Instructions for Data Archiving in SASHA & CIP SSA

Jens Riis-Jacobsen 28/1/2011

What data will be archived?

There are two types of data that will be archived

1. Scientific studies and for these the archive will include for each one a description of study, raw data, analyzed data, reports, and publications.
2. Data related to the management of the program and projects

Purpose of archiving

The knowledge and supporting data generated in a research institution are its most valuable assets, and the management and safeguard of this is a core responsibility. The purpose of the archiving of scientific studies is ensure that the generated data and knowledge is stored in an adequate format and properly documented so that scientists or professionals in the future may utilize the data and knowledge to answer additional questions or apply the knowledge in relation to development work.

Our research is funded by external donors, and we have the obligation to document the good use and management of the resources we receive, and we must therefore archive project management information for each project. More specifically the archiving of project management information helps:

1. Document use and management of project resources
2. Facilitate internal and external monitoring and evaluation of project
3. Ensure institutional memory and accumulation of lessons learned
4. Share project information among stakeholders in the project to improve coordination and efficiency.

How do we in practice do the archiving?

1. The responsible scientist for each study or the project manager must organize the information as described below.
2. The scientist burn a dvd with all the studies and projects under his or her responsibility in the particular year
3. The dvd is submitted to the data manger in SSA, who will revise the structure of the information and acknowledge receipt.
4. A copy is sent to RIU at CIP-HQ for permanent storage.

Note: In 2011 Christmas holidays will only be approved for scientists who have fulfilled their archiving obligations. If not they will spend the holiday period archiving.

Archiving of scientific studies

What is a study?

A study is here defined as whatever each scientist in their field considers a study, examples include:

1. Breeding experiments/trials e.g. a trial series where a list of germplasm is being tested
2. A questionnaire applied to a group of people

Each profession will have a clear idea of what they consider a study, and at least initially we will accept each scientist's definition of a study.

What information should be archived for each study?

The information required for each study will depend on the type of study and the degree of standardization within the field. In for example sweetpotato breeding where a trial is carried out in accordance with the agreed breeding protocol, the description of the study only needs to document each trial with key information. On the other hand for example a non-standard nutrition study may require very detailed description on how the data was obtained and how it was analyzed. The basic rule of thumb is that the archived study should include all that is needed for another scientist to be able to interpret and use the data and knowledge generated. This means that the researcher must ask him/herself if the data is so well structured and documented that it would be easily understandable and useable by another scientist e.g. 5 years from now.

Each study should be stored in a separate folder and this folder should only include information relevant to the study. Each study folder should include:

1. A one page summary of the study providing key information using Dublin Core standard (see annex A)
2. Background folder (optional) with a description of more general considerations, preliminary thoughts on methodology etc.
3. Method folder with a detailed description of the method applied in the study and e.g. in a survey this should include the actual questionnaire applied. It should be easy to identify the final version that was actually applied, and preliminary versions should be clearly separated out.
4. Raw data folder which should obviously contain the raw data. If the data has been revised the folder may contain subfolders with different versions of the raw data
5. Analyzed data folder contains the outcome of the analysis, and if relevant may also contain the statistical scripts used.
6. Reports folder should contain any write up of the analysis and findings
7. Publications folder should contain any published or submitted manuscripts

From the above list folders may be omitted if they are empty.

The above list is mainly relevant for studies like surveys and laboratory experiments. For breeding trials made with CloneSelector the folder list will not work, and a single file contains raw data, method

description, and analysis. In that case a folder with Trials can be created, and all CloneSelector Fieldbooks should be stored in this.

Do I have to archive each year separately or can I accumulate the whole study across the years in one folder?

Even though a study runs over several years you should only create one folder for each study, and keep all the information related to the study in that folder. Each year we archive all studies that have had any change, such as e.g. a publication. So for each study you created a folder structure as described above, and for the duration of the project you simply add information in the relevant parts. This folder is then archived annually for as long as the study remains active including the final period for publications.

Ownership, use rights and due credit for intellectual work

The research has been funded by CIP and donor projects and the data and knowledge generated belongs to CIP. However, the scientists involved in a study will have a period of time where they have exclusive access and use right to the generated data. Either after publications have been made and/or sufficient time has passed the data may be made publicly available for other scientists to work on. If the data sets are used in further work by other scientists and the scientists that generated the original study will be ensured due credit for their work. Specific rules, copy right statements etc will be agreed before any data sets are made public.

Archiving project management information

Each program project should be organized in a folder which should contain all the relevant information for the management of the project. This folder should include the following subfolders

1. Contract information folder has all information related to the original project document and all agreements and addendums signed. This should contain subfolders:
    a. Project Document with the original project document and all annexes
    b. Project agreement, i.e. the signed agreement between donor and responsible institution
    c. Addendums should contain any addition or modification of the original agreement.
    d. Letters of Understanding or subcontracts which the lead institution has signed with other institutions
2. Work plans and budgets should contain for each year the approved work plan and budget, and any modifications these may have undergone during the year. A folder should be created for each year, and if a global work plan exists for the whole project period this should also be in a separate folder.
3. Progress Reports, should contain folders with
    a. Official project reports typically submitted to management of lead institution and subsequently to donor and steering committee.
    b. Internal reports should include all internal reports typically from the different components of the project. It may also include monitoring and evaluation data. Internal reports should be in a folder for each reporting period such as one per year or per semester.

4. Minutes of management meetings, Depending on the size of the project different management committees may exist, and for each one a folder should contain agendas an minutes off all meetings. All files should include the date of the meeting in the file name so that it is easy to find minutes from a certain period.
5. Other meetings, typically a project has a range of meetings to discuss or analyze different aspects and these meetings should be documented in a separate folder.
6. Administrative, financial and procurement information, typically the detailed information in this area is kept by the administrative unit of the institution, and the project folder should only contain overview information such as e.g. inventory of all items purchased with project funds, or standard reports of financial execution

What if it is a large Program with several projects within such as SASHA?

If a project within a program has its own management structure then it is recommended to establish a folder with project management information for each project under the overall program in addition to the program folder. In SASHA this is mainly relevant in the Proof of Concept projects which are implemented in collaboration with one or more external organizations, and this generates a series of management information.

Annex A: Template for key data on study or project

Title (the name given the resource)	
Creator (the person or organization responsible for the content)	
Subject (the topic covered), description (a textual outline of the content)	
Publisher (those responsible for making the resource available)	
Contributor (those who contributed to the study or project)	
Date (when the resource was made available)	
Type (a category for the content)	
Identifier (numerical identifier for the content such as a URL)	
Source (where the content originally derived from)	
Language (in what language the content is written)	
Relation (how the content relates to other resources, for instance, if it is a chapter in a book or part of a larger study)	
Coverage (where the resource is physically located)	
Rights (a link to a copyright notice)	