Making CGIAR Research Outputs Available and Accessible as IPGs[1]

The Independent Review of the CGIAR recently suggested that its Centers should be encouraged to "make their research available and useful for development" – as well as to international science.

Our studies show that many CGIAR research outputs are not widely accessible, thus they are less likely to be taken up and applied in development. Despite their high quality, they are not (yet) international public goods.

An 'availability, accessibility, applicability' framework offers pathways to wider accessibility and research outputs that are thus more likely to be IPGs.

Research organizations like the CGIAR cannot be satisfied just knowing they have produced high quality science. It is essential that the outputs of their research are communicated and put to use, in the village, on the ground, in the lab, or across the negotiating table.

This paper draws on work commissioned by the CGIAR ICT-KM Program to investigate how the information and knowledge needs of the CGIAR's priority stakeholders can be better met. It suggests that, despite substantial efforts and innovation across centers and system-wide added value services such as virtual libraries and data warehouses, much CGIAR knowledge remains hard to see and hard to get.

An 'AAA' – availability, accessibility, applicability - framework is proposed to assist scientists, managers and information professionals develop pathways to improved accessibility for their outputs. And improved chances that they will be put to use

We recognize that research processes undergo several stages between problem identification and dissemination of outputs, and that knowledge sharing and attention to the accessibility of data, information and knowledge are necessary along the whole cycle[2].

This paper focuses on the outputs and results of CGIAR research, how they are documented, shared, disseminated and ultimately applied. We argue that research whose outputs has not been made accessible may not be an

international public good. We also argue that investments to extend the accessibility of outputs will yield a richer harvest for the CGIAR as well as producers and consumers worldwide.

---
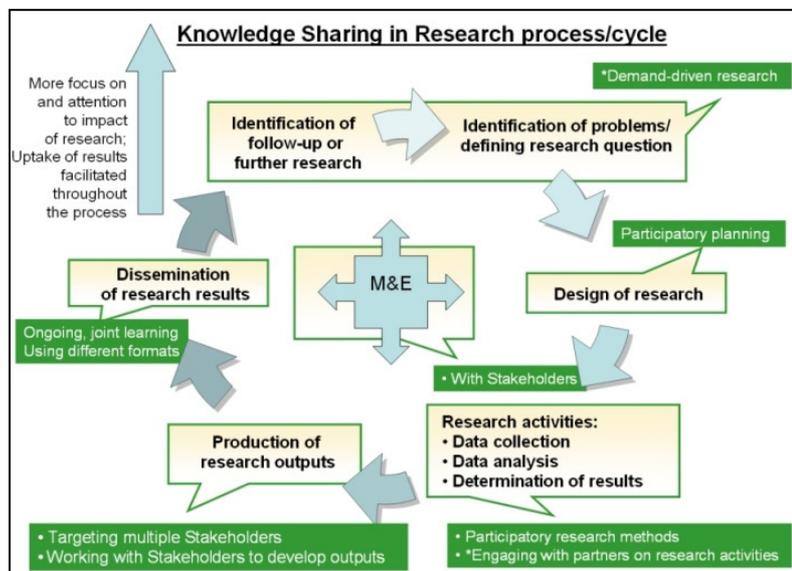
[1] Background note by Peter Ballantyne for Alliance of CGIAR Centers, Global Forum for Agricultural Research (GFAR) and CGIAR Science Council Agricultural Research Public Goods (IPG) Workshop, Maputo, Mozambique, Thursday 27 November 2008. It draws on a longer paper commissioned by the CGIAR ICT-KM Program.
[2] The 'Knowledge Sharing in Research' Project of the CGIAR ICT-KM Program has been looking at the needs and potential of knowledge sharing within and along the research cycle. See: www.ks-cgiar.org

IPGs and the CGIAR

In a recent paper, Robinson and Elliott[3] sketch some changes underway in different CGIAR Centers. The comparative advantage of the Centers is seen to be "in contributing to poverty reduction in the provision of international public goods (or significant regional public goods)."

These IPGs will be produced "in partnership" with other actors in an R for D (research for development) continuum, perhaps as part of an 'innovation systems' approach, certainly with organizations not traditionally considered part of the agricultural knowledge system.

With other actors taking on greater roles in the wider development process, CGIAR research on IPGs needs to generate outputs that other will use to provide local, national and regional benefits. This means that CGIAR products should be easily accessible to 'intermediary' actors who will adapt, improve and apply CGIAR knowledge. It also implies that the capacities of these partners may need to be strengthened and enhanced to ensure they are able to take up and best apply the outputs.

This thinking is captured in a recent Science Council document[4] on the role of the CGIAR as a player in Research *for* Development. This sets out key arguments explaining why the CGIAR should pursue IPGs in its strategies and activities.

The first part of the document is a wide-ranging discussion by Jim Ryan of IPGs in the CGIAR context[5]. He provides a valuable 'working definition' of IPGs:

> "International public goods are taken to mean research outputs of knowledge and technology generated through strategic and applied research that are applicable and readily accessible internationally to address generic issues and challenges consistent with CGIAR goals."

The key here is the identification of knowledge and technology as the main outputs, and their ready international applicability and accessibility as essential features. **Making these outputs accessible (and applicable) is therefore a critical task for the CGIAR**.

The CGIAR as Producer of Public Goods and the vital role of partners

CGIAR Centers produce a variety of 'goods' from their research, capacity strengthening, and partnership activities. Encapsulated as outputs, these goods are typically new knowledge and technology (from research), enhanced skills and capabilities, and stronger institutional arrangements and policy evidence leading to change.

Phil Pardey[6] argues that "most research products are not intrinsically public — technically, non-rival and non-price excludable — or private goods; they fall in the "shades of grey" category, have multiple attributes (with some attributes that are more or less rival or excludable than others), and, above all, can be made more or less public (or not) through policy and practical actions on the part of the CGIAR or others."

[3] Robinson, M. and H. Elliott. 2007. Meta-evaluation of External Program and Management Reviews (EPMRs). Washington DC: CGIAR. www.cgiar.org/pdf/agm07/agm07_epmr_meta_evaluation.pdf

[4] Science Council. 2006a. Positioning the CGIAR in the Research for Development Continuum. Rome: CGIAR. www.sciencecouncil.cgiar.org/fileadmin/user_upload/sciencecouncil/Reports/REPORT_of_R4D_website.pdf

[5] Ryan, J. 2006. "International Public Goods and the CGIAR Niche in the R for D Continuum", IN: Positioning the CGIAR in the Research *for* Development Continuum. Rome: CGIAR Science Council. www.sciencecouncil.cgiar.org/fileadmin/user_upload/sciencecouncil/Reports/REPORT_of_R4D_website.pdf

[6] Science Council. 2006a. op. cit.

This last point is very important as we look at the information and knowledge in the CGIAR. As he reiterates, "it is as much how the CG opts to use its research products, not necessarily something intrinsic in the product itself that determines if the output is more or less a "public good.""

Information and knowledge are not born 'public'; they must be worked upon to ensure they become public, i.e.: available, accessible, and applicable.

Depending therefore on how a research activity is sponsored and designed, the outputs of research may or may not be a 'public good.' Much research by public institutions with public funding is not a 'pure' public good, as when it is designed to benefit a specific community or local problem, when it is commissioned by an interest group for its own use, when it involves public and private interests and proprietary knowledge or technology, or when it is otherwise not intended for broad public use.

The means by which an output is made available, accessible and applicable also determine whether it can become a public good. A classic example is where publicly funded research outputs are disseminated in limited-access scientific journals that explicitly exclude some classes of potential users, or where outputs are only available on web sites that are less accessible to people without good web access. A similar case can be made for outputs made available in one language only, written in a 'scientific' style, published in a proprietary format (such as .doc or .pdf) or with restricted intellectual property licenses. There are many such ways in which the public character of a research output can be limited by decisions taken regarding its dissemination. As the ACIAR[7] recently argued: "Knowledge is a classic club good, available free [only] to those who have the capacity to access (and use) the information."

For the goods to be truly international and public, the information and knowledge[8] outputs produced through CGIAR efforts must be made available, accessible and, in an 'innovation systems' paradigm, be taken up and applied through the efforts of others.

The recent CGIAR independent review argues that the "CGIAR can achieve development impacts where they matter — for farmers, families, and societies — only by being part of an international public goods delivery system." At one end, the CGIAR needs access to global knowledge, at the other, it needs access to "adaptive and delivery agents" and partners who "turn outputs into outcomes and impacts." The strong capacities of the partners are thus vital elements in this delivery system.

The CGIAR as a Global Public Goods 'Infomediary'

Given the large number of potentially public good outputs produced by the collective CGIAR, various information and knowledge based facilities and services (typically applications and services such as information portals, search and dissemination systems, databases, and information modeling and analysis tools) are emerging to maximize access to the public good outputs of the CGIAR. These do more than just provide access. Through various value-adding activities, they become new public services in their own right.

An example is SINGER - the System-wide Information Network for Genetic Resources (http://www.singer.cgiar.org). This facilitates user access to the content of 11 CGIAR genebank collections through a single entry point. While each individual accession in a genebank only has the potential to be a public good, it is the collections put together and documented in the various Centers and the services like SINGER that are the international public goods made available and accessible by the Centers collectively.

---

[7] ACIAR. 2006. ACIAR and public funding of R&D: submission to the Productivity Commission study on public support for science and innovation. www.aciar.gov.au/publication/IAS41
[8] 'information and knowledge' is used as a shorthand for data, observations, literature, software applications etc.

Another is the CGIAR Virtual Library (http://vlibrary.cgiar.org) that provides enhanced access to texts and other materials available in the many CGIAR libraries, and increasingly publication catalogues. It enhances the international public good nature of any CGIAR outputs included in the collections – by making them more accessible; it also enhances the public good nature of any third party materials by also making them easier to identify, if not necessarily to actually obtain. It also provides a system-wide search service for CGIAR staff.

The On-line Learning Resources project (http://olr.cip.cgiar.org) contributes to the 'transfer' of international public goods, knowledge and technologies produced by the CGIAR Centers and provides a platform and approach for sharing these learning resources.

CGMap (http://cgmap.cgiar.org/), the newest system-wide application from the ICT-KM Program, allows easy navigation through information on current and planned research and research-related activities – and their outputs - across the CGIAR.

While SINGER, Agricultural Science and Technology Indicators (ASTI), or FishBase seem to be accepted as IPGs, they are not always valued as such[9].

Triple A Framework to Assess and Guide Actions

In 2005, the CGIAR ICT-KM Program defined IPGs as:

> "data, information, and value-added information and services based on data and information that are:
> - Searchable and located in repositories (electronic)
> - Globally available
> - Open and easily accessible to all"

This is a good starting point. The three bullets particularly focus on the availability and accessibility of the research outputs.

The earlier definition by Ryan is particularly useful as it adds the notion that an IPG output needs to be applicable[10]. This says something about the relevance of the original research, how appropriate and relevant the message is, as well as something about the ways it is presented and packaged and how its re-use is regulated.

> A research output is an IPG when "others should be able to access and benefit from the research" – J. Ryan

To help assess the IPG status of CGIAR research outputs, we propose a simple framework around three notions associated with the 'Availability, Accessibility and Applicability' (AAA) of the research outputs. Without attention to these three, a research output will not become an IPG.

Thus, an IPG could be a CGIAR-produced data, information, or knowledge asset that:
- is described and stored for posterity
- can be easily found and accessed,
- encourages use and re-use of knowledge [and can easily be appropriated],
- is appropriable, accessible, sharable, reusable, available, affordable
- is available, accessible and applicable without restrictions

---

[9] See discussion of the "public good failure problem facing databases" by Robinson and Elliott, op. cit., pp 31-34.
[10] Ryan, J. 2006. "International Public Goods and the CGIAR Niche in the R for D Continuum", IN: Positioning the CGIAR in the Research *for* Development Continuum. Rome: CGIAR Science Council. www.sciencecouncil.cgiar.org/fileadmin/user_upload/sciencecouncil/Reports/REPORT_of_R4D_website.pdf

Such an IPG is:
- created and deposited in formats and systems that allow perpetual access;
- licensed to allow and encourage widespread use and appropriation of the content;
- described and indexed to allow easy finding and dissemination;
- optimized to enhance understanding and encourage widespread use.

An IPG information or knowledge service or tool is:
- organized and structured to provide easy access;
- designed and optimized to share and exchange;
- makes CGIAR-produced and CGIAR-stewarded knowledge available, accessible and applicable without restrictions.

The Three A's

Availability: Research outputs are stored in appropriate open digital formats[11] and described using public metadata standards so they can be found through structured search and access systems. They may not always be publicly available in full. Availability means assembling and storing content so it will be permanently accessible, and describing it in systems so others know, and can find, what outputs have been produced.

Accessibility: Research outputs are publicly available online using accepted public formats and appropriate licenses so they can be queried, viewed, and obtained in full. Outputs are optimized so metadata and full content can be harvested and shared across different platforms and applications, and they can be incorporated into other systems and services. Accessibility means making outputs as easy to find and share and as open as possible, in the sense that others are free to use, reuse, and redistribute them[12], with appropriate acknowledgement and without restrictive legal, technological or financial barriers.

Applicability: Research processes are open and inclusive so that all perspectives and knowledge are taken into account during research design, planning, implementation and communication. Research outputs are customized and/or adapted for easy uptake and use by other actors in agricultural innovation systems, increasing the public benefits derived from the data, information and knowledge produced through research. Applicability means research and innovation processes that are open to different sources of knowledge, and outputs that are easy to adapt, transform, apply and re-use.

How Publicly Accessible are the CGIAR's Research Goods?

In this 'Google generation' a quick general search in Google for pdf files – a likely format for serious research outputs – reveals thousands of CGIAR outputs on different web domains. We could not see how complete these outputs are in terms of the total production of a CGIAR Center, nor how accessible they are.

To better understand what information is actually made available and accessible, we looked more closely at the activities of some Centers to see what is actually produced and what becomes available and accessible. For this, the CGIAR Performance Measurement System provides a good source of information on what a Center says it produces each year, in terms of peer-reviewed literature. Similar data is likely to be available for all Centers so the exercise could be repeated.

---

[11] A format is what enables an application to interpret the raw data contained in a file. Some are more open and accessible than others. Using open formats grants "accessibility and perennity to your data" **-** See http://www.openformats.org/en6 and http://www.okfn.org/iai/wiki/FormatRegistry.
[12] http://www.opendefinition.org

Information from one Center gives us an illustration. Peer-reviewed and refereed publications listed by the Center in its 2006 performance management system report were searched for in different systems and services to assess their availability and accessibility. No attempt was made to assess applicability.

Searching for each of these outputs in different systems revealed:

Many of the text outputs of the Center are available full text online, ranging from 100% of the journal articles to 32% of the book chapters. Most of the Center's own publications are available online. Most are available full text in an internal 'repository' database, however, the necessary permissions to make these publicly accessible are not in yet place.

Aside from the Center's own publications. that are mostly openly available as full text, 22% of articles, 60% of papers, and 32% of book chapters could be found and accessed 'openly' ie, without any kind of login or permission. The book chapter figures are inflated as many 'appear' to be full text in Google, but they are actually available in an incomplete form with some pages missing.

Many of the Center's own research outputs could not be found by searching its library catalogue. Hence the CGIAR virtual library also has incomplete coverage of the Center's outputs. For this exercise, the 'best' place to find Center-produced and authored outputs was Google Scholar.

What could we conclude?

1. Not all of a Center's outputs are internally available in full text. What is available is often not a version than can or should be made publicly accessible.
2. While many products are available online, most, especially those not published by a Center, are not immediately open.
3. Access in developing countries is enhanced (for participating organizations) when the article is published in a journal that is included in services like AGORA.
4. Concentrated efforts to obtain publisher permissions to make articles publicly available before and after publication through institutional repositories, or to pay for open access online, can transform the accessibility of a Center's outputs.
5. The search tools provided by a center to find its own research outputs can be less complete and useful than those of Google scholar.
6. The performance measurement system of the CGIAR has created an incentive for centers to collect information on many types of peer-reviewed research outputs, but not others.

Copyrights

One way that Centers can make their content accessible, and proactively encourage use and re-use, is via the copyright and licenses they use. This is not the place to discuss these issues fully, except to flag that there are significant differences between the Centers, leading to different levels of accessibility.

Frequently, Centers and scientists give some of their rights to other organizations - publishers.for example – and the content becomes even less accessible to people without financial resources. Some Centers are negotiating with publishers to be able to make normally closed-access outputs openly available through their repositories (usually after some time lag). Some, like IRRI, have also moved to adopt more open creative commons licenses.

How IFPRI Makes its Outputs Available and Accessible

A report by the IFPRI Communications Division for the Center's internal review December 2006, provides valuable insights into strategies and tools employed. Some specific initiatives include:

IFPRI's Own Research Outputs are Available through:

- Library catalog online –indexes ALL Center publications and staff outputs, often links to full text
- Publications catalog online – lists Center published outputs, gives links to full text on website

IFPRI's Own Research Outputs are Accessible through:

- Publications catalog is source of metadata for other information systems via 'harvesters'
  - Eg, Google scholar, RePec, WorldCat, AgEcon, CiteSeer
- Agreements with publishers to provide free electronic access to the full text of IFPRI researchers' articles to developing-country partners.
- RSS feeds and blog stories and announcements, wikipedia pages
- Some own publications added to Google Books, Johns Hopkins (for sale)
- Publications can be browsed and ordered via 'shopping cart' application
- More flexible publishing arrangement with the Johns Hopkins University Press, allowing IFPRI to retain regional rights for copublished titles

Using information from the IFPRI website (www.ifpri.org/pubs/articles/articles.asp), the improving accessibility of journal articles and book chapters can be seen:

- 2006   240 items   40% full text open accessible
- 2007   212 items   71% full text open accessible
- 2008 (to February) 24 items   92% full text open accessible

Social Media

In recent years, Centers have begun to use more 'social media' – like blogs and wikis. These are often referred to as 'web 2.0' applications. IFPRI's World Hunger Blog was launched in June 2005. The number of visitors grew from 52 visitors a month in June 2005 to 5,953 visitors a month in October 2006. Page views rose from 119 per month in June 2005 to 7,234 per month in October 2006. In recent years, IFPRI's RSS feed has become one of its most frequently viewed pages, with monthly "hits" increasing by more than 300 percent since January 2006.

Other examples of the use of this web 2.0 in the CGIAR include:

- The ICT-KM Program runs several blogs as vehicles to distribute news and updates.
- The Bioversity news service is produced on a blogging platform; the IRRI, IFPRI and Bioversity libraries have blogs highlighting developments.
- IRRI has loaded thousands of its photos to Flickr – the free photo-sharing space.
- Almost all Centers have introduced RSS feeds on their web sites that allow readers to subscribe to automatic updates on news of each Center.
- WARDA, IRRI and other Centers share more popular – and more applicable - introductions to some of their research work as video's on You Tube.
- Several Centers have descriptive pages on Wikipedia.
- The IRRI-CIMMYT 'good practices in research data management' project uses a wiki.

Making Research Outputs Applicable

The preceding sections have focused on ways that CGIAR research can be made more available and accessible – as preconditions for its application.

There is much thinking and work across the CGIAR on ways to make its research more relevant and more likely to be taken up[13]. Some is more 'downstream' and focuses on packaging knowledge in different ways – extending its accessibility. Others are more 'upstream', focusing on research processes in design and operational phases when the likely relevance, and applicability, of final outputs will be shaped.

Some examples include:

- Bioversity has been working on 'marketing strategies' to help stimulate the uptake of research results. [upstream]
- The motto of ILRI's '**Innovation Works' team**[14] is "Knowledge to action for larger, wider and faster impacts." [upstream]
- The Institutional Learning and Change Initiative (ILAC - www.cgiar-ilac.org) is setting up an 'ILAC Learning Laboratory' in which the leaders of selected collaborative programmes will come together to share their knowledge and experiences, learn from one another, experiment with new approaches for facilitating pro-poor innovation and receive support for change efforts. [upstream]
- WARDA and IRRI have been using video to communicate science with farmers. This 'Zooming-in Zooming-out' approach helps extension service providers create regionally relevant learning materials that are locally appropriate[15]. [downstream]
- The Virtual Academy for the Semi-Arid Tropics (VASAT - www.icrisat.org/vasat) links rural farm communities with researchers, credible intermediaries and markets through an ICT interface and open and distance learning methods. [downstream]

The Next Generation? Pathways to Accessibility

The study shows CGIAR scientists and systems designed around relatively traditional forms of science publishing and communication (with pockets of innovation). Particularly for literature and publications, these outputs are less accessible than is perhaps desirable.

Part of the solution is to move towards the next generation of science communication where individuals and groups are empowered to document and communicate their own activities in different channels and social media. This 'e-science' has traditional peer-reviewed outputs, perhaps managed differently than now. It also gives a 'cloud' of communication and dissemination possibilities for individuals to adapt to their own needs and situations.

Alongside this move to 'science 2.0' – that will transform information, knowledge and communication flows in and around the CGIAR – there is an urgent need to invest in activities and strategies that 'convert' research outputs into IPGs. Some overall areas for action include:

- Promote Discussion and Agreement on an AAA Vision for the CGIAR
- Advocate 'Triple A' Thinking to also Measure Research Performance
- Guide and Help Scientists to Adopt 'Triple A' Thinking
- Establish Complete Scientific Information Repositories
- Leverage CGIAR Publishing Muscle
- Encourage Creative Innovative Experiments
- Make CGIAR Knowledge Visible through Partners

---

[13] See especially the 'Innovation Works' initiative at ILRI and the KS in Research Project of the ICT-KM Program.
[14] See: www.ilri.org/research/Index.asp?SID=213
[15] See: http://iaald.blogspot.com/2008/01/zooming-in-zooming-out-communicating.html

To put some 'next generation' actions[16] in motion, a system-wide effort can be conceived to: 1) bring together and make research outputs widely and openly accessible through system-wide 'InfoSpaces,' 2) add value through 'InfointoChange' activities that compile, transform and deliver information in forms that others can easily apply, and 3) extend the reach and value of CGIAR research by developing strong 'InfoAccess' partnerships with other organizations engaged in agricultural innovation for development.

Through such an information access initiative, the accessibility of 'traditional' CGIAR research outputs in the form of literature, reports, documents, and publications can be enhanced. Non-traditional outputs can also be made more accessible by combining two inter-connected CGIAR domains – publishing and libraries – in a system-wide undertaking to make research outputs widely accessible, and more likely to be taken up by partners in development.

Research InfoSpaces

Each CGIAR Center and constituent entity currently maintains multiple different content collections and services – libraries, publications catalogues, repositories, archives, image databases, dataset databases, expert databases, web site management systems, web sites, etc., and, increasingly, wiki's, blogs, and a growing range of 'social' applications. Much content and metadata describing all this content is also available in or on third party services (publisher systems, google books, etc). In terms of licenses to use it, some of the content is completely open, some is closed, with various stages between.

Making all of this accessible – to the system as well as to partners – requires major re-thinking and re-tooling across the CGIAR. Uniting and integrating the core functions and processes of 'library-like' and 'publishing' activities will help deliver the following:

- Comprehensive indexes of CGIAR research outputs archived in open digital formats
- Complete open and harvestable and searchable repositories of CGIAR cross-center and system-wide research outputs (including co-created content)
    - 'Documents', publications, images, video, learning resources, etc
- Applications and services that use open standards, open applications, open APIs, capable of exchange across systems and platforms
- Content that is licensed to be as openly accessible and re-usable as possible
- Wide dissemination of research outputs, and the messages they contain, across different platforms and systems
- Intelligent search and harvesting tools to access this content

InfointoChange

It is not enough to bring together and make research outputs available and accessible. They need to be taken up and applied by others. The CGIAR is home to many communication, outreach and marketing initiatives. Greater system-wide efforts are needed to build on these, to continue experimenting and innovating, to ensure they add value, and to forge alliances with partners that will extend the use and value of the CGIAR's outputs. This would deliver:

- Added value information awareness, aggregation, dissemination, personalization, visualization, and synthesizing services for external and internal clients, including access to external and internal content, delivered through own and others' services.
- Content and services that are re-usable and can be 're-purposed' across different CGIAR platforms, Centers, communities, and beyond, across organizational boundaries
    - Such as in CGXchange, CGMAP, CGVlibrary
    - That can be included in a CGIAR staff member's personal 'dashboard'
    - That can be customized and personalized, or included in content 'mash-ups'

---

[16] The following directions grew out of an ICT-KM brainstorming and planning session in May 2008.

- Content and services that are web 2.0 enabled, allowing content to travel, be tagged, be tracked and become 'social';
- Innovative metrics and measures of research output 'accessibility' – and perhaps quality – that better fit the CGIAR's ambition to act along the entire Research to Development continuum;
- Fully e-literate CGIAR staff and scientists: communicating their research in accessible ways; intelligently consuming and using scientific information

InfoAccess Partnerships

The CGIAR is part of wider global knowledge and innovation systems serving agriculture and rural development. To globally mobilize and access the products of these, and its own, knowledge systems, effective collaboration and joint ventures need to be dedicated to improving information access.

This can be done in two main ways: First, by building relationships and joining forces with relevant content producers and aggregators, the CGIAR can develop novel products and services that add continuing and collective value, ensuring that CGIAR staff, partners and collaborators have access to the information and data they need. Second, by forming strong partnerships with other 'infomediaries' who will adapt and re-package CGIAR knowledge, data and information, products will be applied and taken up by other actors in the research for development continuum.

Joint actions need to be developed with international, regional and national partners (in terms of outreach and uptake), with specialized infomediaries (in terms of content, access and expertise), and with other access providers (for all kinds of value-adding and value-extending results).

Examples include:
- Joining forces with publishers and FAO to strengthen AGORA as an access point to otherwise 'closed' outputs of CGIAR research
- Working with RAIS and GFAR to better incorporate CGIAR content in their services
- Combining with NAL, Wageningen, FAO and others to negotiate joint access to international journals and databases
- Working with local organizations to make CGIAR content available under very open licenses so it can be adapted and translated for local uses
- Partnering with Microsoft and/or Google to make CGIAR content accessible
- Partnering with several publishers to make CGIAR research outputs widely accessible