and evidence-based information justifying additional interventions. The information system comprises two main components i.e. a database of existing and past projects implemented jointly under the FNSSA partnership, and a knowledge management system.

The ARC organised a project workshop as per task 3.2-1, from 30 September-2 October 2019 in Pretoria. WP3 project partners from Europe and Africa, an HLPD member, and two African experts who are knowledgeable in implementation of the roadmap for FNSSA attended the workshop. The workshop had the objective to identify the information needs of the HLPD to drive the roadmap and assess how the information collected and processed by WP3 can meet this need. The workshop was structured to involve theoretical presentations followed by very concrete practical exercises using the information so far collected by project partners.

Accordingly, the following key questions were to be answered from the workshop; (i) What are the thematic fields involved in the HLPD roadmap? (ii) Does the material collected by WP3 cover all these thematic fields? (iii) What additional information is required to meet the HLPD needs? (iv) Which semantic objects (thesaurus, controlled vocabulary, ontology) seem the most appropriate to achieve the KMS objectives? The workshop resulted in the identification and prioritisation of an inventory of key resources to feed the KMS, definition of some initial concepts and clarification of the relationship between the database and the KMS. Key thematic fields of the roadmap, and potential users of the KMS were also identified.

During the 12th FNSSA Working Group Meeting in Addis Ababa, Ethiopia, WP3 partners consulted stakeholders on the KMS and database development. There was interest and enthusiasm on both the KMS and database, and some requests were made with regards to the type of tool that stakeholders require. The feedback from stakeholders is important and incorporated in the ongoing development of the KMS.

This deliverable includes the report of the Pretoria workshop W3.2-1 and preliminary results to identify terminology (thematic fields, controlled vocabulary) to use in the LEAP4FNSSA KMS. Based on data of the projects in the database (e.g. abstracts in the current database), text mining approaches were applied in order to extract keywords automatically. These keywords highlighted 2 types of semantic information useful for the LEAP4FNSSA KMS: (1) semantic information for classification task, (2) semantic information for indexing. In this context, 2 kinds of evaluations were

done by the workshop participants to evaluate the quality of keywords obtained by text mining with the LEAP4FNSSA data.

The terms of reference for the query interface were to be finalised in a workshop in Budapest Hungary in May 2020. However, due to the COVID-19 situation, the workshop was indefinitely postponed, but the process of developing the KMS using resources so far available and semantic objects obtained from the Pretoria workshop will continue. The report of the Budapest workshop (W3.2-2) is thus not part of this report.

1. Introduction

1.1. Purpose, context and scope of this deliverable

Work Package 3 of the LEAP4FNSSA project aims to provide the core information system for the partnership platform. The information system will help to strengthen the decision-making process of the High Level Policy Dialogue (HLPD) and other stakeholders, through producing information on the progress and achievements of the Roadmap and evidence-based information justifying additional interventions. The information system comprises two main components i.e. a database of existing and past projects implemented jointly under the FNSSA partnership, and a knowledge management system.

This deliverable is work in progress towards development of a KMS for the LEAP4FNSSA project. This report shows the preliminary indicators and initial concepts defined, and the identified resources list and results of analysis of semantic information to feed the KMS. The report of the workshop held in Pretoria and a report on preliminary analysis of semantic information are also shared.

2. Progress towards developing the KMS

2.1 Resources inventory workshop (3.2-1) summary

The ARC organised a project workshop 3.2-1, from 30 September-2 October 2019 in Pretoria. WP3 project partners from Europe and Africa, an HLPD member, and two African experts who are knowledgeable in implementation of the roadmap for FNSSA attended the workshop. The workshop had the objective to identify the information needs of the HLPD to drive the roadmap and assess how the information collected and processed by WP3 can meet this need. The workshop was structured to involve theoretical presentations followed by very concrete practical exercises using the information so far collected by project partners. The following key questions were to be answered from the workshop; (i) What are the thematic fields involved in the HLPD roadmap? (ii) Does the material collected by WP3 cover all these thematic fields? (iii) What additional information is required to meet the HLPD needs? (iv) Which semantic objects (thesaurus, controlled vocabulary, ontology) seem the most appropriate to achieve the KMS objectives? The workshop resulted in the identification and prioritisation of an inventory of key resources to feed the KMS, definition of some initial concepts and clarification of the relationship between the database and the KMS. Key thematic

fields of the roadmap, and potential users of the KMS were also identified. The full report of the workshop is given in Section 3 below.

On 6th November 2019, WP3 consulted key stakeholders on the LEAP4FNSSA database and KMS, as part of the HLPD stakeholder event held before the 12th FNSSA Working Group Meeting in Addis Ababa, Ethiopia. The WP3 presentation was well received by the participants, who showed a lot of interest for the tool and provided relevant feedback for further development. Some of the participants recommended that the KMS and database should be able cope with different information requests from all the types of stakeholders (researchers, farmers, policy makers). Additional requests were made for a system that could allow research in certain areas such as food systems and nutrition, plastic pollution and agriculture and intelligent designs of agricultural system to be linked and prioritized in funding. The feedback from stakeholders is important and incorporated in the ongoing development of the KMS.

2.2 Information on list of resources to feed the KMS

The participants discussed possible sources of information to input into the database. A decision was made that the KMS and the database would only include all projects going back to 2007 when the FNSSA partnership started. In the interim, bilateral projects would not be included. African Union grants projects would be included as they were part of the partnership. Documents for all the projects which are included in the database would be input into the database. These documents were then prioritized in order of their usefulness for the KMS. The inventory of resources to feed the database would include the following in order of importance; project documents and reports, workshop reports, fact sheets, policy briefs, case studies, scientific publications, theses and presentations. The ARC and SIU have collected some of the key resources needed to feed into the KMS. The work done by SLU, WR, RUFORUM and all other WP3 contributing partners in development of the database has made a significant contribution towards the development of the KMS.

Participants also discussed the project's data management plan and how WP3 would ensure compliance with the plan. It was agreed that there should be continuous monitoring and cross checking with the project leaders to ensure that the Data Management Plan is complied with.

2.3 Initial concepts and analysis of semantic information

Based on current data of the projects (e.g. abstracts in the database), text mining approaches have been applied in order to extract keywords automatically. These keywords enable to highlight 2 types of semantic information useful for the LEAP4FNSSA KMS: (1) semantic information for classification task, (2) semantic information for indexing. In this context, 2 kinds of evaluations were done by the workshop participants to evaluate the quality of keywords obtained by text mining with the LEAP4FNSSA data. The results of this exercise are given in Section 4: Semantic information to integrate in the LEAP4FNSSA KMS.

3. Workshop 3.2-1 report detail

3.1 Workshop objectives

The workshop had the objective to identify the information needs of the HLPD to drive the roadmap and assess how the information collected and processed by WP3 can meet this need. The workshop was structured to involve theoretical presentations followed by very concrete practical exercises using the information so far collected by project partners. The following key questions were to be answered from the workshop:

- a. What are the thematic fields involved in the HLPD roadmap?
- b. Does the material collected by WP3 cover all these thematic fields?
- c. What additional information is required to meet the HLPD needs?
- d. Which semantic objects (thesaurus, controlled vocabulary, ontology) seem the most appropriate to achieve the KMS objectives?

3.2 Participants

The workshop was attended by officials from different WP3 contributing partners namely; ARC (Shadrack Moephuli, Petronella Chaminuka, Bridget Murovhi, Mampe Masemola, Walter Maphosa), CIRAD (Thierry Helmer, Mathieu Roche, Pierre Martin), SLU (Agneta Lindsten, Tomas Lunden), SIU (Adam Csorba), WR (Peter van Boheemen). The African Union (AU) commission HLPD was represented by Dr Monica Ebele Idinoba, whilst the Forum for Agricultural Research in Africa (FARA) was represented by Dr Aggrey Agumya. Ioannis Dimitriou of SLU joined via videolink. The ARC invited project partners from the Department of Science and Technology (DST) and National Research Foundation, but due to other commitments, they could not join the workshop as expected. The workshop was organized and hosted by ARC and facilitated by CIRAD.

3.3 Day 1: 30 September 2019

3.3.1 Opening session

The delegates introduced themselves, their organisations and countries, and thereafter the Chief Executive Officer and president of the ARC, Dr Shadrack Moephuli welcomed the delegates to South Africa, and specifically to the ARC. A special welcome was extended to the representatives of the AU commission and FARA. He explained that the ARC is a large organisation with a presence in six of the country's nine provinces, and conducts research in all the agricultural commodities and value chains except sugar cane. He stated that the ARC is excited to be part of the LEAP4FNSSA project and values the partnership with different African and European partners from this project and other past FNSSA projects such as ProIntensAfrica.

3.3.2 LEAP4FNSSA Project, Roadmap for EU-Africa partnership and workshop objectives presentations

The first presentation of the session was an overview of the LEAP4FNSSA project and this was given by Dr Moephuli. The second presentation on the “2016 Roadmap for EU-Africa partnership on FNSSA” was given by Dr Monica Ebele Idinoba. In her presentation she also explained the composition and role of the HLPD Bureau and the HLPD Working Group in implementing the roadmap. The third presentation was on workshop objectives and the role of semantics in our project. Thierry Helmer explained that the workshop was aimed at exploring ideas for designing an effective knowledge management tool that can be used to gather and disseminate information regarding EU-AU partnerships that are aligned to Food and Nutrition Security and Sustainable Agriculture (FNSSA). To achieve this objective, officials agreed on: (i) identifying potential users of the system; (ii) identifying the kind of information to be extracted; (iii) explore options for information generation; (iv) linking workshop outputs with objectives of other work packages.

3.3.3 Presentation of the database and existing semantic tools and resources

Peter van Boheemen, Agneta Lindsten and Tomas Lunden gave a presentation of the database and demonstrated how to search for information on projects, and the different ways in which the projects could be clustered. Pierre Martin presented a knowledge-based system developed by Cirad, named e-CoFIL, that combines piece of knowledge from 3 information systems to obtain a cross-cutting vision of Cirad’s activities in the agricultural sectors with the SDGs. Mathieu Roche made a presentation about text mining approaches in order to extract information in free texts. These approaches based on 3 steps (i.e. crawling, classification, information extraction) have been implemented at Cirad in the context of animal disease surveillance issues (i.e. PADI-Web: <https://doi.org/10.1371/journal.pone.0199960>). This pipeline could be adapted for the KMS pipeline of the LEAP4FNSSA project.

The following discussion points were raised after the presentations:

- How can the database help to synthesise the projects, and is there a way of including information on outputs and impact of the project in the database?
- Who will maintain the database and KMS after the project comes to an end? Maintenance of the database once the project ends could be challenging because of the intensity of the workload in maintaining it and uploading information. Projects and organisations require incentives for them to bring in information to the database.
- This matter needs to be taken up with the project SC at a right time?
- It is important for the project to contribute towards designing an efficient and user friendly database system and adds value to the existing databases which are many.
- The ideal database should contain/share reports of past projects in order guide investments and provide lessons to existing/potential projects.
- The data base can be designed in such a way that one can easily narrow down their search based on geolocation, thematic areas and type of research.
- Given that data collectors rely only on people bringing information, WP3 can also explore the opportunity to invite people to enter information directly into the data base. This however is

likely problematic as it can compromise the quality, validation and moderation of the information on the database.

- The database needs to be accessible- to HLPD members, researchers, policy makers, funders and other stakeholders.
- Currently, The ARC and SZIU (Adam) are responsible for collecting project documents that are considered valuable for the FNSSA roadmap. Given the amount of EU and AU partnership projects, it was agreed that the ARC and SZIU should collect project documents that date back as far as 2007.
- It was explained that the type of project documents that are being collected are in line with the objectives of the LEAP4FNSSA. The collectors of documents rely heavily on the Web for information. It was suggested that they should not limit their search for projects to agricultural science links but rather extend to other disciplines that might have an indirect impact to the agricultural sector.
- The greatest challenge in mining documents is that, some African countries do not have good access to network connectivity, and not much of the information on implemented projects is available online. This also makes it difficult to find reported information regarding EU-AU partnership projects. In such instances, it would also be difficult to disseminate information in such countries.

3.3. 4 Database objectives and intended users

Participants engaged in an exercise to identify potential users of the database and the kind of outputs that would be derived from the database. Table 1 shows the results of this exercise as indicated by the workshop participants.

Objectives and intended users of the database

Potential users of the database	Identify intended objective of the database
Researchers in Africa & Europe R& D funders on different levels Donors General public Open science Journalist Policy makers HLPD Governments Extension officers	Trends on wellbeing of people, information, research focus, funding. Identify and visualise gaps in research and knowledge Identify distribution of research expertise across countries and organisations Information of infrastructure, research capacity and research products. Create maps, graphs, tables, statistics

3.4 Day 2: KMS Pipeline approaches and tools

Pierre Martin made a presentation on Agrovoc and the application of Agrotagger to the WP3 database:

- For collection of documents, the data miners could use the project and donors' website (mostly for ongoing projects) for various reports.
- However, for finished projects, there should be a liaison with the co-ordinator for assistant.
- HLPD will give guidance on which data should be made public/private.
- KMS will not publish any classified / sensitive data.
- KMS will be held as a public domain.
- Analysed data should be qualified so to avoid being biased.

3.4.1 Evaluation of the semantic aspects of the KMS pipeline

Based on current data of the projects (e.g. abstracts in the database), text mining approaches have been applied in order to extract keywords automatically. The method and the results are presented in Section 4.

3.4.2 From database to KMS

- There are different types of databases in the world.
- KMS should be developed in such a way that it is user friendly and the key words build from semantic layers.
- It should also use prediction of words.

Sustainability of the KMS database

Now	After the project
CIRAD is hosting the KMS database now as there are no specific institute /saver that is mandated to host , feed and maintain it	The future of the data base is not well stipulated. A letter should be written to EU for possibilities/ strategies to maintain the database.
KMS should be well developed so that it can still be used by future projects	The challenge should be communicated with the steering committee for guidance

The same issues were raised for the WP3 project database presently hosted by WR.

3.5 Day 3: 02 October 2019

The following issues were scheduled for discussion on day 3: Sources and collecting data strategy, KMS & database users, Linking WP3 project database & the KMs, Hosting project database and KMS sustainability, Data management plan, Adding/linking with other work packages, Communicate first semantic results.

3.5.1 Inventory of resources for the KMS

The participants discussed possible sources of information to input into the database. A decision was made that the KMS and the database would only include all projects going back to 2007 when the FNSSA partnership started. In the interim, bilateral projects would not be included. African Union grants projects would be included as they were part of the partnership. Documents for all the projects which are included in the database would be input into the database. These documents were then prioritized in order of their usefulness for the KMS.

Challenges in collecting documents

- Some projects do not have published documents- Annual reports for (EU-AU) projects have reports.
- Reports may not be on the web but coordinators have them- One-person from the entire project should be the one communicating to the coordinators.
- Publications should be regarded as reports because some project reports are incomplete. Some are more detailed and tells the project life.
- For most EU-AU projects, the most important documents are publications and final reports.
- Some publications are published after the projects ended.
- Publications are outputs for researchers and not all projects do that.

Downloading documents vs references

- It will be difficult to publish or make available documents that were received through private access.
- It would be better if links to documents be used than fully making all project documents available.
- Quality of published article contents or the number of publications per projects.
- Each may have different influence to the potential user.
- Dr Monica Ebele Idinoba suggested that documents should be selected based on the potential users and the impact it had on people.
- Collecting all kinds of documents would lead us into developing a portal. There are +- 700 project of EU-AU – too much data to collect therefore KMS should provide links and not documents.

All documents are to be considered although they are not all of the same priority level.

Type of documents	Priority
Project reports	1
Workshop reports that do not include minutes	1
Fact sheets – they are mostly focused on one aspect of the project	2
Policy briefs	2

Case studies	2
Publication	3
Thesis	3
Presentations, news, posters	4

- Non project specific documents should not be collected as they are not time efficient.
- All documents should be collected however; they should be named and classified differently.
- KMS does not need naming of document, it only needs project ID and title.
- The KMS team will receive files in folders from ARC and SLU and select the ones they deem most important.
- Case studies should be included if it's available and link into Roadmap objectives.

3.5.2 Data management plan

Pierre introduced the data management plan required by H2020 (https://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-data-management/data-management_en.htm) and the last release of the Leap4FNSSA one implemented by WP4 in collaboration with WP3-2. The database and KMS should be included in the interim report, which ARC and SLU will prepare. There should be continuous monitoring and cross checking with Jacques to ensure that the Data Management Plan is complied with by WP3.

3.6 Resolutions of the workshop

1. The database should include all projects going back to 2007 when the FNSSA partnership started.
2. WP3 will have an open consultation within the project and with other partners to solicit inputs into the project database. This will particularly include the WP2 project partners, whose inputs would be sought using a google form.
3. Additional inputs into the database and project clustering would be sought at the Addis meeting in a side event. WP3 should prepare for the Addis meeting.
4. ARC and SIU will be given access to add projects to the database.
5. For engagement of the HLPD Bureau, the questions to solicit inputs into the KMS should first be sent to the working group. The upcoming Addis meeting in November 2019 could also be used as an additional forum for getting inputs regarding the expectations of stakeholders from the KMS.
6. In identifying documents for feeding the database, the following priority would be given.
 - a. Project documents, reports and project evaluations were most important. This also includes workshop reports.
 - b. Policy briefs and fact sheets were second in priority.

- c. Case study reports, student thesis, and publications were classified third in priority.
 - d. Presentations, news blogs, posters were not necessary to collect.
 - e. Non-project specific documents such as reports compiled by FARA based on funding which is not tied to a project – collect if we come across them a decision will be made later.
7. The KMS might have potentially sensitive output and strategic information. Once it is complete the information will be reviewed, and through consultation of the project SC, decisions will be made regarding restrictions to the KMS if necessary. This will be considered in tandem with the project's data management plan.

4. Semantic Information to Integrate in the LEAP4FNSSA KMS

4.1 Objective: Evaluation of the semantic aspects of the KMS pipeline

Based on current data of the projects (e.g. abstracts in the current database), text mining approaches have been applied in order to extract keywords automatically. These keywords enable to highlight 2 types of semantic information useful for the LEAP4FNSSA KMS: (1) semantic information for classification task, (2) semantic information for indexing. In this context, 2 kinds of evaluations have been done by the workshop participants to evaluate the quality of keywords obtained by text mining with the LEAP4FNSSA data. A brainstorming based on the results of the evaluation has been organised by the CIRAD team.

4.2 Material and Method

In order to highlight terminology to integrate in the LEAP4FNSSA KMS (KEOPS: Knowledge ExtractOr Pipeline System), a dedicated terminology to the LEAP4FNSSA project has been extracted with a text-mining tool [Lossio-Ventura et al. 2016] from a corpus dealing with Food and Nutrition Security and Sustainable Agriculture (i.e. abstracts of projects). Different ranking measures of the text-mining tool (i.e. BioTex) have been used for selecting terms: *All* (all terms) with *C-Value*, *All* with *F-TFIDF-C*, *Multi* (multi-word terms) with *C-Value*, *Multi* with *F-TFIDF-C* [Lossio-Ventura et al. 2016].

For evaluating the terminology extracted automatically, a survey (toEvaluate_LEAP4FNSSA_Name_final.xlsx) has been proposed to the 9 participants of the LEAP4FNSSA workshop (WP3). The results (anonymized data) of this evaluation is given in the following datasets: Data_LEAP4FNSSA_WP3_terminology_2019.tar.gz

The name of the datasets in Dataverse is called: LEAP4FNSSA (WP3 - KMS): Terminology for KEOPS.

4.3 Results

Based on the analysis of this survey, CLASSES and GROUPS OF KEYWORDS have been selected.

CLASSES (for classification tasks): Food production, Transformation, Distribution, Consumption

GROUPS OF KEYWORDS (for indexing tasks): Water management, Adaptation to climate change, Crops, Agricultural innovation systems, Smallholder farmer development, Partnerships in agricultural research development, Food security, One health, Research + Training, Project management.

EXAMPLES OF INSTANCES:

Water management: irrigation advisory, groundwater storage, water scarcity, irrigation, water scarcity, ground water storage, water quality, fertigation, recycling, water governance, water harvesting.

Adaptation to climate change: climate change, climate variability, water scarcity, climate variability, sustainable agriculture, groundwater storage, soil fertility, irrigation, water governance.

Crops : Cassava, banana, Fonio, fibre, irrigation.

Agricultural innovation systems: innovation system, knowledge, capacity management, research, impact, innovation support.

Partnerships in agricultural research development: development research, knowledge, capacity, innovation, decision support, agricultural research, scientific research.

Food security: food, food system, food production, market access, high quality, dietary fibre, urban food, livestock, crop, fruit.

Research + Training: capacity, knowledge, research, university, scientific research.

Project management: management, quality, implementation, partnership, work packages.

5. Way forward

To date WP3 partners have managed to organise the resources inventory workshop, identify the list of resources to be fed into the KMS, collect some of these resources required, and he work towards development of the KMS. The Pretoria workshop and monthly meetings for all WP3 partners have resulted in close synergies between development of the database and the KMS. The plan was to have a workshop on text mining (W3.2-2) in Hungary in May 2020, where the terms of reference for the query interface could be finalised. This workshop has been postponed indefinitely due to the COVID-19 situation. WP3 partners will continue to monitor the situation and will organise the workshop at the earliest possible opportunity after the pandemic. The work on both the database and KMS is ongoing even during the global lockdown, and partners will continue to use online monthly meetings as a way of communicating and working together.

References

Juan Antonio Lossio-Ventura, Clément Jonquet, Mathieu Roche, Maguelonne Teisseire: Biomedical term extraction: overview and a new methodology. *Information Retrieval Journal*. 19(1-2): 59-99 (2016).
