

Benchmarking CGIAR Research Outputs for Availability and Accessibility

Meena Arivananthan, Peter Ballantyne, Enrica M. Porcari

ABSTRACT: To better understand ways that CGIAR Centers make their published research outputs available and accessible, a benchmarking study of six Centers was carried out in 2008 and 2009. The study scored several typical 'pathways' that could be used to identify and obtain research outputs published in 2006. Results indicate that these outputs are generally available—they can be identified—in various pathways but that they are much less openly accessible in full text. The results also show differences between the approaches of the six Centers, differences in performance between pathways, and differences in the accessibility of different types of outputs.

RESUMÉ: Pour mieux comprendre comment les centres du GCRAI rendent leurs publications des résultats de recherche disponibles et accessibles, une étude des standards de six centres a été exécutée en 2008 et 2009. L'étude a marqué plusieurs «chemins d'accès» typiques qui pourraient être utilisés pour identifier et obtenir la publication des résultats de recherche en 2006. Les résultats indiquent que ces produits sont généralement disponibles—ils peuvent être identifiés—sous divers

chemins d'accès mais qu'ils sont beaucoup moins accessibles publiquement en plein texte. Les résultats montrent aussi des différences entre les approches des six centres, différences de performance entre les chemins d'accès, et différences d'accessibilité de différents types de produits.

RESUMEN: En el 2008 y 2009 se hizo un estudio comparativo de seis centros del Grupo Consultivo para la Investigación Agrícola Internacional (CGIAR) para conocer más a fondo los medios que dichos centros utilizan para asegurar la disponibilidad y accesibilidad de los resultados de investigación que publican. El estudio calificó diversos 'medios' típicos que se podrían utilizar para identificar y obtener resultados de investigación publicados en el 2006. Los resultados indican que dichos resultados están, en general, disponibles—es decir se pueden identificar—en diferentes medios pero que su accesibilidad como texto completo es mucho menos abierta. Los resultados también muestran diferencias entre los enfoques de los seis centros, diferencias de desempeño entre los medios y diferencias en la accesibilidad de diferentes tipos de productos.

Introduction

Knowledge generated by scientists in the CGIAR plays an important role in delivering solutions for the poor—sustainable agricultural growth to help reduce poverty. Established in 1971, the CGIAR, an alliance of 15 agricultural research centers with 8,000 researchers and technicians in 200 locations, has a strong legacy, producing impacts in agricultural production, germplasm improvement and collection; and policy, with notable examples like the Green Revolution in South Asia, New Rices for Africa and Quality Protein Maize.^{1,2}

A recent independent review of the CGIAR stressed the need for centers to make their research available and useful for development science.³ With a mandate to create international public goods from their research, the CGIAR is aware that research outputs need to be made widely available and accessible and shared with individuals and partner organizations who may apply and deliver on-the-ground impacts.

For these international public good outputs of research to have impact, each research output:

- Needs to be helped to 'travel' across boundaries
- Should be described and stored for posterity
- Should be easily found and be accessed
- Needs to be configured and licensed to be easily shared and re-used
- Has to be as affordable as possible

In short: they need to be Available, Accessible and Applicable without restrictions.⁴

Traditionally, CGIAR research outputs are disseminated through close collaborative efforts with a wide range of research partners—international, regional, national and local. They are usually shared via scientific journals, conferences, books, networks and other traditional communication methods.

In recent years, the CGIAR Science Council has promoted and rewarded the production of 'high impact' articles in academic journals. Collaboration is also evolving and is now multi-pronged, and includes not only national agricultural partners, but NGOs, universities, the private sector and many other players. The knowledge generated is no longer the property of an exclusive scientific audience.

These knowledge products include data, improved germplasm, training programs, international best practices; policy and management advice; information systems; models and technologies. To become international public goods, these products or outputs need to be made available and accessible, increasing the potential for them to be applied by an increasingly diverse range of partners and users.

Building on an earlier initial analysis,⁵ in 2009 the CGIAR ICT-KM program launched a benchmarking exercise with six CGIAR Centers. The aim was to assess the actual availability and accessibility of their published research

outputs. The study also aimed to take stock of current pathways⁶ used by the six Centers to disseminate their research outputs, identifying promising ways to enhance future efforts in this area.

Methodology

For the purpose of this study, we used the definitions below based on the ICT-KM Triple-A framework⁷:

- Availability: research outputs stored in open digital formats and described using public metadata standards so they can be found through structured search and access systems;
- Accessibility: research outputs publicly available online so they can be queried, viewed and obtained in full.

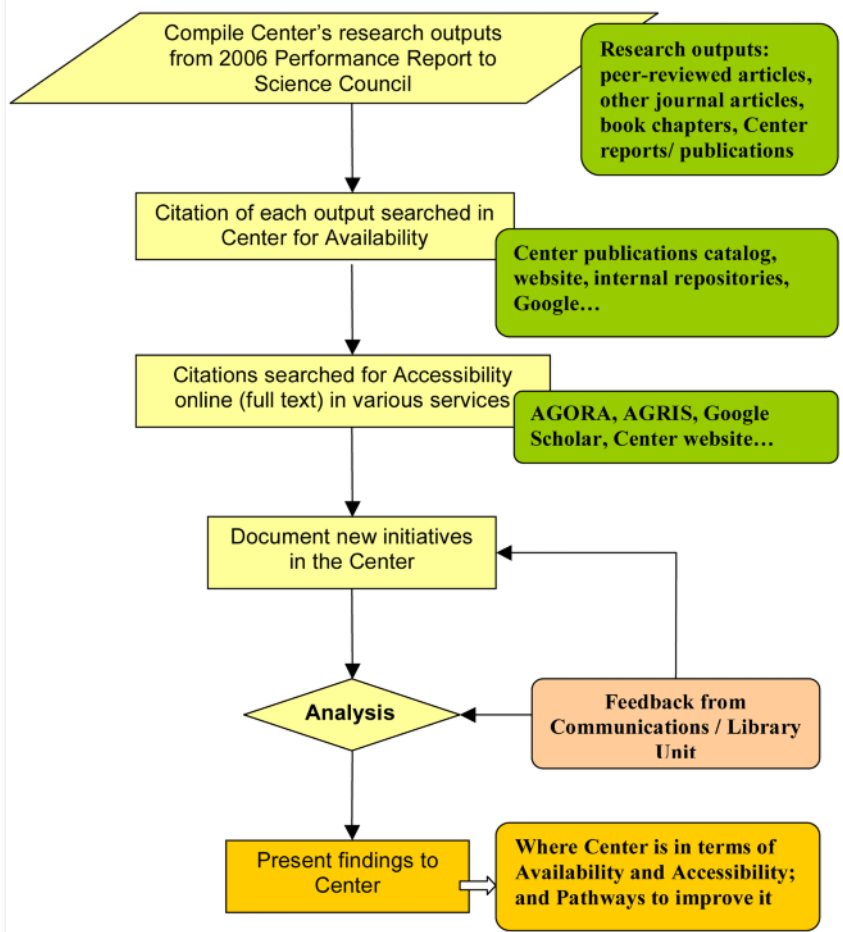
The study used a simple bibliometric approach initially developed and tested by Peter Ballantyne for the ICT-KM Program in 2008. The starting point (see Figure 1) was to compile a list of published research outputs for each of the six Centers. The 2006 Performance Report submitted by each Center to the CGIAR Science Council was used as an ‘official’ list. In total, 1088 research outputs were identified, comprising peer-reviewed journal articles indexed in Thomson/ISI, externally published books and book chapters, articles in other journals (not indexed by Thomson/ISI) and publications published by each Center.

The availability and accessibility of each of the outputs was tested by searching for them in a range of standard ‘pathways’ used by each Center (Table 1). It should be noted that not all Centers use all the pathways; and that other potential pathways were identified but were not tested across all the Centers due to lack of access (eg CAB database) and time.

TABLE 1 – Availability and Accessibility Pathways

AVAILABILITY (citation/ abstract)	ACCESSIBILITY (full text)
Center library catalogue	Center website
Center publications catalogue	AGORA database
Center ‘institutional archive’	CGVlibrary search engine
Center website	AGRIS database
Center web search function	Google Scholar search
Google general search	Google Books search

FIGURE 1 – Methodology to study the availability and accessibility of research outputs



An online search was carried out for each output and if the citation or abstract was found, it was marked as being available through the respective pathway, be it the Center’s library catalogue or Google. For a research output to be marked as accessible, it would have to be available in full text for public use. Access would not require any password or subscription fee.

Data was collected in the same way an external user would try to access a document. Insider know-how was deliberately avoided. This was an important aspect to maintain since the study hinges on the accessibility of research outputs to individuals outside the CGIAR.

Results

The study provided each Center with an overview of where they are in terms of research output dissemination and how and where they could improve this. Data collected from each Center was shared with the Center’s information or library unit. They were also requested to provide any additional ways in which they disseminated their research outputs. A comprehensive analysis was then shared with each center with recommendations for future steps.

TABLE 2 – Overall Availability and Accessibility per Center and Pathway

Center	Availability pathways					Accessibility pathways									
	center library catalog	center publications catalog	Institutional archive	FT institutional archive	listed center website	FT center website	FT online	FT online open	FT online agora	center web search indexed	CGIAR VLibrary	AGRIS catalog	Google scholar	Google all	FT online Google books
1	100	95	80	59	100	19	75	43	38	0	4	13	32	70	17
2	77	9	45	45	38	13	59	49	32	27	30	8	48	63	34
3	27	31	91	75	31	35	63	51	26	31	6	40	53	87	8
4	99	20	19	23	93	50	91	61	26	94	90	100	61	94	38
5	100	100	0	0	100	41	76	71	30	99	100	0	53	87	2
6	92	17	92	27	78	24	51	40	27	79	47	3	9	77	8
Average	83	45	55	38	73	30	69	53	30	55	46	27	43	80	18

Availability pathways
 Accessibility pathways
 Average of 6 Centers for a pathway
 Center scores above the average

Table 2 gives an overview of the results for each pathway, for each Center. The average score for each pathway across all the centers is highlighted in yellow.⁸ The percentages highlighted in green show values above the average. From this table, Centers 4 and 5 are out in front, scoring above the average for the six Centers on 11 of the 15 pathways. Outputs of these two Centers are more accessible across the pathways studied than the others.

Many of the Centers make their research outputs readily available through their library catalogue, institutional archive or website. Based on the average, 69% of their journal articles, book chapters and own reports/publications were available in full text (FT) in the catalogues or on websites, however only 53% were accessible to anyone outside of the center via the CGIAR Virtual Library, AGORA,⁹ AGRIS¹⁰ or Google.

Figure 2 shows the same data for selected pathways, with more availability pathways on the left and more accessibility pathways on the right. Overall, Centers are

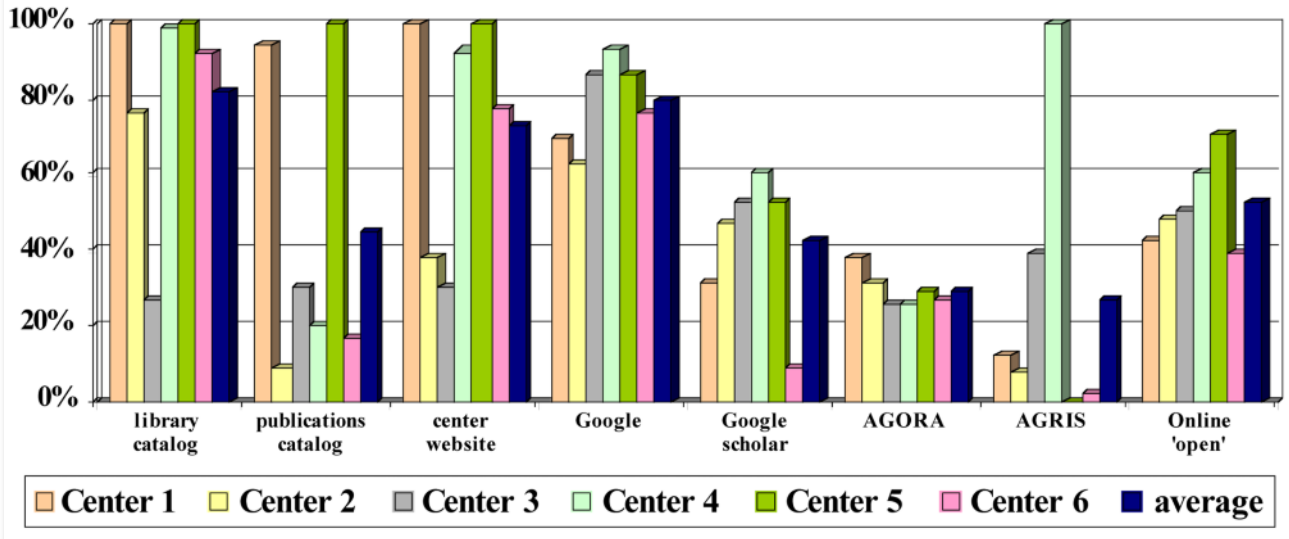
doing better at making their outputs available than they are at making them accessible.

With regard to availability, the library catalogue, publications catalogue and Center website scored on average 83%, 45% and 73% respectively. Two of the Centers scored 100% availability of their 2006 inputs in these pathways. Only 19% of Center 1 research outputs are available full text via its website while Center 5 scores 41%. It is clear that Centers make different use of different pathways.

General Google search services as a pathway to accessibility cast a wide net by tapping into countless websites including open access journal sites and Center and project websites. Searches made here lead to citations, abstracts or full text. They also lead to accessibility through Google Scholar and Google Books. The 80% average score shows that Center outputs are well-indexed by Google; Center 4 scores highest with 94% of its outputs accessible through Google.

Accessibility of outputs in a specialized service like

FIGURE 2 – Availability and Accessibility in Selected Pathways, per Center



AGRIS however reveals another pattern. Center 4 research outputs are 100% accessible in AGRIS, while none could be accessed from Center 5. Since contribution to AGRIS is voluntary and requires additional efforts from each Center, this is a pathway where up to 100% accessibility could be possible, depending on the commitment of each Center and their valuation of the extra services provided by AGRIS.

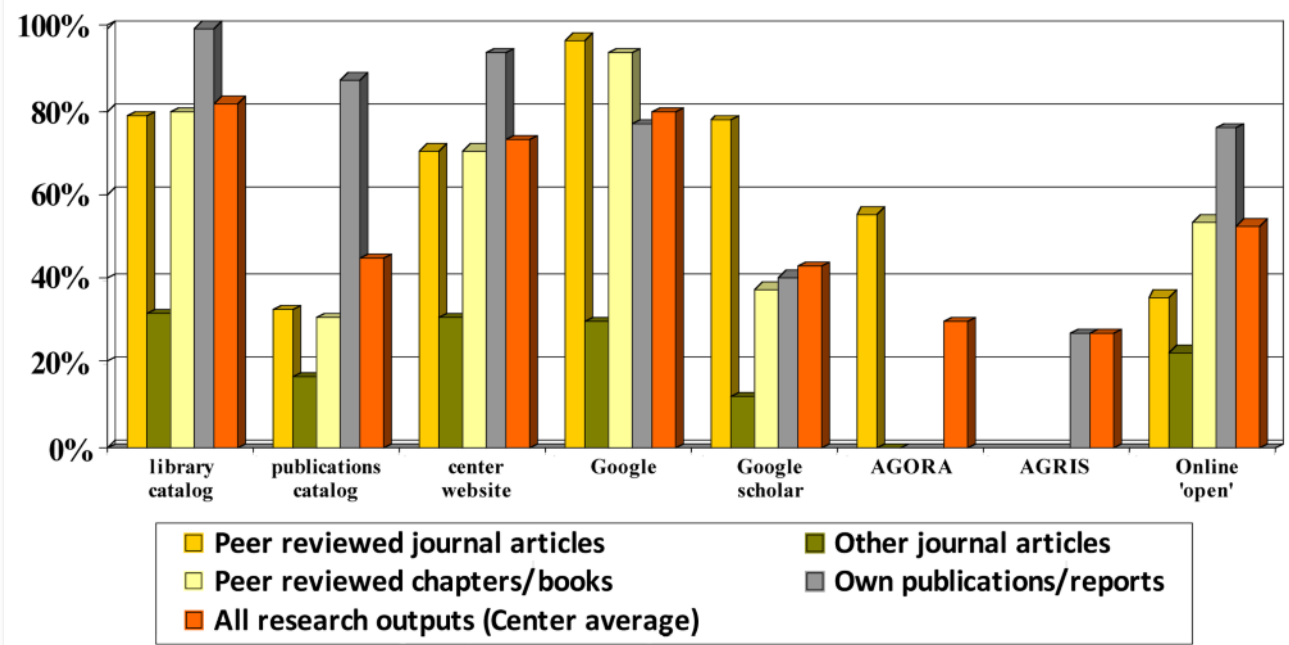
Most telling perhaps is the score for outputs being 'online open.' Here Centers score between 40% (Center 6) and 71% (Center 5) for the full text 'open' accessibility of their research outputs, with an average of 53% for all the 6 Centers.

Figure 3 below gives a more detailed look at the availability and accessibility of different types of research outputs across the various pathways.

Peer-reviewed journal articles, book chapters and books are subject to copyright by publishers and as such, full accessibility is a major issue. While these are available as citations and abstracts via the Center website or Google, the percentage of full text accessibility is low with less than 40% for peer-reviewed journal articles and 54% for peer-reviewed books and chapters.

AGORA is an accessibility pathway for journal articles and is targeted to eligible institutes in developing countries. The results show that close to 60% of the Centers'

FIGURE 3 – Availability and Accessibility by Type of Output



articles are accessible in full through AGORA to this developing-country community, as opposed to less than 40% of journal articles to global audiences. We do not know the extent to which a journal's accessibility affects journal selection choices of authors; an article in an 'AGORA' journal has a 20% higher accessibility score, for researchers in the eligible institutions.

A Center's (own) publications and reports score highly in terms of availability in library catalogues, publications catalogues and on Center websites, yet some 24% of them do not yet seem to be accessible online in full, on any site. Since these outputs are totally under the control of a Center, it is reasonable to expect them all to be available AND fully accessible on the Internet.

Conclusion

With some exceptions, published outputs of the 6 Centers are available (indexed) across several different pathways.

Library and publications catalogues seem particularly strong and widely used. External services like Google Scholar index much of the Centers' outputs, and are particularly strong on external mainstream journals and books.

It is not difficult to identify online the journal articles that are so revered and valued by the CGIAR as a measure of scientific quality. It is quite another challenge to actually get access to these articles, particularly for people with limited online access or financial resources.

Accessibility, in terms of getting to the full content without login restrictions is much lower. Reports published by the Centers score highest in this regard, journal articles score lowest. For eligible users, AGORA bridges the gap to access journal articles somewhat.

For all types of outputs, improvements can be achieved to a greater or lesser extent by adopting and investing in some promising pathways.

These include:

- Capture all scientific publications in a library or similar unit, ensuring that they are properly indexed.
- Deposit full text outputs in a state-of-the-art repository that is harvested and indexed by other services.
- Make sure all Center-produced reports and publications are available and publicly accessible full text in an institutional repository or web platform.
- Pay attention to copyright and licenses for all outputs; paying particular attention to any that are published by third parties in journals, books or other formats. Negotiating open access, republishing and reuse rights increases their accessibility.
- Make authors aware of 'quality-accessibility' choices and tradeoffs so they can find the right balance between outputs that are more or less accessible, and concrete ways they can maximize both.
- Promote open content, open access and open licenses across the institute.

- Promote learning and exchange across professionals working with information, publications and communication in different Centers, passing on good practices that work.
- Use social media¹¹ to promote research outputs in different formats and with wider networks.

Notes

1. We acknowledge the assistance of Maria Garruccio (Bioversity International), Michael Hailu (ICRAF), Edith Hesse (CIAT), Reinhard Simon (CIP), Petr Kosina (CIMMYT) and Helen Leitch (WorldFish), for their support and cooperation throughout this study.

2. See: Snapshot of CGIAR Impacts brochure[o] www.cgiar.org/pdf/cg_impact_brochure_may2005.pdf

3. CGIAR Independent Review Panel. 2008. Bringing Together the Best of Science and the Best of Development. Independent Review of the CGIAR System. Report to the Executive Council. Washington, DC: CGIAR. http://www.cgiar.org/pdf/agmo8/agmo8_independent_review_synthesis_report.pdf

4. See the CIARD initiative for more information: www.ciard.net

5. Ballantyne, P.G. 2008. Making CGIAR Research Outputs Available and Accessible as IPGs. Paper for CGIAR Agricultural Research Public Goods Workshop, Maputo, Mozambique, 27 November. http://ictkm.cgiar.org/document_library/program_docs/ICT-KM%20AAA_complete.pdf

6. The term 'pathway' is used to describe any system, service or tool that is designed to help someone identify, locate, obtain and re-use a research output. From a producer perspective, a pathway is also a service, system or tool designed to share and disseminate a research output. Examples of pathways include: a library catalogue, Google search service, Google Scholar, Scirus, a specialized database such as AGRIS, a blog, or a creative commons license. They are normally used in combination by individuals and information centers to achieve certain objectives.

7. See the CGIAR ICT-KM Program 'triple-A' initiative: <http://ictkm.cgiar.org/what-we-do/triple-a-framework/>. A series of related articles and stories is on the ICT-KM and IAALD blogs: <http://ictkm.wordpress.com/tag/aaa/> and <http://iaald.blogspot.com/search/label/aaa>

8. A pathway is 'scored' according to how well it performs when searched for a specific output. Finding all journal articles from 2006 in the library catalogue gives a score of 100%. Finding none of the identified outputs in a pathway scores zero. Scores and percentages mentioned were right at the time the different Center analyses were done. They are likely to have changed based on improvements and other actions taken by the Centers in the meantime.

9. AGORA (Access to Global Online Research in Agriculture) set up by FAO with major journal publishers enables developing countries to access extensive scientific library collections. Providing a collection of 1278 journals to institutions in 107 countries, membership is limited to countries listed as either band 1 or 2 in the FAO list.

10. AGRIS contains over 2.5 million bibliographic references which provide access to international literature covering agricultural sciences and technology, including grey literature. It is

part of the CIARD (Coherence in Information for Agricultural Research for Development) initiative, in which the CGIAR, GFAR (Global Forum on Agricultural Research) and FAO collaborate to create a community for efficient knowledge sharing in agricultural research and development.

11. See the ICT-KM Program's series on social media tools: <http://ictkm.cgiar.org/2009/07/29/social-media-the-next-revolution/>

About the Authors

MEENA ARIVANANTHAN is Knowledge Management and Sharing Officer with the ICT-KM Program of the CGIAR. — m.arivananthan@cgiar.org

PETER BALLANTYNE is President of IAALD and Head of Knowledge Management and Information Services, International Livestock Research Institute (ILRI). — p.ballantyne@cgiar.org

ENRICA M. PORCARI is Chief Information Officer and leads the ICT-KM Program of the Consultative Group on International Agricultural Research (CGIAR). — e.porcari@cgiar.org