

Interest Group on Agricultural Data (IGAD) Pre-Meeting P14

At Helsinki, Finland October 22, 2019

The IGAD Pre-Meeting P14 will focus on IGAD/RDA Recommendations: Implementing Good Practices in Agriculture

Since its inception in 2013, the <u>Interest Group on Agricultural Data</u> (IGAD) has grown in community strength to over 170 members, and is now one of the Research Data Alliance's most prominent Thematic Groups. IGAD is a domain-oriented group working on all issues related to global agriculture data. It represents stakeholders involved in managing data for agricultural research and innovation, this includes those producing, aggregating and consuming data. IGAD also promotes good practices in research with regards to data sharing policies; data management plans; and data interoperability. In addition, IGAD is a forum for sharing experiences and providing visibility to research and work in agricultural data.

Many thanks to the University of Helsinki for their support in hosting the IGAD Pre-Meeting! The location is:

Room Sali 2, Metsätalo, Unioninkatu 40, Fabianinkatu 39, Helsinki FINLAND <u>https://tilavaraus.helsinki.fi/en/city-center/metsatalo-unioninkatu-40/metsatalo-room-2</u>



Looking forward to meeting you in Helsinki

Co-chairs of the Interest Group on Agriculture Data (IGAD), for the IGAD Task Force 2019 Imma Subirats Coll (Food and

Imma Subirats Coll (Food and Agriculture Organization of the United Nations, Italy) Patricia Rocha Bello Bertin (EMBRAPA, Brazil) Cynthia Parr (National Agricultural Library. USDA Agricultural Research Service, US)

IGAD Task Force 2019

Leanne Wiseman (Griffith Law School, Griffith University, Australia) Taina E. Pihlajarinne (Faculty of Law, University of Helsinki, Finland)

Agenda

Location: Room Sali 2, Metsätalo, Unioninkatu 40, Fabianinkatu 39, Helsinki (Finlandia) 22 October 2019

Time	Session
08:30	Registration starts
9:00 - 9:10	Welcome addresses. Opening remarks Imma Subirats (FAO of the United Nations)
9:10 - 9:30	Keynote: TBA André Laperriere (<i>Executive Director, Global Open Data for Open Data and Nutrition</i> <i>GODAN</i>)
Urgent needs for agricultural data Moderated by: TBA	
9:30 - 9:45	Agricultural Production, distribution and the demands for pesticides in Nigeria Michael Adedotun Oke (Department of Agricultural and International Development of Michael Adedotun Oke Foundation) The demands for all forms of agricultural inputs such as seeds, fertilizers, pesticides all over the world are increasing, but there must be guided agricultural policies related to the usage, supply, regulation and policy framework so that it can benefit the farmers, agricultural players at large. This paper focuses on the needs to explore the use of pesticides in preservation, storage systems for Onions and Bambara nut, because here in Nigeria there was a bumper harvest of onion, which was recorded in Sokoto and this year Bambara was translated into a great income and booms for the farmers. The use of the different agrochemicals are based on the individuals purposes in storage, preservation, although the
	usage must be guided with principles, economic factors, safety of production, distribution systems, food quality and security, understanding the agronomic principles, social factors towards sustainable agricultural systems. Here in Nigeria the different Bambara nut is stored either in special coated sacks or pesticide's treated normal sacks, some weevils like those that destroy beans also attack Bambara nut, so to store effectively one must be careful to not damage the seeds. Seasonally, many people are engaged in the local processing of Bambara nuts. Girls are involved in the processing, distribution and selling of the crops, it is traded in almost every weekly market in the North. Meanwhile the onion farmers in Nigeria and dealers have made passionate appeals for provision of Morden storage facilities and more support from government in the form of soft loans and others such as fertilizers, pumping machines and pesticides.
9:45 - 10:00	Title TBA Dyana Sari (<i>University of Tribhuwana Tunggadewi</i>)
	We are now facing a situation on the one hand, food shortages, in other parts of the world facing the disposal of food still in good condition. On the one hand, a group of humans do not know what to eat for tomorrow, while in other places, humans are busy taking care of increasing higher food standards, even though we are on the same planet. Why don't we share what has become an advantage in a place, to fulfill it in a place that is in dire need. The problem is just how to move from a place that is abundant to places that are lacking

	and are in need. Technology can certainly answer it, for no reason it cannot because of distance, space and time. Humans can survive in space, of course those who live on earth must also be more able to survive. But why do people still not finish it, because too many things are taken care of, the consequences are all wasted and those who suffer still suffer. If capital becomes the problem, it is still not difficult to solve it. Many people want to help, many organizations want to raise funds. The problem is now, we want or do not implement it. For the poor and starving to be helped immediately without disrupting the interests of people who are abundant. The world is only one for all so that someone may not feel very hungry while others throw it away. The first way is to gather information, where food is abundant and wasted for one reason or another. If only to maintain price stability, it is not necessary to dispose of it, but immediately contact the collecting agent to immediately process it into food, no matter how. Of course by the way that excess food will not cause food prices to fall. Therefore, there are parties who can immediately be contacted to take and process it. Of course all of this needs immediate treatment, such as an emergency, so that the wasted food can be immediately saved and can be immediately processed so that it is not damaged and can be utilized. Likewise, if a food industry has more food than wasted, why not give it before it is wasted due to expiration, for example. In order to be consumed immediately by those who need it. Not all innovations must find something new but by saving food that is usated, it is also an important innovation for saving humanity in need.
10:00 - 10:15	Open Local Geo-Farmer's Map Paul Kasoma (<i>Youths in Technology and Development Uganda (YITEDEV</i>))
	In the 1980s, millet farms used to be infested by swarms of small usually black caterpillar like worms eating off the tender millet leaves. Farmers never sprayed their farms to get rid of the worms because the worms were not capable of causing a devastating effect to the crop since their life span was very short. Since then, Uganda had never witnessed a more devastating attack like the one witnessed in 2017 by the fall army worm. Farmers looked on helplessly as the worm wreaked havoc across the country destroying maize farms at the speed of wind. Responsible authorities were caught off guard. In a desperate move to rescue farmers, advisory information was being delivered through ineffective communication channels like TVs, Radios, and Newspapers. The information could reach a very marginal number of farmers because there was no effective social mobilization framework under functional communication structure that could help to rally all segments of the society to take appropriate action towards mitigation. The GPS and GIS have advanced quite well and attracted significant use in the agricultural sector lately. Open source platforms are essential in crowd-sourcing data, enhancing research, promoting inclusive participation, and innovation in different sectors. However, in agriculture, mapping has for long been around natural resources like land, water sources, vegetation, arable, and un-arable land etc., leaving out smallholder farmers, and farmers' organizations in the mapping ecosystem. This has created a big gap between smallholder farmers and service delivery. However, without a comprehensive farmers' map, effective service delivery still remains a big challenge. We are seeking to utilize the existing open source technologies by capturing the coordinates and locations of different farmers and farmers' organizations and categorizing them basing on the type of crops they grow for easy identification, access to markets, networking, and for effective data and information flow, and service delivery. The map will b

10:15 - 10:30	Mobilizing Capacity Development for Bridging the Digital Divide in Agriculture Foteini Zampati, Suchith Anand (GODAN)
	Currently, nearly 800 million people struggle with debilitating hunger and malnutrition and can be found in every corner of the globe. That's one in every nine people, with the majority being women and children. The solution to Zero Hunger lies within existing, but often unavailable, agriculture and nutrition data. Open Data offers a great potential for innovations from which the agricultural sector can benefit decisively due to a wide range of possibilities for further use. The Global Open Data for Agriculture and Nutrition (GODAN) Initiative supports proactive sharing of open data to make knowledge on agriculture and nutrition available, accessible and usable, in an effort to deal with the urgent challenge of ensuring world food security. A solution to Zero Hunger already lies within existing, but often unavailable, agricultural and nutritional data. Mostly in developing countries smallholder farmers are not harnessing the power of data and must overcome challenges that need to overcome: first, to gain access to relevant data and services provided by others and, second, to make sure that any data they share does not actually weaken their positions. Improving data access in agricultural methods and farming realities. Today, our society is globally connected and so is our food system, we source our products from all over the world and agriculture is making significant impacts in local communities. A large share of the worlds farmers could benefit from bridging the digital divide (Jellema et al 2015, Berdou and Miguel Ayala, 2018). To be able to bridge this digital divide we need to strengthen the capacity to access and reuse data. It is important that the digital solutions developed are designed to allow all farming communities to maximize their potential; this is especially true in countries with very low literacy levels and limited knowledge of digital technologies, yet where the untapped agriculture potential remains among the highest in the world. In this presentation, we will see on how we can bridge the d
10:30-11:00	Coffee Break
Agricultural standards and semantics Moderated by TBA	
11:00 - 11:15	The Agrisemantics recommendations to improve data interoperability in agriculture Sophie Aubin (INRA), Caterina Caracciolo (FAO of the United Nations), Brandon Whitehead (CABI)
	The Agrisemantics working group focussed on the use of semantics as a tool to achieve better data interoperability, especially in the area of agriculture. During our talk, we will introduce the recommendations produced by the group, and illustrate our plans for their maintenance and extension. We will also hint at a few examples of adoptions already ongoing and invite the audience to share ideas and suggestions for further refinement and adoption of the Agrisemantics recommendations.
11:15 - 11:30	Chinese Agricultural Thesaurus and its application on data sharing & interoperability Xuefu Zhang (<i>Chinese Academy of Agricultural Sciences</i>)
	We aim at exploring an efficient way to share and interoperate agricultural data. In this presentation, firstly we will introduce Chinese Agricultural Thesaurus. Then introduce the architecture and key technologies for semantically processing and opening Chinese agricultural literature data, and method to integrate with AGRIS of FAO to interlink and mashup multilingual agricultural data semantically based on KOSs (the Chinese Agricultural Thesaurus and AGROVOC and their mappings). The results show it is helpful for different

	language-background people to discover international multilingual resources easily and efficiently.
11:30 - 11:45	How Agriculture DWG of OGC and IGAD RDA can cooperate on Agriculture Interoperability Karel Charvat (Plan4all, Wirelessinfo, Agriculture DWG of OGC)
	Most of agriculture related information have direct or non direct spatial character. There is a large variety of data that can be used for agriculture management. this data can be extremely large and heterogeneous. As example we can mentioned: Filed data, Soil data, Yield data, Earth Observation, Climate data, Sensors including online sensors on machinery, A digital elevation model (DEM), Machinery tracking and Land Use. The big problem is heterogeneity of data and low level of usage of standards for farming data. There exist large number of standardisation effort, but problem is, that these efforts are usually not synchronised among different initiatives. The idea if this presentation is to look, how standardisation effort of Open Geospatial Consortium (Agriculture DWG) and IGAD can by synchronised. During presentation will be presented results of discussion of OGC Agriculture DWG in Leeuwen and Banff. During Leeuwen meeting next topics was discussed as key issue for standardisation:Semantic, Ontologies and Linked Data (not only Open Linked Data) – in this direction seems to be important build harmonized models for soil and parcels (plot information). JSON formats, WFS 3.0 and other activities has to bee consider; Trust and security of information and methods of data sharing; New visualization methods including 3D; Metadata standard and linkage of metadata and data using new standards GeoDCAT IP, etc., how to search metadata using search engines; IoT technology and standardization in this domain; Maps as object in relation to An Open Format Linking Geospatial Web Services and Information; Mobile access and citizens science methods or VGI methods; Cloud and HPC computing; FAIR data principles; EO is one from key sources of information for agriculture; ISOBUS and standards for machinery monitoring. During Leeuwen meeting was also identified number of large projects, which can contribute to this standardization effort DataBio, IoF2020, EUXDAT, O4Agri, Demeter, CYBELE, SIEUSOIL and STARGATE The main target of pres
11:45 - 12:00	Intelligent Plant Data Linkage: A View from History, Philosophy and Social Studies of Science
	Sabina Leonelli (University of Exeter) The intersection of historically and sociologically informed philosophy of science, including existing work on data practices and modelling tools, fosters understanding of both the technical and the social conditions under which data can be mined and reused. Such understanding can help towards bridging the current gaps between cutting-edge data science solutions and domains of application, and produce strategies and tools for data linkage that take advantage of the latest computational innovations while at the same time being usable and useful to data users in different sectors and different parts of the world. Building on extensive experiences in the study of Open Data systems in biology and biomedicine, this talk will promote a discussion around which contributions the field of history, philosophy and social studies of science could make towards the development and governance of an intelligent system for plant data linkage, with the goal to promote a global, context-sensitive and sustainable knowledge base for research on food security and related environmental challenges. These may include: (1) a systematic understanding of the global network of infrastructures and standards developed to enhance data interoperability, as well as the diversity of normative visions and criteria used to identify, classify, mine and visualise data; (2) an analysis of the historical roots of the various components of these networks, in an effort to explain current differences and specificities; (3) a critical overview of the role of collections in contemporary crop science, the relation between germplasm collections and digital databases, and the ethical and social issues raised by data collection and monitoring; (4) a framing of the relation between conservation and innovation within

	precision agriculture; (5) support towards developing semantics that can leverage the diversity of participants in agricultural knowledge production, while learning from comparable experiences in biomedicine and public health.
12:00 - 13:00	Lunch
	Implementing FAIR principles Moderated by: TBA
13:00 - 13:20	GO FAIR Food Systems Implementation Network Ben Schaap (GODAN)
	A GO FAIR Food Systems Implementation Network has been set up to advance a global data ecosystem for agriculture and food by implementing FAIR data and services. For better reuse of data, we need to achieve a shared understanding of how we describe and publish data, in a semantically grounded manner. The purpose of this implementation network is to support the implementation of FAIR principles in agri-food sciences, in providing guidelines, tools, methods, etc. with specific efforts towards achieving semantic interoperability. To this end, the Food Systems IN will boost the adoption and implementation of recommendations from existing initiatives such as RDA working groups, GODAN working groups, and W3C. We will work on the following objectives: 1 To advocate for FAIR data principles in data sharing policies. 2 To foster the continued implementation of FAIR principles based on existing recommendations and if needed support to create new ones. 3 Facilitate agreement on the use of vocabularies, standards and protocols. 4 To disseminate best practices to a large community of practitioners.
13:20 - 13:40	The Global Long-term Experiments in Agriculture Network (GLTEN) Metadata Portal Richard Ostler (Rothamsted Research)
	The GLTEN was launched in May 2018 with the aim of establishing a collaborative international network for long-term agricultural experiments (LTEs). The GLTEN now represents over 50 LTEs across the world and covers diverse and contrasting climate and environmental regions, crops and land management systems. To improve the visibility of LTEs and LTE datathe GLTEN has developed a searchable online metadata portal. The portal does not host data, but instead provides metadata about LTEs and directs users to the LTE data owners preferred data access process. The metadata portal uses a minimum information checklist approach to capture key facts about LTEs including design and treatment factors; cropping system; site characterisation including location, landscape, soil and climate properties; administrative information including data access & licensing, contacts and associated organisations; recorded measurements and observations and data availability. To ensure interoperability with related systems, the checklist is mapped to schema.org in line with DataCite mappings and metadata are annotated using existing publicly available ontologies and controlled vocabularies. Metadata are published using CCO licensing and can be queried using the portal API.
13:40 - 14:00	COPO: Extending the frontiers of "FAIR" Data in Agriculture Anthony Etuk (<i>Earlham Institute</i>)
	Modern scientific advances are increasingly reliant on data. In the agricultural sector, for instance, much of today's research and innovation are driven by the integration of ever-increasing amounts of heterogeneous data from disparate sources. Whereas the awareness of the FAIR principles for Findable, Accessible, Interoperable and Reusable data has increased, challenges still exist to implement these principles for storing, managing, analysing and distributing data. Collaborative Open Plant Omics (COPO) is a system that attempts to address some of these challenges by enabling scientists to describe their data

	and research objects using community-sanctioned metadata, and then use public or institutional repositories to share it with the wider scientific community. Data generators are encouraged to use rich semantic metadata when publishing research objects, thus allowing data consumers and tools to find, aggregate, and analyse data which was previously invisible or not shared at all. COPO supports a number of data standards or schemas including the CG Core 2.0 schema (https://agriculturalsemantics.github.io/cg-core/cgcore.html), implemented for use by CGIAR centres. CG Core 2.0 is a schema produced by the CGIAR Big Data Platform to collect information on socio-agronomic studies, thus harmonising data input between CGIAR centres. These centres store their data on locally hosted, publicly accessible digital repositories based on one of three repository software types, i.e. DataVerse, dSpace or CKAN. COPO user interface wizards allow metadata to be added to research objects (e.g., datafiles) with guided help and natural workflows. Metadata can be suggested, that might be appropriate for data based on past submissions and similar workflows. By building upon existing standards to push the state of the art in agricultural data dissemination and mining, COPO seeks to ease the burden of sharing and reusing data by the community. COPO is entirely open source and is freely available on GitHub. A production instance of the platform for use by the community as well as more information can be found at copo-project.org.
14:00 - 14:20	Grassroots Infrastructure: An interoperable data repository for plant science Xingdong Bian, Simon Tyrrell, Robert P. Davey (<i>Earlham Institute</i>) Integrative research requires extensive multi-level approaches to enrich and expose data and workflows so that informatics infrastructures can process them effectively. As part of the Designing Future Wheat (DFW) project, the Grassroots Infrastructure has been developed at the Earlham Institute (EI) to consolidate data and analyses, facilitating consistent approaches to generating, processing and disseminating public datasets in the plant sciences. It is also part of the Wheat Initiative Wheat Information System (WheatIS) project, formalising the infrastructure as the federated UK WheatIS node involving partners from the University of Bristol, the European Bioinformatics Institute, Rothamsted Research, and the John Innes Centre. Its lightweight reusable software stack comprises: an iRODS data management layer to provide structure to unstructured filesystems and WebDAV APIs exposed via EIRods-DAV; interfaces to interact with local or cloud-based analysis platforms; an Apache web server layer to deliver content and provide access to public programmatic interfaces; services such as: BLAST searches on multiple databases across different sites, systems for storing and searching field trial experiments, a mapping tool showing pathogen samples with temporal and spatial data, and adding API layers to the SeedStor application by the Germplasm Resources Unit allowing it to make seamless connections to research objects stored on our new DFW data portal. To make the data as reusable as possible, all data is marked up using appropriate ontologies. The Grassroots Infrastructure can be run locally or packaged in virtual containers and deployed on a variety of hardware thus representing a decentralised system, allowing information generators to retain control over their resources but allowing interconnected resources to access each other consistently. We are currently working on va

	It is fully open source and available on GitHub. More information can be found at https://grassroots.tools
14:20 - 14:40	Enabling FAIR data to support innovation in agriculture Ruthie Musker, Kathryn Reynolds (<i>CABI</i>); Deborah Yates <u>, Pauline L'Henaff</u> (<i>ODI</i>)
	CABI and ODI are leading a two-year project on enabling FAIR data access and sharing in agricultural data ecosystems. The project, funded by the Gates Foundation, is a continuation of a pilot project, in which we developed data flow maps in India and Ethiopia to highlight challenges in FAIR data compliance and implementation in Gates Foundation funded programmes. Our primary goals for the project are to create tools, guidelines and learning programmes that will support the management, sharing and governance of FAIR data across agricultural data ecosystems for programme officers and grantees. Our presentation will give detail into some clear examples of FAIR data implementation in agricultural data ecosystems in India and Ethiopia, including a Soil and Agronomy Data Sharing Policy in Ethiopia, and data sharing agreements in India. In our presentation, we will explore how we have implemented FAIR, what has worked (or not), and what methodologies exist behind our FAIR tools. Although the aim of our project is to provide tools and resources, we want to leverage the IGAD community to ensure we build upon and align with existing FAIR data research and resource development.
14:40 - 15:00	Coffee Break
	Changing Ag Data Management Culture Moderated by: TBA
15:00 - 15:20	Effective Open Data Use in Agriculture And Nutrition; Outcomes, Lessons And Recommendations From GODAN Action Chipo Msengezi (CTA) The Global Open Data for Agriculture and Nutrition (GODAN) initiative focuses on building high-level policy and public and private institutional support for open data. This focus on advocacy, convincing stakeholders in agriculture and nutrition of the potential benefits of open data, is vital but, on its own, is not enough. GODAN Action supports the wider GODAN initiative by focusing on the capabilities of data users, producers and intermediaries to engage effectively with open data in the agriculture and nutrition sectors. This 3.5-year project is coming to an end in December 2019. The team conducted a lot of research on open data standards, on impact assessment and produced tools and training material. It includes work in the area of land data and weather data. The research has led to key recommendations for key agriculture open data stakeholder groups who include data publishers, data managers, standard developers, data users, evaluators, development organisations, policymakers and trainers. These recommendations relate to the need for addressing discovery issues for land and nutrition data, by improving the use of metadata; the need for promoting the integration of evaluation and elements of evaluation during and beyond the lifetime of a project and the outcomes of open data training approaches and
	experiences of the trainers network. The project partners will be able to share lessons and point participants to useful tools such as the geospatial service recently created, a framework for measuring the impact of open data and resources for training.
15:20 - 15:40	Improving agricultural data management practices: raising the bar together Cynthia Parr (USDA National Agricultural Library)
	The United States National Agricultural Library (NAL) is currently engaged in a number of coordination efforts that involve both changing the culture and improving technology around FAIR agricultural data management. In part, this is part of ongoing work to improve and sustain the Ag Data Commons, the US Department of Agriculture's system for

	cataloging and storing datasets created with USDA funding, and the open source DKAN Science platform behind it. However, there is no one system that will improve consistent data management practice due to the diverse needs of agricultural research communities. Therefore, NAL has been working with US agricultural university libraries, discipline-specific data managers, and other stakeholders to promote the application of existing recommendations, and to identify gaps and barriers to adoption and form collaborations to accelerate progress. For example, a Data Management Plan review service will be most effective if we coordinate efforts across libraries and research offices, and assist funders in training those who review plans. We initiated a series of workshops aimed first at researchers and then at extension agents, Driving Innovation through Data in Agriculture (DIDag) to increase awareness of new policies and to begin addressing challenges in specific communities. We are participating in the development of small unmanned aerial systems research data standards. Finally, we plan to leverage the central cataloging features of the Ag Data Commons to provide key services to a growing group of agricultural genomics databases, AgBioData, as they coordinate their FAIR data practices. Through the RDA Interest Group on Agricultural Data we can ensure that emerging practices and promotion strategies are shared effectively across global communities.
15:40 - 16:00	Developing Best Practice Principles for Ag Data Management Leanne Wiseman* and Jay Sanderson** (Griffith Law School)
	Agricultural data must be managed like any other asset. This requires strategy and best management practice. For the full value to be ascertained from Ag Data, those organisations that collect, manage and share Ag data should have a clear direction and sense of what is to be achieved and how best to achieve that goal. While many different industries are developing Data Strategies to guide the creation of an enabling environment for data sharing, many have yet to embed best management practice for data. Best Management Practice should not only ensure that the technical and technological capabilities of data management and sharing are addressed but also that the legal, ethical and security concerns of the farmers and agricultural industries who contribute the data, are also considered as part of best management practice. The focus of this paper is to discuss the development of Best Management Practice principles for Australian agricultural industries. In line with RDA/IGAD recommendations, the Best Management Practice principles emphasize the importance of transparency and trust in the way that ag data is collected, managed and shared.
Open Forum: The future of IGAD Moderated by: TBA	
16:00 - 16:15	The History of the Research Data Alliance Interest Group on Agricultural Data Imma Subirats (FAO) and Cynthia Parr (USDA)
16:15 - 17:30	In this session we want to hear from the community on what you want or need from the Interest group on Agricultural Data. RDA is seeking outcomes from the work of its interest and task groups. Please be prepared with your answers to the questions: 1) What benefits have you or your organization gained through participation in IGAD? 2) How can IGAD help make a difference in the future? We will use ideas from this discussion to report to RDA leadership and to develop an action plan for future activities.
	Closure of the Pre-meeting and moving to the RDA Plenary Meeting Session



Food and Agriculture Organization of the United Nations