



CGIAR Open Access and Data Management Implementation Guidelines

Final version

July 2014



Contents

Section 1: Introduction	3
1.1: About the Guidelines	3
Section 2: Roles & Responsibilities.....	4
2.1: Members of the CGIAR Consortium (the Centers)	4
2.2: Consortium Office	4
2.3: Open Access Implementation Working Group	5
2.4: The Data Management Taskforce.....	5
2.5: CGIAR Partners.....	5
Section 3: Open Access and Data Management Plans & Implementation	6
Section 4: Next Steps.....	9
4.1: Phase 1 – Planning and Coordinating	9
4.2: Future Phases	9
ANNEX: Guidelines for Open Access Plans	10
A: Technical Infrastructure	12
A.1: Suitable Repositories	12
A.2: Interoperability	13
A.3: Metadata	14
A.4: Limited Internet Connectivity.....	14
B: Storage, Copyright/Open Licenses, and Translations	15
B.1: Data Storage and Preservation for Future Use	15
B.2: Copyright and Open Licenses	15
B.3: Translation	16
C: Administration of Open Access & Data Management.....	16
C.1: Strategy and implementation	16
C.2: Staffing	17
C.3: Financial Administration	17
D: Assessment, Impact, and Review	18
D.1: Incentives.....	18
D.2: Assessment and review.....	19
D.3: Tracking Impact and Uptake.....	19

Section 1: Introduction

1.1: About the Guidelines

In November 2013, all 15 members of the CGIAR Consortium unanimously endorsed the Open Access and Data Management Policy (the “Policy”)¹, a policy designed to make final CGIAR Consortium information products – including publications, datasets, and audiovisual materials – Open Access (“OA”).

The Open Access Guidelines (“OA Guidelines”) address a first phase of Consortium-wide implementation, one focused on **planning** and **coordinating** efforts throughout the CGIAR Consortium, to a point where all members of the CGIAR Consortium (“Centers”) have active Open Access implementation plans in place. The goal is to provide guidance and support to make this first milestone possible by the end of 2014. Furthermore, the OA Guidelines are premised on the following key principles which emerged during consultation with Research Centers in the months since the Open Access and Data Management Policy was passed:

- The primary responsibility for implementing Open Access must lie with the Centers.
- The Consortium Office’s key contribution will be the overall policy, supported by suggested, generic implementation guidelines that describe the basic principles to be followed in Center-developed implementation plans.
- Centers should develop their own specific and verifiable implementation guidelines and plans, according to the Consortium Open Access and Data Management Policy and this set of generic guidelines. Such plans should be tailored by each Center to their specific research programs, resources, and previous experience with Open Access and Data Management.

Thus, these OA Guidelines are intentionally broad in nature and are designed to offer as much flexibility as possible for Centers in planning for and preparing their own implementation plans and general approaches to supporting Open Access and Data Management. Even so, Centers’ plans – and, in particular, the resulting infrastructure – need to meet certain minimum criteria in order to work together across the Consortium.

Since the OA Guidelines are so broad, more detailed guidance and additional (optional) recommendations will be produced and offered through an evolving set of online resources. These resources, which will collectively be referred to as the “OA Support Pack,” will be a combination of new resources produced in response to requests for guidance by Research Centers; adaptable or re-usable materials developed by Research Centers; and materials produced by external organizations. The Support Pack will be available via <http://open.cgiar.org> by September 30, 2014.

Two notes regarding the scope of these Guidelines:

1. **Final information products.** The emphasis of the Open Access and Data Management Policy is on final research outputs – those information products (regardless of format) that are “stable” and unlikely to undergo further change (e.g., post-publication materials, datasets collected over the life of a project that has ended, etc).
2. **Data Management and Open Data.** These Guidelines focus on data within the context of Open Access – in other words, making final versions of research outputs (including data sets, analysis tools, survey instruments, models, summaries of data, maps and spatial products) openly and freely accessible for use and re-use by others. Data Management is addressed in as much as it affects making data OA.

¹ CGIAR Open Access and Data Management Policy is available at <http://bit.ly/1g8qioD>.

Section 2: Roles & Responsibilities

2.1: Members of the CGIAR Consortium (the Centers)

CGIAR Consortium Centers have primary responsibility and accountability for day-to-day implementation of Open Access – for instance, establishing Center-specific policies and procedures; managing and curating repositories; establishing policies and procedures for scientists that encourage, support, and reward deposits of research outputs and data in appropriate repositories; ensuring that Center-specific implementation efforts are consistent with the Open Access & Data Management Policy; and reporting metrics via agreed-upon channels.

As an early part of the implementation process, Centers should each prepare an Open Access and Data Management Plan. Section 3 of these Guidelines provides an overview for developing these Open Access and Data Management Plans; further details are included in the Annex.

CGIAR Research Program Lead Centers are responsible for ensuring that program participants and partners are compliant with the Open Access & Data Management Policy and Guidelines through their Program Participant Agreements (PPAs). CRP Lead Centers must appropriately allocate sufficient resources to allow for the implementation of Open Access. CRPs may choose to unilaterally adopt the Lead Center's Open Access plan, or to develop their own, provided it adheres to the core basic requirements of the policy and these guidelines.

Furthermore, Centers are responsible for ensuring that all relevant future agreements and contracts, including confidentiality, partnership, collaboration, development, licensing, distribution, material transfer agreements, employment contracts, and grants, comply with the Policy and with the CGIAR Intellectual Assets Guidelines.²

2.2: Consortium

The CGIAR Consortium, on behalf of the CGIAR Consortium members (the Centers), is focused on advancing Open Access and Open Data at a system level, providing Centers and CRPs with the policy framework for Open Access and Data Management, as well as the system-wide coordination, aggregation, and sharing of tools, seed funding, resources, and advocacy needed by members of the Consortium to implement OA and data management.

The Consortium serves in a consultative capacity by facilitating meetings, bringing in external resources, and acting as a champion and advocate for Open Access and Open Data internally and externally as CGIAR strives to become a leader in open knowledge for agriculture research.

The Consortium shall:

- Comply with the CGIAR Open Access and Data Management Policy, including developing a Consortium Office Open Access and Data Management Plan for its own outputs and products that could serve as a template for other CGIAR entities;
- Facilitate the process to develop, maintain, assess, and revise the Open Access and Data Management Policy and Implementation Guidelines;
- Facilitate the process to develop, maintain, assess, and revise the CG Core Metadata Schema and relevant vocabularies adopted at the Consortium level;
- Develop, compile and aggregate data, benchmarks, milestones, and metrics at the Consortium level;

² See the Implementation Guidelines for the CGIAR IA Principles on the Management of Intellectual Assets (<http://bit.ly/1nkHhdC>) for more details.

- Compile and make available tools, resources, and content for the Open Access Support Pack;
- Develop new tools to bring together content or harvested metadata across CGIAR repositories (and partner repositories, when/where applicable);
- Develop new tools to analyze or visualize aggregated data from repositories across the CGIAR Consortium (including support for other value adding products and services derived from or leading to better use of open data and content);
- Provide aggregated reporting to Open Access funders and other interested parties;
- Highlight and showcase emerging best practices from within the Consortium;
- Negotiate contracts and Memoranda of Understanding related to Open Access that Centers might benefit from (e.g., publisher agreements that minimize OA fees);
- Promote CGIAR OA principles in relevant initiatives and organizations related to Open Access and Open Data, such as AATP, GODAN, and others; and
- Communicate OA activities with relevant Center staff, involving them in negotiations, priority setting, product testing etc.

2.3: Open Access Implementation Working Group

The Open Access Implementation Working Group (OAIWG) has been established to help create the enabling environment for Open Access implementation³ and consists of Knowledge Managers from Centers, CRPs and CO representatives. It will help to oversee and guide the implementation of the CGIAR Open Access Policy between 2014 and 2018 and for managing appropriate communications around this Policy. See the Terms of Reference for the OAIWG for full details.

2.4: The Data Management Taskforce

The CGIAR Data Taskforce (DMTF) will take a lead role in coordinating Open Access implementation with a focus on data management issues.⁴ The DMTF is made up of Data Managers or their equivalent from Member Centers and CRPs, along with CO representatives. In particular, the Data Management Taskforce will provide oversight of data standards and protocols and will be responsible for defining the appropriate standards and interoperability protocols to be implemented and applied across CGIAR Open Access repositories. See the Terms of Reference for the DMTF for full details.

2.5: CGIAR Partners

Information products produced by lead Centers and participating Centers (including partners) in CRPs are subject to the Policy on all new contracts established since the adoption of the Policy. A contract may have been entered into which contains restrictions on, for example, sharing the data under a research and/or development project or under a commercialization endeavor. Future agreements should be carefully negotiated to ensure that any such restrictions are limited in duration, territory and/or field of use, if applicable, and fully justifiable by reference to the CGIAR Principles on the Management of Intellectual Assets (i.e., in particular articles 6.2, 6.3 and 6.4)⁵ and these Guidelines. For more details on this, work with your Center's IP focal point.

³ See the Terms of Reference for the CGIAR Open Access Implementation Working Group for additional details.

⁴ See the Terms of Reference for the CGIAR Data Management Task Force for additional details.

⁵ Article 6.2 is on "Limited Exclusivity Agreements"; Article 6.3 is on "Incorporation of Third Party Intellectual Assets" and Article 6.4 is on "Intellectual Property Rights", all part of the CGIAR IA Principles.

Section 3: Open Access and Data Management Plans & Implementation

According to Article 4.1.9 of the Open Access and Data Management Policy: *“Open Access and Data Management Plans should be prepared in order to ensure implementation of this Policy. Such Plans shall, in particular, outline a strategy for maximizing opportunities to make information products Open Access.”*

Each CGIAR Consortium Member Center should develop a Center-specific Open Access and Data Management Plan detailing how Open Access and data management will be implemented and supported. Plans should address all of the elements specified by the Policy. **Likewise, all aspects of implementation – particularly in terms of the technical infrastructure including repositories, metadata, and interoperability – should comply with the minimum parameters set forth in the Policy and these Guidelines.** See Table 1 for a high-level overview of these minimum parameters and the Annex for details. Numbers in parentheses refer to articles from the CGIAR Open Access and Data Management Policy.

Centers are expected to, at minimum, adopt, and ideally, enhance and exceed these requirements. Open Access Plans should be updated regularly to reflect current thinking as Centers move forward with implementation. More detailed guidance and examples will be forthcoming as part of the OA Support Pack.

Table 1: Essential Elements to Include in Open Access Plans

Essential Elements for Plans	Minimum Requirements
Scope	
Openness (4.1.1)	General statement regarding interpretation of openness, what is covered in the plan, brief synopsis of current OA situation/needs/challenges, how any repositories or exchange systems include partner contributions and provide them access, areas to address in the future.
Information Products (4.2)	Plans should address how all types of information products addressed in the Policy are collected, stored, and disseminated; differences in treatments based on type of information products and format of information products. Types of information products to be addressed include: peer-reviewed journal articles; reports and other papers; books and book chapters; data; data summaries (e.g. maps, indicators); video, audio, and images; computer software and models; and associated metadata.
Technical Infrastructure	
Suitable repositories (4.1.2)	Repository systems should meet current industry standards for interoperability and metadata. Plans should address which repositories are in use for each type of information product (see below), associated URL(s) for the repositories, and identify which repository platforms/systems are in use. Where information or data is stored or published in systems that may not be ‘repositories’, plans should show how these meet Open Access principles and requirements, including standards of interoperability and use of metadata.
Interoperability (4.1.3)	Plans should address which interoperability protocols and standards are adopted in any repositories as well as any other digital or web-based aggregation or harvesting systems and services; how content is transferred or interpreted between systems (internally or externally); and any new or emerging tools, protocols, or frameworks being tested or considered for adoption.

Metadata (4.1.3)	Centers and CGIAR system units should demonstrate that they are working towards adopting the CG Core Metadata Schema. Plans should confirm that they have adopted CG Core and offer details on any additional metadata schemas and taxonomies which are in use and how they are applied.
Limited internet connectivity (4.1.8)	Plans should address how the Center and system units are providing access to content when internet access is limited.
Data storage and preservation for future use (4.1.4)	Plans should address storage and preservation issues such as storage redundancy; storage formats and preservation mechanisms; use of persistent identifiers; recommended & accepted file formats.
Copyright and open licenses (4.1.5)	Plans should address recommendations on which open licenses to use for different types of information products or if the Center adopts recommendations offered by the CO.
Translation (4.1.7)	Plans should address how open licenses allow for re-use such as translations of information products and ways in which translations are encouraged.
Administration of Open Access & Data Management⁶	
Strategy and implementation	Plans should address short-term and long-term strategic goals, objectives, and priorities for the Open Access and data management programs as well as implementation timelines.
Processes and workflows	Plans should detail workable processes to acquire and deposit material into repositories, manage/maintain repositories, assure quality control etc.
Staffing	Plans should address which department(s) or position(s) have day-to-day responsibility for Open Access and data management, which group(s) have oversight for Open Access and data management, and the general structure of these groups. Plans should address recommended workflows for depositing materials into repositories.
Financial administration	Plans should address financial administration, funding, and major expenditures for supporting Open Access and data management support.
Assessment, Impact, and Review	
Incentives and professional expertise (4.1.6)	Plans should address ways in which creators of information products are encouraged to comply with the Policy and how compliance is tracked, measured, and reviewed.
Assessment and review (5)	Plans and progress should be updated on a yearly basis. Plans should address mechanisms for internal review.
Tracking impact & uptake (5)	Plans should address which metrics are being collected and how they are interpreted in order to understand usage, impact, and uptake of materials disseminated through Open Access.

Content should be deposited in full as soon as possible after an item is complete or in its final form. Plans should address timelines for depositing information products into repositories. Plans’ deposit schedules should be consistent with these guidelines or indicate shorter timelines than those presented here. See **Table 2** for details **Error!**
Reference source not found.

⁶ Strategy, staffing, and financial administration are not directly referenced in the OADM Policy, but are necessary for successful execution of Open Access and data management and so should be addressed in Centers’ Plans.

Table 2: Deposit Schedules

Types of Information Products	Deposit Schedule
Peer-reviewed versions of journal articles	Ideally, at the time of publication Latest: 6 months from publication ⁷
Self-published journals, books, reports etc.	Immediately
Reports and other papers	As soon as possible Latest: within 6 months of completion
Externally or commercially published books and book chapters	As soon as possible Latest: within 6 months of completion
Data and data sets	As soon as possible Latest: within 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 ⁸
Video, audio, scientific images	As soon as possible Latest: within 6 months of completion
Photographs	As soon as possible Latest: within 6 months of completion or publication
Computer software/applications/code	As soon as possible Latest: within 6 months of completion
Metadata	As soon as possible Latest: before or on publication of the information product
Core/corporate governance documents appropriate for public consumption (e.g., financial reports, board agendas and minutes, annual reports, as appropriate)	As soon as possible

Automated exceptions/extensions. Certain types of information products (in particular data collected pursuant to hypothesis-driven research) may take longer than 12 months to clean, analyze and publish. Thus, 12 months should be seen as the aim, with 24 months as the long-stop date for making such data Open Access. During the implementation phase, the best timing of disclosure of information products will be identified.

Exceptions. The general principle is to make information products Open Access, but that is always “*subject to the legal rights and legitimate interest of stakeholders and third parties, including intellectual property rights, confidentiality, sensitivity (including price and politically-sensitive information), farmers’ rights and privacy.*”

Exceptions also include aspects referred to in Articles 6.2, 6.4, and 6.4 of the CGIAR IA Principles.

Effective date. The Policy is effective as of 02 October 2013. Only final information products produced after the effective date are covered under the Policy.

⁷ Researchers are strongly encouraged to work with publishers to secure appropriate permissions. However, if a publisher’s contract prevents compliance, appropriate copies should be deposited into the repository immediately, but be blocked from public access until after the embargo is lifted. Furthermore, if an article is not able to be deposited into a repository due to a publisher’s restrictions, a metadata record should be added to the appropriate CGIAR repository and should include a link to the publisher’s website.

⁸ The timeframe for dissemination of data is as soon as possible after data collection has been completed, and in any event within 12 months, although this may vary according to the type or nature of the results, among other factors.

Section 4: Next Steps

4.1: Phase 1 – Planning and Coordinating

The emphasis for 2014 will be on planning and coordinating efforts across Centers. Much of this work is focused on collecting and aggregating information to discover what is already in place throughout the Consortium and bringing all Centers up to minimum compliance with the Policy.

Activities included in Phase 1 for 2014:

- Gather baseline information about current practices: inventory established repositories, adopted interoperability protocols, Centers’ focal points. (March – May)
- Form two new groups designed to support Open Access implementation: Open Access Implementation Working Group (OAIWG) and Data Management Taskforce (DMTF) (March-May)
- Share baseline information about the current status of Open Access and Data Management as gathered via the Inventory (May/June)
- Through the OAIWG and DMTF, set minimum requirements for metadata for CG repositories (June/July)
- Begin sharing detailed guidance on a variety of topics through the OA Support Pack (June - September)
- Begin work on an advocacy plan and advocacy materials (June-September)

Furthermore, all Centers should be starting to work with researchers to collect and describe final information products. Ideally, these products will be immediately deposited into suitable repositories. For Centers without such repositories in place at this point, information products should still be collected and described. Workflows can be established even before repositories are in place.

In order to advance progress with Open Access as quickly as possible and have Members in compliance with the Policy, Centers and the Consortium should aim to complete the following tasks by these deadlines:

Activity	Deadline
Report baseline metrics via the CGIAR inventory of current practices (survey)	30 April 2014
Develop and submit Center-based Open Access and Data Management Plans	15 December 2014
Launch suitable Open Access repositories – or adapt existing repositories	15 December 2014
Incorporate metadata recommendations into new/existing repositories	15 December 2014

4.2: Future Phases

Once all Centers have developed their Open Access plans, have established workflows in place to acquire and deposit content into repositories, and have suitable repositories in place, Consortium-level efforts will transition to aligning the infrastructure across Centers, increasing the visibility and usage of information products, ensuring compliance of the Policy, and assessing the impact of the Policy.



ANNEX: Guidelines for Open Access Plans

Guidelines for Open Access Plans

Since CGIAR will be implementing Open Access in a distributed environment, it is important that each the approach meets the minimum criteria outlined in these guidelines, and is in line with resource levels, technical capabilities and infrastructure, and the exact nature of its information products.

Plans will be shared with the Consortium Office and the OAIWG and DMTF in order to foster the exchange of knowledge and spark new ideas for implementation. Plans should also be made publicly available in order to support knowledge exchange within the broader agricultural research community.

Plans should be reviewed and updated yearly to reflect progress, the adoption of best practices, and the incorporation of new technologies. The first round of implementation plans, will be due to the Consortium Office in December 2014. The Consortium Office plan is envisaged earlier, to provide an option for Centers to use it as a template. For the first phase of implementation, plans will be used as the reporting mechanism to help establish initial progress.

Furthermore, plans will be used to inform review of the Open Access and Data Management Policy and its implementation. According to Article 5 of the Policy: “*The CGIAR Consortium Office will carry out an evidence-based review of the implementation of this Policy on an annual basis. The reviews will be used to devise appropriate institutional tools and guidelines for the implementation of this Policy.*” Plans for the review will be developed in the coming months.⁹

Each plan should include a general statement regarding any Center--specific interpretation of openness, current OA situation, workflows to address content deposition into repositories or other systems, maintenance of these platforms, quality control etc., what is covered in the plan, and areas included in the policy that are not yet supported by the Center.

The following types of information products should be addressed in each plan:

- Peer-reviewed journal articles (Article 4.2.1)
- Reports and other papers (Article 4.2.2)
- Books and book chapters (Article 4.2.3)
- Data and databases (Article 4.2.4)
- Video, audio, and images (Article 4.2.5)
- Computer software/applications/code (Article 4.2.6)
- Metadata (Article 4.2.7)
- Core/corporate governance documents appropriate for public consumption

Some Centers might elect to support different types of information products in different ways. For instance, a Center might use DSpace to collect, archive, and disseminate peer-reviewed journal articles, Dataverse for socioeconomic datasets, and Github for computer software. Other Centers might use a single repository and offer the same services for all openly-accessible information products. At a minimum, plans should include an overview of how the different types of information products are supported.

Additional guidance for specific types of information products will be addressed through the Support Pack.

⁹ As part of the Implementation Roadmap, the Open Access Implementation Working Group and the Data Management Taskforce will both be involved in helping to devise the evidence-based review of implementation and establishing M&E metrics.

A: Technical Infrastructure

In order to comply with the CGIAR Open Access and Data Management Policy, each CGIAR Consortium Member Center and system unit must use suitable, interoperable, standards-compliant Open Access repositories in order to provide access to the body of information products and associated metadata produced by members of its community. Many of the Centers and units already have their own repositories – and some Centers have multiple repositories adopted for different purposes – leading to a complex, distributed network of repositories throughout the CGIAR community.

In terms of the technical infrastructure, plans should address:

- Repository systems (Article 4.1.2)
- Interoperability (Article 4.1.3)
- Metadata (Article 4.1.3)
- Limited internet connectivity (Article 4.1.8)

The Technical Requirements included here are the minimum necessary requirements in order to create an interoperable network of repositories and present open knowledge in a coherent, meaningful way across repositories, Research Centers, and CRPs. Furthermore, adoption of widely-used systems, protocols, and standards will make it possible to continue to enhance the infrastructure in the future and incorporate new tools designed to maximize visibility, discoverability, re-use, and uptake of CGIAR information products.

A.1: Suitable Repositories

According to Article 4.1.2 of the CGIAR Open Access and Data Management Policy: “*Stable, permanent, Open Access repositories shall be utilized, to enable users and other sites and search engines to access or locate information products, including application programming interfaces (APIs) or other mechanisms enabling those information products to be available from the CGIAR website and associated web-based products. Preference should be given to existing repositories to minimize the number of repositories in use (and the interoperability challenges presented by multiple incidences of repositories.)*”

All of the information products covered in the Policy must be deposited into an Open Access repository which will then provide the necessary infrastructure to archive and disseminate the body of knowledge captured in CGIAR’s information products.

System design must be scalable, flexible, and facilitate extraction of data in multiple formats and for a range of uses as internal and external needs change, including potential uses not accounted for in the original design. In general, this will involve the use of standards and specifications in the system design that promote best practices for information sharing, and separation of data from the application layer to maximize data reuse opportunities and incorporation of future application or technology capabilities, including the ability to export information in RDF.

At a minimum, all repositories should be OAI-PMH compliant or expose metadata through standard APIs.

Publication and General-Purpose Repositories

Many Centers and CGIAR units are already using DSpace for publication repositories. DSpace is an open-source, standards-compliant, widely-adopted repository platform which meet the specified criteria. Centers or units not wishing to maintain their own publication repositories, or wishing to collaborate and cost-share should consider using

CGSpace¹⁰, a shared installation of DSpace that serves several CGIAR entities and CRPs, or hosting their own installation of DSpace. Other possibilities and recommendations will be provided by OAIWG if any Centers do not expect to have a workable repository in place within the next few months.

Data Repositories

At the November 2013 CGIAR Data Standards Summit, 5 main types of data related to agricultural research were identified: (1) socio-economic; (2) spatial; (3) genetic; (4) genomic; and (5) germplasm. A major data type of relevance to Centers that is missing from these streams is agronomy trial data (particularly that focusing on management). Effective immediately, the Data Standards Task Force will begin working to coordinate and standardize data management practices around these areas, recognizing that in the NCBI suite of platforms, there are already good options for genetic and genomic data repositories in particular. Data from each of these areas has highly-specialized characteristics and will require different types of stewardship, and, in many cases different platforms.

Many Centers are already using DSpace (primarily publications), Dataverse (primarily socioeconomic data), or AgTrials (primarily breeding and agronomy trial data) as repositories. Genetic and genomic data from Centers is also deposited into NCBI databases. All these repositories are or will soon be open-source, standards-compliant, and widely-adopted systems which meet the specified criteria.

Individuals supporting data repositories are strongly encouraged to work through the Data Management Taskforce when encountering new data management challenges. Since research programs and subject areas touch multiple Centers, it is expected that others are encountering – or will encounter – similar challenges. Moving forward, it would be beneficial for Centers to work together to support emerging areas (e.g. the need to effectively archive and make accessible agronomy trial datasets).

At a minimum, plans should:

- Address which repositories/platforms are in use and for what purposes they are being used – e.g.: Dataverse for primarily socio-economic data sets, DSpace for peer-reviewed and other publications, Open Journal System (OJS) for CGIAR-published journals; and
- Include URLs to the repositories.

Further guidance on repository systems and data management will be forthcoming as part of the Support Pack.

A.2: Interoperability

As a result of this distributed environment, repositories must be syntactically and semantically interoperable. This means that at a minimum, repositories should all be OAI-PMH compliant and comply with the CG Core metadata schema. Other considerations for using standard APIs can be considered by the Data Management Taskforce.

Interoperability is possible by adopting commonly-implemented technical protocols, standards, and vocabularies. Although the interoperability landscape continues to evolve, using widely-adopted repository systems, metadata schemas, ontologies, and vocabularies make it easier to incorporate new tools, protocols, and initiatives as they are released.

¹⁰ CGSpace home: <https://cgspace.cgiar.org/>. CGSpace is a collaboration of several Centers and CRPs and is hosted by the International Livestock Research Institute (ILRI).

At a minimum, plans should address:

- Which mechanisms are in place to enable cross-system transfer of content or metadata and how these mechanisms are being used;
- How content is transferred between systems (internally or externally); and
- Any new or emerging tools, protocols, or frameworks being tested or considered for adoption in the coming 1-2 years.

Examples of interoperability protocols or tools used in connection with Open Access and Data Management include:

- AGROVOC Linked Open Data API to use AGROVOC terms from within DSpace
- OAI-PMH enabled to allow harvesting of metadata by OAIster

Further guidance on interoperability and examples of protocols will be forthcoming as part of the OA Support Pack.

A.3: Metadata

The CG Core Metadata Schema (CG Core)¹¹ will be a common framework for CGIAR Consortium Member Centers, CRPs, and other entities to present and share metadata in consistent ways across the network of CGIAR repositories. CG Core is based on Dublin Core, a widely-used metadata standard, with a limited number of additional elements specific to the CGIAR environment.

All CGIAR Consortium repositories should adopt the CG Core. It is not intended to replace existing metadata schemas used by Centers, which include additional domain-specific details and vocabularies. It is expected that most repositories will include a crosswalk to map existing element sets to CG Core elements.

Plans should confirm adoption of the CG Core elements within repositories. In addition, plans should address:

- Other metadata schemas in use and how they are applied, and
- Vocabularies in use and how they are applied.

See “CG Core Metadata Schema & Guidelines” for current details and guidelines of CG Core. Further guidance on applying CG Core will be forthcoming as part of the OA Support Pack.

A.4: Limited Internet Connectivity

Research Centers are encouraged to design repositories and websites in ways that support low-bandwidth and mobile users, and others with limited internet connectivity, without compromising quality of the information products. At a minimum, plans should address how Centers support individuals with limited internet connectivity and steps taken to optimize for low-bandwidth connections or provide alternate versions of information products that require minimal data in order to download.

Recommendations related to accessibility, low bandwidth access, and mobile access will be included in the OA Support Pack.

¹¹ At the time of this writing, CG Core Metadata Schema is in development as a draft. The Data Management Taskforce will take ownership for CG Core after the taskforce has become operational.

B: Storage, Copyright/Open Licenses, and Translations

Plans should address three elements related to the treatment of information products:

- Data storage and preservation (Article 4.1.4)
- Copyright and open licenses (Article 4.1.5)
- Translations (Article 4.1.7)

B.1: Data Storage and Preservation for Future Use

Challenges associated with digital storage and preservation are twofold. First, a digital object must be preserved in such a way that the digital bits which comprise the object are able to be accessed in the future – i.e. ensuring the object itself is intact, that it is stored in such a way that its integrity is maintained. In this regard, it is recommended that Centers follow standard best practices for data storage and security such as redundancy and the use of SSL access to servers and systems.

The second aspect of storage and preservation is ensuring future usability of digital objects – i.e. the ability to interpret, understand, and use stored information. Digital preservation may require technologists to routinely migrate files from one format to another as formats become obsolete. Even so, technologists cannot guarantee future usability of proprietary formats.

As a result, in order to promote future use, Centers should encourage content creators to only deposit into repositories and other OA platforms information products which are in standard formats (e.g. PDF, Office Open XML, or plain text for documents; PDF, CSV, or XLS for data; MP3, MP4, MOV etc. for audio/video files, etc.)¹²

Centers might elect to offer different levels of preservation to different categories of information products based on uniqueness and importance.

At a minimum, plans should address:

- Steps taken to ensure secure storage of data such as storage redundancy;
- Acceptable or recommended file formats for storage of each type information product;
- Preservation mechanisms currently adopted;
- Steps taken to deal with file format obsolescence.

Additional guidance from the Data Management Taskforce will be included in the OA Support Pack.

B.2: Copyright and Open Licenses

Article 4.1.5 states that “*Suitable open licenses shall be used that recognize the legal rights to information products and encourage their use and adaptation.*”

The license conditions upon which information products are made Open Access may vary depending on the nature of the information products and the need to limit or restrict access or usage rights to certain audiences and users. No single license is appropriate for all research projects.

¹² List of Open File Formats as maintained in Wikipedia: http://en.wikipedia.org/wiki/Open_file_format

At a minimum, plans should:

- Offer Center-specific guidance or recommendations on adoption of open licenses;
- Provide recommendations for authors to ensure the originating Center maintains rights to translate key works;
- Identify which types of open licenses are commonly used, for what types of information products, and in what systems they are used;
- Identify ways in which re-use of information products is encouraged; and
- Identify a process to collect copyright and open licensing questions that arise in relation to the Open Access and Data Management Policy in order to share knowledge across Centers and inform development of Consortium-wide recommendations in this area.

Additional guidance for researchers, particularly in terms of working with publishers and copyright agreements, will be issued as part of the Support Pack.

B.3: Translation

The Policy indicates in Article 4.1.7, “*Translations of key documents and other media into pertinent languages are encouraged. All versions should be deposited in suitable repositories and made Open Access.*”

At a minimum, plans should address:

- Tools embedded into repository and search functions to allow content to be discovered in and translated into relevant languages;
- Ways in which the Center has encouraged translations of key documents;
- Processes used to identify and track translations of key documents;
- Metrics used to monitor usage of translations of key documents through repository downloads; and
- Ways in which the Center is encouraging translations of key information products, particularly those which are targeted to reach specific beneficiaries.

Additional guidance will be included in the Support Pack.

C: Administration of Open Access & Data Management

Implementing Open Access and data management will require time, infrastructure, financial resources and human expertise. Article 4.1.1 of the Policy states: “*Best efforts shall be used to make information products Open Access...*” ‘Best efforts’ is an intentionally high standard, which requires resources to be secured and expended in meeting the challenge of Open Access. Best efforts should include:

- Ensuring adequate staffing and expertise to support Open Access and data management
- Securing necessary financial resources

C.1: Strategy and implementation

Support for Open Access and data management comes in many forms and levels, so no single model is appropriate for all Centers. It is recommended that Centers develop plans for Open Access and data management in a way that is in keeping with the Open Access and Data Management Policy and these guidelines and aligned with: current research agenda and priorities, technical infrastructure, staffing levels and areas of expertise, budget, and recent experience providing support for Open Access and data management.

While the Policy is designed to make final CGIAR Consortium information products available via Open Access, this transition process will take time. Article 2 of the Policy addresses this transition process: *“Implementation of and compliance with this Policy...will be phased over a transition period. The transition period runs from the effective date of the Policy for an initial period of 5 years, with comprehensive implementation by the end of 2018.”*

Further, not all CGIAR Consortium information products will be of equal value. Article 4.1.1 of the Policy states: *“Open Access arrangements should consider the characteristics of the information product, their potential impact, the level of data processing required, and whether the information products generated are within the scope of this Policy. Some judgment therefore needs to be made over the information products that will be made Open Access.”*

Centers – particularly those which do not have open access repositories in place or have not been making information products openly accessible in this manner – are encouraged to approach implementation as a staged process.

At a minimum, Centers’ plans should include a section related to strategy which addresses:

- Priorities for implementation
- Goals and objectives for the time period covered by the Policy
- Timelines and key milestones for the time period covered by the Policy

C.2: Staffing

Open Access and data management support requires a mix of operational support and oversight/coordination. Furthermore, Open Access, data management, and the technical environment are constantly changing and will require ongoing professional development to support.

At a minimum, plans should address:

- Departments or individuals serving as the focal point for operational support for data management
- Departments or individuals serving as the focal point for operational support for Open Access
- A general workflow for depositing different types of information products into repositories
- Departments or individuals providing oversight or coordination of Open Access and data management
- Ways in which gaps in professional expertise will be addressed

Additional guidance such as sample job descriptions for data managers will be issued in the OA Support Pack.

C.3: Financial Administration

Open Access and data management will require financial investments by Centers. Examples of anticipated or potential expenditures include:

- Article processing charges (APCs): charges levied by some Open Access journals and some “hybrid” closed-access journals that offer authors the ability to secure an open license for a particular article
- Repositories: software (if using software that is not free/open source), hardware, server space, processing space
- Staff: hiring new data managers, ongoing professional development for staff supporting Open Access and data management
- New tools: licensing of new tools to enhance discovery and encourage usage of Open Access content

Centers should begin to budget for Open Access and data management costs in future funding proposals, including the forthcoming second call for CRPs. As an interim step, the CGIAR Consortium is working towards making dedicated funds available to support Open Access and data management; it is intended that these funds will be allocated based on guidance from the OAIWG and DMTF. However, this is not a sustainable solution. Centers and researchers need to begin to budget for Open Access costs such as article processing fees in funding proposals. Alternatively, Centers may wish to consider alternate options such as recommending that authors publish in Open Access journals that do not levy article processing fees or determining what constitutes reasonable fees.

At a minimum, plans should address:

- Budgets for infrastructure and appropriate staffing
- Budgets for professional development
- Fees for Open Access publishing
- Other major expected expenditures

Additional guidance such as examples of mechanisms to support fee-based Open Access publishing will be forthcoming in the Support Pack.

D: Assessment, Impact, and Review

The CGIAR Consortium will continue to work with experts across funders, private sector, academia, and society to develop meaningful metrics to assess usage, impact, and uptake for Open Access based on iterative learning and experimentation. New feedback, best practices, and overall expertise will be incorporated into the development of future Open Access objectives, milestones, and metrics. As such, it is expected that plans for assessing impact and uptake will evolve.

As a starting point, Centers' plans should address:

- Incentives (Article 4.1.6)
- Assessment and review of Open Access and data management at a programmatic level (Article 5)
- Tracking impact and uptake (Article 5)

Consortium-level plans for assessing and reviewing impact will be addressed in 2014-2015 as part of a second phase of implementation. During the first phase of implementation, Centers will report on their existing methods for incentivizing compliance.

D.1: Incentives

Since Open Access and data management is mandated through the Policy, all CGIAR Consortium researchers and other creators of information products are required to comply. Centers should encourage and reward sharing of information products (including data) through appropriate Open Access channels. Likewise, due consideration should be given in individual appraisals for compliance with the policy but also for taking extra steps to encourage discovery and uptake of information products.

At the minimum, plans should address:

- Ways in which researchers and authors/creators of information products are encouraged to comply with the Policy.
- Review processes or measure that reward scientists who are in compliance with the Policy.

Furthermore, Centers are encouraged to share any sample text from employment contracts or appraisal forms which reference (directly or indirectly) Open Access or data management.

Specific examples will be collected and shared via the Support Pack.

D.2: Assessment and review

Measuring successful implementation of the Policy will require both quantitative and qualitative information due to varying stages of information management maturity across CGIAR and the evolving nature of successful partner and public engagement. Centers are encouraged to develop their own methods for tracking progress of Open Access.

At the Consortium level, baseline information will be collected in 2014 and will be updated annually thereafter to track progress with implementation across Centers. Article 5 of the Policy states: “*The CGIAR Consortium Office will carry out an evidence-based review of the implementation of this Policy on an annual basis.*” The Policy itself will then be reviewed in 2015 and every two years thereafter.

At a minimum, plans should address mechanisms for internal review.

D.3: Tracking Impact and Uptake

As the scholarly communication landscape evolves through developments in Open Access and interoperable technologies, research institutions are developing and testing new ways to assess the impact of research outputs and track usage. “Altmetrics,” or new types of metrics which are only possible in the digital environment, are still experimental yet allow new ways to glean insight into how individuals interact with bodies of research or individual information products. While specific tools are still experimental, the scholarly community is embracing a shift towards new types of metrics to track usage in quantitative ways and trace the path of uptake in qualitative ways.

Centers are encouraged to experiment with new ways of measuring, assessing, and tracking research outputs. It is expected that once a cohesive infrastructure is in place, the Consortium will collectively test various tools and metrics to track usage and establish a common understanding for interpreting these new metrics.

At the minimum, plans should address:

- Which metrics are being collected, how these metrics are collected, and how they are interpreted in order to understand usage, impact, and uptake of materials disseminated through Open Access

Additional resources will be forthcoming through the Support Pack to help address altmetrics and other emerging ways to track impact and uptake.