# CIP's Open Data & Data Management Policy

### 1.1 Introduction

CIP regards the results of its research and development activities as international public goods and is committed to their widespread dissemination and use to achieve the maximum possible access, scale, and scope of impact and sharing of benefits to advantage the poor, especially farmers in the developing countries. CIP, as the owner of all intellectual property produced, is responsible for the proper management of the outputs of all of its research and development activities. Data that results from a research activity is considered a public good. This policy regulates CIP's creation, processing, curation and preservation of data. This Policy has been developed in connection to and is consistent with the CGIAR Principles on the Management of Intellectual Assets[1], the CGIAR Open Access and Data Management Policy[2], the CGIAR Open Access and Data Management Implementation Guidelines, the OECD Principles and Guidelines for Access to Research Data from Public Funding (OECD, 2007)[3], and the UNESCO Universal Declaration on Bioethics and Human Rights (UNESCO, 2005)[4].

The implementation of this Policy is supported with two companion documents, CIP Research Data Management Guidelines and Procedures, and CIP Open Access and Data Management Implementation Plan.

### 1.2 Corresponding Framework Policy

In November 2013, all 15 members of the CGIAR Consortium unanimously endorsed the Open Access and Data Management Policy[2]. CIP regards this policy as the Framework Policy for Open Data and Data Management within de center.

This policy and the corresponding Implementation Guidelines are premised on the following key principles:

1. Open Access means immediate, irrevocable, unrestricted and free online access to all research outputs defined as information products, by any user worldwide.
2. Open Access applies to the following indicative types of information products ("information products"): peer-reviewed journal articles; reports and other papers; books and book chapters; data and databases; data collection and analysis tools (e.g. models and survey tools); video, audio and images; computer software; web services (e.g. data portals modeling on-line platforms); and metadata associated with the information products above.
3. The primary responsibility for implementing Open Access lies with the Centers.

[1] http://bit.ly/IAPrinciples

[2] http://bit.ly/OAPOLICY

[3] http://www.oecd.org/science/sci-tech/38500813.pdf

[4] http://unesdoc.unesco.org/images/0014/001461/146180e.pdf

4. Data underpinning peered reviewed publications must be made openly accessible as soon as possible, the latest being 12 months of completion of data collection or appropriate project milestone, or within 6 months of publication of the information products underpinned by that data.

5. The policy applies to all information products delivered from November 2013 onwards.

### 1.3 Application and Scope

This policy applies to all data generated by CIP and CIP's partners and collaborators within the framework of a CIP project starting November 2013. All scientific research outputs need to be verifiable by underlying data. The kinds of research data that are covered by this Policy include final data, as well as metadata, as applied to socio-economic, genetic, phenotypic, spatial, and environmental observations. Data includes also algorithms used to transform and analyze factual observations.

The objectives of this policy are:

1. To ensure that data, produced by research projects is of high quality and is readily available to CIP and the global research community to support research findings. This is best achieved through use of a common repository and use of standards in documentation and data formats. Details on standards and formats are defined in CIP's Data Management Guidelines and Procedures.

2. To ensure that data is readily and fully accessible in compliance with what is stated in the CGIAR Open Access and Data Management Policy

3. To support a culture of best-practice research data management in CIP by providing the resources (platforms, protocols, supporting processes) and incentives to the scientists.

4. To encourage harmonization (when applicable) so that data from separate research activities can be brought together to enrich the research process and institutional learning.

5. To inform CIP and partners of their obligations under the Policy.

### 1.4 Policy Statement

1. CIP is committed to enable and consolidate an Open Access culture within its research community. Research data management is a priority for CIP, and as such it will receive institutional support. CIP is committed to strive for the highest standards throughout the research data lifecycle (creating, processing, analyzing, preserving, giving access and re-using)

2. CIP will provide the necessary resources, including through infrastructure (e.g. data repository, archives and backups), services (e.g. to support data standardization and documentation), advice (e.g. statistical advice on experimental planning) and training, to achieve research data management consistent with this policy.

3. Research data management is a shared responsibility and a team effort. Researchers and supporting teams will need to work in partnerships to implement this Policy.

4. CIP staff are responsible for following this Policy and for implementing it using as a guide the CIP's Data Management Guidelines and Procedures document.

5. Research proposals must include a Data Management Plan and the associated costs to ensure compliance with this Policy. They must specifically address and budget for data collection, management, integrity, confidentiality, anonymization, prior informed consent, retention, sharing and publication.

6. Researchers must be familiar with donor Open Access policies and must apply either CIP standards or the donor (funding agency) standard, whichever more stringent.

7. For the purpose of publication, sharing and open access a distinction is made between raw data and final data.
    a. Records of both raw and final data must be registered, stored and preserved for institutional availability.
    b. Final data can be a research output in itself or can be data associated with a distinct research output.
        - Data as a research output must be made available within 12 months of completion.
        - Data associated with a distinct research output must be made available within 6 months of publication of the research output.
8. Intellectual property rights associated with research data created within the framework of collaborations and partnerships with third parties, will normally reside jointly with CIP and the third party. The ownership of intellectual property, and the Research Data Management Plan, will be specified in the relevant agreement (LOA, Consultancy Agreement, Intern Agreement, etc.)
9. All data must be accompanied by appropriate metadata and associated documents for archiving. The assigned metadata must be consistent with the CGIAR Core Metadata Schema and metadata standards for specific data types, as stated in CIP's Research Data Management Guidelines and Procedures in companion with its annexes.

### 1.5 CIP Data Management Guidelines and Procedures

CIPs Data Management Guidelines and Procedures are based on the Data Life Cycle[5] within the Project Life Cycle. It consists of six phases: creating, processing, analyzing, preserving, giving access, and re-using data. Phases 1 to 5 are addressed as an integral part of the Guideline[6]. The Guideline has 8 annexes and 1 checklist[PA(1)].

### 1.6 CIP Open Access and Data Management Implementation Plan

The CIP's Open Access and Data Management Implementation Plan has been developed pursuant to the CGIAR Open Access and Data Management Policy (adopted 2013) and the CGIAR Open Access and Data Management Implementation Guidelines (adopted 2014[PA(2)]).

### 1.7 The Data Workflow

Both the publications workflow and the data workflow, which are normally linked, operate within CIP's Project Lifecycle. CIP's strategy is to embed these lifecycles as much as possible in existing processes within Phases 5 to 7 in the Post-Award section of the Project Life Cycle. Figure 1 shows the workflow for data to identify, track, and make it openly accessible.

[5] http://www.data-archive.ac.uk/create-manage/life-cycle.

[6] http://cipotato.org/resources/

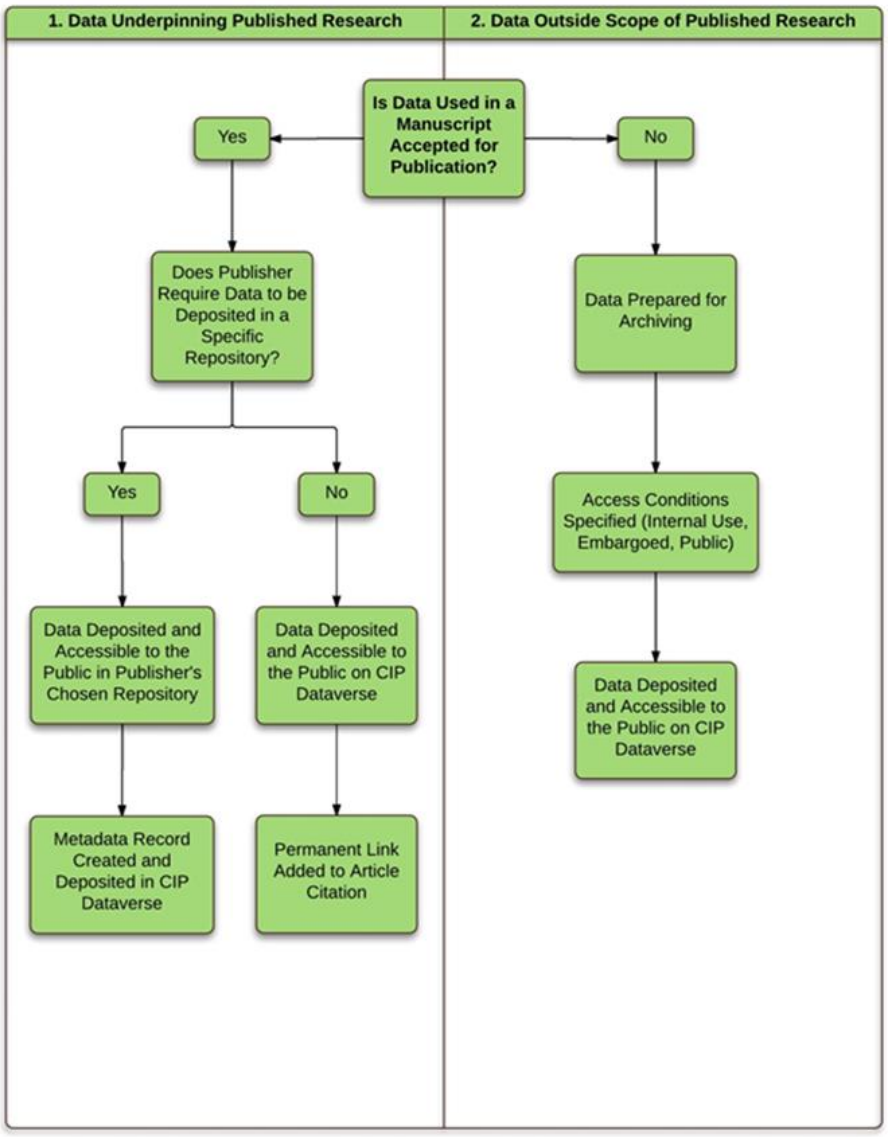


**Figure 1.** CIP Workflow for data.

### 1.8 Roles and Responsibilities

Researchers as project leaders are responsible for quality of data that results from their research activities and for making it openly accessible in accordance with this policy. This includes the data produced by partner institutions.
The Research Informatics Unit is responsible for providing support on data management related processes, including but not limited to: elaboration of a Project Data Management Plan, curation of data, definition of Metadata in accordance with the CGIAR Core Metadata Schema, upload of final data sets to the institutional repository of data. This unit is also responsible for the curation and maintenance of the institutional repositories.

**1.9 References**

CGIAR (2015). Open Access & Open Data Support Pack. https://sites.google.com/a/cgxchange.org/oad-support-pack/home. Accessed on April 23, 2015

Harvard (2014). Retention of research data and materials.http://osp.fad.harvard.edu/content/retention-of-research-data-and-materials. Accessed on April 23, 2015

OECD (2007). OECD Principles and Guidelines for Access to Research Data from Public Funding. http://www.oecd.org/science/sci-tech/38500813.pdf. Accessed on April 23, 2015

UNESCO (2005). Universal Declaration on Bioethics and Human Rights.http://unesdoc.unesco.org/images/0014/001461/146180e.pdf. Accessed on April 23, 2015