

Open Access Management
PROPOSED STRATEGY for Phase II of the
CGIAR Research Program on Climate Change, Agriculture and Food Security (CCAFS)
Submitted to CGIAR Consortium 31 March 2016

Planning for and implementing OA/OD

In accordance with the [CGIAR Open Access and Data Management Policy \(OADM\)](#), CCAFS is mandated to produce international public goods and ensure that they are open via FAIR principles – that is, they are Findable, Accessible, Interoperable and Re-usable to enhance innovation, impact, and uptake. CCAFS has developed its [Data Management Strategy \(DMS\)](#) to enable the program to fulfil its obligations with respect to making information and data products supporting documentation from its research globally available.

CCAFS aims to providing a “one-stop-shop” for its information and data products generated and expects to attract data contributions from scientists working in related areas even if not directly funded by CCAFS. With development outcomes in mind, the program will increase accessibility, visibility and usability of scientific outputs by a global community.

Goal

The goal of CCAFS Data Management Strategy (DMS) is for CCAFS information and data products to be available for long-term use by partners and the scientific community.

Guiding principles

In defining the DMS, we have adopted the following principles:

- It adheres to internationally recognized standards such as FAIR principles, Dublin Core metadata standards and interoperability protocols.
- It has to be easy to implement and any burden to researchers that is generated from its implementation must be balanced by the benefits that the researcher will get from making information and data products available, and by the support that CCAFS will provide.
- It should not affect the autonomy of scientists to carry out their research; the strategy ensures the independence and creativity of scientists in the collection of data that is relevant to the research objectives.
- Ethical use and sharing of personal and private data, and visibility and credits to data generators need to be ensured, like accessibility and adherence to international standards for data storage.

Objectives

Overall, CCAFS DMS guides the creation of an enabling environment where scientists and partners produce and share high quality data outputs, while at the same time enables a variety of data management procedures and practices at project level. Through creating “portals” specifically designed for common types of datasets scientists can publish their data. CCAFS supports scientists and CGIAR Centres to produce well-managed and documented information and data products that are easy to use long-term.

Critical issues and anticipated challenges

Issues that are critical to the implementation of CCAFS DMS include:

- Establishing a process - A clear process for data sharing and management must be established, from legal agreements through to operating and reporting principles. This

conveyor belt is implemented by CCAFS through an online ICT planning and reporting system that identifies the information and data products that are being generated, and ensures that products are made publicly available within the timeframes agreed upon with partners.

- Supporting compliance - Support and encourage the design and implementation of data management plans and repositories that enable projects to comply with the CGIAR OADM Policy.
- Enabling a data culture - Implementing this strategy requires a significant cultural shift among program participants. Appropriate incentives and penalties should be established to promote data sharing. Metrics on data sharing from each program participant should be used as a criterion for performance evaluation and rewards or penalties.
- Among the conditions to facilitate the establishment of a more conducive data culture CCAFS will:
 - ✓ Support program partners in the process of submitting data to suitable repositories;
 - ✓ Work with existing CCAFS repositories to enable interoperability;
 - ✓ Highlight benefits to researchers to be derived from data sharing such as increased visibility, potential for increased collaboration and publication, and reputation; and
 - ✓ Make available statistics about data downloading and use for evaluation, program planning, informing the CCAFS research agenda and promoting our scientific partners.

Technical considerations and operations

Scope

CCAFS DMS looks at making “Data+” available in public archives while research outputs such as papers and publications resulting from analysis of primary data are shared through alternative portals.

In the CCAFS DMS we use the term “Data+” to indicate the actual data generated by the research process once it has been cleaned and is considered of good quality, and the documentation that will enable the use of these datasets in the future. This includes but is not restricted to documents about the methodology for data generation/collection, computer programs used for data processing, quality assessment, and any metadata that helps building a description of the context in which the data has been originated.

The operational objectives of CCAFS DMS are:

1. Guide CCAFS in designing and implementing support mechanisms to reach the goal of the DMS;
2. Make available high quality Data+ to potential users now and well into the future;
3. Encourage appropriate levels of standardization, adoption of international standards and harmonization so that data from separate research activities can be brought together to enrich our understanding of processes, outcomes and impacts in the areas of the world where CCAFS works;
4. Procedures for maintenance and backup includes long-term preservation of data in standardized accessible data formats, converting formats when needed due to software upgrades or changes, safe-keeping of data in a secure environment with the ability to control access where required. In addition, CCAFS follows regular data back-ups, to ensure that information products are maintained, curated, and archived “into perpetuity”.
5. Promote the production of FAIR outputs:
 - a. **Findable:** Data and metadata should be richly described to enable attribute-based search;
 - b. **Accessible:** Data and metadata should be retrievable in a variety of formats that are sensible

to humans and machines using persistent identifiers;

- c. Interoperable:** The description of metadata elements should follow community guidelines that use an open, well defined vocabulary; and
- d. Reusable:** The description of essential, recommended, and optional metadata elements should be machine processable and verifiable, use should be easy and data should be citable to sustain data sharing and recognize the value of data.

Supporting mechanisms

Supporting mechanisms will be necessary for the implementation of this strategy. These include:

1. Through the Data Management Support Pack, CCAFS provides guidelines for making data available in such a way as to respect the trust that information providers have placed in CCAFS scientists;
2. Creating, maintaining and supporting portals that make data publication easy when CCAFS considers it necessary (see Annex Table 9);
3. The Program Participant Agreements (PPA) established with CGIAR Centres and other partners stipulate that data is to be made freely available and set up the timeframes for data publishing by scientists involved in CCAFS research activities:

"The Contracted Party agrees to publicly share any data and/or models generated as a result of activities under this Agreement through CCAFS's data portals as soon as practically possible, but no later than twelve (12) months of generation for meta data and twenty four (24) months for other data and/or models. Such data portals include, but are not limited to, the CCAFS agricultural trial data repository (www.agtrials.org), the Adaptation and Mitigation Knowledge Network (www.amkn.org), the CCAFS climate data portal (www.ccafs-climate.org), the CCAFS Research Data on Dataverse (dataverse.harvard.edu/dataverse/CCAFSbaseline) and the repository of Agricultural Research Outputs (www.cgspace.cgiar.org). Access to the data should be fully granted to the CCAFS Knowledge and Data Sharing unit at CIAT."

Annex Table 9. Identification of repository or platform housing information products for indicative datatypes

Indicative Data type	Repository or Platform		Interoperability Protocol ^(*)
	Name/s	URL/s	
Climate data (incl. downscaled GCM data - condensed and disaggregated spatial data, useful documents and links).	CCAFS Climate	http://ccafs-climate.org/	OAI-PMH implemented
Agricultural trials (database on the performance of agricultural technologies at sites across the developing world)	AgTrials	http://agtrials.org/	OAI-PMH implemented
Climate projection models (to identify sites whose current climate realities are similar to the possible future climates of a reference location)	CCAFS Analogues	http://www.ccafs-analogues.org/	OAI-PMH to be implemented
Socioeconomic (incl. CCAFS village	CCAFS Dataverse	https://dataverse.	OAI-PMH /

baseline study, CCAFS organizational baseline study, gender household survey, etc)		harvard.edu/dataverse/CCAFSbaseline	SWORD implemented
Agricultural Crops (incl. environmental classifications and stress patterns of growing environments for agricultural crops)	TPEI	http://www.ccafs-tpe.org/tpe/	OAI-PMH to be implemented
Publications, Data & Tools (incl. reports, journal articles, working papers, datasets, tools, etc.)	CGSpace	https://cgspace.cgiar.org/handle/10568/3530	OAI-PMH / SWORD implemented
One-Stop-Shop (incl. all information products)	CCAFS Website	https://ccafs.cgiar.org/	OAI-PMH to be implemented
Data information outputs reported by CCAFS researchers	CCAFS Planning & Reporting	https://activities.ccafs.cgiar.org/	OAI-PMH implemented

(*) Thus far, interoperability protocols are being implemented at metadata level and are envisioned for CCAFS Phase II to be implemented at data level.

Coordination and decision making

The development and implementation of the CCAFS DMS is coordinated by the CCAFS Knowledge and Data Sharing team at CIAT. Decisions are closely coordinated with the CCAFS PMC as well as the monitoring, evaluation and learning team (MEL). Strategies and other key decisions go to ISC for approval. To ensure the adherence to common standards CCAFS collaborates closely with the Knowledge and Data Management teams of CIAT, System Office, CGIAR Centres and key partners (e.g. University of Reading).

Strategic elements of coordination and decision making

In order to achieve the objectives set out above, at the program level CCAFS will:

- Discuss, define and adopt the CGIAR OPDM;
- Ensure the CCAFS DMS and the CCAFS Intellectual Asset Management Policy are aligned;
- Negotiate and coordinate actions with the System Office of CGIAR , as well as with CGIAR Centres that are part of CCAFS;
- Include the required elements of these policies into the contracts established with Participating Centres – e.g. CCAFS Program Participant Agreements (PPAs);
- An implementation plan needs to be put in place;
- Support and resource mechanisms to receive and archive data;
- Provide resources (e.g. guidelines, webinars, and manuals) to help strengthen the capacities of CCAFS staff, CGIAR Centres and partners in implementing open access open data standards.

In order to fulfil Participating Centre contractual obligations under the PPA agreements, CCAFS expects that the CGIAR Centres will do the following:

- Allocate sufficient resources to allow for the implementation of the CCAFS DMS;
- Utilize and help build the CCAFS DMS Support Pack as key mechanism for the implementation of the CCAFS DMS;
- Ensure their information and data products comply with the CGIAR OADM Policy.

Promoting CCAFS DMS Implementation

CCAFS will develop and implement an internal communications strategy that deals with a spectrum of users and activities, including:

- Training Open Access Ambassadors from CCAFS Flagships, regions and Centres: these will be focal points for ensuring that projects within their unit are complying with OADM guidelines and are also responsible for coordinating submissions to CGSpace and ensuring correct reporting within CCAFS portals. This group will act as a community of practice to share knowledge, lessons, and questions, and enable discussions on open access and open data with the wider network of CCAFS scientists.
- Presence on CCAFS intranet and website of open access guidelines, tips and tools: the intranet will have a section where the DMS implementation plan is summarized and essential links are provided. It will also link to an updated list of CIAT publications.
- Annual statistics report of CCAFS information and data products: this report, produced in January of each year, will include metrics of the top publications, analytics of websites, platforms and databases, and other insights. This will allow scientists to monitor use and uptake of their products, and guide decisions on value for money.
- Content from CCAFS repositories automatically fed to CCAFS website to showcase information and data products and raise awareness of the repository and its uses.
- Communications staff share new products and publications via social media and other channels in order to enhance dissemination and generate enthusiasm for information and data products.
- Participate in a CIAT-led community of practice of data managers and focal points for data management. Also participate in CGIAR-wide communities of practices on Knowledge Management, Open Access Implementation Working Group and Data Management Task Force.
- Inclusion of Open Access and Open Data targets in internal performance management indicators, to monitor progress on targets.
- Develop awareness amongst key CCAFS people of CCAFS policies and guidelines.

Narrative for required resources

CCAFS has dedicated resources and capacities for the implementation of the CGIAR OADM within CCAFS. These are brought in by the CCAFS knowledge and data sharing coordinator and a small team consisting of a data manager, two web developers (back-end) and an interface/ interoperability (front-end) expert facilitating the conceptual, operational and technical infrastructure. The team is supported by the communications and knowledge management unit, as well as the MEL team platforms. Moreover, CCAFS relies on data management staff in the CGIAR Centres, especially with respect to quality assurance, data curation, data standards and capacity building.

In addition to human resources, IT infrastructure is key to achieve good data management. The IT infrastructure that CCAFS has put in place and relies on is made up of a mix of data servers and cloud space.

CCAFS recognizes the need to establish an enabling environment where scientists and partners produce and share high quality data outputs, while at the same time making information and data products available for long-term use by partners and the scientific community

The implementation of CCAFS's DMS relies on all CCAFS units and requires, as mentioned above, a team with appropriate experience and skills. CCAFS will build as much as possible on CGIAR Centres' and partners' existing open access and open data policies and data quality assurance systems that are already in place, and is developing partnerships for this purpose.

A preliminary, indicative budget is shown in Annex Table 10.

Annex Table 10. Budget

Human, technical and other resources	Annual amount 2016 – 2017 (transition period)	Annual amount 2018+ (after 2 nd round of CRPs in effect)	Explanatory notes
Technology			
Data Repository	2,500	2,000	Above mentioned Repositories
Publications Repository	10,000	10,000	CGSpace annual support and maintenance fee
Hardware/storage (cloud etc)	1,000	1,500	Server Upgradation and Cloud - (AWS instances) (Server + Backup)
Bandwidth	-	-	Covered under staff costs
Programming/development	-	-	Covered under staff costs
Annual maintenance fees	2,000	2,000	Shared with CIAT
Website development related to repositories	-	-	Covered under staff costs
Operations and Travel	10,000	10,000	Operations and Travel
Staffing			
Staff salaries – Open Data Data Management	400,000	350,000	(**) Knowledge and Data Sharing Coordinator – 1 FTE Technical Manager – 1 FTE Data Manager - 0.5 FTE Interface/ interoperability (front-end) expert - 1 FTE Web developers (back-end) - 2 FTE
Data Quality/Curation	-	-	(**)
Staff salaries – Open Access publications Information management	-	-	Covered under data management above
Staff salaries – IP/Legal (in support of OA-OD)	-	-	(*)
Staff salaries – IT (in support of OA-OD)	-	-	(*)
Membership Fees			
Altmetrics provider(s)	-	-	
ORCID (unique researcher IDs)			
Publisher-based institutional memberships			
<Other>	50	50	Permanent Identifier Prefix – Handle
Other Expenses			
TOTAL	425,550	375,550	

(*) No budget since CCAFS relies on Lead Centre infrastructure

(**) Costs are complemented by the other GI - CRPs where the joint online ICT platform is concerned.