



Workplan 2017

Submitted by Elizabeth Arnaud based on meetings reports with the CGIAR Ontology working group, Divseek Semantics working Group, Planteome, Wheat Data Interoperability Working group

Workplan Coverage April 15 – December 31, 2017 (Q2 – Q4)

Context and objectives

One of the aspirations of the Big Data in Agriculture Platform is to build greater collective action across CGIAR and its partners around some key topics in order to connect and build organizational momentum and share information and ideas. A key part of this will be achieved through the Communities of Practice (CoPs). The objective of a CoP under the CGIAR Big Data in Agriculture Platform is to facilitate, oversee, and communicate collective action on a particular topic across CGIAR and its Big Data in Agriculture partners.

Given the Multi-disciplinary nature of the CGIAR Big Data For Agriculture, the Ontology Community of Practice (CoP) will span the knowledge domains of importance for the platform with the objective of establishing best practices, guidelines in the selection, use and application of semantics for data harmonization at the collect and storage levels, for data interoperability and data discovery following the FAIR principles (Figure 1). The CoP will contribute to make the best practices and recommended ontologies a long-term reference resource for the CGIAR Big Data in Agriculture in particular and for other platforms. The INSPIRE projects will offer opportunities to test the power of the ontologies and best practices for combining data sets to address a research question.

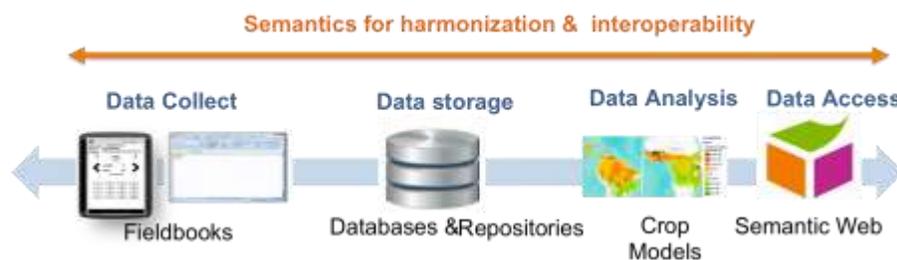


Figure 1: semantics are needed from the data collection to the data access

It will be crucial for this CoP to liaise with the four other CoPs of the Big Data platform for capturing their needs, issues and opportunities, and providing support to find solutions. Members of the Ontology CoP who are also members of the other CoPs will ensure the liaison.

The CoP aims at bringing together CGIAR centers, partners in agricultural research, academia, and private sector partners for creating a critical mass of expertise to tackle to major issues related to semantics for FAIR data in agriculture. It will build on the existing [Crop and Agronomy Ontology CoP](#), on

the CGIAR Ontology working group and existing similar working groups in other initiatives. It will include partners having expertise in ontology design, ontology-driven applications (fieldbooks, databases, annotation tools, query tools, etc) and semantic web technologies.

The CGIAR Ontology Working Group identified the necessity to have a CGIAR Concept/Metadata hub that would feed keyword picklists for the data annotation in repositories and also for the production of surveys/fieldbooks



Figure 2: In colour, knowledge domains addressed by the Big Data platform. In grey, knowledge domains addressed by CGIAR Excellence in Breeding and Genebank platforms.

Products

1. **A map of available expertise in the CoP and identification of additional expertise needed to address Big Data for Agriculture requirements** - Identification of expertise of the CoP members and what expertise we need to attract to optimize the scope and use of the ontologies for Big Data in Agriculture, e.g. Scientists involved in the development of the Sustainable Development Goal Ontology (SDGIO) and the Food Ontology (FoodOn). Also, technical expert advice is needed on the use and adaptation of ontologies to address the requirements of a Big Data (volume, velocity, variety) as well as expertise in ontology engineering for data mining, knowledge extraction, decision-making models, Knowledge Management will be invited to join the CoP – *Timeline: Q2*
2. **Matrix of existing standard vocabularies /ontologies with their guidelines/tutorials, mappings solutions between ontologies & current gaps for knowledge domains of interest to the Big Data in Agriculture** - This action was identified by the CGIAR ontology working group with the objective of accessing a comprehensive global ontology concept server for Agriculture, Food

Security, Nutrition for Health, Ecosystem services. We can connect to the inventories currently performed by existing semantic working groups like Agrisemantics RDA, DivSeek, Planteome, GACS and extend as necessary. The results of the survey launched by the socio-economic CoP will be a key element to use. – *Timeline: Q3*

3. **Ontology modeling questions will be debated to propose solutions** – Modeling the agricultural knowledge to produce quality fieldbooks or surveys, metadata sets and to support performing query systems is not trivial. For example, modeling the crop traits is not always an easy task particularly for representing the plant response to a biotic and abiotic stress, for the scoring for a farmers' preferred varietal trait, for variation of the agronomic parameters under unexpected event (flood, wind, pests), for the variables measuring socio-economic and cultural indicators, etc. *timeline: Q4*
4. **Definition of the key features of an ontology hub as a service to ontology-driven annotation tools and ontology-driven query systems** – Based on their lessons learned from data annotation experience, the CoP members will articulate what is an ideal semantics resource for getting FAIR agricultural data and be able to use the resulting semantic graph to extend a query to discover data. This will be based on exchange of experience in the use of existing tools, including cutting edge tools of the semantic web, query graph, etc *Timeline: Q4*
5. **Provide expert advice on demand to the other CoPs and identify the capacity building/training needs**– *Timeline: Q4*
6. **Compilation of the Data-driven Agronomy CoP 's perspectives on the Agronomy ontology and the fieldbook builder prototype** - An agronomic fieldbook builder driven by the new Agronomy Ontology will be prototyped in a specific project and the valuable feedback and perspective of the CoP on Agronomy will be collected. *Timeline : Q4*

Coordination and Communication

1. **Constitution of a light CoP Advisory Group of 6 representatives** – The adequate composition will be defined and AG members invited, securing the representation of the four other CoPs. This group will support the CoP coordinator in monitoring progress towards the objectives, will identify any additional needs and expertise required, and identify possible alternative strategies when an agreed product cannot be delivered as planned.
2. The CoP will use the communication space and tools provided within the website of the CGIAR Platform for Big Data in Agriculture. The existing Crop Ontology Community Web site will be maintained updated with news from the Big Data Platform as long as it is useful . *Timeline: Q2*
 - a. An online discussion group will be settled
 - b. Ad hoc videoconferences

Conference

1. **Big Data Convention:** the Ontology CoP will organize a dedicated session with possibly keynote speakers at the upcoming Platform Convention, 18-22 September in CIAT, Cali, Colombia, to progress on the listed products. This convention will also offer a venue for assessing some

progress made regarding the topics identified during the 2016 PhenoHarmonIS workshop. It will include hands-on sessions for practicing tools for data annotation, brainstorming session on ontology modeling issues, research axis and as needed, hackathons on semantics, tools and data.

Timeline: Q3

- 2. CoP will contribute to developing the content of the 2018 PhenoHarmonIS workshop and motivate sponsors.**
- 3. The CoP members will share the list of meetings, conferences, hackathons of interest**