Celebrating an open access knowledge and information repository for CGIAR research
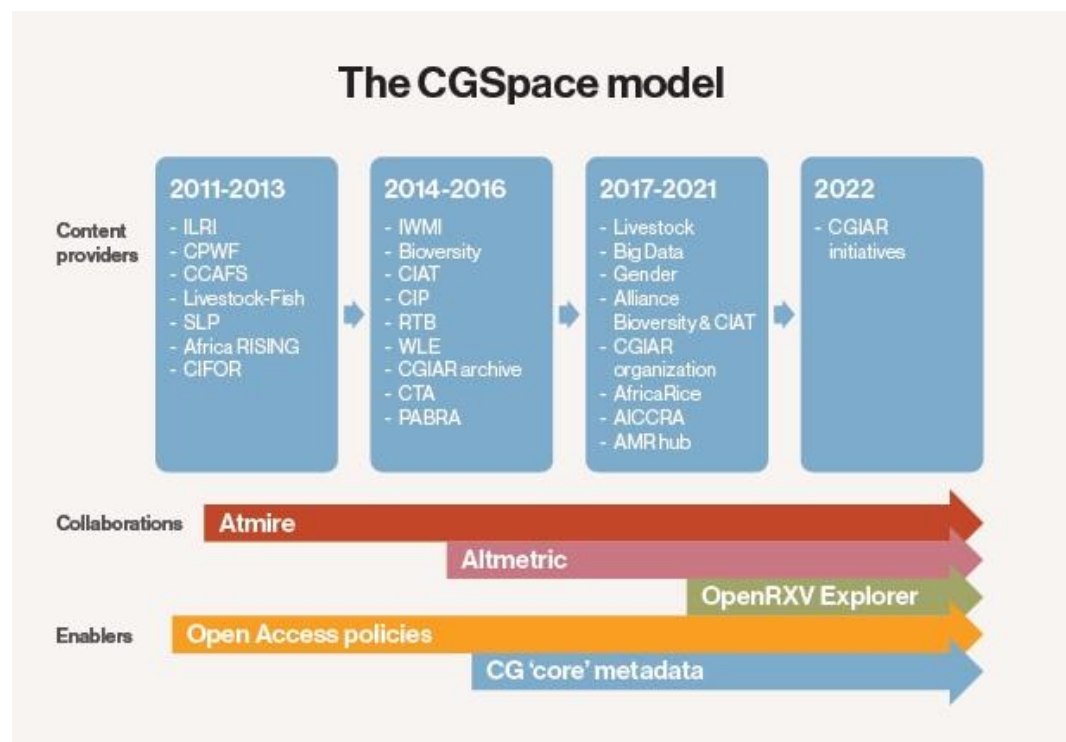
Posted on: 25 Oct 2022

by: **Peter Ballantyne, Abenet Yabowork, Alan Orth and Michael Victor**

In late 2009, the International Livestock Research Institute (ILRI) established a digital repository to act as a complete open archive of the information and publications generated through its research. This was driven by ILRI's aim to have its knowledge travel by making it open access, by publishing it in full, by giving it permanent identifiers, and by providing reliable access through robust open repositories and other communication channels. The goal was to have the information taken off unreliable websites, given permanent identifiers , and made easier to find and share through new digital and web services.
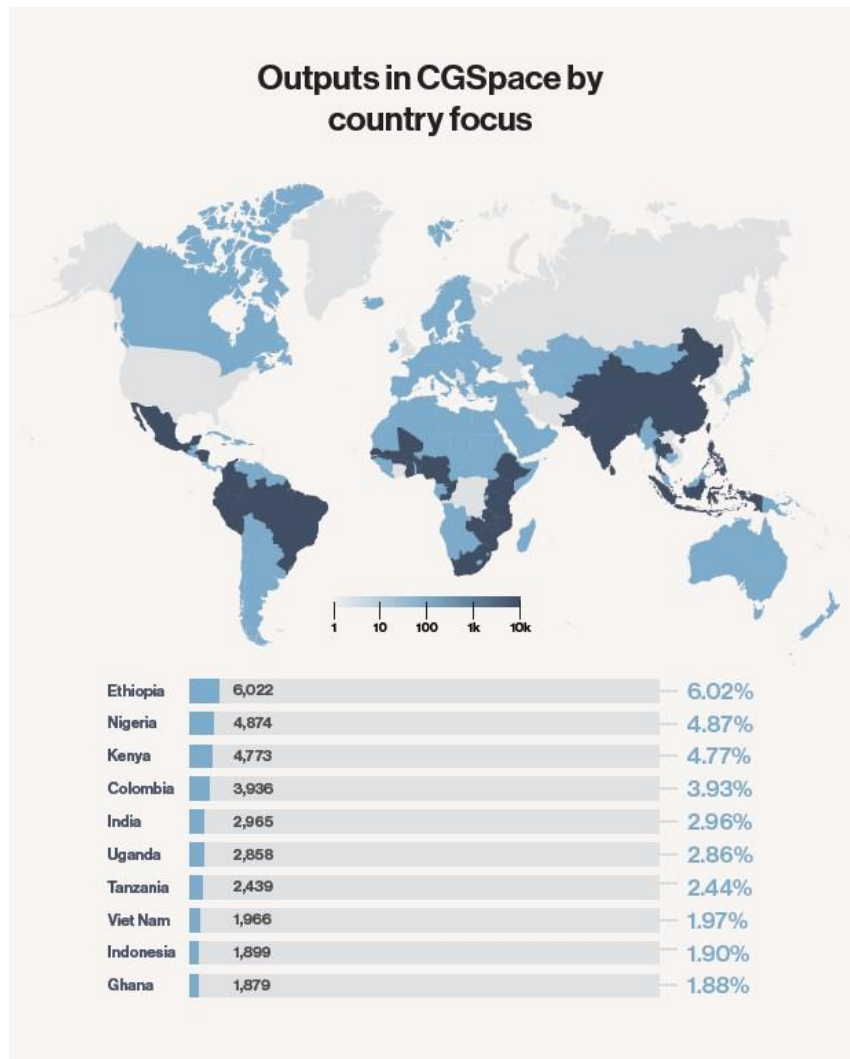
Over time, the repository has evolved from its initial ILRI focus into a collaboration involving seven CGIAR research centers, several CGIAR research programs and platforms and other initiatives and projects. The CGSpace repository is now the largest single collection of CGIAR-associated research outputs (read about CGSpace origins and early developments).

The CGSpace model

Content providers	2011-2013	2014-2016	2017-2021	2022
	- ILRI - CPWF - CCAFS - Livestock-Fish - SLP - Africa RISING - CIFOR	- IWMI - Bioversity - CIAT - CIP - RTB - WLE - CGIAR archive - CTA - PABRA	- Livestock - Big Data - Gender - Alliance Bioversity & CIAT - CGIAR organization - AfricaRice - AICCRA - AMR hub	- CGIAR initiatives

Collaborations: Atmire, Altmetric, OpenRXV Explorer

Enablers: Open Access policies, CG 'core' metadata

Last month, the repository added its 100,000th unique item, reaching an important milestone on the route towards making CGIAR knowledge widely accessible.

Reflecting the collaborative nature of most CGIAR science, the research outputs in CGSpace involve co-authors from thousands of national and international organizations – including all CGIAR entities and partners in national agricultural research systems. The outputs also cover most regions of the Global South, particularly sub-Saharan Africa, South and Southeast Asia and Latin America, which have been the main focus of much CGIAR research.

Today, the repository is a direct resource for people to explore and discover many types of CGIAR research outputs – articles, books and reports as well as posters, presentations and multimedia. While most are aimed at scientists, there are resources and briefs designed for policy and decision makers as well as manuals, guides, brochures and other training resources for extension and advisory services.

Outputs in CGSpace by country focus

Country	Outputs	Percentage
Ethiopia	6,022	6.02%
Nigeria	4,874	4.87%
Kenya	4,773	4.77%
Colombia	3,936	3.93%
India	2,965	2.96%
Uganda	2,858	2.86%
Tanzania	2,439	2.44%
Viet Nam	1,966	1.97%
Indonesia	1,899	1.90%
Ghana	1,879	1.88%

The repository content 'travels' in different ways as well. It is channelled across websites and other services bringing the content to where researchers and information seekers are. CGSpace content is harvested and found through Google Scholar, through online knowledge services like GARDIAN and CORE, via Altmetric, and on websites such as the Alliance of Bioversity and CIAT, ILRI, or the CGIAR Gender Impact platform.

To give information seekers additional options, the CGSpace team recently joined forces with partners to create an open explorer application that provides more structured search options as well as ways to visualize and download search results. This service is able to search across several DSpace repositories (including ICARDA and WorldFish) and will hopefully be extended in the future to include others. This structured discovery is helped by the collaboration with other CGIAR centers and partners on metadata and terminology standardization – using 'CG core' metadata and constant quality checking and curation to make the content as consistent as possible.

It does not only serve public audiences. CGSpace has been used to support internal performance and output reporting of ILRI scientists and it is used by CGIAR research programs (2010–2021) and the current One CGIAR research initiatives to record evidence of outputs included in performance reports.
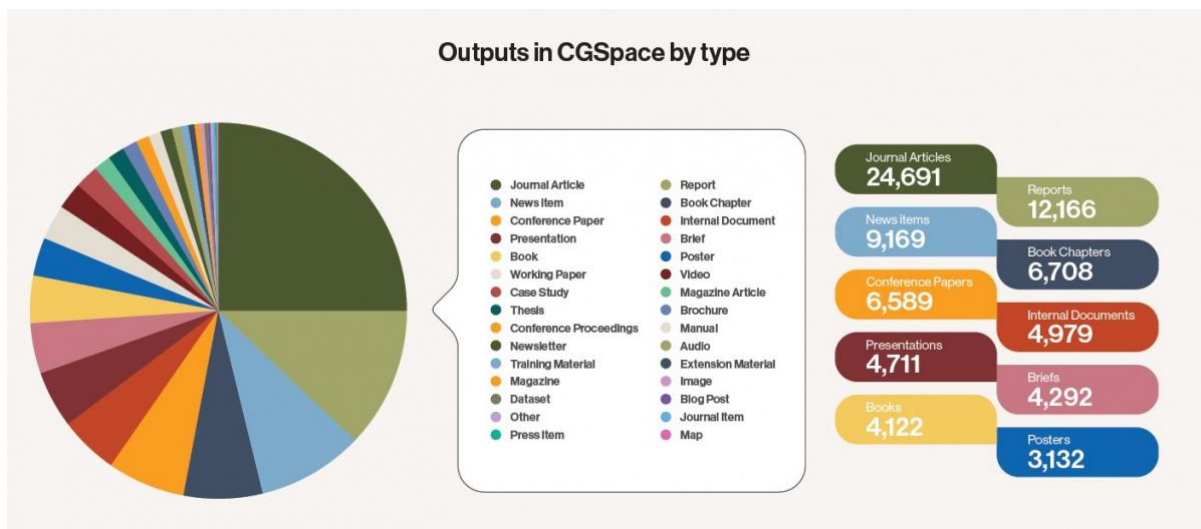
The immediate future is bright as CGSpace currently serves as the joint repository for all research outputs of the new CGIAR regional and research initiatives and impact platforms.

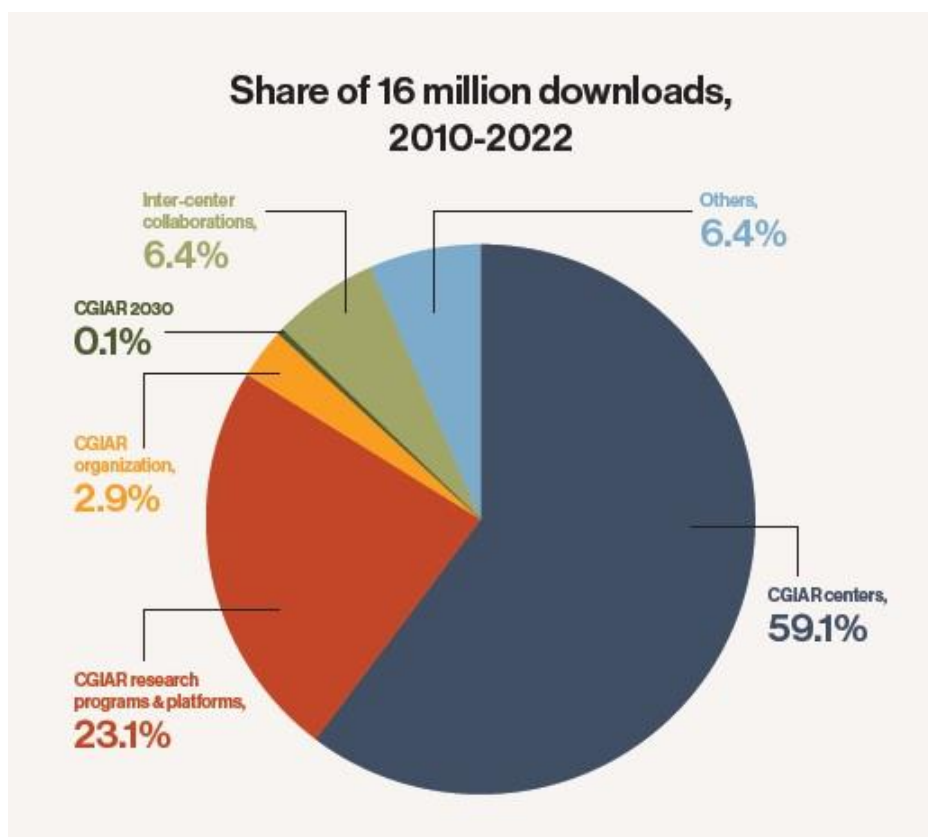As we reflect on the 100,000 milestone, the back end of CGSpace shows up some interesting indicators.

First, around 2010–2011, when momentum on the repository started, ILRI and other CGIAR centers adopted various open access policies and, looking at accessibility data then and now, we can see around 90% of all CGSpace products today are open access in some form, versus around 75% or lower 10 years ago. The repository has been an important part of that drive towards open access, providing reliable publishing channels for diverse products, unique and permanent identifiers, and a means to promote the outputs to different audiences.

Accessibility of CGSpace content

Publication date	Total items	Openly accessible
2020-2022	12,120	91%
2015-2019	22,323	85%
2010-2014	18,895	76%
2005-2009	11,517	59%
2000-2004	11,797	68%
1995-1999	8,100	60%
1990-1994	6,373	57%
1985-1989	4,901	64%
1980-1989	2,077	65%

Second, while the Altmetric data shows high attention scores mainly for articles and books, the download data for items in CGSpace show that the most popular items include more practical, extension-oriented manuals as well as more scientific outputs. The practical outputs are as useful as the usual scientific articles and books.

Outputs in CGSpace by type

Legend:
- Journal Article
- News Item
- Conference Paper
- Presentation
- Book
- Working Paper
- Case Study
- Thesis
- Conference Proceedings
- Newsletter
- Training Material
- Magazine
- Dataset
- Other
- Press Item
- Report
- Book Chapter
- Internal Document
- Brief
- Poster
- Video
- Magazine Article
- Brochure
- Manual
- Audio
- Extension Material
- Image
- Blog Post
- Journal Item
- Map

- Journal Articles: 24,691
- Reports: 12,166
- News items: 9,169
- Book Chapters: 6,708
- Conference Papers: 6,589
- Internal Documents: 4,979
- Presentations: 4,711
- Briefs: 4,292
- Books: 4,122
- Posters: 3,132

Third, the repository has recorded around 14 million item views and 16 million file downloads over the past 12 years. Many of the item views are for outputs where the article or product is physically online elsewhere, so, while a download is not reported the user is sent to the individual item. The outputs of center research make up about 60% of the downloads, the CGIAR research programs and platforms that operated from 2012–2021 accounted for a further 23% of the downloads. The data show that some authors and outputs are much more downloaded than others, but no obvious pattern across centers or programs.

Share of 16 million downloads, 2010-2022

- Inter-center collaborations, 6.4%
- Others, 6.4%
- CGIAR 2030, 0.1%
- CGIAR organization, 2.9%
- CGIAR centers, 59.1%
- CGIAR research programs & platforms, 23.1%

As we look forward, the repository will undergo a further transformation to meet the emerging demands of the One CGIAR and its partners, hopefully retaining its focus on open content, using open code and software, working in collaboration and with industry experts, and delivering high quality products and services that can be easily and efficiently consumed and used by others and put to work to advance the transformation of food, land, and water systems in a climate crisis.

View Infographic: *CGSpace - an open access knowledge and information repository for CGIAR research*