# INTEREST GROUP ON AGRICULTURAL DATA (IGAD) MEETING

**Gaborone[1] (Botswana). 5-6 November 2018**

## DATA COLLECTION IN AFRICA
## STATE OF THE ART, CHALLENGES AND SOLUTIONS

Formed in 2013, since its inception the Interest Group on Agricultural Data (IGAD) has grown in community strength to over 200 members, becoming one of the RDA's most prominent Thematic Groups. IGAD is a domain-oriented group working on all issues related to global agriculture data. It represents stakeholders in managing data for agricultural research and innovation, including producing, aggregating and consuming data.

Beyond this IGAD promotes good practices in research with regard data sharing policies, data management plans, and data interoperability, and it is a forum for sharing experience and providing visibility to research and work in agricultural data.

One of IGAD's main roles is to serve as a platform that leads to the creation of domain-specific Working Groups. To date, five successful Working Groups (WGs) have been formed under IGAD, these are: Wheat Data Interoperability, Rice Data Interoperability, AgriSemantics, On-Farm Data Sharing and Capacity Development WGs. The groups played an active role at the last RDA Plenaries, in Berlin (Germany), Montreal (Canada) and Barcelona (Spain).

### OBJECTIVES

- To promote good practices in the research domain: data sharing policies, data management plan, data interoperability
- To provide a platform for networking and cross-fertilization of research ideas in data management and interoperability
- To solicit and promote interactions and projects among the major international institutions and groups worldwide which work on agricultural research and innovation
- To achieve data interoperability

### EVENTS AT RESEARCH DATA ALLIANCE PLENARY MEETINGS

The RDA holds its plenary meetings every six months at different locations around the world. Built around the RDA Working and Interest Group meetings the sessions are designed to consolidate the RDA's work and that of its interest groups in building social and technological bridges towards open data sharing. IGAD's role since its inception at the RDA Gothenburg has expanded with every

[1] Venue to be announced

successive Plenary. The outcomes of the Wheat Data Interoperability Group have proved to be a success and the challenge is now on to replicate those successes with the Rice Group.

IGAD has used the Plenary meetings as a place to reach out and forge new alliances with other groups, as well as creating new offshoot groups aimed at specific solutions. During Plenary sessions IGAD has hosted a wide array of speakers and discussions and it continues to work alongside major international initiatives such as GODAN, CGIAR, CTA, EMBRAPA, CAAS, FAO of the UN, INRA, CiMMYT, CODATA, NARIs/NARS, among others.

- RDA P11. March 2018 in Berlin (Germany) Invitation and agenda, Insights
- RDA P10. September 2017 in Montreal (Canada) Invitation and agenda, Insights
- RDA P9. April 2017 in Barcelona (Spain) Invitation and agenda, highlights
- RDA P8. September 2017 in Denver (USA) Highlights
- RDA P7. March 2016 in Tokyo (Japan) Agenda, results, highlights
- RDA P6. September 2015 in Paris (France) Insights, Results
- RDA P5. March 2015 in San Diego (USA) Agenda
- RDA P4. September 2014 in Amsterdam (The Netherlands) Agenda
- RDA P3. March 2014 in Dublin (Ireland)
- RDA P2. September 2013 in Washington DC (USA)

## NEW IGAD MEETING IN GABORONE (BOTSWANA) AT THE RDA P12, NOVEMBER 2018

The new edition of the RDA Plenary Meeting, the biannual meeting of the research data community, will be held on 5-8 November 2018 in Gaborone, Botswana. In this edition, the 12th RDA Plenary Meeting will be included in the International Data Week (IDW 2018), and combined with SciDataCon 2018. To a greater degree than at IDW 2016 in Denver, Colorado, the events will be integrated, such that each day will feature an inspiring and engaging range of activities.

**With the focus on "Data collection in Africa: State of the art, challenges and solutions", IGAD members are organizing a series of events November 5 and 6, 2018, during the IDW 2018.**

The conclusions of these activities will be shared at the 12 RDA Plenary Meeting at the Agricultural Data Interest Group Session.

On November 5, the IGAD community members will organize a one-day meeting dedicated to the theme "How agriculture in Africa can benefit from data collection". During this day, IGAD would like (i) to explore data collection practices in Africa and challenges associated to it; (ii) to discuss possible benefits of data collection for farmers and (iii) to share straightforward and low cost data collection methods.

November 6 will be dedicated to a GODAN Action Workshop associated to IGAD. Open data can benefit a great variety of actors tackling food security and nutrition challenges in developing countries. However, the full realisation of these opportunities requires insights in how to close the digital divide between open data and the potential users. GODAN Action therefore is an innovative project researching and making practical application of its findings to enable effective use of open data in agriculture and nutrition. The project explores three work streams: (1) improving knowledge

and capacity on (usage of) standards, by providers and intermediaries; (2) understanding impact, to be able to efficiently monitor and measure impact and to be able to improve the potential impact of new initiatives; (3) capacity building as a mechanism to leverage and extend the potential of networks. The project has set up a programme to build the capacity of potential stakeholders to both understand the potential of open data for agriculture and nutrition and to engage with it practically.

GODAN Action, IGAD and partners propose two parallel workshops for (1) researchers and (2) infomediaries (technology specialists, librarians, journalists, communication officers etc.). The aim is to enable participants to discover, use and describe the benefits of open data, and how they impact agriculture and nutrition development.

**PARTICIPATION**

Would you like to attend and/or active participate in these events? Please write an e-mail to AIMS@fao.org.

We are looking forward to meeting you in Gaborone,

***Co-chairs of the Interest Group on Agriculture Data (IGAD), for the IGAD Task Force 2018***

*Esther Dzale-Yeumo (INRA, France)*
*Patricia Rocha Bello Bertin (EMBRAPA, Brazil)*
*Armando Stellato (University of Tor Vergata, Italy)*
*Imma Subirats Coll (Food and Agriculture Organization of the United Nations, Italy)*

**IGAD Task Force (Botswana 2018)**

Centre International de la Recherche Agronomique pour le Développement / International Centre for Agricultural Research for Development (CIRAD)
Committee on Data for Science and Technology (CODATA)
Food and Agriculture Organization of the United Nations (FAO)
Global Open Data in Agriculture and Nutrition (GODAN) and GODAN Action
Institut national de la recherche agronomique (INRA)
International Maize and Wheat Improvement Center (CIMMYT)
Technical Centre for Agricultural and Rural Cooperation (CTA)

# Agenda

**1st IGAD Pre-Meeting Day**
**Venue:** Botswana Innovation Hub, Plot 69184, Block 8, Science and Technology Park Gaborone, Botswana
*5th November 2018*

| Time | Session |
|---|---|
| 08:30 | Registration starts |
| 09:00 - 09:30 | **Welcome addresses. Opening remarks**<br>Esther Dzale and Patricia Bertin (IGAD co-chairs)<br>Chipo Msengezi (CTA, The Netherlands)<br>Andre Laperriere (GODAN, United Kingdom)<br>Pascal Bonnet (CIRAD, France) |
| 09:30 - 10:30 | **Panel Session on Data collection in Africa: State of the art, Challenges & Solutions**<br>Moderated by Pascal Bonnet (CIRAD, France) |
| 10:30 - 11:00 | *Coffee Break* |
| 11:00 - 11:30 | **Open Land Use for Africa (OLU4Africa)**<br>*By Karel Charvat (Czech Center for Science and Society/Plan4all, Czech Republic)*<br><br>The Plan4business project (2012-2014) has identified a gap in land use data availability, especially outside big cities, in suburban and rural areas. The Urban Atlas of the European Environmental Agency covers only cities above 100,000 inhabitants. So for example in the Czech Republic, only 13 cities are covered. For the rest, there is the Corine Land Cover, which can be used only for regional and national analysis. The lack of land use data on local level led to an idea of combining data from various sources and of different levels of detail into a seamless map. This idea has been picked up by the SDI4Apps project and turned into a pilot application (Pilot IV: Open Land Use Map through Volunteered Geographic Information). The innovative aspect of the pilot is not only in the methodology of combining data into a seamless database, but also in using crowdsourcing for data collection and update. An important aspect is that data are available as open data.<br><br>Open Land Use 4 Africa project integrate newly identified data sources into the map. One of the sources is the Africover project which has done land cover mapping of particular African countries. The data is available for download for 10 African countries. Also it was found GHSL (Global Human Settlement Layer) dataset – from which areas of residential land use has been identified. It is also available for the download. Moreover, it is possible to use other tags from OSM to identify land use except the 'landuse' tag. For example, 'natural', 'amenity', 'leisure' tags etc. All in all in the list of priorities according to the data spatial resolution will be 1) features from OSM 2) features from Africover 3) features from GHSL. The current state of the map can be seen here: https://goo.gl/Rn32HD |

| | |
|---|---|
| 11:30 - 12:00 | ***Opening access to data: how to balance open access and privacy rights***<br>*By Leanne Wiseman (Griffith University, Australia) and Jay Sanderson, (USC, Australia)*<br><br>The adage of 'open data' is that data are freely accessible as long as "personal data" is secured and does not lead to identification of individuals. Recent developments in Privacy law, such as the passage of the GDPR in Europe, has focussed the attention around the world on the way in which 'personal data' is being currently being managed and shared. This presentation will discuss what is defined as "personal data" under the GDPR, the rights being given to EU residents to control the use of their data and what safeguards (e.g. anonymization) need to be put in place to ensure compliance/accreditation. While not all agricultural data is 'personal data' there is certainly the need for a clearer understanding within the agricultural community of what types of agricultural data may also contain or become 'personal data.' Distinguishing between 'personal data' and non-personal data is an important first step in determining the most appropriate policy, regulatory and practical approach to managing the opening of access to agricultural data in Africa. Understanding what is meant by "personal data" is also necessary to ensure that the risks and benefits of data sharing are fully contemplated, and is needed to build trust in open data systems. Indeed, education and capacity development of small lot farmers is vital so that trust can be developed in the farmer-researcher-donor relationships. Good ag data management practices are integral to the success of open data initiatives in agriculture and nutrition in Africa. |
| 12:00 - 13:00 | *Lunch Break* |
| 13:00 - 13:30 | **Role of competence centers (research) and open science movement in the framework of digital innovation hubs in agriculture, example of #DigitAg convergence lab**<br>*By Pascal Bonnet (CIRAD)*<br><br>If **open research and public open data** are an essential fuel towards a digital transformation of agriculture, this is not the only component that triggers such transformation. Firstly there is a need for creating an **ecosystem** of interactions between business and research entities. In Europe, **Digital Innovation Hubs** DIH are nowadays organizing to provide **services** to **agricultural sector** and **ICT business** (farmers, agribusiness, advisors, governments, innovation brokers, startups, investors, ICT developers...). To pursue this purpose **DIHs** need also to **facilitate access** of the private sector to **competence centers** (universities, research and technology organizations, incubators) who provide various competences in **digital** technologies and know how, offer access to various **infrastructures** and research platforms and open "information" (data, algorithms, models, thematic knowledge). Moreover, **competence centers** in digital agriculture have to make sure that a high level of hybridization and interaction between computer science and more thematic science (agronomy, sociology..) exists.<br>The presentation will focus on the formation of the **#DigitAg convergence laboratory** (standing for digital agriculture http://www.hdigitag.fr/en/ ) in Occitania region France, which is an attempt to boost interaction and interdisciplinary work between various scientific laboratories engaged in digital agriculture research. The presentation will focus on the current **#DigitAg** PhD program by scrutinizing the topics that are being addressed so far, to show the new networking this has created |

| | |
|---|---|
| | between the two domains of digital and agricultural science, and how far elements of open science (open data, open source..) are being mobilized. |
| 13:30 - 14:00 | **Data4Ag: Farmers benefit from data and digitalisation**<br>*By Chipo Msengezi (CTA)*<br>Farmers' organisations (FOs) and agribusinesses are important players in providing services to farmers in ACP countries. However, farmers' organisations and agribusinesses face a limited performance of delivering better services to farmers. These lead to inefficient production and inadequate data access, handling and sharing. To improve the livelihood of farmers, there is a need to support organisations serving those farmers with aggregated data for better decisions making. Having a more efficient management system for supporting the agribusiness clearly can improve access to inputs, crop husbandry advice and access to markets.<br><br>The presentation will discuss CTA's Data4Ag programme which addresses four components to support FOs in their approach to digitalisation using a data strategy that positions them to extend data services for the benefit of smallholder farmers. These key components are: (1) Field proofing and farmer profiling to deliver better services to smallholder farmers; (2) Scaling up data services for farmers through extensive consultation with farmers' organisations; (3) Working in collaboration with stakeholders to raise awareness and inform policy on data management issues in order to support and improve farmer registration and data management by FOs; (4) Capacity building on data management particularly focussed on shared data. |
| 14:00 - 14:30 | **Agricultural data collection and management in Burkina Faso: state of the art and issues**<br>*By Philippe Yanogo, consultant in agriculture and environment, CEO GENITEC entreprise*<br>In Burkina Faso, food production is mainly based on crops and animal husbandry. The first sector concerns essentially cereals like sorghum, millet, maize then subsidiary tubers like yam and sweet potato. The last gather cattle, goat and sheep. Climate is typically dry with 3 types of agro-ecological zones. Thus, from south to north one can note followings zones: South-Sudanese, North-Sudanese and Sahelian areas. Crops production is tributary to rain which last 3-4 months per year. Climate change events like flood and drought disturb frequently the production. Three main rivers cross the territory draining away 2/3 part of drain water to the Atlantic Ocean by Ghana's coast. The other part is draining away to Niger River. In short, hydrologic network is comparable to a funnel whose collected water is drain out of the territory.<br><br>Burkina's agriculture sector faces many kinds of problems among which:<br>-Degradation of natural resources due to instability of rainy season, excessive exploitation of lands, recurrent events of extremes climate phenomena (floods, droughts) that threaten hydrologic network and watershed,<br>-Reduction of diversity of the ecosystems,<br>-Increase of pests' attacks associate with agro-chemical pollution,<br>-Deficit of organization at government's and farmers' levels to control access to agricultural markets (supplying and sells),<br>-Deficit of organization of agricultural actors to access to knowledge.<br>To face all these challenges, we need to collect and manage data in order to |

| | |
|---|---|
| | produce more relevant information and knowledge systems. For example, we need more and better data and knowledge related to:<br>-Agricultural meteorology and climate change to prevent or launch alert about precipitations previsions , extremes events and manage water reserves (develop complementary irrigation regulation, extend drop to drop practice, practices of water conservation, etc.),<br>-Dams and rivers watersheds management to improve and increase dry season production,<br>-Soils fertility management to enhance its productivity,<br>-Pests control and alert to protect crops and environment,<br>-Market information system for improve the income of the farmers |
| 14:30 - 15:00 | **AgroFIMS, the Agronomy Field Information Management System**<br>*By Céline Aubert, Marie-Angélique Laporte and Medha Devare (CGIAR, Bioversity International, France)*<br><br>To fully understand the implications of varying factors within any cropping system, it is essential to combine results of field management practices with information on crop phenotypes. Integration of prebreeding, breeding and agronomy data does not, however, happen as frequently as it could. One reason is that agronomic trial data are often collected, described, and stored in inconsistent ways. This situation impedes data comparison, mining and interpretation for meta-analysis, as well as data reuse in models and decision-support tools. Hence, the use of standardized metadata and variables to capture and describe the data is necessary.<br><br>The use of ontologies with fieldbooks has proven to be a successful way to harmonize breeding data, such as with the Crop Ontology and Integrated Breeding Platform (IBP). Ontologies proffer further advantages for data mining and exploration, including the possibility for making inferences, and connecting with Linked Data in other related domains (e.g. nutrition and health). Similar tools for agronomy trials are needed. AgroFIMS: (Agronomy Field Information Management System) consists of a web application displaying variables that are concepts in the Agronomy Ontology, relevant crop ontologies, and other related ontologies. These variables are sorted into a series of modules representing the typical cycle of operations in agronomic trial management, such as land preparation, irrigation, weeding, soil fertility, weather and soil parameters, biotic and abiotic stress observation and control, and more. The fieldbook application allows the creation of data collection sheets using the same base set of semantics, including variables, terminology, units, protocols, etc., enabling standardization in data collection and description and linkages with breeding and other related data. In addition, AgroFIMS will reduce error and ease data collection via an in-development mobile app that will allow digital data collection in the field; it is also possible to adapt AgroFIMS to other digital data collection platforms such as Open Data Kit (ODK).<br><br>AgroFIMS is built on the open source statistical software Rand R-Shiny web framework; both enable users to collect, analyse and communicate results for different types of experiments using interactive web interfaces. The tool has been developed on the Highly Interactive Data Analysis Platform (HIDAP) by a team at the International Potato Center (CIP) who have used the platform and ontologies to develop a breeding fieldbook. Ontology experts and agronomists at Bioversity |

| | |
|---|---|
| | International, the International Food Policy Research Institute (IFPRI), and Ontocale SRL have also been key contributors to AgroFIMS. An AgroFIMS prototype was subject to user testing by agronomists in September 2018 and garnered interest and enthusiasm. A first release will be used by early adopters in 2019 for field testing to capture data during the 2019 cropping season.<br>1 AgroFIMS GitHub: https://github.com/AGROFIM |
| 15:00 - 15:30 | *Coffee Break* |
| 15:30 - 16:00 | **Showcasing the improvement and uptake of Dataverse through global collaborations and networks**<br>*By Richard Fulss, Jesús Herrera de la Cruz (CIMMYT, Mexico), Sonia Barbosa (Harvard University, USA), Jonathan Crabtree (ODUM Institute, USA)*<br><br>The Global Dataverse Community Consortium (GDCC), initiated and represented by members of the IQSS at Harvard University, the Odum Institute for Research in Social Science in the US and by DANS in the Netherlands, promoted the creation of an international melting pot (GDCC) where many interactions at global level are conceptualized. This will be presented and discussed in the session with a specific focus on the current African context including potential use cases as well as an imaginable support infrastructure concept for research data management with Dataverse.<br><br>One selected showcase will demonstrate how the International Maize and Wheat improvement Center (CIMMYT), which is part of the Global Dataverse Community Consortium, utilizes and promotes Dataverse in the international agricultural research area. Thanks to the global partnership network of CIMMYT, including its offices in several African countries, the use of Dataverse has been expanded. It was implemented in collaboration with another research organization in Africa. This could be a use case for adoption not only in the African but in the global context as well. |
| 16:00 - 16:30 | **Farmers' rights on data and ownership issues**<br>*By Foteini Zampati (Global Open Data for Agriculture and Nutrition (GODAN) and Kuratorium für Technik und Bauwesen in der Landwirtschaft (KTBL))*<br><br>There is no doubt that data-driven agriculture has increased the agricultural production and productivity, reduced the risk and improved resilience in farming, brought more economic and efficient use of natural resources, and helped farmers in decision making. Nowadays, with the exceeding technology and innovations, more agricultural technology providers (ATPs) enter the market, focusing more on aggregating farmers' data that subsequently led the farmers raising concerns about data ownership, privacy and security.<br><br>Farmers need to feel and be engaged in the decision process of how collectors will use their data, they need assurances of their privacy and control, they seek transparency in their interactions with providers, they would like to have access to data and to receive also benefits of their data. Mostly in developing-countries smallholder farmers are not harnessing the power of data and must overcome challenges and risks to ensure that investments benefit them. In this case, two are the main 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.<br><br>Because of the above mentioned situation, important questions and issues have arisen:<br>● Who owns data?<br>● Who is entitled to the value of the data?<br>● How will that data be used or potentially shared?<br>● What about data protection? what do we mean by the farmers' s rights to data?<br>● What is the state of recognition of these rights in national and international level?<br>● What's the role of GDPR in the agricultural sector?<br>● How should these rights be implemented in local and international laws, guidelines and policies and how can they be protected?<br>● What should be done to include farmers in the mechanisms of data (collection, evaluation, transmission, use)<br><br>These issues of course aren't new to people in the agricultural sector. But there is a big need today to address them right and quickly to ensure farmers rights. |
| 16:30 - 17:00 | **IGAD: Identifying training gaps in agricultural data management and interoperability - Findings of the IGAD Capacity Development for Agriculture Data WG**<br>*By Suchith Anand (GODAN), Chipo Msengezi (CTA), Karna Wegner (FAO)*<br><br>At the RDA9 IGAD Pre-meeting in Barcelona, a strong interest was identified for developing a working group (WG) that looks for synergies in capacity development in the context of agricultural research and its potential contribution to the realization of the Sustainable Development Goals (SDGs). As a result, the IGAD Capacity Development WG was formed and endorsed by the Research Data Alliance (RDA) in August 2018. The aim of this WG is to develop synergies between existing education and training activities and agricultural science needs by performing a landscape assessment to identify existing gaps and training requirements within Interest Group on Agricultural Data (IGAD) WGs and related groups. A particular focus will be on sharing knowledge about training initiatives and technologies, reducing digital divides so that researchers and practitioners in developing countries can benefit. It will also empower the existing collaboration between IGAD, GODAN and GODAN Action. One of the first activities of the WG is to perform a landscape assessment to identify existing gaps and training requirements within the context of IGAD network. The Capacity Development WG will present the first results of this landscape assessment in the IGAD pre-meeting in Botswana. These results will be discussed with the experts that will be present during the RDA 12th IGAD pre-meeting and circulated within the IGAD community. After a report of the landscape assessment is created, a case study to identify good practices for capacity development and advocacy will be developed as a next step. |
| 17:00 - 17:30 | **2017 Global Agricultural Science Research Fronts Analysis**<br>*By Zhang Xuefu, Sun Wei and Xu Qian (Chinese Academy of Agricultural Sciences)* |

| | |
|---|---|
| | We can find research fronts by continuously tracking the most important scientific research papers in the world, and analyzing the cited patterns and clustering of papers, especially the frequent co-citation of highly cited papers into clusters. A research front is formed when a cluster of highly cited papers are cited together to a certain degree. Base on the bibliometrics method, we selected 14 research fronts in 2017 from the ESI database of agricultural science, animal and plant science in Web of Science relying on the related experts. We also analyzed the structure and layout of the 2017 global agricultural science research fronts and etc. |
| 17:00 - 17:30 | **Wrap up and Closure of the day**<br>**Details about the 2nd Pre-Meeting Day** |

**2nd IGAD Pre-Meeting Day**
**Venue:** Botswana Innovation Hub, Plot 69184, Block 8, Science and Technology Park Gaborone, Botswana
*6th November 2018*

| Time | Session | Facilitator |
|---|---|---|
| 08:30 - 09:00 | Registration | |
| 09:00 - 09:30 | Introductions<br>Welcome and overview | Chipo |
| 09:30 - 10:30 | **Understating Open Data**<br>*Defining Open Data*<br>*Data responsibility* | Charles |
| 10:30 - 11:15 | **Licensing Open Data** | Mercy & Charles |
| 11:15 - 12:15 | **Publishing Data Open**<br>*Frameworks for sharing and exchanging data (Five Stars, FAIR principles)* | Mercy |
| 12.15 - 13.00 | **Use case:** | **Boniface** |
| 13:00 - 14:00 | Lunch | |
| 14:00 – 15:30 | **Cleaning data for use and sharing** | Charles Boniface |
| 15:30 – 15:45 | Break | |
| 15:45 – 17:00 | **Exploratory data analysis** | Charles Boniface |
| 17:00 – 17:15 | **Closure of workshop** | |