Open Access and Open Data at CGIAR: challenges and solutions

Medha Devare, Megan Zandstra, Abby Clobridge, Michelle Fotsy, David Abreu, Elizabeth Arnaud, Paul Baraka, Enrico Bonaiuti, Srinivasarao Chukka, Ibnou Dieng, Kate Dreher, Sufiet Erlita, Henry Juarez, Soonho Kim, Jawoo Koo, Usman Muchlish, Martin Müller, Leroy Mwanzia, Jane Poole, and Salman Siddiqui with other data, knowledge, and information managers across CGIAR.[i]

Abstract

CGIAR is a global research partnership of 15 geographically and scientifically diverse Centers dedicated to reducing poverty, enhancing food and nutrition security, and improving natural resource management. The Centers are charged with accelerating innovation to tackle challenges at a variety of scales from the local to the global. This requires data to be findable, accessible, interoperable, and reusable (FAIR) and inter-linked where relevant. CGIAR Centers have made strong progress in implementing publication and data repositories; however, many of these still represent silos whose contents are not generally easily discoverable or inter-linked (e.g., agronomic trial data with socioeconomic or adoption data in the same geographies). In the absence of such interoperability-mediated discovery, "open" is of limited utility. The overall goal is for CGIAR's trove of research data and associated information to be indexed and interlinked through a demand-driven cyberinfrastructure for agriculture, ensuring that research outputs are discoverable by humans and machines, and reusable via appropriate licensing to enhance innovation, uptake and impact. There are challenges to achieving this goal, not only across CGIAR, but for the agricultural domain in general. Among the foremost hurdles is that "open" tends to remain an unfunded mandate, making it difficult to operationalize effectively. Further, there is still significant concern on the part of scientists about making data open – largely centered around issues of trust, time, and quality – resulting in repositories frequently exposing metadata rather than the data sets themselves. While the ability to find metadata about resources qualifies as improvement, it continues to impose barriers to data access, discoverability, integration, and analysis, without which complex challenges to global agriculture development cannot be effectively addressed. CGIAR is addressing the urgent need to create a data sharing culture and enabling environment for Open Access

and Open Data (OA/OD) that includes projects planning for OA/OD and allocating funds to support it, in parallel with the technical infrastructure mentioned above. While the technology necessary to enable FAIR outputs exists, achieving success implies data provider and consumer trust and buy-in, agreement and adherence to interoperability standards and/or mapping across varied approaches, and compliance with guidelines (including those on citation and licensing governing content reuse).

Keywords: agricultural research; CGIAR; Open Access; Open Data; research partnerships

Introduction

CGIAR[ii] is a global research partnership of 15 Centers primarily located in developing countries (see Table 1), working in the agricultural research for development sector. Research at these Centers is focused on poverty reduction, enhancing food and nutrition security, and improving natural resource management to address key development challenges. It is conducted in close collaboration with local partner entities, including national and regional research institutes, civil society organizations, academia, development organizations, and the private sector. Thus, the CGIAR system is charged with tackling challenges at a variety of scales from the local to the global; research outputs are however often not easily discoverable and "dark data" is common (Heidorn, 2008), often residing on individual laptops, not being well described, indexed, or stored to be accessible and usable by the wider scientific community.

Innovating in this space and enhancing research impact increasingly depends upon enabling the discovery of, unrestricted access to, and effective reuse of the publications and data generated as research outputs by Center scientists. CGIAR Centers have made strong progress implementing publication and data repositories that meet minimum interoperability standards; however, work is still ongoing to enable consistent and seamless information discovery, integration, and interoperability between related outputs. Echoing Allemang and Teegarden (2016), Centers are therefore working to create an enabling environment for open access to research outputs, including a system-wide CGIAR Open Access and Data Management Policy implemented in 2013 (CGIAR, 2013), and in parallel with new technical infrastructure to demonstrate the full value of CGIAR research and accelerate new kinds of analyses and discovery. An Open Access, Open Data (OA/OD) initiative funded by the Bill and Melinda Gates Foundation (led from the CGIAR System Office in Montpellier, France) was tasked with coordinating implementation across all Centers, and tackled foundational challenges towards this goal by identifying key needs, and addressing technical and cultural issues.

Table 1: CGIAR Centers and headquarter locations

Center	Headquarters
Africa Rice Center (AfricaRice) www.africarice.org	Ivory Coast
Bioversity International www.bioversityinternational.org	Italy
International Center for Tropical Agriculture (known by its Spanish acronym CIAT for *Centro Internacional de Agricultura Tropical*) www.ciat.cgiar.org	Colombia
Center for International Forestry Research (CIFOR) www.cifor.org	Indonesia
International Maize and Wheat Improvement Center (known by its Spanish acronym CIMMYT for *Centro Internacional de Mejoramiento de Maíz y Trigo*) www.cimmyt.org	Mexico
International Potato Center (known by its Spanish acronym CIP for *Centro Internacional de la Papa*) www.cipotato.org	Peru
International Center for Agricultural Research in the Dry Areas (ICARDA) www.icarda.org	Lebanon
International Crops Research Institute for the Semi-Arid Tropics (ICRISAT) www.icrisat.org	India
International Food Policy Research Institute (IFPRI) www.ifpri.org	USA
International Institute of Tropical Agriculture (IITA) www.iita.org	Nigeria
International Livestock Research Institute (ILRI) www.ilri.org	Kenya
International Rice Research Institute (IRRI) www.irri.org	Philippines
International Water Management Institute (IWMI) www.iwmi.cgiar.org	Sri Lanka
World Agroforestry Centre (previously known as the International Centre for Research in Agroforestry, ICRAF) www.worldagroforestrycentre.org	Kenya
WorldFish www.worldfishcenter.org	Malaysia

There are several hurdles to achieving these "open" goals at CGIAR and elsewhere, foremost of which is the fact that "open" is still largely an unfunded mandate. In the context of research publication, institutions struggle to meet the costs associated with author-pays OA publication models (Bonaccorso et al., 2014). Shifting publication costs from readers to authors lightens library budgets, but adds to the costs of research projects. Resources must be made available to either pay OA article processing charges (APCs), or incentivize authors to support free OA journals, foregoing more prestigious "hybrid" publications (Van Noorden, 2013).

Costs and buy-in associated with building and sustaining institutional repository programs is a challenge faced by universities and research institutions globally, particularly when budgets for such tools do not exist (Burns, Lana & Budd, 2013). In the CGIAR context, research publications and data are often co-produced among Centers and research programs. There is therefore an imperative to enable interoperability among CGIAR repositories, to maximize discoverability among co-producers, and to support open goals as a team. This requires both enabling technology (Chan, 2004) and standardized metadata schemas (Park & Tosaka, 2010) across all Centers. Among Centers with diverse mandates, budgets, and pre-existing data and information management infrastructures, this poses a significant challenge, requiring extensive collaboration.

Understanding the gaps and challenges that CGIAR's 15 Centers grapple with was an important first step in its strategy to support OA/OD, accomplished via an online survey and informal feedback collected through regional OA implementation workshops, both conducted in 2015. A follow-up focus group brainstorming session centered on issues related to OD in 2016. It was clear from the OA/OD assessment that almost every Center was taking steps towards making publications and data open, and towards better documenting and curation of other research outputs. However, it was also evident that momentum could be gained through a consistent approach to human and technical infrastructure, policies, standards, and interoperability, among other issues. In order to ensure that the OA/OD "scaffolding" being erected across CGIAR could interlink and pay off in the long term, it was critical the OA/OD initiative provide coordinated support to Centers and the cross-cutting CGIAR Research Programs (CRPs) through which they collaborate.

The assessment and feedback indicated that there is still significant concern among scientists about making data open, largely centered around issues of trust, time, and quality. These types of concerns are echoed in the literature (Borgman, 2012; Mauthner & Parry, 2013). Researchers must be able to rely upon institutional assurances that their data will be shared in accordance with ethical and legal guidelines. They must also grow accustomed to sharing their data openly, facing potential scrutiny associated with data quality and accepting that other researchers may

publish off data which may have taken years to collect (Tenopir et al., 2011).

The technology necessary to overcome access and integration issues exists (Kitchin, 2014); these concerns can be addressed by enabling the harvesting of content from diverse platforms, repositories, and databases; integrating different types of resources and disciplines in meaningful ways; seeing patterns in and mining data ("big" and small); and seamlessly leveraging it for visualization, analysis, and support in decision-making processes. However, achieving success implies data provider and consumer trust and buy-in to a sharing culture, as well as agreement and adherence to standards for metadata, vocabularies, and data itself (and/or mapping across varied approaches), and compliance with guidelines (including those on citation and licensing governing content reuse). It is therefore critical that CGIAR's solutions for successful OA/OD address these issues, make progress on adopting common standards, and build cross-walks across differing schemas. Through the OA/OD initiative, and the ambitious new Big Data and ICT Platform (2017), CGIAR is developing the infrastructure and capacity that will ensure that research outputs are FAIR – findable, accessible, interoperable and re-usable (Wilkinson et al., 2016) – enhancing CGIAR's impact through data sharing and reuse to generate new and innovative solutions to global agricultural challenges. This infrastructure will initially be piloted within the CGIAR system and harvest only CGIAR Center repositories and relevant platforms before being made available for global use. Lessons will be drawn from Allemang and Teegarden (2016), who highlight ways in which a data-sharing infrastructure might be built for agriculture by focusing on stakeholder engagement, data sourcing and handling, and sharing and collaboration frameworks. The infrastructure will be extended to the agricultural research domain in general following the CGIAR pilot phase.

This paper will outline the approach used to identify the specific contexts and challenges faced by CGIAR Centers in the pursuit of OA/OD for development. A comprehensive survey, regional workshops and a brainstorming session were used to ascertain the environments under which CGIAR Centers manage their data and information resources. The results of these approaches were then synthesized in order to identify ways in which Centers may move forward as a group to achieve open access goals.

Method

Effective implementation of OA/OD across CGIAR Centers required a clear understanding of the needs and capacity of each of the system's 15 Centers with respect to managing research outputs for openness. A broad and multi-faceted online survey was developed and administered in 2015 with the help of a knowledge management consultant using the SurveyMonkey online

survey and analysis tool. The purpose of this survey was not to conduct a research study, but to provide a snapshot of needs and gaps in terms of how information from different data streams was being managed. Understanding this will clarify key points of intervention in implementing OA/OD across CGIAR and other organizations and academic institutions (some of whom have expressed interest in learning from our experiences). For the CGIAR system, data types typically include, but are not limited to, agronomic, breeding, genomic, natural resources management, geospatial, and/or socioeconomic disciplines. One representative for each of the fourteen responding Centers and two (out of fifteen) CRPs completed this 58-question survey in consultation with relevant Center colleagues, which provided an indication of: the OA publications landscape across CGIAR; data management and quality assurance practices; how other research outputs are handled; and gaps and needs in human and technical resourcing and in providing enabling environments for OA/OD.

Another intervention towards effective OA/OD across the distributed CGIAR system involved the organization of three regional workshops focused on using an implementation plan template[iii] developed by the CGIAR System Office[iv] as the starting point for consistent OA/OD operationalization, including budgeting. These workshops were held in Kathmandu, Nepal, for Centers in Asia; and at two of the Centers – the International Livestock Research Institute (ILRI) in Addis Ababa, Ethiopia, for those in Africa, the Middle East, and Europe; and the International Center for Tropical Agriculture (CIAT) in Cali, Colombia. Three to four data, information, communications, knowledge specialists, and/or legal/Intellectual Property (IP) personnel attended on behalf of each Center, and were generally able to leave the workshops with a draft implementation plan in hand.

A brainstorming session was held in mid-2016 as a follow-up for CGIAR data managers to engage on specific issues. Each Center has at least one data manager who also serves as a focal point for the OA/OD and big data initiatives through their participation in CGIAR's Data Management Task Force (DMTF). For the brainstorming exercise, the twelve data managers in attendance at the 2016 DMTF annual meeting were asked to respond individually on post-it notes to questions in three areas pertaining to open data: (1) aspirations, (2) challenges, and (3) solutions. These notes were then posted on walls and clustered by the group into broader themes. Specific questions formulated by the OA/OD Initiative Lead at the CGIAR System Office for each of the three areas were as follow:

Aspirations
● Where do you see CGIAR and your Center in the next 3–5 years in terms of Open Data, and in terms of being able to leverage big data capabilities?
● What can we realistically accomplish in three years? (The caveat here being that the

CGIAR Open Access and Data Management [OADM] Policy [CGIAR, 2013] goes into full effect in 2018, and that Centers have varying levels of capacity and budget to comply with this objective.)

Challenges
● What are the top 3–5 challenges in achieving your aspirations?
● What are some stumbling blocks?

Solutions
● What are some solutions to overcome these challenges – including possible actions from your Center, the CGIAR System Office, and others?

Key findings from the survey and the brainstorming exercise are presented below.

Results

Results from the survey, brainstorming, and anecdotal evidence from the regional workshops suggest that Centers are operationalizing OA/OD in different ways, demonstrated by the wide range of approaches, priorities, and workflows. The open-ended questions at the end of the survey elicited vastly different, but useful reactions. Respondents were assured confidentiality; Center identities have therefore been masked in this paper.

Policies, plans and workflows
● Eight Centers reported having an OA/publications policy separate from the CGIAR OADM Policy (CGIAR, 2013), while ten Centers reported having a separate OD/data management policy.
● Of those Centers with separate publications and/or data policies, only two explicitly stated timelines in line with the OADM Policy. One Center with a 2012 (pre-CGIAR) policy is currently revising the document, and expects to implement timelines consistent with the OADM Policy.
● All responding Centers and CRPs indicated that their CRP personnel and partners are expected to comply with whichever OA/OD policy the Center uses, indicating good uptake of CGIAR's OA/OD policy in general.
● Nearly all of the survey respondents (13) indicated they were using or planned to use the implementation plan template provided by the CGIAR System Office as a starting point to guide OA/OD operationalization. Three other respondents indicated they were unsure. These responses are consistent with the plans that have been shared with the System Office up to the time of

writing.

● Workflows, capacity, and day-to-day responsibility for OA/OD operations varied widely among CGIAR Centers and CRPs. Libraries, knowledge centers, and research data management/research support units are involved in routine workflows for publications. For data, some Centers have a dedicated data management or similar unit, while others rely on their library or knowledge management (KM) team for support. A few Centers referred to a geographic information system (GIS) specialist or unit for support in GIS data streams. Communications departments were mentioned by some Centers/CRPs, and a few Centers highlighted support by their legal team or IP focal point.

● Nearly all respondents indicated that they have data management workflows in place or under development. Of the two respondents indicating otherwise, one of the two had begun to develop guidelines by the time of writing.

To summarize, the OA/OD initiative has contributed to general alignment with the CGIAR OADM policy by Centers. While there are similarities in the policies, plans and workflows employed by Centers and CRPs, it is clear that cross-cutting support across an entity as complex and spread out as CGIAR requires both flexibility and breadth in order to achieve OA/OD goals.

Open Access to publications and publication repositories

● The percentage of peer-reviewed publications in Center repositories that are Open Access without restriction, primarily via Creative Commons Attribution (CC BY) licenses, is growing. Two Centers reported that over three quarters of the peer-reviewed publications within their repositories were fully downloadable without restriction. Most Centers/CRPs indicated that around half (40-60%) of the peer-reviewed publications in their repositories are fully downloadable without restriction. The remaining publications are available as metadata records only, with access provided by article publishers upon payment only.

● Most Centers are moving towards DSpace or other standards-compliant, interoperable OA repository platforms such as EPrints, Invenio, CONTENTdm, or Koha.

● Reasons for selecting particular repository platforms varied, but open source, good interoperability, and adoption by other Centers were cited as the top reasons (Figure 1).

● The Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH) is the primary interoperability protocol in use by Center publication repositories. OAI-PMH works with DSpace and Dataverse, the two main publication and data repositories across Centers.

● Nearly all Centers/CRPs report using the Dublin Core metadata standard and/or plan to implement/map to the recently released CG Core metadata schema v.1.0[v].
These results suggest that Centers have achieved varying levels of 'openness' with regard to their research publications, but that they aspire to many common goals, tools and standards is encouraging.

Open Data and Data Repositories

● Nearly all of the responding Centers/CRPs indicated that they had a data repository in place or in development; several indicated having more than one repository (e.g. for geospatial vs. other kinds of data).

● Most Centers have adopted Dataverse as the platform for at least some of their repositories. Two others are using CKAN. "Good interoperability" was the most-cited reason for selecting a particular system, with "open source" being selected with only slightly lower frequency (see Figure 2).

● In terms of data streams, agronomy data, socioeconomic data, and plant breeding data were reported as having more than 1000 openly accessible data sets across responding CGIAR Centers/CRPs; this may, however, be misleading as just one Center is typically responsible for a majority of the data sets in any particular discipline, and the quality and reusability of the data and associated annotation can be highly variable. For instance, 85% of agronomy data sets were in the AgTrials (CGIAR Research Program on Climate Change and Agricultural Food Security [CCAFS], 2013) repository, and a majority of the plant breeding data reported was also from one Center repository (see Figure 3).

● The responsibility for upload/deposit of data sets was more or less evenly split between data managers and scientists. Workflows to streamline and ease deposit and license choice are being developed at most Centers. While Centers typically try to use the least restrictive CC licenses, in some instances (e.g. data on germplasm development) more restrictive or custom licenses are employed.

● Practices for data quality and/or data cleaning were not consistent among Centers, and responsibilities for these also varied (see Figure 4). Several Centers indicated that no one was specifically responsible for assessing data quality or cleaning data. Five other Centers indicated that researchers, project leaders, or science theme leaders were responsible, while six indicated this as the responsibility of a specific unit.

● OAI-PMH was referenced by seven of the responding Centers/CRPs as being in use with their data repositories. Likewise, most indicated that Dublin Core and/or the CG Core metadata schema based on Dublin Core would be adopted.

● Several respondents indicated that GitHub was the repository of choice for applications and software.

Variations in data management practices and data repository use among Centers poses a significant challenge. Although 'interoperability' is cited as a primary goal for the implementation of data repositories among Centers, the current state of data management support and open data implementation demonstrates that CGIAR-wide interoperability goals will likely require stronger advocacy and particularly innovative approaches.

Current Practices and Culture

● Only one Center indicated having a centralised OA fund to pay article processing charges (APCs) when publishing in OA journals. Another noted that APCs are split between the Center's KM unit and research divisions. Two Centers indicated in the survey – as others have stated in their OA/OD Implementation Plans – that researchers were being encouraged to incorporate budgets for OA in new project proposals (anecdotal evidence suggests however that this is sometimes met with resistance, pointing to a need for advocacy).

● While researchers are being encouraged to budget for OA in new project proposals, only a few Centers indicated that OA/OD data and information managers were being consulted during project planning to ensure consistent implementation.

● Responses were quite varied to questions around researcher awareness of and compliance with OA/OD requirements. Many Centers indicated that they hoped to incorporate OA/OD into researcher performance evaluation and work plans in the coming years, but taking concrete steps towards this goal will involve leadership buy-in. Several Centers also suggested that support, education/advocacy, and incentives would help the organization move forward.

Opportunities and concerns: technology, culture, and budget

Common concerns cited were:

● Lack of funding, capacity, process -- encompassing staff time, staff resources, and infrastructure.

● Lack of Open Access culture/attitude/awareness among researchers.

● Lack of researcher incentives, and OA/OD not being part of researchers' performance evaluation criteria.

Several Centers/CRPs offered suggestions and ideas for raising awareness about and incorporating OA/OD into the CGIAR organizational culture:

● Improve Center capacity to support OA/OD – financial support needed so this is no longer an unfunded mandate.

● Work with senior leadership to ensure buy-in of Open Access awareness and implementation strategies.

● Conduct training/workshops and webinars with Centers and partner researchers ("data clubs" and "data talks" to raise awareness, advocate for, and provide the "how to").

● Develop easy-to-use guidelines, templates, workflows, and informational materials to support adoption of OA/OD practices.

● Survey researchers to understand their needs/gaps relating to OA/OD.

● Work with human resources teams to conduct research data management sessions during staff induction.

● Recognize and reward scientists who publish in OA journals and share data (e.g. via social media) and nominate OA/OD champions as speakers during Science Week, OA/OD week, and other events.

● Include the OADM Policy in CRP branding materials and guidelines.

● Include OA/OD (e.g. data collection and management) in work plans and proposals, and ensure accountability via scientist evaluations; consider penalties for non-compliance,

● Include published data (and not just publications) in performance reviews.

● Compile and share OA/OD success stories as accompaniment or follow-ups to training and communication efforts to motivate researchers.

Many of the issues in the 2015 survey were also raised by CGIAR data managers in attendance at the 2016 annual meeting, although the targeted nature of the brainstorming resulted in more detailed exploration of key areas in managing open data across CGIAR. Figure 5 provides the results of the clustering of ideas arising from the brainstorming around aspirations, challenges, and solutions related to open data at CGIAR.

Discussion and Conclusions

Centers were split between having their own OA/OD policies and using the CGIAR OADM Policy (CGIAR, 2013). Anecdotal evidence from regional workshops indicates that those with separate OA/OD policies opted for this option in order to gain stronger buy-in from Center leadership and/or researchers; other Centers indicated the CGIAR OADM policy was sufficient. Potential discrepancies between Center and CGIAR policies are being addressed through dialogues with individual Centers and advocacy activity. While workflows, capacity, and responsibilities for OA/OD operationalization vary among CGIAR Centers, there is evidence that these elements are largely present, particularly for data management. This evidence simplifies appropriate intervention, support provision, and compliance tracking.

In recognition of Centers' expressed need for shared tools, best practice guidelines, and examples to assist with OA/OD implementation, an OA/OD Support Pack (CGIAR, 2015) has been developed as a resource, which brings together technical and cultural resources, including exemplar policies, templates, workflows, guidance on licensing and publishing while retaining copyright, data management plans, metadata schemas, advocacy materials and talking points in support of OA/OD, terms of reference for OA/OD-related positions, and more. This resource currently exists as an open web-based platform, where documents and links can be added and

curated over time in response to Center needs. The need for a more attractive and user-friendly
design is being considered, and will likely be developed in 2017.

To increase momentum towards OA/OD implementation across CGIAR, each Center (as the
enduring entity and ultimate steward of resources) has developed or is developing a clear,
practical, and actionable Open Access and Data Management Implementation Plan which
complies with CGIAR's OADM Policy. The template for this implementation plan is being
widely employed, and encourages consistency among Centers. In addition to facilitating
collaborative work on the implementation plan, the brainstorming session and workshops
conducted as part of this initiative also enabled shared learning on a variety of issues. These
included interventions to incentivize OA/OD, ways to ensure that OA/OD managers and staff
were engaged throughout the project planning and budgeting cycle, and awareness-building and
advocacy approaches.

As reported in the results, each Center has standards-compliant publications and data repositories
with open source and good interoperability cited as the top reasons for choosing a particular
platform. The adoption of repositories responsive to industry standards and adoption of the CG
Core metadata schema make it possible for the planned CGIAR Big Data and ICT Platform
infrastructure to harvest and aggregate metadata and associated files from these repositories, as a
prelude to harvesting global platforms of relevance. Infrastructural work to improve the quality
and annotation of data, and inter-linkages across CGIAR tools and platforms is already
underway through substantial in-kind contributions from staff at several Centers. This work
includes the following activities:

● Development of a reference ontology for agronomic trial management which will enable
linkages between breeding and agronomy data, be available as linked open data, and be of added
value to a variety of applications and users. This activity is led through a sub-grant by a team
based at Bioversity, one of the CGIAR Centers, and includes active engagement of scientists
from CGIAR, Agricultural Model Intercomparison and Improvement Project (AgMIP)[vi], French
Agricultural Research Centre for International Development (CIRAD)[vii], and French National
Institute for Agricultural Research (INRA).[viii]

● The agronomy ontology will inform the agronomy field book that will address data
harmonization/standardization early in the data life cycle (i.e. at the collection stage). It is
envisioned that this will encourage agronomists at CGIAR and partner entities to create data
sheets and collect data with consistent metadata, terminology, scales, and methodologies; as well
as enabling them to store these temporarily in a database with view/edit permissions. Application
program interfaces (APIs) to easily deposit final field book data in key repositories is also
envisioned; these improvements will help address data quality issues at the point of collection.

● Through the OA/OD initiative, CGIAR has collaborated with and provided leadership on OA/OD issues relating to the agricultural domain, working closely with stakeholders such as the Food and Agricultural Organization (FAO) of the United Nations (on a road map for agri-semantics efforts), AgMIP (on data harmonization and interoperability), and the Global Open Data for Agriculture and Nutrition (GODAN) initiative (on impact assessment and advocacy efforts).

CGIAR is in the unique position of guiding and supporting 15 very diverse and geographically dispersed centers through a significant cultural and technical shift in the way publications and data are managed and shared. The OA/OD advocacy and implementation work described in this paper is expected to continue and be further developed, with the technical and cultural needs and gaps articulated by CGIAR Centers being addressed via activities proposed as part of the new CGIAR Big Data and ICT Platform (2017). Assuming the platform continues to receive anticipated levels of funding, OA/OD efforts across CGIAR will receive support, and it is envisaged that the platform—primarily by demonstrating a compelling value proposition for the sharing of resources—will assist in promoting a change in organizational culture as well as improving infrastructure to enable discovery, reuse, and innovative new analyses of CGIAR research outputs.

References

Allemang, D. and Teegarden, B. (2016). A global data ecosystem for agriculture and food. Global Open Data for Agriculture and Nutrition. CABI, UK.

Bonaccorso et al. (2014). Bottlenecks in the open-access system: Voices from around the globe. Journal of Librarianship and Scholarly Communication, 2(2), 1.

Borgman, C.L. (2012). The conundrum of sharing research data. Journal of the American Society for Information Science and Technology, 63(6), 1059-1078.

Burns, C.S., Lana, A., & Budd, J.M. (2013). Institutional repositories: exploration of costs and value. D-Lib Magazine, 19(1/2), 1-17.

CGIAR. (2013). CGIAR Open Access and Data Management Policy. 2 October 2013. Available from http://hdl.handle.net/10947/4488.

CGIAR. (2015). Open Access & Open Data Support. Website. Available from https://sites.google.com/a/cgxchange.org/oad-support-pack/home.

CGIAR Big Data and ICT Platform. (2017). Leveraging CGIAR data: Bringing Big Data to agriculture and agriculture to Big Data. Available from http://library.cgiar.org/bitstream/handle/10947/4450/2.%20Big%20Data%20platform%20 CGIAR%20Resubmission.pdf.

CGIAR Research Program on Climate Change and Agricultural Food Security (CCAFS). (2013). AgTrials: Global Agricultural Trial Repository and Database. Available from http://www.agtrials.org/. Cali, Colombia.

Chan, L. (2004). Supporting and enhancing scholarship in the digital age: the role of open access institutional repository. Canadian Journal of Communication, 29(3).

Heidorn, P.B. (2008). Shedding light on the dark data in the long tail of science. Library Trends 57(2): 280-299.

Kitchin, R. (2014). The data revolution: Big data, open data, data infrastructures and their consequences. Sage.

Mauthner, N.S., & Parry, O. (2013). Open Access digital data sharing: Principles, policies and practices. *Social Epistemology*, 27(1), 47-67.

Park, J.R., & Tosaka, Y. (2010). Metadata creation practices in digital repositories and collections: Schemata, selection criteria, and interoperability. Information Technology and Libraries, 29(3), 104.

Tenopir, C., Allard, S., Douglass, K., Aydinoglu, A. U., Wu, L., Read, E., ... & Frame, M. (2011). Data sharing by scientists: practices and perceptions. PloS one, 6(6), e21101.

Van Noorden, R. (2013). The true cost of science publishing. Nature, 495(7442), 426.

Wilkinson, M.D., Dumontier, M., Aalbersberg, I.J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.W., da Silva Santos, L.B., Bourne, P.E. and Bouwman, J. (2016). The FAIR Guiding Principles for scientific data management and stewardship. Scientific data, 3.

About the author(s)

David Abreu, International Center for Tropical Agriculture (CIAT), Cali, Colombia. Email: d.abreu@cgiar.org

Luz Marina Alvare, International Food Policy Research Institute (IFPRI), Washington, DC, USA. Email: l.alvare@ifpri.org

Elizabeth Arnaud, Bioversity International, Montpellier, France. Email: e.arnaud@cgiar.org

Peter Ballantyne, International Livestock Research Institute (ILRI), Addis Ababa, Ethiopia. Email: p.ballantyne@cgiar.org

Paul Baraka, World Agroforestry Centre (ICRAF), Nairobi, Kenya. Email: p.baraka@cgiar.org

Marc Bernard, Africa Rice Center (AfricaRice), Abidjan, Côte d'Ivoire. Email: m.bernard@cgiar.org

Enrico Bonaiuti, International Center for Agricultural Research in the Dry Areas (ICARDA), Amman, Jordan. Email: e.bonaiuti@cgiar.org

Srinivasarao Chukka, International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Hyderabad, India. Email: s.chukka@cgiar.org

Abby Clobridge, FireOak Strategies (formerly Clobridge Consulting), Durham, NC, USA. Email: aclobridge@fireoakstrategies.com

Lea Delos Reyes, International Rice Institue (IRRI), Los Baños, Philippines. Email: l.delosreyes@irri.org

Medha Devare, CGIAR System Management Organization, Montpellier, France. Email: m.devare@cgiar.org

Ibnou Dieng, Africa Rice Center (AfricaRice), Abidjan, Côte d'Ivoire. Email: i.dieng@cgiar.org

Kate Dreher, International Maize and Wheat Improvement Center (CIMMYT), El Batán,

Mexico. Email: k.dreher@cgiar.org

Sufiet Erlita, Center for International Forestry Research (CIFOR), Bogor, Indonesia. Email:
s.erlita@cgiar.org

Cecilia Ferreyra, International Potato Center (CIP), Lima, Peru. Email: c.ferreyra@cgiar.org

Michelle Fotsy, CGIAR System Management Organization, Montpellier, France. Email:
m.fotsy@cgiar.org

Maria Garruccio, Bioversity International, Roma, Italy. Email: m.garruccio@cgiar.org

Chandima Gunadasa, International Water Management Institute (IWMI), Colombo, Sri Lanka.
Email: c.gunadasa@cgiar.org

Jesús Herrera de la Cruz, International Maize and Wheat Improvement Center (CIMMYT), El
Batán, Mexico. Email: j.herrera@cgiar.org

Henry Juárez, International Potato Center (CIP), Lima, Peru. Email: h.juárez@cgiar.org

Suan Pheng Kam, WorldFish, Penang, Malaysia. Email: s.kam@cgiar.org

Soonho Kim, International Food Policy Research Institute (IFPRI), Washington, DC, USA.
Email: s.kim@cgiar.org

Jawoo Koo, International Food Policy Research Institute (IFPRI), Washington, DC, USA. Email:
j.koo@cgiar.org

Vika Kovariansi, Center for International Forestry Research (CIFOR), Bogor, Indonesia. Email:
v.kovariansi@cgiar.org

Martina Mascarenhas, International Center for Tropical Agriculture (CIAT), Cali, Colombia.
Email: m.mascarenhas@cgiar.org

Usman Muchlish, Center for International Forestry Research (CIFOR), Bogor, Indonesia. Email:
u.muchlish@cgiar.org

Martin Mueller, International Institute of Tropical Agriculture (IITA), Ibadan, Nigeria. Email:
m.mueller@cgiar.org

Jacqueline Muliro, World Agroforestry Centre (ICRAF), Nairobi, Kenya. Email:
j.n.muliro@cgiar.org

Leroy Mwanzia, International Center for Tropical Agriculture (CIAT), Cali, Colombia. Email:
l.mwanzia@cgiar.org

Kenneth Oraegbunam, International Institute of Tropical Agriculture (IITA), Ibadan, Nigeria.
Email: k.oraegbunam@cgiar.org

Alan Orth, International Livestock Research Institute (ILRI), Nairobi, Kenya. Email:
a.orth@cgiar.org

Amalia Perochena, International Potato Center (CIP), Lima, Peru. Email: a.perochena@cgiar.org

Jane Poole, International Livestock Research Institute (ILRI), Nairobi, Kenya. Email:
j.poole@cgiar.org

Carlos Quiros, International Livestock Research Institute (ILRI), Nairobi, Kenya. Email:
c.f.quiros@cgiar.org

Kotnana Ramesh, International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Hyderabad, India. Email: k.ramesh@cgiar.org

Joel Ranck, International Potato Center (CIP), Lima, Peru. Email: j.ranck@cgiar.org

Itzel Saldívar, International Maize and Wheat Improvement Center (CIMMYT), El Batán, Mexico. Email: i.saldivar@cgiar.org

Moh. Agus Salim, Center for International Forestry Research (CIFOR), Bogor, Indonesia. Email: m.salim@cgiar.org

Trushar Shah, International Institute of Tropical Agriculture (IITA), Ibadan, Nigeria. Email: tm.shah@cgiar.org

Salman Siddiqui, International Water Management Institute (IWMI), Colombo, Sri Lanka. Email: s.siddiqui@cgiar.org

Simone Staiger Rivas, International Center for Tropical Agriculture (CIAT), Cali, Colombia. Email: s.staiger@cgiar.org

Stanley Tan, WorldFish, Penang, Malaysia. Email: s.tan@cgiar.org

Marimagne Tchamba, International Institute of Tropical Agriculture (IITA), Ibadan, Nigeria. Email: m.tchamba@cgiar.org

Marco van den Berg, International Rice Institue (IRRI), Los Baños, Philippines. Email: m.vandenberg@irri.org

Indira Yerramareddy, International Food Policy Research Institute (IFPRI), Washington, DC, USA. Email: i.yerramareddy@cgiar.org

Megan Zandstra, International Center for Tropical Agriculture (CIAT), Cali, Colombia. Email: m.zandstra@cgiar.org

Thomas Zschocke, World Agroforestry Centre (ICRAF), Nairobi, Kenya. Email: t.zschocke@cgiar.org

---

[i] See "About the authors"

[ii] http://www.cgiar.org/about-us/

[iii] http://bit.ly/2oe1CGN

[iv] The System Office is the CGIAR management/administrative entity.

[v] https://docs.google.com/viewer?a=v&pid=sites&srcid=Y2d4Y2hhbmdlLm9yZ3xvYWQtc3VwcG9ydC1wYWNrfGd4OjRjRjNDQ4NTg3OTU5N2YxZTM

[vi] http://www.agmip.org/

[vii] http://www.cirad.fr/en/home-page

[viii] http://www.inra.fr/en