**CGIAR Research Program on
Climate Change, Agriculture and Food Security (CCAFS)**

CCAFS[1] - Data Management Strategy

June 2015

Contents

Introduction

CCAFS is mandated to producing international public goods and has developed this Data Management Strategy (DMS) to enable the programme to fulfil its obligations with respect to making data and the relevant supporting documentation from its research activities available to the world community.

The Program Participant Agreements (PPA) established with CGIAR centres and other partners stipulate that data is to be made freely available and sets up the time scales 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 metadata 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."

The aim of the Data Management Strategy (DMS) is to guide the creation of an enabling environment where scientists and partners are able to produce and share high quality data outputs throughout CCAFS, while at the same time enabling a variety of data management procedures and practices at project level. This is achieved through creating "portals" specifically designed for common types of data where scientists can publish their data and by the provision of guidance and support to scientists and CGIAR Centres to facilitate producing well-managed and documented datasets that are easy to use both now and in the future.

Guiding principles for this strategy are:

- accessibility,
- ease of use,
- ethical use and sharing of personal and private data,
- provision of support for data generators,
- ensuring that credit and visibility go to data generators,
- adherence to international standards for data storage.

CCAFS aims to providing a "one-stop shop" for data generated by its research activities and expects to attract data contributions from scientists working in related areas even if not directly managed or funded by CCAFS. It will increase accessibility and visibility of scientific outputs to a global community for adding even more value to the products of CCAFS research with development outcomes in mind.

In this strategy 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, as well as 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 collection/generation, computer programs used for data manipulation and data processing, data quality assessment, and any metadata that helps in building a description of the context in which the data has been originated.

In defining this strategy, we have adopted the following principles:

> *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 his/her data 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.*

Goal

The goal of this DMS is for CCAFS data products to be made available for long-term use by partners and the scientific community.

Objectives

The objectives of this strategy are as follows:

1. To guide CCAFS in designing and implementing support mechanisms to reach the goal;
2. To make available high quality *Data+* to potential users now and well into the future;
3. To 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; and
4. To 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.
   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.

Scope

This DMS looks at making *Data+* available in public archives. It does not include research outputs such as papers and publications resulting from analysis of primary data. CCAFS has created alternative portals to share this type of information.

Supporting mechanisms

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

1. Providing 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 consider it necessary;

Strategic Elements

Programme Level

In order to achieve the objectives set out above, the CCAFS programme needs to:

- Based on Consortium level policies, discuss, define and adopt a data sharing and data ownership policy and Intellectual Property policy;
- Negotiate and coordinate actions with the Consortium Office of CGIAR (CO), as well as Participating 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;

Centre Level

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

- Allocate sufficient resources to allow for the implementation of the DMS;
- Utilise the provided support package for the implementation of the DMS;
- Ensure their data research outputs comply with the CGIAR Open Access Policy

Implementation

Three key elements are essential to the implementation of this strategy:

1. Establishing a *process*

   A clear process for data sharing and management must be established, from legal agreements through to operating and reporting principles. CCAFS is implementing a planning and reporting system to enable the program to identify the data products that are to be generated and ensure that these products are made publicly available within the timeframes agreed with partners.

2. Supporting *compliance*

   Support and encourage the use of data repositories that enable projects to comply with the CGIAR Open Access Policy.

3. 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 criteria for measuring performance, reward or apply penalties. Among the conditions to facilitate the establishment of a more conductive data culture CCAFS must:

- 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
- Make available statistics about data downloading and use so as to be able to use this information as a planning tool for the programme to promote CCAFS's research agenda and that of our scientific partners among the global audience.