## CCAFS Data management guidelines

## Writing Good Research Protocols

This guide will suggest a systematic and complete recording of planning information, in the form of a protocol, to ensure that the project process – especially the analysis stage – can be efficient and effective, and that the legacy of the project does not languish unusable after the exercise has ended. This will include protocols for activities and at the project level. It will not prescribe a format for protocols but stress the importance of developing adequate protocols for any research activity.

## Data Ownership and Authorship

This guideline will look at issues of ownership and Intellectual Property. It will consider all parties who might have a claim to the data and include an example template of a data ownership agreement. The guide will also discuss authorship listing both the benefits and responsibilities associated with authorship. It stresses the need to establish ownership agreements right at the beginning of the research process and discusses some common difficulties in working with some specific types of partners.

## Budgeting and Planning for Data Management

This guide will emphasise the need to budget for data management activities both in terms of time and money and ensuring the right skills are available for the tasks.

## Creating a Data Management Plan

This will include a template of the data management plan and will link to the budgeting guideline. This should include plans for data quality assurance, archiving, storage, etc. This will also include roles and responsibilities for data management tasks to ensure that they are allocated to suitable, named individuals.

## Storing Numerical and Non-numerical Information

This guide will include suggestions on formats for numerical data, coding for text data as well as storing images, video clips and audio files.

## Transition from *raw* to *primary* data

This guide will describe the process of turning *raw* data, i.e., data as they are collected, into data that are ready for analysing and archiving.

## Data Quality Assurance

This guide will be partly concerned with data assurance throughout the research process but will also include suggestions for verifying and validating based, for instance, on local knowledge.

## Meta-Data

Mention will be made of some of the popular initiatives for meta-data (DDI etc.) but the guide will concentrate on the What, When, Where, Who, How questions to fully document the data. The guide will come with a template for meta-data which will also include project and contact information.

## Principles for Archiving and Sharing Data

This guideline will explain the benefits of archiving data. This will include Intellectual Property (IP) linking back to the data ownership guideline. It will also include ethical aspects ensuring anonymity and confidentiality of information providers. Consent forms will be mentioned pointing out that these forms must detail the uses that will be made of the data and that the data must not be used for any other purpose. This will include a section on GPS. The guide will highlight the difference between a public archive and a central data and document store for project information.

## Creating and using a Data & Document Store (DDS)

This guide will give examples of a DDS and how it could be managed. The guide will include references to video demonstrations showing how such a store might be managed.

## Data Types and Portals

This guide will list the portals available for CCAFS data together with the types of data accepted into each of these portals - for example, the AgTrials portal and the data that should be archived there.

## Dataverse (an example of a data archive)

This guide will use the CCAFS Baseline Dataverse as an example of an archive detailing the structure of the Dataverse and the process for creating studies and uploading files. It will include a template for the cataloguing information and will link to a couple of video demonstrations about Dataverse.

## D-Space (an example of how documents can be shared)

This guide will be similar to the one above but will discuss sharing documents on D-space.

## Terms of Reference for Data Managers at Project Level

This guideline will list the terms of reference for those responsible for data management project and activity levels. The intention is to help Principal Investigators in defining the tasks, responsibilities and skills of those who may be recruited to fulfil a data management role.

# Templates:

In addition to the guidelines there will be some templates/forms to accompany them:

* Project activity description.
* Project activity phase II.
* Metadata – a detailed list of the meta-data information needed to accompany the data archive.
* Quality ranking – this will help determine whether the archive from a project/activity is of gold, silver or bronze standard.
* Data and documents to submit for archiving (Checklist).