Evaluation of CGIAR Platform for Big Data in Agriculture: Annexes

December 2021

LINK TO FULL EVALUATION REPORT

Contents

Annex 1: Methodology

Annex 1.1 Overall Approach

The evaluation approach followed a mixed methods design (qualitative and quantitative) to collect data and assess the CGIAR Platform for Big Data in Agriculture (hereinafter the Platform)'s achievements and outcomes. Definitions of mixed methods in evaluation borrow many of the concepts from mixed methods in research and social sciences (Mertens, 2017).

"Mixed methods research is a type of research in which a researcher or team of researchers combines elements of qualitative and quantitative research approaches (e.g., Use of qualitative and quantitative viewpoints, data collection and analysis, inference techniques) for the purposes of breadth and depth of understanding and corroboration" (Johnson R.B. and Turner, 2007)

Among the quantitative methods used, the evaluation team designed and administered an online survey to reach a wide range of stakeholder groups (data users and partners). The qualitative methods consisted of semi-structured interviews, document analysis, and two case studies (individual winner of the Inspire Challenge and a Community of Practice). Both qualitative and quantitative methods complemented each other in ways that together brought credible evidence to answer the evaluation questions.

The evaluation approach and methods tended to embrace a systems-thinking approach to capture interlinked issues in this innovative program. A mixed methods approach was chosen to capture the complexity in different dimensions, based on the five dimensions of complexity presented by Bamberger *et al* (2016). These elements can indeed be recognized in this evaluation: (1) **Embeddedness and Nature of the System**: the nature of context in which the Platform operated related to the history of CGIAR and its members and programmatic components (Centers, CGIAR Research Programs [CRPs], CGIAR System…) and (2) **Institutions and Stakeholders:** the diversity of stakeholders involved with different capacities, priorities, governance, funding, and implementation models, and (3) **Causality and Changes:** complexity is also linked to the fact that the Platform aims to stimulate a cultural change inside CGIAR by embracing new digital tools and adopting new standards in data management and stewardship.

Qualitative inquiry tended to be exploratory using open questions and snowball sampling (interviewees suggesting other people to meet the evaluation team). Data collection methods were reviewed after better familiarization with the context and the type of evidence available (see data collection methods). Thus, the exploratory nature of the inquiry assessed, to the extent possible, any systemic and transformational changes among CGIAR Centers as well as among Communities of Practice (CoPs) and external partners. The aim is to understand if there are any intended or unintended changes (for example new collaborations, policy changes, raised awareness, new internal and external capacities, etc.,) and identify elements that show the extent to which the Platform fostered or contributed to those changes.

1.1.1. FAIR Guiding Principles:

The assessment of data management and stewardship followed the 'FAIR Guiding Principles' (Wilkinson et al., 2016). The guidelines intend to improve the findability, accessibility, interoperability, and reuse of digital assets. A diverse set of stakeholders – representing academia, industry, funding agencies, and scholarly publishers – came together to design and jointly endorse this concise and measurable set of principles that are referred to as the **Findable, Accessible, Interoperable,** and **Reusable** (FAIR) Data Principles. The Platform adopted the Netherlands Institute for Permanent Access to Digital Research Resources (DANS) Metrics for FAIR compliance and developed the GARDIAN Guide for FAIR Data. The rationale for this new guide is that FAIR principles are not orthogonal and have not been designed for automated machine-based evaluation.

> **Box 1:** The FAIR Guiding Principles for Scientific Data Management and Stewardship (Wilkinson et al., 2016).
>
> **To be Findable:**
>
> F1. (meta)data are assigned a globally unique and persistent identifier
>
> F2. data are described with rich metadata (defined by R1 below)
>
> F3. metadata clearly and explicitly include the identifier of the data it describes
>
> F4. (meta)data are registered or indexed in a searchable resource
>
> **To be Accessible:**
>
> A1. (meta)data are retrievable by their identifier using a standardized communications protocol
>
> A1.1 the protocol is open, free, and universally implementable
>
> A1.2 the protocol allows for an authentication and authorization procedure, where necessary
>
> A2. metadata are accessible, even when the data are no longer available
>
> **To be Interoperable:**
>
> I1. (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.
>
> I2. (meta)data use vocabularies that follow FAIR principles
>
> I3. (meta)data include qualified references to other (meta)data
>
> **To be Reusable:**
>
> R1. meta(data) are richly described with a plurality of accurate and relevant attributes
>
> R1.1. (meta)data are released with a clear and accessible data usage license
>
> R1.2 (meta)data are associated with detailed provenance
>
> R1.3. (meta)data meet domain-relevant community standards

Annex 1.2 Data Analysis

The evaluation matrix forms the main analytical framework in this Platform evaluation. It sets out how each evaluation question and evaluation criteria are addressed, and breaks down main questions into sub-questions, mapping them against the indicators, data collection and analysis methods, and/or lines of inquiry, and sources of information. It ensured that all data collected are analyzed and triangulated, and supports the identification of evidence gaps. As such, the evaluation matrix ensures that the evaluation design is robust, credible (reducing subjectivity in the evaluative judgment), and transparent. The evaluation team revised the matrix based on an in-depth desk review and the feedback provided by peer reviewers, the Platform management team, and the CAS team. The evaluation matrix can be consulted in Annex 2.

The validation of results and quality assurance (QA) relied on triangulating data and findings from different sources and using different methods. This approach also helped in ensuring to the extent possible transparency, independence of judgment, and minimization of bias. A two-stage analysis was conducted: Three Component Studies (CS) covering the three main Platform clusters of activities (Modules) were conducted concurrently and serve as the main input for the final evaluation report. In addition, internal and external peer reviews help to strengthen the soundness of the articulated findings, lessons learned, and recommendations. Presentation of findings uses verifiable evidence and robust inference pathways from evidence to conclusions and from conclusions to recommendations.

In line with CGIAR evaluation policy and international quality standards (American Evaluation Association's [AEA's] Guiding Principles for Evaluators), the evaluation approach ensures the integration of the following principles: participatory, learning-oriented, utilization-focused, and gender-responsive. Participation implies that the continuous involvement of different stakeholders throughout the evaluation process leads to conclusions and recommendations that are more widely accepted, and thus more likely to be acted on, and more likely to lead to the envisaged outcomes. To the extent that their time allowed, the Platform management team was invited to participate actively in the evaluation through the review of the inception report, data collection instruments (online survey protocol and its administration), and in the interpretation of the results. Stakeholder groups consulted were inclusive of stakeholders' categories and subcategories (see stakeholders' mapping) and special attention was given to gender balance and youth inclusion among the interviewed persons. Likewise, data collection was conducted in a way that ensured understanding, respect, and confidentiality of stakeholders' views and perceptions (see interview guide in Annex 3).

Finally, this Platform evaluation was designed as an entirely desk-based exercise; it followed a pre-determined process, guided by the validated Terms of Reference. The evaluation study considered findings and information elaborated under the framework of internal monitoring and evaluation (M&E) efforts across the Platform implementation and previous efforts of synthesis of evaluative evidence in CGIAR.

CAS Secretariat's guided and quality assured the evaluation process and the CAS team contributed by reviewing the inception report, data collection instruments and participating in weekly meetings to comment on the early results of the component studies results (22/9/2021).

Annex 1.3 Data Collection Methods

Data collection methods were designed to answer the evaluation's questions first under the framework of three Component Studies (CSs); each study covered a cluster of activities corresponding to the three Platform modules: ORGANIZE (CS1), CONVENE (CS2), and INSPIRE (CS3). The studies followed the same analytical framework around the evaluation criteria and questions outlined in the evaluation matrix, however, not all the component studies benefited from the same team level of effort by subject matter experts (there was more focus on ORGANIZE and CONVENE), the INSPIRE study was designed to build on the findings of the recent internal review conducted by the Platform team. Some data collection methods have been used exclusively for a CS (i.e., testing the GARDIAN platform) while others were used across the three studies (online survey, interviews of the Platform management team, Center focal point interviews…).

Annex 1 - Table 1: Key features of the Three Component Studies

	CS1- ORGANIZE	CS2- CONVENE	CS3- INSPIRE
Focus	Data generation, access, and management.	Collaboration and conventions around big data and agricultural development.	Inspire Challenge competition and how big data can deliver development outcomes.
Main Target Group	Data users Data managers	Members of Communities of practice (CoPs), participants to the Conventions, and Capacity development beneficiaries	Competition candidates
Cross-cutting themes	Gender, Youth, Climate Change (CC), Capacity Development (CapDev)		
Study design	Mixed quantitative and qualitative	Qualitative	Qualitative
Data collection methods*	Platform Statistical Analytics Interviews Stakeholders' survey	Interviews Stakeholders' survey	Interviews Descriptive case study Complementing the 2021 Inspire Challenge Review
Sampling technique	Representative and purposeful	Purposeful	Purposeful

* Here the main method for each CS is specified but all CSs used and triangulated with other data collection methods results.

1.3.1 Desk Review

The desk review focuses on existing internal and external documentation. This exercise started at the inception phase to obtain an initial understanding of the Platform's structure, governance and management, implementation, and an initial appraisal of output achievements. The desk review covered several types of internal and external documents: strategies, evaluation reports, annual reports, etc; and the 2021 Synthesis results. The list of consulted documents can be found in Annex 7. From the desk review, the evaluation team assembled the list of stakeholders and their categories (see Inception Report). This helped the team to ensure the participation of most of the categories of stakeholders. Desk review and research continued throughout the evaluation process and with a view to identifying preliminary answers to the evaluation questions.

Document Analysis and Use of Evaluative Evidence

Like other analytical methods in qualitative research, document analysis requires that data be examined and interpreted in order to elicit meaning, gain understanding, and develop empirical knowledge (Corbin & Strauss, 2008; see also Rapley (2007) cited in Bowen (2009). The evaluation team selected and conducted a keyword search to assess the extent to which issues around Open Access /Open Data (OAOD) principles have evolved within CRPs and Centers (i.e. semantic standards, and ontologies, analytic environments, etc.). This can involve the identification of changes of scientific culture embracing the Platform and principles therein, evaluated from the evidence of the recent scientific practices and outputs. This can be quantitative and qualitative and is focusing on reflecting on the level of adoption of the principles and the tools available (via the Platform).

The evaluation team (and CAS) identified six key documents dated after 2017 (after the launch of the Platform) to be used to contrast with preliminary findings. These are listed in the table below.

Annex 1 - Table 2: Mapping of evaluative evidence related to the Platform

#	Topic/Evaluand	Title	Year Published	Type
1	INTELLECTUAL ASSETS PRINCIPLES	Review of CGIAR Intellectual Assets Principles-2017	2017	Review
2	OPEN ACCESS/ OPEN DATA	Review of CGIAR's Open Access/Open Data Policy and Implementation Support-2018	2018	Review
3	BIGDATA	CGIAR Platform Performance Management Standards (PMS) Pilot Assessment Report, Big Data	2019	Assessment
4	DIGITAL STRATEGY	Toward a digital one CGIAR: Strategic research on digital transformation in food, land, and water systems in a climate crisis	2019	Assessment
5	INSPIRE CHALLENGE MODULE	CGIAR Platform for Big Data in Agriculture - Inspire Challenge Review (2017-2020)	2021	Review
6	CGIAR RESEARCH PROGRAMS	2021 Synthesis of Learning from a Decade of CGIAR Research Programs	2021	Synthesis

1.3.2 Qualitative: Key Informant Interviews

The team conducted 51 Key Informant Interviews (KIIs) via videoconference, apart from one face-to-face interview[1] (see Annex 4 Interviewees Profile). The team did not conduct Focus Group Discussions due to differences in time zones and the limited availability of many stakeholders. The identification of interviewees tended to be as inclusive as possible of the categories and subcategories identified in the stakeholder mapping and then the evaluation team used a snowball approach to identify individuals and adjust the gender imbalance as women were underrepresented. We approached interviews as discussions loosely structured around a small number of key issues. The aim was to encourage stakeholders to talk

---

[1] Seizing the opportunity of the presence of one of the evaluation team members in the same country as an inspire challenge winner.

freely about what they consider important. Most of the interviews were conducted by three evaluation team members (a lead facilitator, observer, and co-lead) and a note-taker. KII results were used to triangulate the quantitative and qualitative data collected through desk research and the online survey.

The analysis of qualitative information collected from interviews followed the evaluation matrix structure and levels of inquiries. In addition, the team quantified stakeholders' sentiments about the worth of four key components of the Platform (Figure 1 below). This exercise was intended to strengthen consistency across the three Component Studies (not all team members participated in all the interviews) and address any potential expert bias in interpreting the results.

Annex 1 - Figure 1: Key informants sentiments summary about the key components of the Big Data Platform (Source: KII notes)
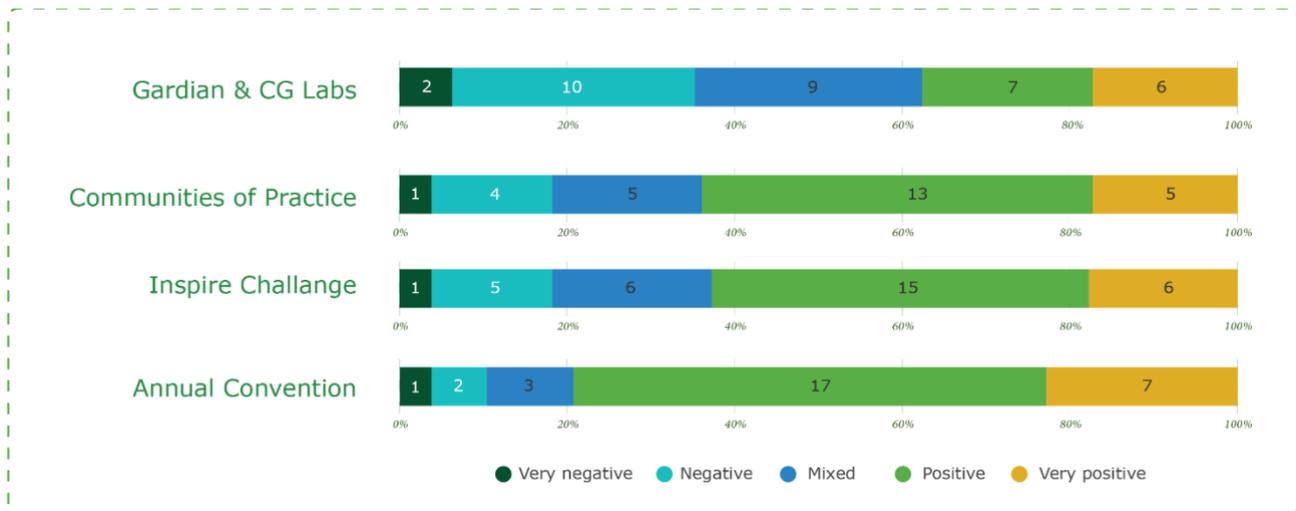
1.3.3   Qualitative: Two Short Case Studies

The evaluation team conducted two case studies: one on the Ontologies CoP and one on an individual winner of an Inspire Challenge award. These two cases were purposefully selected to showcase the diversity of outputs and their related outcomes, failures, and successes, as well as harvesting unintended outcomes and challenges. The selection of the CoP ontologies was motivated by the fact that preliminary interviews and desk research showed that it had achieved outstanding outcomes, while the choice of the Inspire winner was simply motivated by the opportunity of meeting the winner in person. The study about the individual experience will not be included in the final report to avoid any identification of the person who preferred not to be cited. The analysis framework will follow the key criteria and inquiry levels of the evaluation matrix.

1.3.4   Quantitative: Platform Statistical Analytics

In relation to data collection, the usage analytics came from Google Analytics on GARDIAN (available from March 2020), bigdata.cgiar.org analytic summaries (provided in September 2021), GARDIAN Platform analytics in terms of achievements, and content analysis of the MySQL dump of the metadata database provided in August 2021. There was also a corpus of the keyword searches on GARDIAN (provided in September 2021). Membership lists to CoPs were also provided (anonymized). A full report of the Google Analytics usage of GARDIAN was initially performed in August 2021 and is available in the CS1 report. Some of these results were contrasted in the key findings with other summary analytics. This data collection method was complemented by interviews and an online survey that targeted users of the above-mentioned tools.

1.3.5   Quantitative: Online survey

The evaluation team conducted an online survey targeting all stakeholder groups of the Platform and partners including CoP members. The evaluation team used SurveyMonkey software for survey administration and analysis between September 20-30 2021. The survey was designed in a way that respondents were directed to a set of specific questions based on the type of engagement with the Platform. The survey was sent by the Platform management team to comply with GDPR data privacy

regulations: contact lists cannot be shared with the evaluation team. Two versions of the survey were released, one in English and one in Spanish (9 out of 110 responded in Spanish) to the following groups:

- All subscribers of the Platform Newsletter (2803 subscribers)
- Participants of the annual conventions – those who opted in for correspondence
- Participants of the Inspire Challenge (applicants, finalists, and non-finalists)
- External partners (Academia, private companies, NGOs...)
- Internal partners (CRPs, CGIAR Centers, projects.)
- Users of CG Labs.

The detailed results of the online survey can be found in 5.

Annex 1.4 Phases of the Evaluation

1.4.1 Inception Phase

The inception phase was dedicated to fine-tuning the evaluation plan and methodology and gaining a first understanding of the Platform program. An induction meeting took place via videoconference on July 20 2021 between the evaluation team and the CAS Secretariat.

As presented above, the inception phase focus was on the following elements:

- Preliminary project theory model(s); refinement of the evaluation questions, elaboration of evaluation methodology including quantitative and qualitative approaches through an evaluation framework ('evaluation matrix').
- A stakeholder analysis identifying key stakeholders, networks, and channels of communication;
- Program timeline and key outputs achievements based on Platform annual reports and the 2016 proposal.
- Division of roles and responsibilities between the evaluation team members.
- People to be interviewed and surveys to be conducted and a debriefing and reporting timetable.

As a requirement to finalize the inception report, a consultation was arranged between CAS Secretariat, the evaluation team, and peer-reviewers to interrogate the evaluation approach and methodology and enhance the evaluation matrix.

1.4.2 Data Collection Phase

The data collection phase was meant primarily to collect data from desk research and stakeholder consultations (Key Informant Interviews, Focus Group Discussions, online surveys, etc.). The evaluation team collected the evidence according to the plan, completed its analysis, and prepared a preliminary list of findings and conclusions.

1.4.3 Reporting Phase

In the reporting phase, the evaluation team prepared a presentation of preliminary findings to debrief the CAS Secretariat and the Platform Management team and to seek validation, factual corrections, and feedback. The team developed three Component Studies and a draft evaluation report for CAS Secretariat comments and factual corrections. Under CAS Secretariat's guidance, the report was reviewed by a team of external peer-reviewers. With the feedback from relevant stakeholders, the evaluation team finalized the evaluation report considering comments according to the team's judgment.

1.4.4 Management Response

During this phase, CAS Secretariat will liaise with the Platform management and other stakeholders as deemed necessary in light of ONE CGIAR reform and new operational structures. The management response will be published on the CAS Secretariat website.

1.4.5 Dissemination

The evaluation report, the executive summary, and the evaluation brief, and other knowledge products along with the management response, will be published on the CAS Secretariat's website. In line with the dissemination and knowledge management strategy developed at the inception phase, tailored presentations will be made to targeted stakeholders and learning events organized with internal and external stakeholders.

Annex 1 - Table 3: Phases of the evaluation

Evaluation Phase	Tasks	Outputs	Responsible	Dates2021
Preparatory	Draft evaluation ToR /ToR Revisions	Final evaluation ToR	CAS Secretariat	9 July
	Selection of consultants from the vetted roster	Evaluation team contracts		
Inception	Onboarding and briefing of the external evaluation team	PPT	CAS & Evaluation team lead	21 July
	Development of the inception report with the evaluation matrix	Draft inception report with evaluation matrix	Evaluation team	26 July
	Introduction consultation with the Platform management, and validation of the Inception Report	PPT	Platform Management	28 July
	Peer review of the methodology and approach.	Final inception report and evaluation matrix	Evaluation Team	30 July
Inquiry	Desk review	Survey result notes	Evaluation Team	1- 11 October
	Survey			
	Interviews	Interview notes		
	Data triangulation for developing Module component studies	3 Module Component Study reports		
Reporting	Analysis and report development	Detailed report outline for feedback to CAS		
	Validation workshop with Platform management	PPT	Evaluation Team and CAS Secretariat	26 October (TBC)
	Submission of draft Platform evaluation report	Draft Platform evaluation report	Evaluation Team CAS with peer reviewers	1 November
	Report review by CAS, peer-reviewers and key stakeholders as needed.	Compiled feedback by peer-reviewers and key stakeholder groups.		5 November
	Integrating CAS and peer-review feedback into discussion version of the report.	Draft final report		5 November – 4 December 4
	Presentation of Draft final Report to SIMEC for feedback	Draft discussion version of the final report, PPT	CAS Secretariat with selected SMEs	

Evaluation Phase	Tasks	Outputs	Responsible	Dates2021
	Revision of the final report integrating SIMEC's feedback	Revised draft discussion version of the Final Report	Evaluation Team	
	Presentation of Final Report to System Council	Draft final report. PPT	CAS Secretariat/Evaluation Team	
		Final report	Evaluation Team	10 December
Management Response	Project Coordination, Monitoring, and Performance Unit liaise to obtain Management Response coordinated by Project Coordination, Monitoring, and Performance Unit.	Management response	Platform Management, liaise with CAS Secretariat, Platform Project Coordination, Monitoring, and Performance Unit.	December 2021
Dissemination	Development of knowledge products and knowledge management in line with the dissemination strategy for the evaluation.	Evaluation briefs and knowledge products.	CAS Secretariat/evaluation team where necessary.	December onwards

Annex 2: Revised Evaluation Matrix

The evaluation team revised the matrix based on an in-depth desk review and feedback provided by peer reviewers, the Platform management team, and the CAS team

Key Evaluation Questions	Sub-Questions	Indicators /Evidence	Data collection methods
		Relevance	
1. To what extent are the Platform's objectives relevant to the needs of its internal and external partners and stakeholders, including end-users in target groups?	1.1 Were the Platform design and approaches aligned with Centers, partners' and end-users' priorities and capacities?	- Internal and external Partners' opinions about alignment of the Platform objectives with partners' priorities and capacities. - Internal Stakeholders indicate that the Platform rationale is coherent with applications of big data in agriculture research for development. - Evidence indicates that the design was appropriate to allow synergies with 3 other CGIAR platforms - Stakeholders confirm that the Platform is relevant in comparison to what is/was already available in other scientific groups in agriculture or other domains.	- KIIs/FGDs (Partners & Platform G&M team) - Online Surveys (Partners including CoP members) - Documents Analysis/ Synthesis of Evaluative Evidence
	1.2 To what extent have cross-cutting themes (Gender, Youth, Climate Change, Capacity Development) been considered in Platform design?	- Number (and timeliness) of guidelines/strategies about the integration of cross cutting themes. Also, level of implementation (based on action plans). - Coherence between Platform design and CGIAR cross cutting themes strategies and guidelines. - Number (and evolution over the years) of specialized partners engaged by the Platform to strengthen relevance and effectiveness of cross cutting themes integration across program levels: design, implementation, monitoring and evaluation.	- Desk Review - Documents Analysis - KIIs/FGDs - Case studies (Inspire Projects) - Publications relating to these themes, enabled by the Platform

Key Evaluation Questions	Sub-Questions	Indicators /Evidence	Data collection methods
	1.3 How flexible is the Platform's design and mechanisms to local and evolving constraints including COVID-19 Pandemic?	- Stakeholders confirm that the Platform design allow for flexibility to achieve future and further development of the Platform / implemented technologies are flexible enough to allow upgrading and evolution depending on new technologies and new desired features, and new concepts. Decisions were taken and implemented in a timely fashion to respond to the evolving context, needs, including COVID-19 Pandemic (examination of the timeline of decision-making process and its implementation).	- KIIs/FGDs - Online Surveys Desk Review / Documents Analysis
		Efficiency	
2. Have resources (funds, human resources, time, expertise etc.) been allocated strategically and timely to achieve Platform outcomes?	2.1 How adequate has the technical, institutional, and administrative support from the Platform's CGIAR internal partners been?	- Platform management's testimonies about partners' support: strengths and weaknesses. - Partners' opinion about Platform capability to manage resources and partners (agile management).	- KIIs/FGDs (Partners & Platform G&M team) - Online surveys Desk review
	2.2 How efficient was the implementation: use of resources, timeliness?	- (%) of planned outputs achievement across the three modules. - Number and length of delays to achieving outputs. - Availability. timeliness for decision making and quality of progress reports. - Stakeholders' feedback about the quality of outputs in relation to each objectives and targets for each module. - Stakeholders' satisfaction about their level of participation of achieving in contributing to the objectives and the planned outputs.	- Desk Review - Document Analysis - KIIs/FGDs (Platform G&M)
	2.3 Were the resources allocated to integrate cross cutting themes sufficient to reach the desired outcomes?	- (%) of budget and resources (expertise, staff time…) exclusively allocated to address integration of cross cutting themes.	- Desk Review (Budget & Workplans) - KIIs/FGDs

Key Evaluation Questions	Sub-Questions	Indicators /Evidence	Data collection methods
		Effectiveness	
3 To what extent did the Platform achieve progress towards outcomes?	3.1 To what extent did the Platform achieve **progress on the outcomes** noted in the proposal?	- Evidence showing that outputs have led to planned (unplanned) changes. - Stakeholders' perception about Platform evidenced effects on the digital agriculture ecosystem. - Evidence showing no discriminated outcomes based on gender, age, regions, sector… - Evidence shows that the Platform enabled Centers to comply with CGIAR's Open Access and Open Data Management (OA/DM) Policy: e.g. improvements in organizational data policy and its implementation. increased familiarity with OADM policy. - Stakeholders' testimonies and results of the synthesis of evaluative evidence about limiting and contributing factors. Stakeholders' opinion about how effective has the Platform been in identifying opportunities for targeting digital innovation in pursuit of policy and institutional reform globally. Projects initiated/supported by the Platform show contribution to digital innovations for research and/or delivery of research.	- Desk Review/Document analysis (*Externally generated documents*) - Online Surveys (Partners, CoP members + CGIAR) - KIIs/FGDs - Results from GARDIAN and SCiO re: open data and publications over time, licensing regimes over time...
	3.2 Has the Monitoring, Evaluation and Learning (MEL) system facilitated (or not) achievement?	- Evidence showing that Platform Monitoring system has informed timely and agile decisions making list of threats & risks/opportunities that have been identified and addressed. - Evidence showing that Platform Evaluation has informed about the worth and merit of Platform design and implementation. - Evidence showing that Platform Learning processes have enhanced design, implementation, and collective learning.	- Desk review/Document analysis (*MEL reports, Annual reports, data sharing agreement templates, resources available via GARDIAN and Big Data web page -- e.g. responsible data guidelines*) KIIs/FGDs

Key Evaluation Questions	Sub-Questions	Indicators /Evidence	Data collection methods
4. How effective has the Platform been in building digital capabilities and partnerships supporting CGIAR research?	4.1 To what extent has the Platform accelerated partners' progress towards better data and knowledge management and stewardship?	- Platform demonstrates **increased performance** over the years with regards to: Searches for data, joint publications based on co-curated data, invitations to host panels at conferences, expressions of interest in online courses via the big data platform. - Evidence shows that number of **jointly published articles** and citations of co-designed research assisted by the platform has increased over the years; Number of publications acknowledging the Big Data Platform. - Stakeholders confirm that outputs have contributed to enhancing CGIAR and partners' capacity to deliver big data management, analytics, and ICT-focused solutions to CGIAR target geographies and communities. - Evidence shows that the Platform has strengthened capacities to store and maintain sequence data according to the FAIR principles. - **GARDIAN** Performance analytics of a representative sample of publications (articles and data sets) shows: a sustained positive progress against GARDIAN FAIR METRICS and that the progress is balanced among Centers, regions and sectors/research topics (CRPs). - Evidence shows positive progress on: **GARDIAN** traffic, data use, users' satisfaction, increasing evolution of number of articles and data sets published. - Users' feedback is positive about the time needed to find the data looked for, and how convenient/practical it is when uploading/creating a dataset (and its metadata).	- Online Surveys - Desk Review/Document analysis - KIIs/FGDs - Platform analytics (GARDIAN, CG Labs)

Key Evaluation Questions	Sub-Questions	Indicators /Evidence	Data collection methods
		- Evidence about incentives and capacity building efforts to reduce gaps to foster meta data publication in **GARDIAN** in support of cross-cutting themes of gender and youth empowerment. - Stakeholders' feedback about relevance and effectiveness of third parties' providers in **GARDIAN** around data gathering, data co-curation and data analysis and management through ability of the platform to integrate applications of artificial intelligence, SMS machine and remotely sensed data. - Evidence shows that the Platform enabled external partners to have their data searchable alongside CGIAR data via GARDIAN: the number of external repositories made discoverable via GARDIAN. - Evidence of big data platform optimizing repetitive tasks in agriculture research for development through exploration of open source software applications (e.g. R and Python).	
	4.2 To what extent has the Platform enabled CGIAR to engage with the **wider Agriculture data (and innovation) digital ecosystem** in terms of both depth of engagement and reach?	- Evidence shows that the Platform enabled (or not) the expansion of external engagement in technical communities of practice to stay abreast of digital innovations related to CGIAR research domains. - Number and type of new partnerships initiated by the Platform that have strengthened CGIAR engagement with the wider Big Data communities. Among them (%) and types[2] of specialized in cross-cutting themes. Conventions:	

---

[2] CSOs/NGOs, Government, Private Sector, Universities…

Key Evaluation Questions	Sub-Questions	Indicators /Evidence	Data collection methods
		- Evidence that the Platform **conventions** have opened the way to build partnerships that leverage CGIAR expertise to shape the future of digital agriculture.	
		- Number of regional/national partnerships and associated resource commitments reported to be initiated thanks to the **conventions** and Stakeholders' testimonies about new partnerships actual and/or potential effects on fostering positive outcomes	
		Communities of Practice (CoPs)	
		- Evidence show a positive and sustained Progress over the years on expansion of **CoPs**: members, type of members, sectors, gender balance, youth, regions, Centers involved in co-designed research and joint publications.	
		- **CoPs** members' satisfaction is high about their engagement.	
		- Uptake and use of CoPs' outputs and reports by CGIAR and CoP members	
		- Evidence shows solid [3] networks' building capabilities of the Platform and leadership around digital agriculture	

---

[3] Principles in evaluating the worth and merits of networks: Durability of Network Institutions: As Scheuch (1993) notes, working together in a successful network causes a specific form of dependency that stabilizes the network and helps develop new joint goals. The evaluation has to check whether the network members have already developed such a strategic dependency. Production of trust: All networks have to produce faith in cooperation and ensure the willingness of every single member to deliver the necessary and promised goods and services. Without such belief in trustworthy cooperation on the part of all members, the network cannot survive. Evaluations should carefully investigate whether the production of trust is working fairly well, or whether there are any threats or risks to it. Institutionalization of network rules: All networks - and all kinds of corporations - need rules that have to be at least informally agreed upon by all network partners. To protect the network from being overtaken by external actors, at least some of these rules (e.g. membership, access to shared resources, use of infrastructure, etc.) must be institutionalized. One task of evaluating a network is to check on the development of such basic rules and to what extent they have become institutionalized (including by sanctions and controls). Strategic dependency: As Scheuch (1993) notes, working together in a successful network causes a specific form of dependency that stabilizes the network and helps develop new joint goals. The evaluation has to check whether the network members have already developed such a strategic dependency.

Key Evaluation Questions	Sub-Questions	Indicators /Evidence	Data collection methods
		topics of interest, data repositories and protocols that support linked data and analysis. National Agricultural Research and Extension System (NARES) - Evidence that NARES requested platform capabilities to pilot monitoring tools - Evidence of NARES requesting data, models or case studies supported by platform capabilities,	
	4.3 New knowledge and innovations: To what extent has the Platform contributed to digital innovations for research and delivery of research through the Inspire initiative?	- Stakeholders feedback about the Inspire Challenge as opening ways to handle big data relevant to agriculture for the benefit of poor smallholder farmers. - Number of examples (e.g. upscale) that inspire how big data can deliver development outcomes. - Stakeholders feedback about opportunities and challenges that have influenced the results. - Stakeholders feedback about process weaknesses and strengths.	- Platform analytics - KII/FGD - Desk Review/Document analysis - Online surveys - Online course curriculum
	4.4 To what extent has the Platform catalyzed the development of **new digital methods for research or delivery of research at CGIAR**?	- Evidence shows new /improved digital methods and innovations such as use of sensors for water or crop yield studies initiated by the Platform for research or delivery of research. - Evidence of data reuse strategies, including uptake and use of Platform data and tools disaggregated by stakeholder groups. - Evidence of platform incubating project proposals based on center co-design - Systematized and established cross cutting capabilities to use CGIAR data	- Document Analysis/Synthesis of Evaluative Evidence. - KIIs/FGDs - Online surveys

Key Evaluation Questions	Sub-Questions	Indicators /Evidence	Data collection methods
		- Part of CGIAR research and development outputs clearly mention using the Platform.	
		- Part of the literature in the domains addressed by CGIAR and the Platform that make reference to the Platform usage.	
		- Evidence shows the Platform investments equipped CGIAR with new cross-cutting capabilities to use its data to address commonly posed research questions regarding agriculture, climate, and food systems.	
		- Evidence shows that the Platform investments make more data available for agricultural analytics, and facilitate the use of these data.	
		CRPs and Centers increased use of data available for analytics.	
	4.5 To what extent has the Platform helped change CGIAR culture and practice regarding responsible, **ethical data** collection, management, and analysis?	- Evidence that the platform instigated the development ethics framework for data collection, generation, sharing and analysis.	- Platform analytics
		- Uptake of Platform tools by internal and external stakeholders.	- KII/FGD
			- Desk Review/Document analysis
		- Changes in CGIAR Centers and CRPs practices that are related *to a* responsible use of data collection, management, and analysis. (Reference: CGIAR policy on OA/OD and CGIAR Intellectual Asset Policy)	- Online surveys
		- Evidence that NARES or CGIAR Centers developed cost-effective and time sensitive data collection approaches based on innovations in research methods made possible by the platform.	
		- Big data ethics framework covers the use of sensors, remote sensing, machine data and artificial intelligence.	

Key Evaluation Questions	Sub-Questions	Indicators /Evidence	Data collection methods
	4.6 What outputs from the Platform target enabling CGIAR to manage potential **legal** or **reputational risk** regarding data privacy and security?	- Stakeholders' perception of Platform outputs' relevance and effectiveness in reducing the legal and reputational risk regarding data privacy and security.	- Platform analytics - KII/FGD - Desk Review/Document analysis - Online surveys
	4.7 To which extent are the Platform's **quality control** mechanisms improving (or not) the discoverability of data?	- Evidence shows that the Platform improved the discovery, visibility, annotation and management of data in accordance to FAIR principles.	- Platform analytics - KIIs
Sustainability			
5 To what extent are the Platform products and communities positioned to be effective in the future, seen from the perspectives of scientists and of the end users of digital agriculture products and innovations?	5.1 To what extent do the internal and external stakeholders value the Platform and seek continuity of its programmatic elements?	- Internal and external Stakeholders' opinion about the relevance of the Platform and value their engagement with the Platform modules and that capacities built in by partners ensure sustainability of results. - Level of involvement/ collaboration /adoption etc. between the Platform and initiatives on standards, platform initiatives Open Science[4]	- KIIs/FGDs - Desk Review/Document analysis
	5.2 To what extent does the Platform position CGIAR with a leadership voice in digital agriculture in the eyes of its international partners?	- Stakeholders' opinion about the CGIAR preparedness to play a leadership role in the digital agriculture landscape. - Sustainability and quality of the Online course curriculum and delivery capacity Stakeholders testify to the overall inclusiveness of the platform in data searches, open sources applications and searchability.	

---

[4] Such as EOSc, Research Data-Alliance , GEOSS (link with GeoGLAM for example), OGC and OGC/DWG in agriculture, W3C groups, GODAN, and recent or ongoing EU projects such as AgInfra.eu (now aginfra plus).

Key Evaluation Questions	Sub-Questions	Indicators /Evidence	Data collection methods
6 To what extent would the Platform outputs outlive the existence of the Platform in relation to the Initiatives of One CGIAR?	6.1 What Platform-generated insights, products, and communities have contributed to the One CGIAR reform/reorganization?	- Platform-generated artefacts, policies, products, communities, and approaches that have been integrated into One CGIAR? - Lessons learned to facilitate the translation of Platform's outputs and outcomes to CGIAR's way of working 7- Making the Digital Revolution Central?	

Annex 3: Key Informant Interview Guide (combined)

The evaluation team designed three interview guides (for management and governance members, internal and external partners, and end-users of outputs). Given the semi-structured nature of the interview and the high overlap between stakeholder categories (e.g., an interviewee can be a member of the steering committee, an external partner, and a user of one of the tools created by the Platform at the same time) the guide was combined and facilitators did not strictly follow the list of questions or the order.

Interview Guide

Good morning/afternoon and thank you for taking the time to meet. My name is […] and I'm a member of the external evaluation team who has been asked to evaluate the CGIAR Platform for Big Data in Agriculture from 2017 to mid-2021, to learn about your experiences and perspectives, and to make recommendations for the next phase of the Program/One CGIAR Initiatives. [My colleague(s) XX is/are also on the call.]

We will hold your answers and comments in strict confidence and interview minutes will be stored in CAS Secretariat[1] folders and they will not be shared with any third party including the Platform team. If there's anything you'd like to say but are particularly concerned about its sensitivity, just let us know. We will take note of your concerns. We will not name you, and we will not quote your words directly. But it will be important to hear your perspectives – positive, negative, and mixed – to be able to offer useful recommendations for any future action.

Your participation is voluntary – that is, you may decline to answer any question, or to participate at all.

Do you have any questions?

May we have your permission to begin?

Please introduce yourself, and tell us what type of engagement you had with the Platform, for how long you've been involved / in partnership?

Please tell us what you expected of the Platform when you began your engagement, and to what degree your expectations were met. *[Probe any unmet expectations]*

What do you think about the results the Platform achieved? *[Probe quality, sufficiency, expectations for the next years]*

[Question to be asked only to members of the management and governance bodies] What was it like to work with/be involved in the Platform – the technical side, and the administrative and leadership sides? Are you aware of any difficulties that could or should be improved for the future?

Have you heard/what do you think about the relevance of the Inspire challenge competition/ the annual conventions/ CoPs?

Were there any difficulties that affected the results? What were these, and what effects did this have on the Platform, in your opinion?

To what extent has the Platform's outputs (i.e., GARDIAN) improved the discoverability of CGIAR data?

During your experience have you faced issues related to Climate change Gender, Youth related data, and analysis?

Have you participated in any capacity development event organized by the Platform?

How flexible are the Platform's design and mechanisms to local and evolving constraints including the COVID-19 Pandemic?

[Question to be asked only to Internal Partners, e.g., centers, CRPs…] To which extent do you think the Platform succeeded in managing resources and partners?

Your entity/ Organization has its own goals and priorities. How well do the Platform objectives support you in reaching them? And how best to proceed in the future?

In your experience, what is the challenge of using Data analytics with reference to your work in agriculture?

[Question to be asked only to Internal Partners, e.g., centers, CRPs…] To what extent has the Platform helped change CGIAR culture and practice regarding responsible, ethical data collection, management, and analysis? *[Ask for examples]*

[Question to be asked only to Internal Partners, e.g., centers, CRPs…] Are you aware of any CGIAR product to manage potential legal or reputational risk regarding data privacy and security?

[Question to be asked only to Internal Partners, e.g., centers, CRPs…] What are the most relevant indicators to monitor and evaluate the Platform (or similar initiative) performance and achievements in the coming years?

To what extent has the Platform enabled CGIAR to engage with the wider Agriculture data (and innovation) digital ecosystem in terms of both depth of engagement and reach?

To what component or products of the platform would you give priority if the platform has to be re-conducted or reshaped under one of the One CGIAR initiatives?

Do you have any other comments or issues you'd like to discuss before we finish?

Thank you very much for your time and your responses. We will continue to be available via email if you have any other thoughts you'd care to share. Again, many thanks, and I wish you the best in your endeavors.

Annex 4: Profile of Stakeholders Consulted in interviews

The team sent interview invitations on August 31st and conducted them in two rounds, from September 2nd until October 6th. About 140 invitations were sent with a response rate of about 40%. In total 51 interviews were conducted, one of which included two stakeholders, hence, the total number of interviewees is 53. The majority of the sampled respondents were men (34 men versus 18 women). Most interviewees were in a leadership or management role within the organization/research center/company they work for. 38 respondents worked for CGIAR Research Centers, mainly from IFPRI and the Alliance of Bioversity International and CIAT, or the CGIAR System Organization.

Annex 4 - Table 1: Number of interviews conducted

Item	Number
Total invitations sent	140
Confirmed interviews	51
Interviews conducted	51
People interviewed	53

Annex 4 - Figure 1: Interviewees' Gender (n=53)
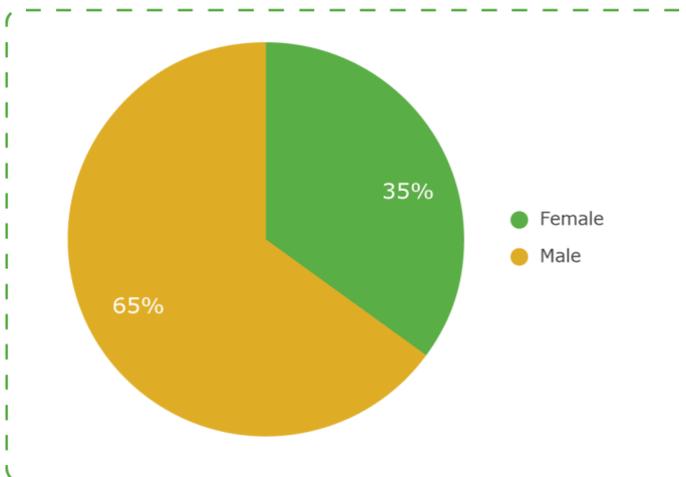
Annex 4 - Figure 2: Interviewees' place of work (n =53)

- CGIAR
- Private Sector
- Donor
- Academia/Research
- International organization
- Other

73%
9%
6%
8%
2% 2%
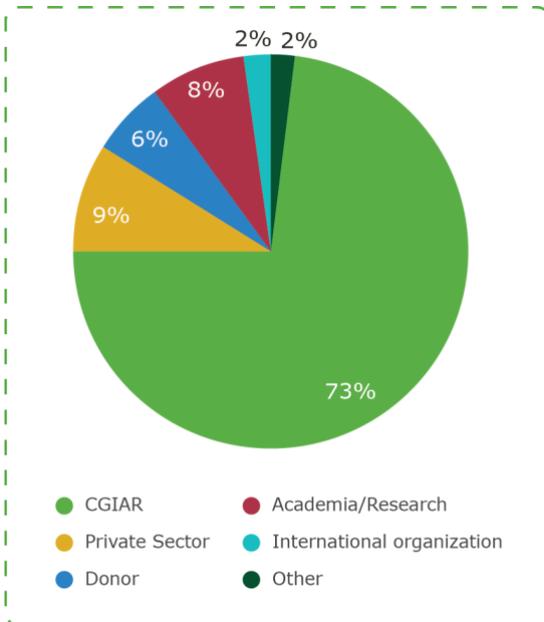
Annex 4 - Figure 3: Interviewees' role/position within the company/organization/Research Center they work (n=53)

- Leadership, member of a Governance body
- Management and administrative staff
- Other
- Scientists/Researcher
- Consultant

36%
23%
27%
8%
6%
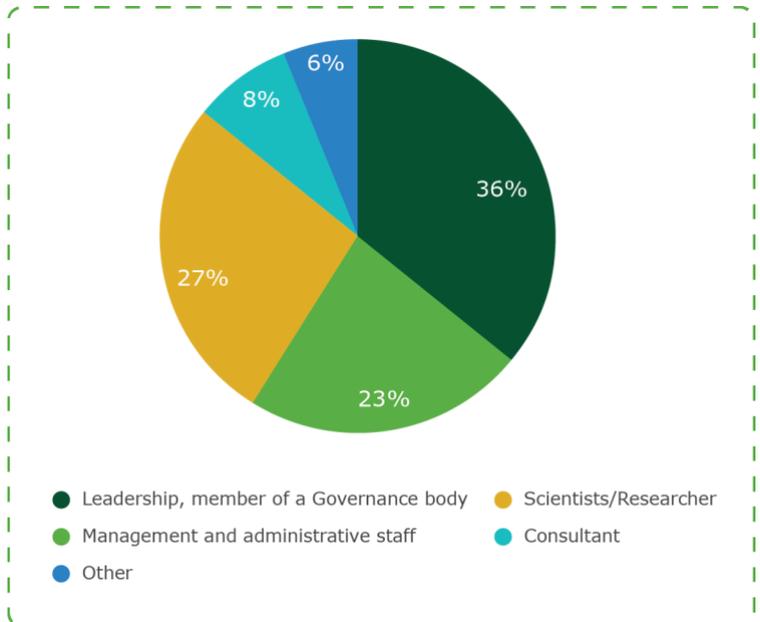
Annex 4 - Figure 4: CGIAR Centers/Platforms where the 37 interviewees from CGIAR work
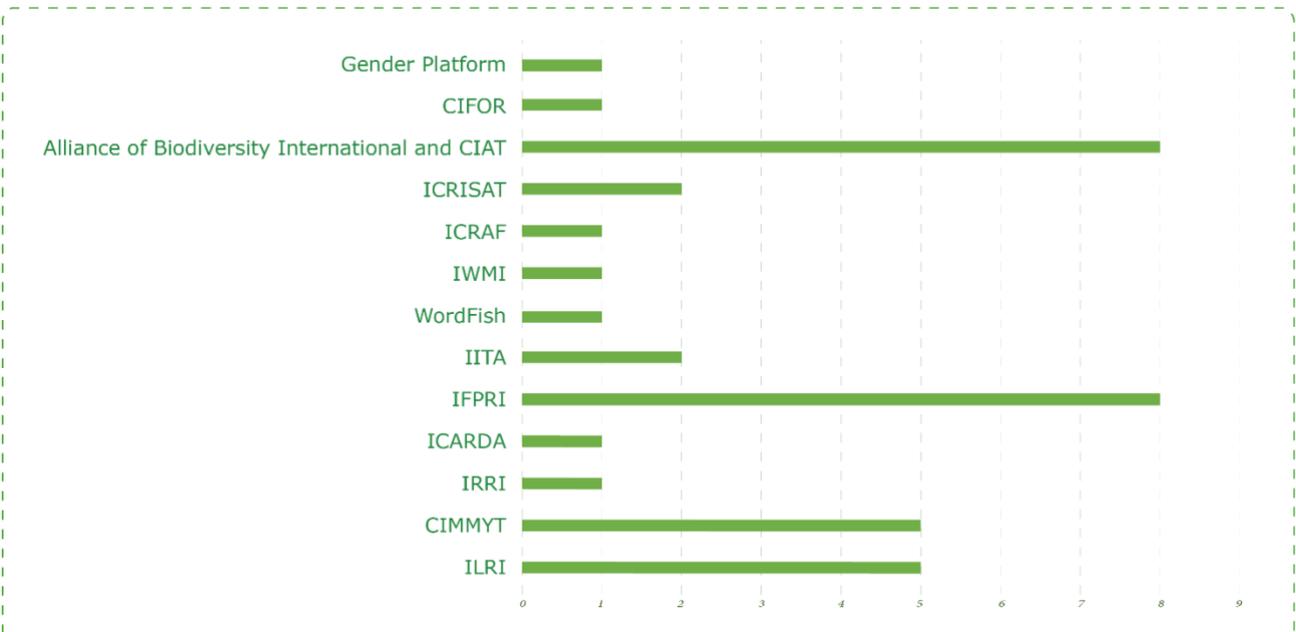
Annex 5: Online Survey

5.1 Introduction to the Survey

The online survey was released on 20 September and closed on 30 September 2021. The survey was designed in a way that respondents were directed to a set of specific questions based on their respective types of engagement with the Platform. Two versions of the survey were released, one in English and one in Spanish (9 out of 110 respondents responded in Spanish). The survey was sent by the Platform management team to comply with General Data Protection Regulation (GDPR) privacy requirements. This meant that contact lists could not be shared with the evaluation team.

The survey was sent to the following groups:

- All subscribers of the Platform Newsletter (2803 subscribers)
- Participants of the annual conventions – those who opted in for correspondence
- External partners (Academia, private companies, NGOs, etc.)
- Internal partners (CRPs, Centers, projects.)
- Users of CG Labs

110 responses were received, four of which were incomplete, hence, the evaluation team considered 106 responses for analysis. The evaluation team asked questions regarding respondents' profiles and their type of engagement with the Platform, as well as their opinions regarding the level of satisfaction with the Platform's products and about its relevance, effectiveness, and sustainability. A 5-point Likert scale was used; respondents could use a range from 1 (strongly disagree) to 5 (strongly agree) to express their opinion regarding a specific statement. For some questions, the evaluation team asked the satisfaction rate using a similar scale, with 1 being "not satisfied" and 5 being "fully satisfied". All questions were compulsory, except for optional questions. For most questions, the evaluation team allowed the answer "I don't know" to guarantee that the completion of the questionnaire could run smoothly. The number of respondents per question changes given the type of engagement respondents had with the Platform. The evaluation team conducted a descriptive analysis using all quantitative questions. For the open-ended questions, the evaluation team identified between 1 and 3 keywords for each respondent and reported the most used ones. The invitation email and full questionnaire are included in this section.

5.2 Survey Results

5.2.1 Respondents' profile

106 complete responses to the survey were received. The first part included general questions regarding the respondents' profile and was compulsory for all respondents, hence, 106 responses to all questions were analyzed in this section. Most respondents were male (66%) compared to a lower rate of 32% for females, while the remainder (2%) did not say **(see Figure 1)**. Most respondents were between 25 and 44 years old **(Figure 2)** and worked either in international organizations, including CGIAR Research Centers, or in universities and other research institutes **(Figure 3)**.

Annex 5-Figure 1: Gender distribution of respondents (n=106)

- Female
- Male
- Not sure/rather not say
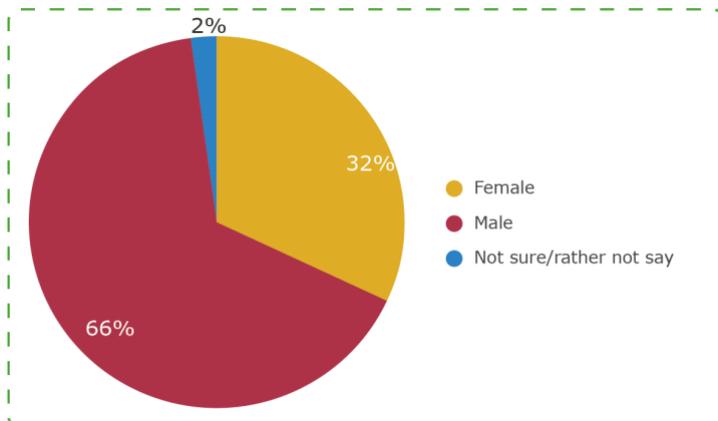
Annex 5- Figure 2: Age distribution of respondents (n=106)

- 18-24
- 25-34
- 35-44
- 45-54
- 55-64
- 65+
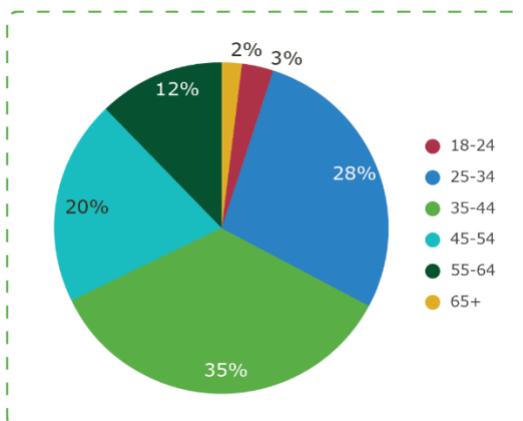
Annex 5 - Figure 3: Distribution of respondents' main field of work (n=106)

- Governmental and public institutions
- International Organizations, including CGIAR research Centers
- National and International NGOs, Foundations, not-for-Profit Sector
- Private sector(including farm, small-medium enterprises)
- Universities and research

The evaluation team divided the stakeholders into two groups, internal to CGIAR (44 respondents) and external to CGIAR (62 respondents). The first group included those respondents that belonged to a CGIAR Center or a CGIAR Research Program (CRP), while the other respondents were considered external stakeholders. More than 35% of the respondents were researchers or scientists. However, within the group of stakeholders internal to CGIAR, only 30% were scientists or researchers with the majority being ICT staff/data managers. However, the converse was true for external stakeholders. 39% were scientists or researchers while only 15% were ICT staff/data managers **(Figures 4 and 5)**. Finally, almost half of the sample had a Masters' degree and over a third (34%) had a PhD **(Figure 6)**.

Annex 5-Figure 3: Distribution of respondents by their main role in their place of work – all sample (n=106)
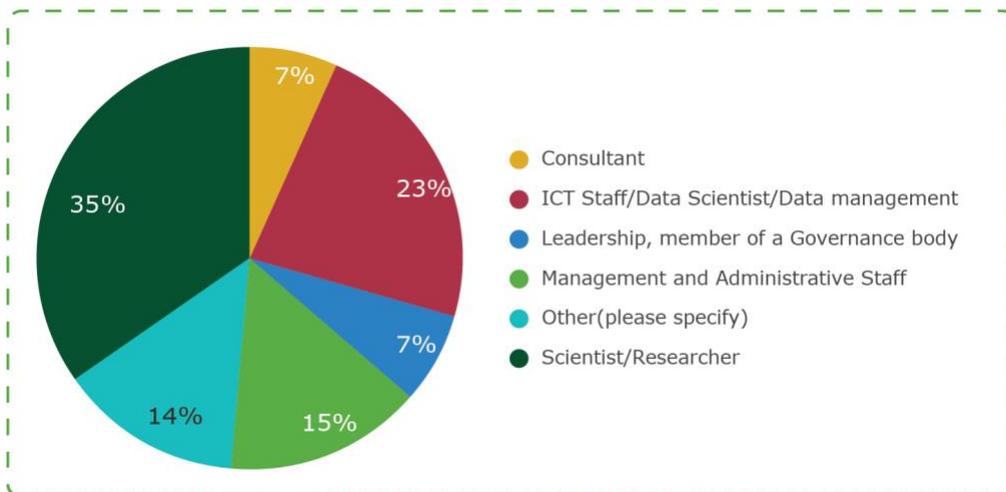
Annex 5-Figure 4: Distribution of respondents by their main role in their place of work – by internal to CGIAR vs external stakeholders (n=106)
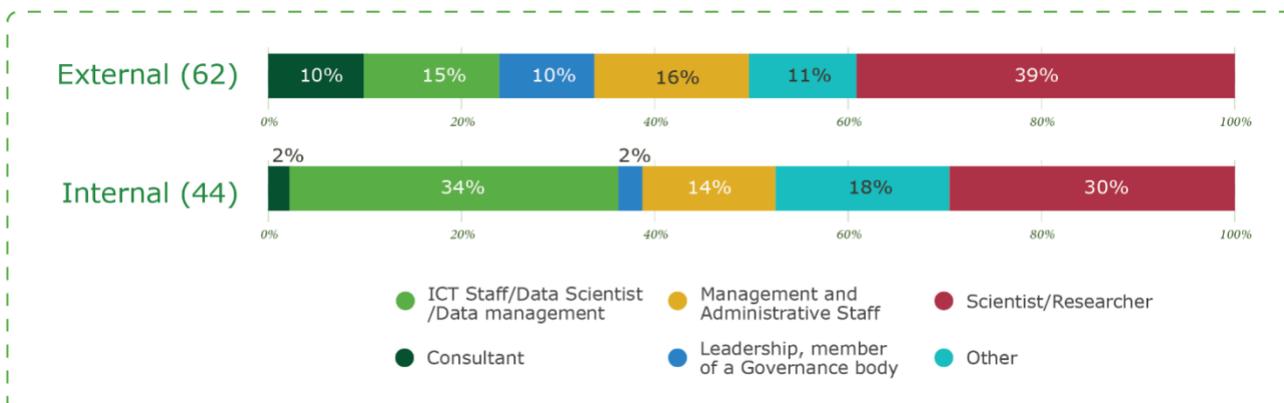
Annex 5- Figure 5: Highest level of education of respondents (n=106)

Legend:
- Bachelors or equivalent
- Master's/Graduate
- PhD/Doctorate
- Post-secondary school qualifications (non-tertiary)

Values shown: 2%, 15%, 34%, 49%
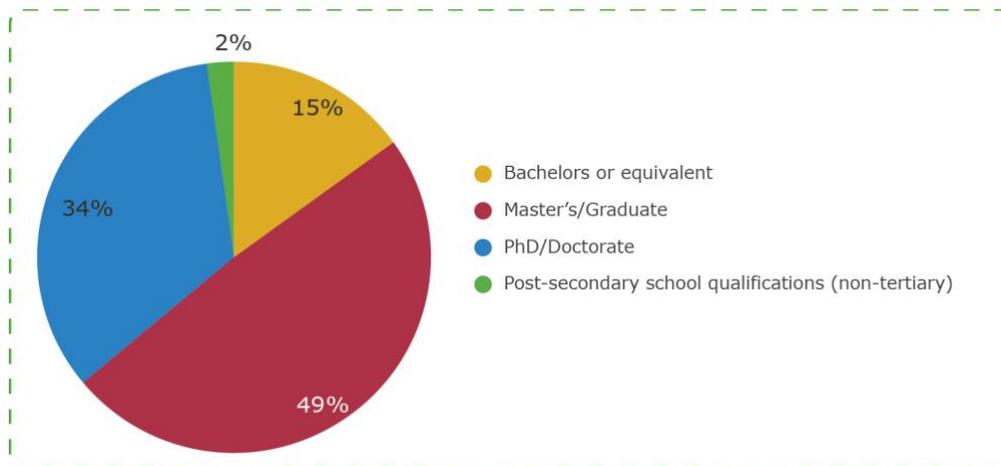
5.2.2. Respondents' relationship with the Platform

This section covers the respondents' relationship with the Platform, and how and when they have engaged with it. The main question was "How do you engage with the CGIAR Platform for Big Data in Agriculture?" and based on their responses, respondents were addressed to more in-depth questions. Respondents could select more than one option and identify as being part of one or more categories (e.g. CGIAR Center and User of GARDIAN). 58% of respondents identified as external to CGIAR Centers or CRPs **(Figure 7)** while most internal stakeholders were employed by a CGIAR Research Center and/or participated in a Community of Practice (CoP) **(Figure 8)**. No employees of ICARDA or ICRAF responded. Neither did staff who worked on the following CRPs or Platforms: Livestock; Forests, Trees and Agroforestry; Water, Land, and Ecosystems; and the Genebanks Platform.

Annex 5-Figure 6: Distribution of respondents by stakeholder group (n=106)

Legend:
- Internal to CGIAR or CRPs
- External to CGIAR or CRPs

Values shown: 42%, 58%
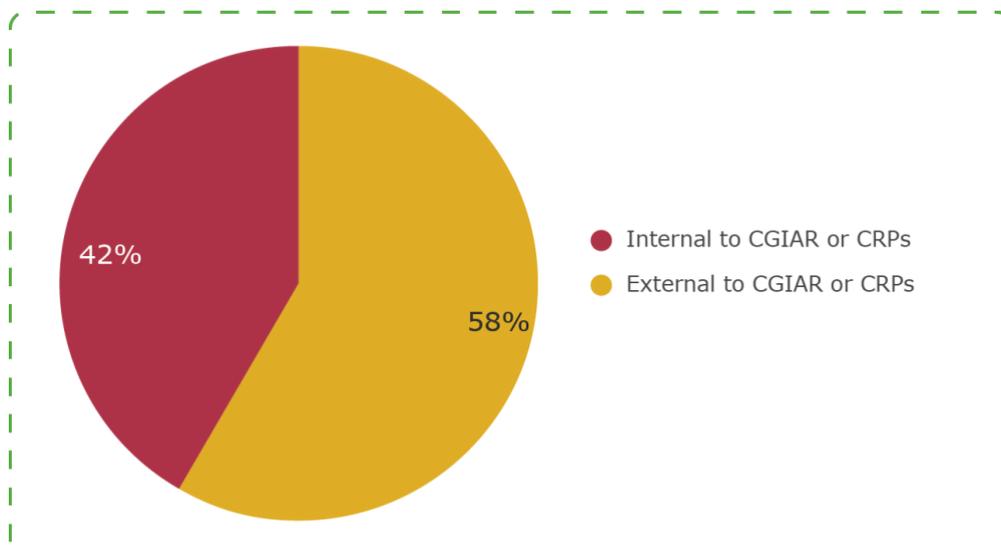
The majority of the respondents had participated in at least one Annual Convention and used GARDIAN at least once **(Figure 8)**. Those respondents employed by a CGIAR Research Center were mostly from IITA, CIMMYT, and the Alliance of Bioversity International and CIAT **(Figure 9)** . Among the 36 respondents who participated in one or more CoPs, the majority were part of the Information and Data Management CoP **(Figure 11)**. Only 10 respondents engaged with the Platform as external partners, namely those

who were not employed by CGIAR Centers or CRPs but had a direct partnership with them[5]. One of these ten respondents also declared that they "had never engaged with the Platform", hence s/he was not redirected to more in-depth questions specifically addressed to external partners. The evaluation team considered the remaining nine as external partners including two from research institutes, three from academia, three from the private sector, and one from an international organization.

Annex 5- Figure 7: Respondent distribution by type of engagement with the Platform as a percentage of the total sample (n=106)
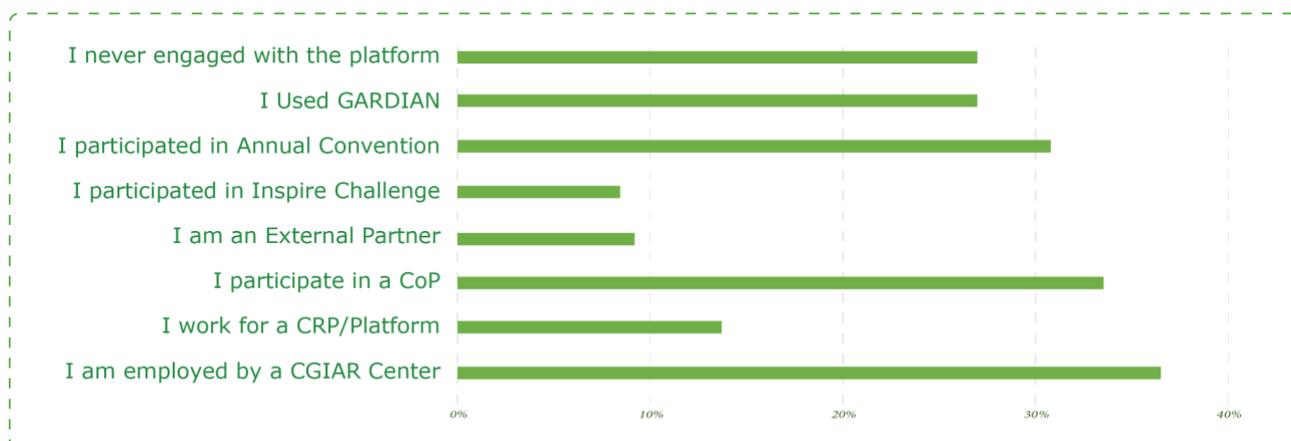
Annex 5- Figure 8: Distribution of respondents by employment in CGIAR research centers (n=39)

---
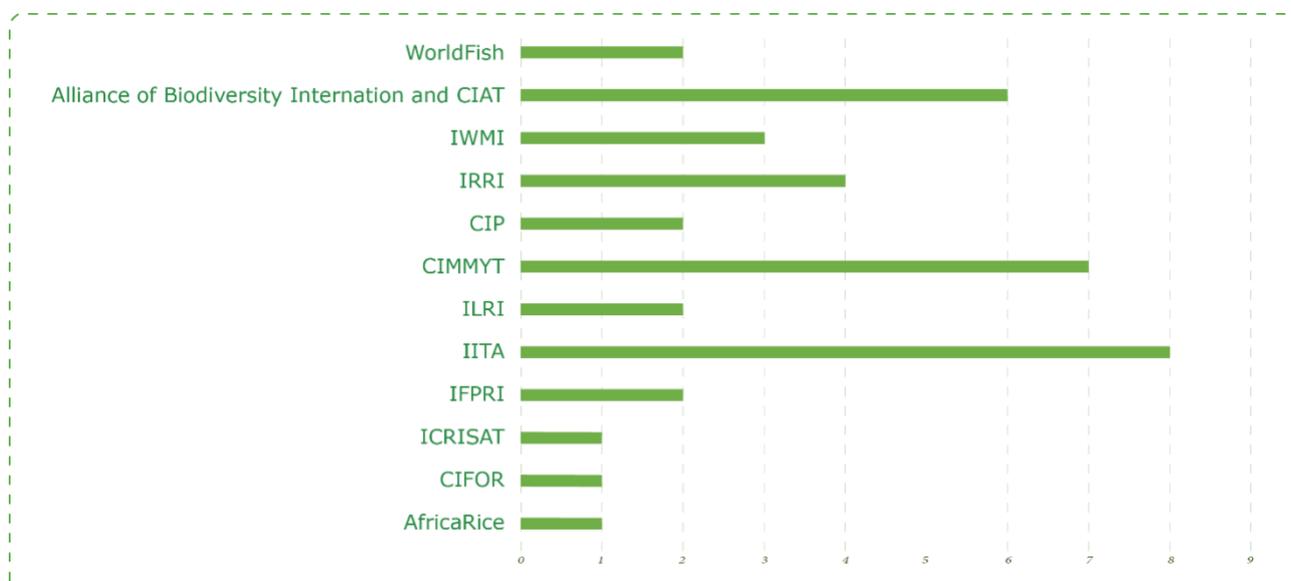
[5] Other respondents external to CGIAR engaged with the Platform as Inspire Challenge Participants, Participants to the annual conventions, GARDIAN users or members of a CoP.

Annex 5 - Figure 9: Distribution of Respondents who work for one or more CRP or Platform (n=15)
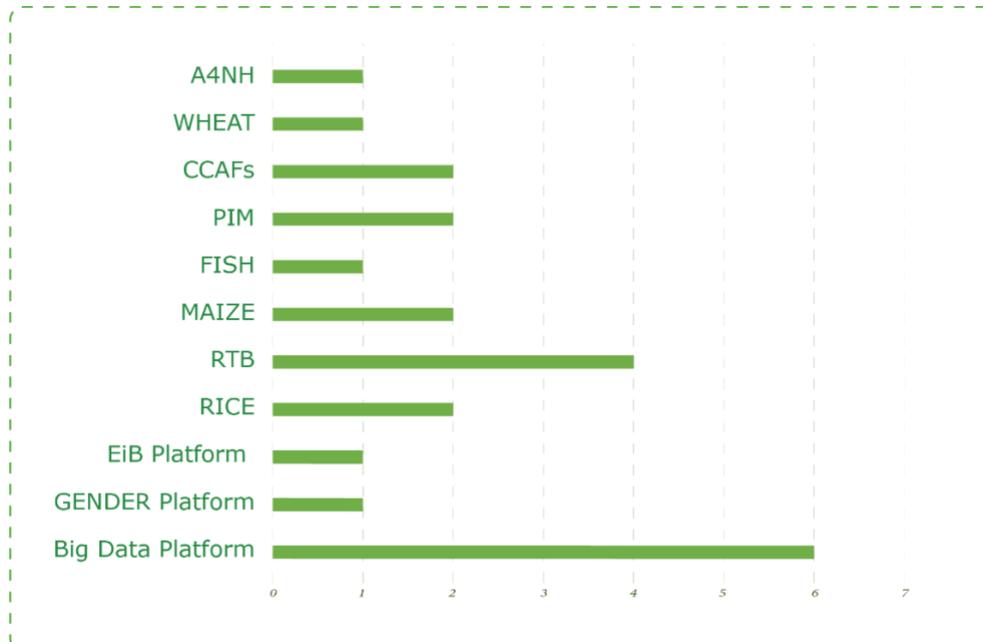
*Note: Most Respondents worked for more than one CRP or Platform simultaneously; hence, although there were 15 respondents, the results sum up to 23 rather than 15.

Annex 5 - Figure 10: Comparison of respondents that participate in CoPs (n=36) as a proportion of the survey respondents' vs proportion belonging to specific CoPs.
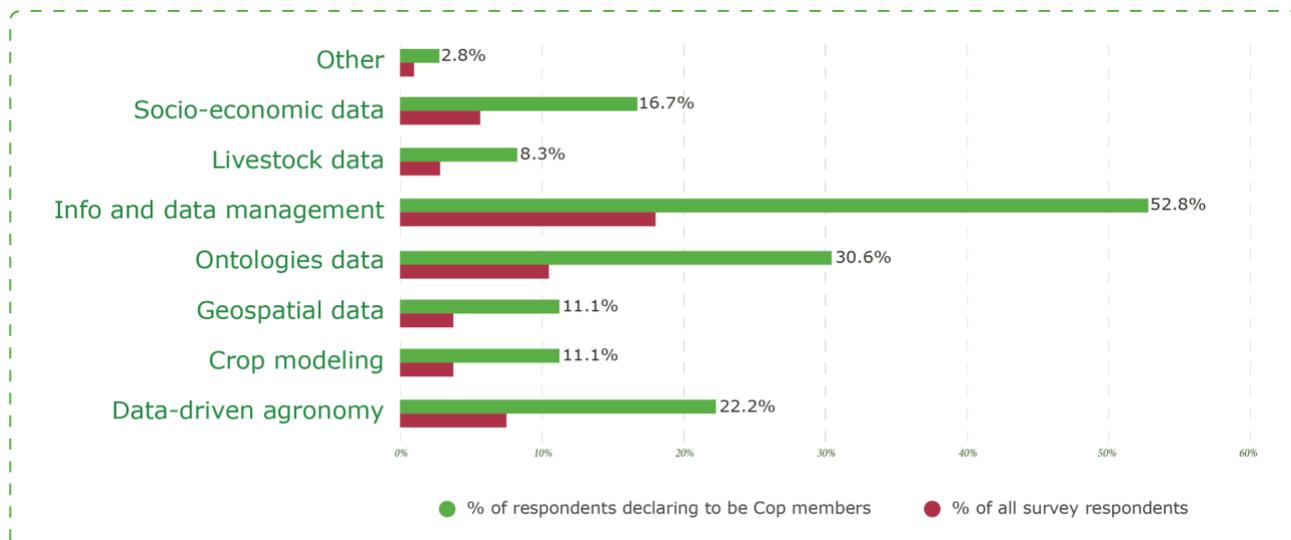
Among the 102 people who answered the question regarding their level of knowledge of the Platform's mandate, more than 38% ranked their knowledge as 3 (average) on a scale from 1 (None) to 5 (Very good). Results were similar for both internal and external stakeholders, although among internal stakeholders there was a higher percentage of people who stated they had a good knowledge of the Platform's mandate **(Figure 12)**.

5.2.3. Stakeholders' Opinion about the Sustainability, Relevance, and Effectiveness of the Platform

5.2.3.1 Sustainability and the Future of CGIAR

In this section, the evaluation team asked questions about the future of CGIAR and its sustainability. **Figure 13** shows the main answers to the question: "What is your dream scenario for One CGIAR e-research and data-driven impact?". The evaluation team identified a maximum of three keywords used by respondents and divided them into seven macro-areas: Data, Collaboration, Budget, Reachability and Impact, Harmonization, Inclusion, and Learning. Inside each area, represented in the figure as a circular diagram, the diagram shows the keywords and the number of times they appeared. Words related to the Data area were the most common (they appeared 23 times) and the most used word was "Data accessibility" which appeared nine times. Learning was another very important area with "Knowledge exchange" as the main word used, appearing five times. Some words were grouped into more than one macro-area, like "Data harmonization" that was grouped into both the Data and the Harmonization areas. In this case, the evaluation team reported the keywords in the intersection between the two circular diagrams.

Annex 5 - Figure 12: Keyword visualization by macro area: Sustainability: Dream Scenario for One CGIAR (Open question)
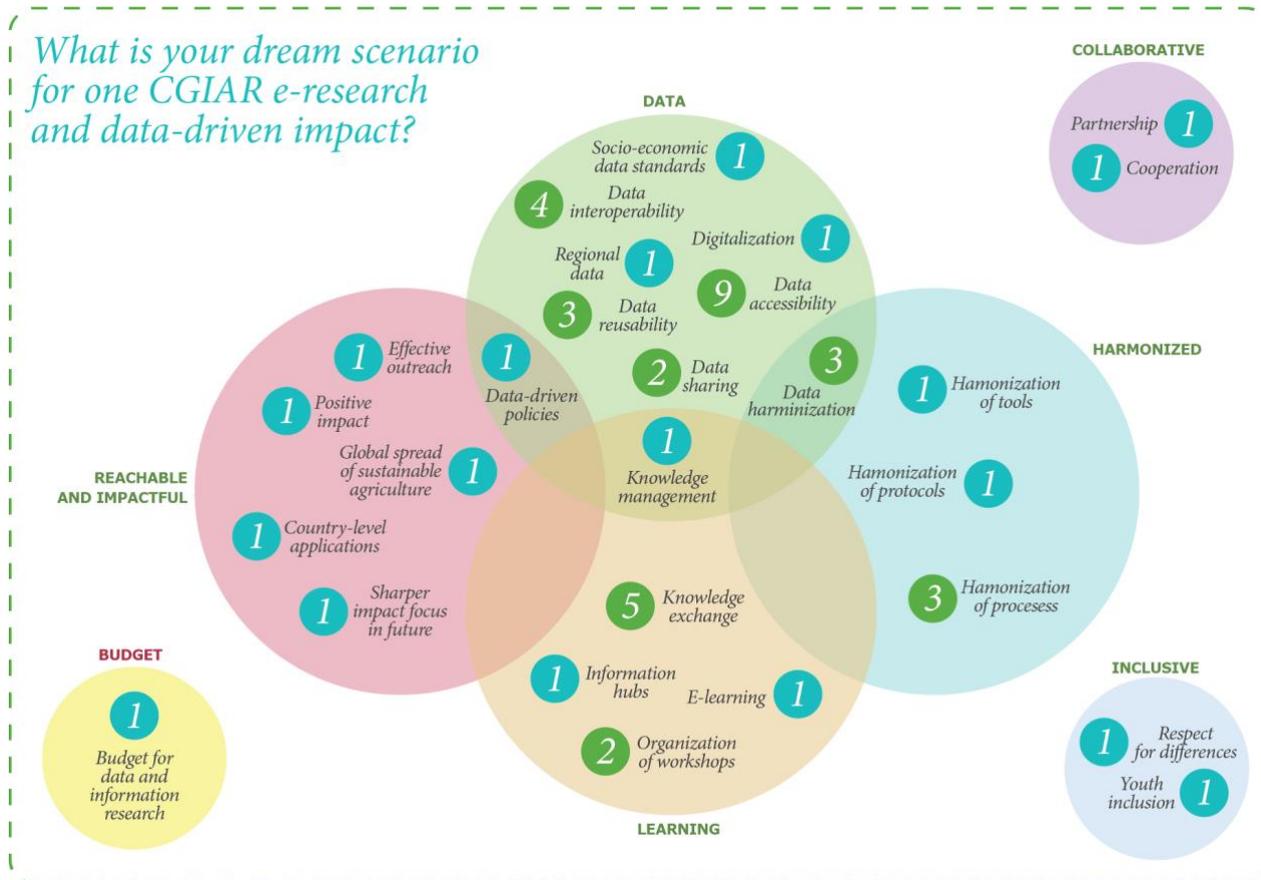
Numbers represent the number of times keywords appeared. Circles represent the macro-area in which keywords were grouped.

Furthermore, according to **Figure 14**, the majority of respondents (67%) claimed that CGIAR is prepared to take on a role of leadership in the international digital agriculture landscape.

Annex 5- Figure 13: Sustainability – Future of CGIAR – How would you rate CGIAR's preparedness to take on a role of leadership in the international digital agriculture landscape? (n=59)
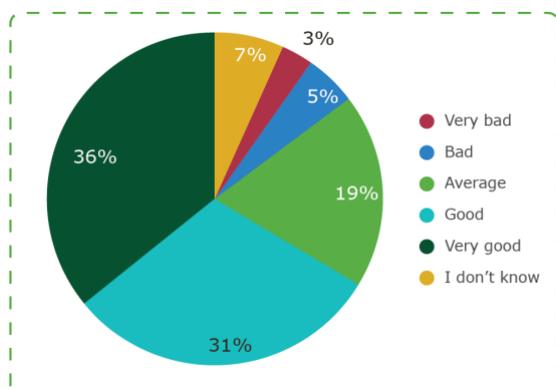
Figure 15 presents answers to the question "What makes CGIAR ready/not ready to take on such a leadership role?". The evaluation team grouped the answers based on the score provided in the previous question. Mainly, critics were concerned about CGIAR's willingness to take a firm stand on Open Data and Open Science, as well as its inability to engage with a broader group of scientists, the lack of a clear long-term financial plan, and the lack of clear governance. On the other hand, the main strengths highlighted by respondents were CGIAR's instructional and technical expertise in agricultural research and development and its commitment to bringing positive change for farmers and stakeholders. Some respondents considered CGIAR's human resources, network, and technological know-how as important.

Annex 5- Figure 14: Sustainability –CGIAR's preparedness to take on a role of leadership in the international digital agriculture landscape? (open-ended question)

Future of CGIAR	What makes CGIAR ready/not ready to take on such a leadership role?
	"CGIAR has the instructional and technical capacity for agricultural research and development, and this makes CGIAR ready to take on such a leadership role." "CGIAR seems well-organized, straightforward, and has a clear idea of what it wants to do, so it is ideally placed." "Adequate resources, including technical expertise and commitment towards bringing positive change for farmers and stakeholders." "Access to all the centers across agriculture. It has vast experience on a global level." "Experience, networks, historical and ongoing data collection efforts. Potential to produce standardized, open datasets representing large temporal and spatial scales." "It's human resources and technological know-how." "CGIAR's extremely rich wealth of data on a multitude of research areas, and the teams of professionals that stand with those data." "Continuous improvement, learning, collaboration, and innovations." "It already has the tools plus the proactive CoPs."
	"Quite fragmented. There is need for more COP-wide messages and engagements." "Bringing new members to discover a lot in what they lack skills." "CGIAR has a vast range of expertise and knowledge from staff who come from many areas of the agricultural domain. Most people within the CGIAR are willing to share that knowledge/expertise and work in multidisciplinary teams towards an objective that will aid the specific stakeholder they are aiming to help. Staff are willing to learn from one another and work together to develop great new tools, products, and policies that will be about positive change within the agricultural domain." "Great potential to do so but products are rarely out scaled outside the CGIAR network and region of operation." "We have significant data assets and longstanding experience in agricultural research drawing on a global workforce." "Lack of government and mainstream attention." "Its consistency in stakeholders' management and engagement."
	"Lack of a clear long-term financial plan that could support a well-defined working plan, transversal and executed by professional that have time and resources to accomplish it." "Pros: lots of smart scientists and innovative capacity. Cons: project cycles don't encourage innovation and leave data work underfunded"

Future of CGIAR	What makes CGIAR ready/not ready to take on such a leadership role?
	"Risk that CGIAR will implode in coming years" "IRRI as part of CGIAR doesn't even have a data scientist or experts in ML, AI." "Due to its re-organization, the CGIAR is not ready to present visionary and disruptive initiatives."
	"In terms of leadership on publications and data management, Big Data/GARDIAN is the front runner for CGIAR, and big challenges await if the One CGIAR transition will imply harmonizing all publications and data libraries. Big Data/GARDIAN is, I think, uniquely well placed to help overcome those challenges." "I believe the Platform still is a small Community of Practice confined to few centers and scientists and has not engaged widely with the broader group of CGIAR centers and Programs"
	"There is no willingness to take a firm stand on open data and open science"

5.3 Module 1 – ORGANIZE

Twenty-nine (29) respondents claimed to be users of the GARDIAN Portal. However, one of them had also answered "I never engaged to (sic) the Platform", hence, was not asked more in-depth questions regarding his/her use of the Platform. 28 respondents answered more detailed questions regarding their use of the GARDIAN portal. Most of them made occasional use of it and only one person stated that s/he used it daily, since its use was directly related to the work that s/he conducted **(Figure 16)**. Eighty-two percent of the respondents claimed to use the portal for professional/non-academic research work **(Figure 17)**. Nine in 10 (89%) of respondents learned about GARDIAN through their CGIAR network and 50% made their work available in it, although 10 respondents out of 13 said that they did not know whether the interest in their work increased through GARDIAN or not.

Annex 5 - Figure 15: Frequency and purpose of using the GARDIAN portal (n = 28)
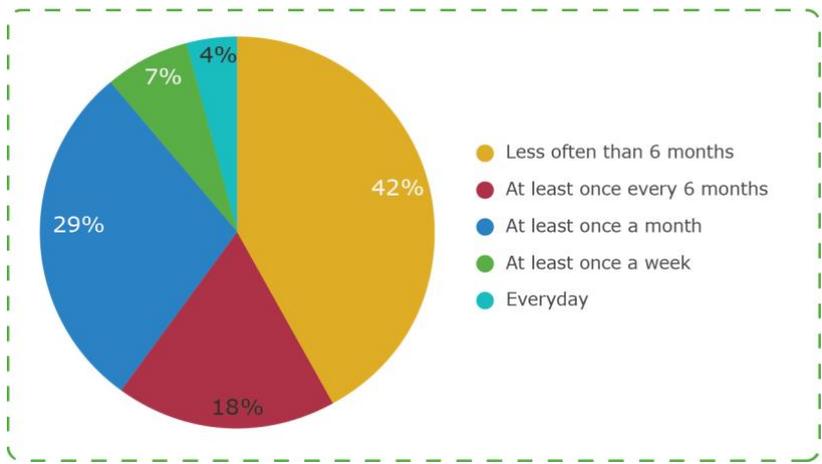
Annex 5 - Figure 16: Distribution of respondents' main reason for using the GARDIAN portal (n =16)
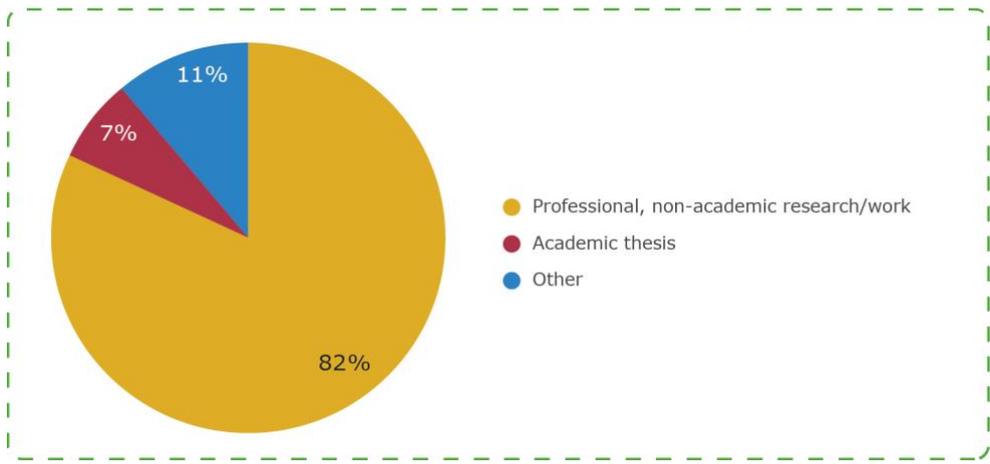
Sixteen respondents answered the optional open question regarding their use of the GARDIAN portal. The majority used it to search for data and also to promote the Platform and raise awareness about its functionalities within their team. **Figure 18** reports some of the main words stated by respondents.

Next, respondents were asked about their perception regarding the easiness to find, access, interoperate and reuse data through GARDIAN. They were asked whether they agreed or disagreed with the following statements: (1) Using GARDIAN I can easily find metadata and data I am looking for; (2) Using GARDIAN I can easily access the datasets I found; (3) Using GARDIAN I can easily integrate datasets I found with other data; (4) Using GARDIAN I can easily reuse the data I found **(Figure 19)**. These questions were also asked as part of the assessment of the **effectiveness** of the Platform.

Annex 5 - Figure 18: Respondents' perception of easiness to find, access, interoperate, and reuse data through GARDIAN (n=26)
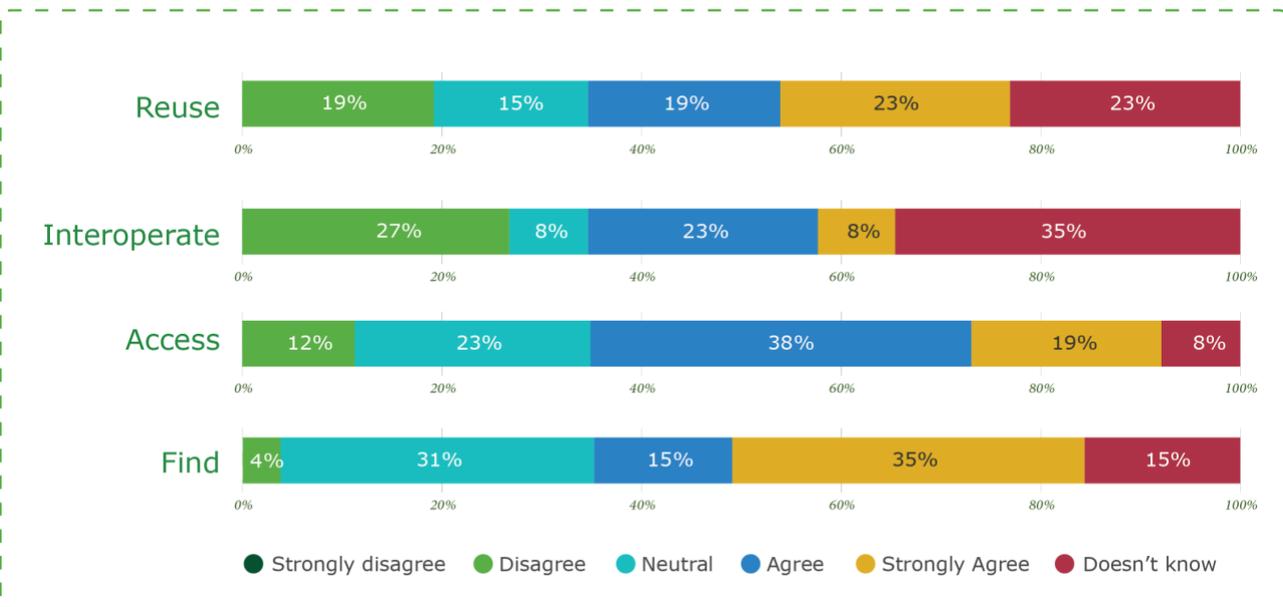
5.3.1 Relevance

As mentioned above, respondents were asked to answer questions related to the **Relevance, Effectiveness, and Sustainability of the Platform.** Respondents had to agree or disagree on a scale from 1 (strongly disagree) to 5 (strongly agree) to specific statements.

Seventy-five respondents answered with regards to their agreement or disagreement with the statement "The Big Data Platform's products, analytical tools and/or activities I engaged in add value to my work. **Figure 20** shows that 40% of respondents strongly agreed with the statement. Similar results are reflected also among internal and external stakeholders (47% of stakeholders external to CGIAR and 35% of internal stakeholders strongly agreed with the statement). Among external partners, 7 agreed with the statement, 1 disagreed and 1 was neutral. In response to the question specifically addressed to staff of CGIAR centers of CRPs: "The Platform's products and activities address my new and evolving needs", most of the sample strongly agreed with it and results showed no statistically significant difference among men and women **(Figure 21)**.

Annex 5 - Figure 19: Relevance - The Big Data Platform's products, analytical tools and/or activities I engaged in add value to my work (n=75)
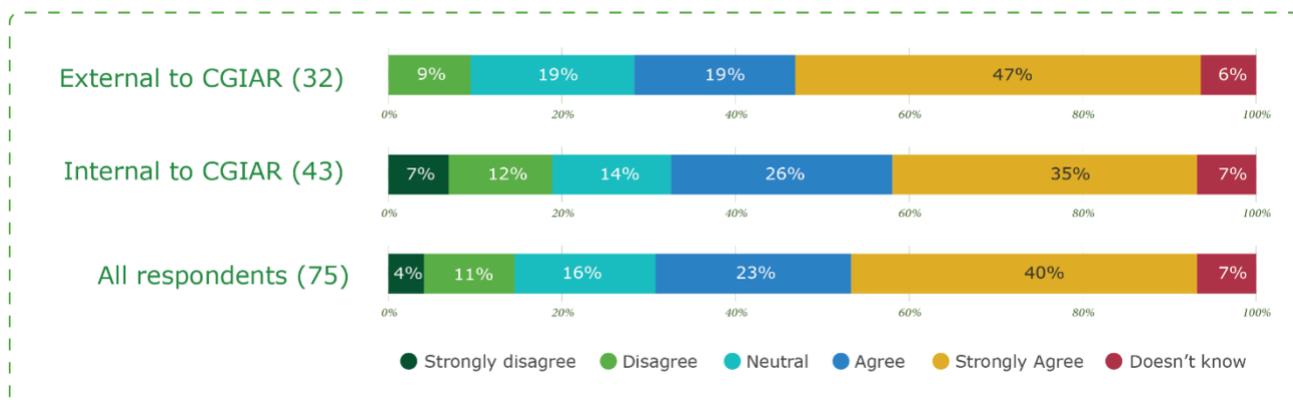
Annex 5 - Figure 20: Responses about: "Relevance of the Platform's products and activities to addressing my new and evolving needs." (n=43)
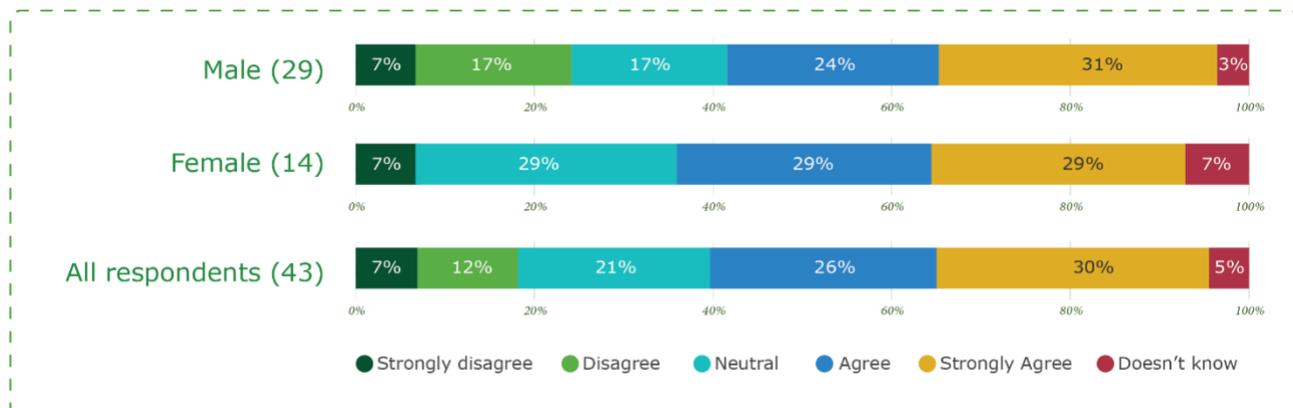
5.3.2 Effectiveness

The survey section containing questions about effectiveness was the densest. When asked a question regarding the time efficiency of the search function in the Portal, out of 26 respondents, 11 agreed with the fact that data search via GARDIAN is time-efficient compared to other similar portals, 7 did not know, 6 were neutral and 2 disagreed **(Figure 22)**. Then, the evaluation team asked specific questions regarding the ability of the Platform to satisfy specific needs. The questions followed the same format and were asked as statements to which respondents could either agree or disagree using a 1 to 5 scale. Most

respondents strongly agreed with the statement "When I engaged with the Platform, I felt that my specific needs at a professional level were met" **(Figure 23)**. Results were disaggregated by gender and are similar for female and male respondents, although a higher percentage of women strongly agreed with the statement compared to men. Among the 9 external partners, 6 strongly agreed, 2 agreed and 1 was neutral.

Annex 5 - Figure 21: Distribution of responses: Effectiveness – Data search via GARDIAN is time-efficient compared to other similar portals I use (n=26)
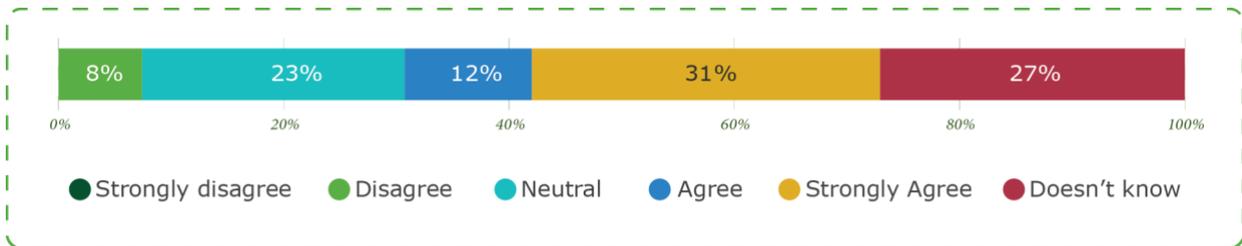
Annex 5 - Figure 22: Effectiveness – When I engaged with the Platform, I felt that my specific needs at a professional level were met – whole sample and by gender (n=43)
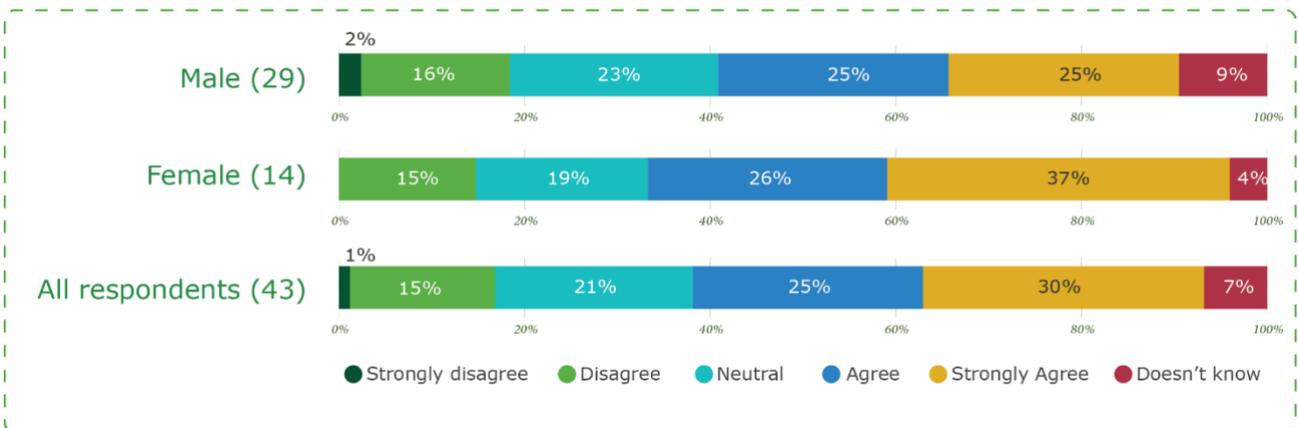
In the figures below, answers to different statements related to Platform's effectiveness are reported.

Annex 5 - Figure 23: Distribution of responses: Effectiveness – Increased ability to discover publications and to use data (n=39)

CGIAR Centers increased their ability to discover agricultural publications through the platform

Doesn't know	5%
Strongly Agree	26%
Agree	31%
Neutral	21%
Disagree	8%
Strongly disagree	10%

CGIAR Centers increased their ability to make effective use of data for analytics thanks to the Platform's support and products

Doesn't know	18%
Strongly Agree	23%
Agree	21%
Neutral	23%
Disagree	10%
Strongly disagree	5%

Legend: Strongly disagree · Disagree · Neutral · Agree · Strongly Agree · Doesn't know

Annex 5 - Figure 24: Distribution of responses: Effectiveness – Ability of the Platform to foster digital innovation and accelerate progress toward methods for agricultural research (n=98)
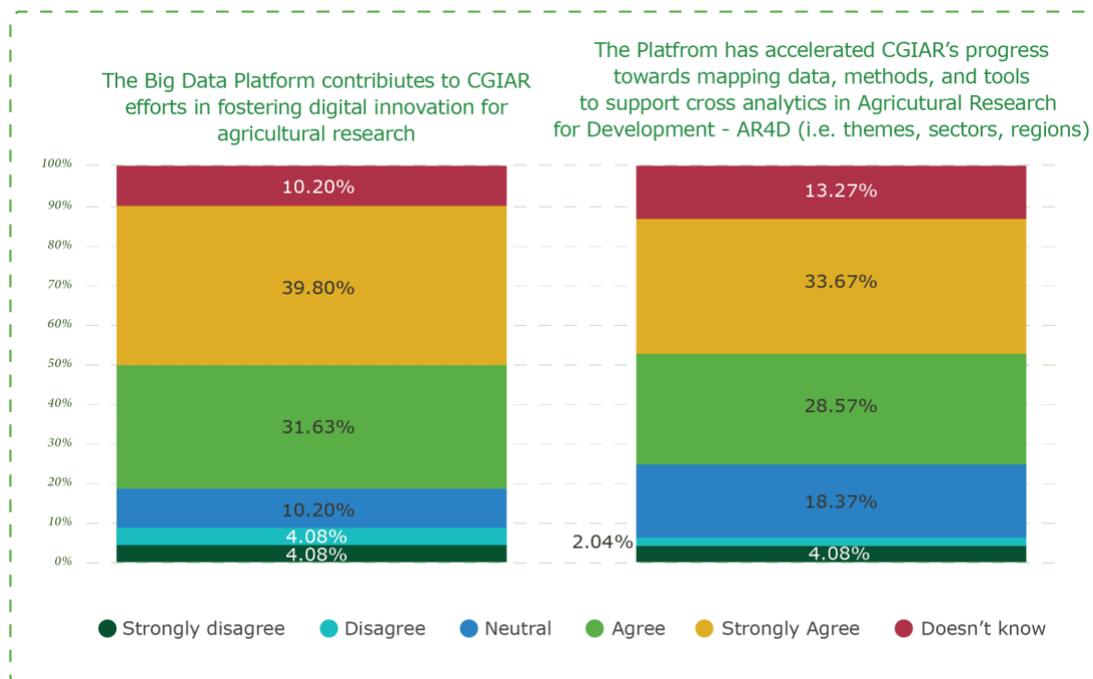
The Big Data Platform contribiutes to CGIAR efforts in fostering digital innovation for agricultural research

Doesn't know	10.20%
Strongly Agree	39.80%
Agree	31.63%
Neutral	10.20%
Disagree	4.08%
Strongly disagree	4.08%

The Platfrom has accelerated CGIAR's progress towards mapping data, methods, and tools to support cross analytics in Agricultural Research for Development - AR4D (i.e. themes, sectors, regions)

Doesn't know	13.27%
Strongly Agree	33.67%
Agree	28.57%
Neutral	18.37%
Disagree	2.04%
Strongly disagree	4.08%

Legend: Strongly disagree · Disagree · Neutral · Agree · Strongly Agree · Doesn't know
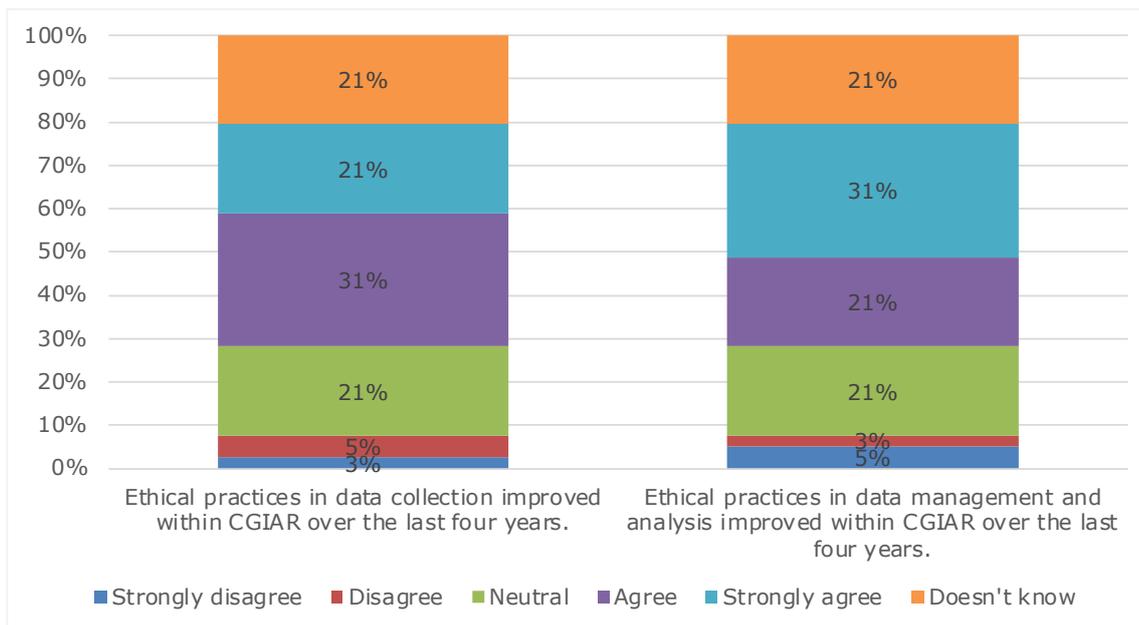
5.4 Module 2 – CONVENE

Among the 106 respondents, 33 participated in the Annual Convention at least once and most of them declared that their expectations were met. 31 respondents answered more detailed questions regarding their participation **(Figures 27 and 28)**. 17 respondents participated only in one of the four conventions, 7 in two conventions, 6 in three conventions, and one participated in all four conventions. 9 respondents participated only in the online Convention, 9 participated only in one of the conventions in-person, and the remaining 13 participated in both an online convention and at least one in-person. In general, the most attended convention among respondents was the one held online in 2020, which is coherent with the fact that it was the convention with the highest number of total participants due to its easy accessibility and free participation **(Figure 27)**.

Annex 5 - Figure 26: Distribution of respondents' participation in the Annual Convention by year (n =31)
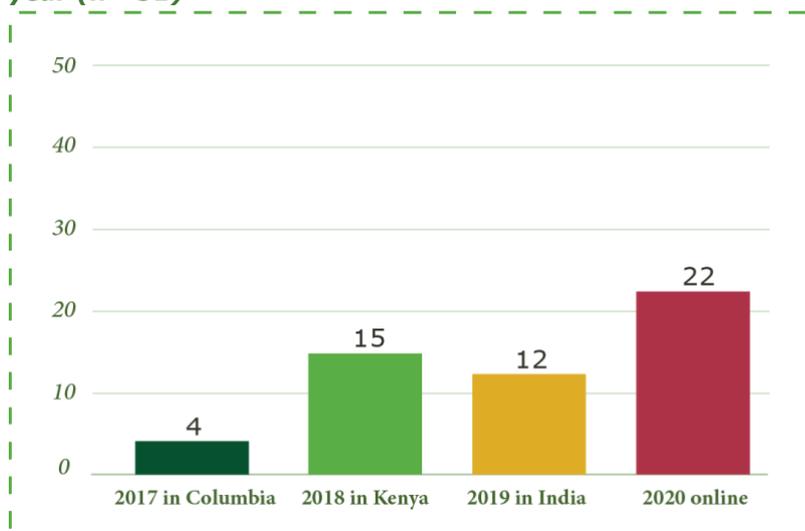
Figure 28 below, presents respondents' views about "To which extent were your expectations met?" with a rating from 1 to 5. Most respondents declared that their expectations were met. Results are similar for both people who participated only online or only in-person. However, 31% of participants who went to both kinds (online and in-presence) declared that their level of satisfaction with the events was average.

Annex 5 - Figure 27: Distribution of respondent's level of satisfaction with the Annual Convention (n=31)
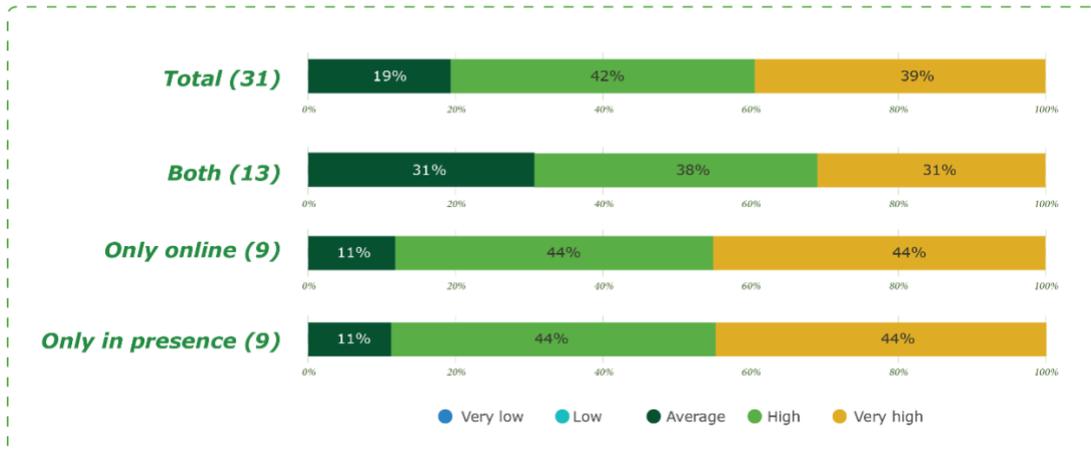
Out of 33 participants to the annual conventions, 21 people answered the open question "What was the added value of the conventions to your work?" and most of them used words like "networking", "partnership", "learning" and "knowledge sharing" **(Figure 29)**.

Annex 5 - Figure 28: Opinions of the value-added value of the Conventions to respondents' work (n= 49)
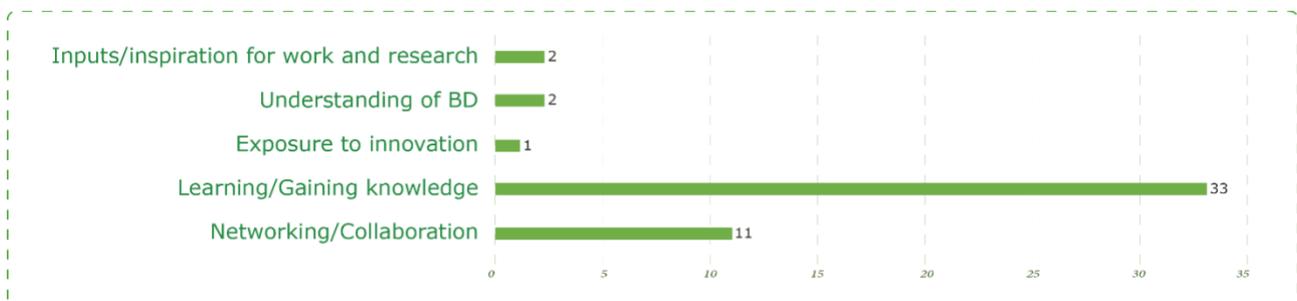
5.4.1 Effectiveness

Thirty-six respondents stated that they were members of the Communities of Practice (CoPs). Twenty-two were part of one CoP, 10 were members of 2 CoPs, 3 respondents were part of 3 CoPs, and 1 respondent was a member of 5 CoPs. (Figure 30).

Annex 5 - Figure 29: Total number of CoPs in which respondents participate (n=36)
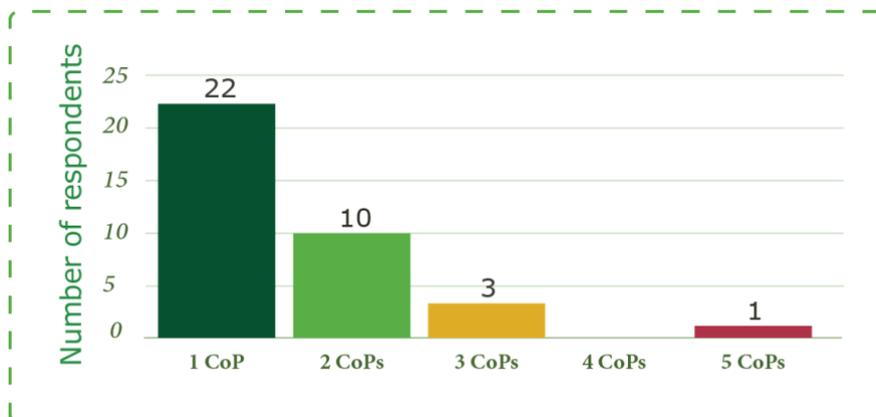
Out of the 36 respondents involved in CoPs, the majority (31) answered more detailed questions regarding their satisfaction. 68% felt that their expectations with the engagement in the CoP were met (either agreed or strongly agreed), 19% were neutral, 10% were not satisfied and the remaining 3% did not know (Figure 31). In terms of effectiveness, less than half (about 43%) of the respondents strongly agreed that "Through the Platform, it is possible to engage with a wider agriculture data and innovation network" (Figure 32).

Finally, 95 respondents answered the question of whether they had attended any courses and seminars provided by the Platform. 58 responded that they had attended the Platform's seminars and course (out of which 9 were external partners), half (50%) were satisfied and 21% highly satisfied with the courses **(Figure 33)**. Among external partners, 3 were highly satisfied, 4 satisfied and 1 neutral.

Annex 5 - Figure 30: Effectiveness – Expectations with the Communities of Practice were met (n=31)
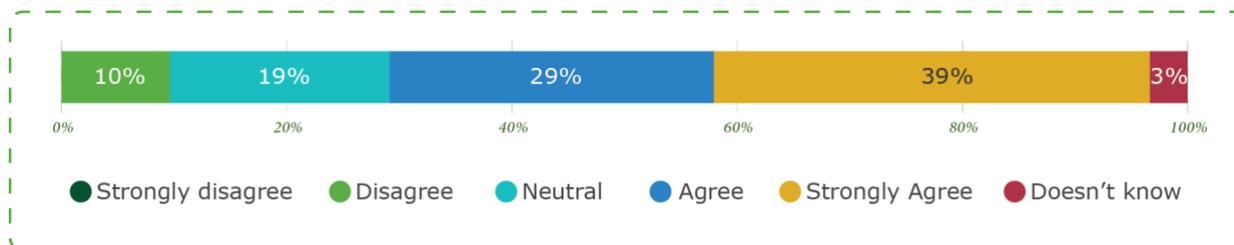
Annex 5 - Figure 31: Effectiveness – Through the Platform, it is possible to engage with a wider agriculture data and innovation network (n=49)
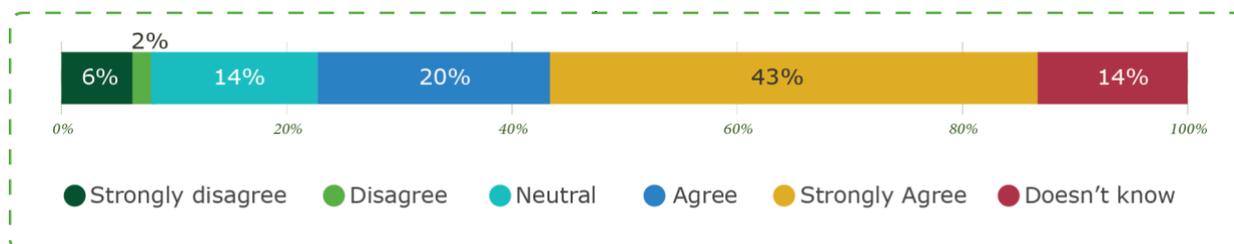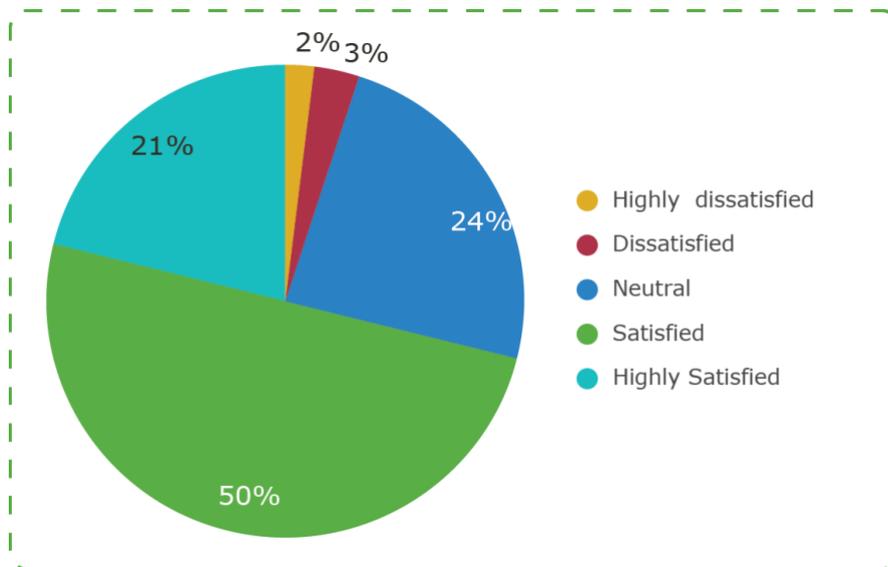
5.5. Module 3 – INSPIRE

Only 9 respondents out of 106 participated in the Inspire Challenge with only one selected as a finalist. On a scale of 1 to 5, this person gave a ranking of 3 to the fairness of the selection process. Among those not selected, only 3 out of 8 received feedback regarding their application.

Annex 5 - Figure 33: Year of first participation in the Inspire Challenge (n=9)
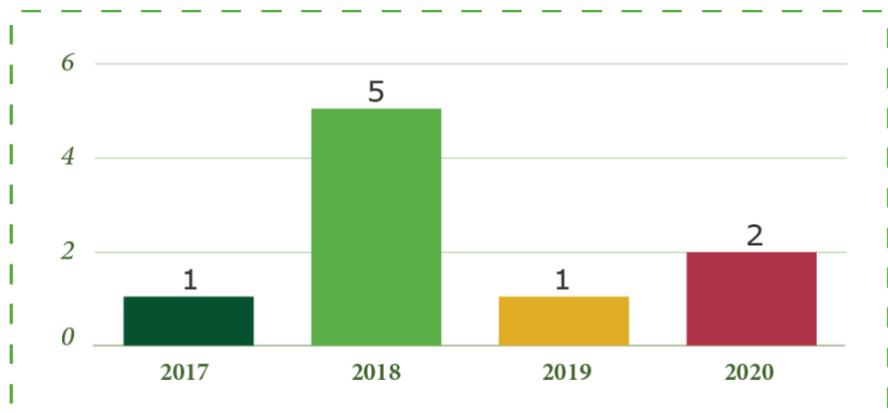
5.6 Effectiveness

Fifty-three respondents answered questions regarding the ability of the Inspire Challenge Grant Process to make significant contributions to digital innovations for agricultural research and development. 59% agreed or highly agreed with the statement, 17% were neutral, 21% did not know and the remaining part disagreed or strongly disagreed **(Figures 35)**. Notably, among the 9 participants of the Challenge, five either strongly or simply agreed with the statement, while one (1) disagreed, one (1) was neutral and one (1) did not know.
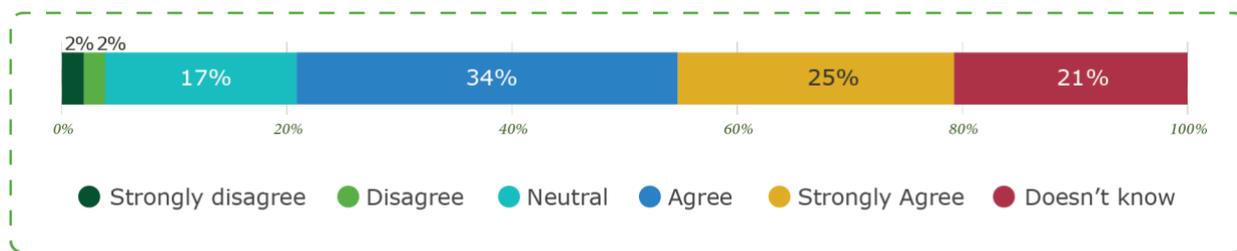
5.7 Survey Limitations

One of the main limitations of the survey is the limited number of respondents in selected groups, which prevented a more in-depth and disaggregated analysis. The evaluation team compared answers among age groups, gender, main field of work, and main role at work but differences were not statistically significant for any of the answers. In most cases, it was not possible to conduct tests for statistical significance due to the small sample size (less than 30 people per group). This limitation was greater for those questions that were addressed only to some specific stakeholders rather than to the whole sample. Comparisons were conducted using a Survey Monkey tool that automatically calculates differences between groups and runs tests to assess their statistical significance.

For gender and age, tests using the software STATA were run. First, the evaluation team conducted the Shapiro-Wilk test for each variable to assess the normality of distribution and then the Lavene's test to check for homogeneity of variance. Where the sample size was more than 30 respondents per group, parametric (t-test) and non-parametric (Kruskal-Wallis) tests were conducted to assess whether the differences in means were statistically significant. Pairwise differences between groups were also assessed when there were more than two groups. In the case of age, the evaluation team compared first, all groups and then divided the sample into two bigger groups, "Generation 1" including people aged 18-44, and "Generation 2" including people older than 44. Including the common Pearson chi-squared and the likelihood ratio chi-squared, the evaluation team also checked the statistically significant difference by gender and each per each answer category (for example strongly disagree, strongly agree, etc.).

Differences between male and female respondents were not statistically significant for any of the questions asked. In the case of "age" the evaluation team found a statistically significant difference at 5% significance level between generation groups for the statement "The Big Data Platform's products, analytical tools and/or activities I engaged in add value to my work". Differences in average rankings for other statements were not statistically significant at 5% significance level.

Finally, a core limitation was the extremely limited number of respondents who had participated in the Inspire Challenge. This did not allow the evaluation team to make conclusions on respondents' views about Module 3.

5.8 Invitation E-mail (sent in English and Spanish)

Dear BIG DATA Community,

Please receive this survey on behalf of the Evaluation Team of the CGIAR Big Data in Agriculture Platform

The CGIAR Advisory Services Shared Secretariat (CAS Secretariat), per its mandate and approved Workplan, commissioned our external evaluation team to conduct the evaluation of the CGIAR Platform for Big Data in Agriculture. This evaluation will serve the dual purposes of accountability and learning, towards One CGIAR.

If you have interacted with the platform, or simply have an interest in the subject, we would like to know

your opinions by filling the online survey.

The survey should not take more than 5-10 minutes of your time. Please submit your responses as soon as possible and no later than **Thursday, September 30th, 2021, at 11.59 p.m. Anywhere on Earth (UTC -12).** All responses will be kept anonymous and confidential.

If you have any questions or concerns about the survey, please do not hesitate to contact the evaluation team:

Ibtissem Jouini: i_jouini@evalchange.com Stefania Sellitti: S.Sellitti@cgiar.org

Sincerely,
The CGIAR Platform for Big Data in Agriculture Evaluation Team

5.9 Questionnaire

Respondents' profile

Question	Question Type	Answers
Gender	Multiple choice	a) Female b) Male c) Non-binary d) Rather not say/Not sure
Age	Multiple choice	a) Less than 18 b) 18-24 c) 25-34 d) 35-44 e) 45-54 f) 55-64 g) 65+
Nationality	Drop down list	Include all nationalities
Current/Usual Country of Residence (No-Covid location)	Drop down list	Include all nationalities
What is your main field of work?	Multiple choice	a) Universities and research institutes b) Private sector (including farm, small- medium enterprises) c) International Organizations, including CGIAR research centers d) National and international NGOs, Foundations, not-for-profit sector e) Governmental and public institutions f) Other _____
What's your role in your place of work?	Multiple choice	a) Leadership, member of a governance body b) Management and Administrative Staff c) Scientists /Researcher d) ITC Staff/Data scientist/Data management e) Consultant f) Other _____
What is the highest level of education you have completed?	Multiple choice	a) Post-secondary school qualifications (non-tertiary) b) Bachelors or equivalent c) Master's/Graduate d) PhD/Doctorate

Engagement with the Platform

In this section we would like to learn more about your background and type of engagement with the CGIAR Platform for Big Data in Agriculture.

#	Question	Question Type	Answers
A1	How do you engage with the CGIAR Platform for Big Data in Agriculture? *Select all that apply*	Checkboxes	a) CGIAR Center - *Select if you are/have been a staff member of one of CGIAR Centers* b) CGIAR Research Program (CRP) or Platform - *Select if you are/have been a researcher or staff of one of the CRPs/Platform* c) Member of CoPs - *Select if you are a leader, coordinator, or member of a Community of Practice* d) External partner - *Select if you are a partner of the CGIAR Platform for Big Data* e) Participant to the Inspire Challenge - *Select if you ever applied (whether you were selected or not) to the Innovate Challenge* f) Participant in the Annual Conventions - *Select if you participated to one or more of the Platform Annual Conventions* g) User of GARDIAN - *Select if you accessed and used the GARDIAN Platform or your publications are in the Guardian Platform* h) *I have not engaged with the platform_____*
A1.1	In which CGIAR Center/Alliance/ Organization do you work?	Checkboxes	a) AfricaRice b) CIFOR c) ICARDA d) ICRISAT e) IFPRI f) IITA g) ILRI h) CIMMYT i) CIP j) IRRI k) IWMI l) Alliance of Bioversity International and CIAT m) ICRAF n) WorldFish o) CGIAR System Organization
A1.2	In which CGIAR Research Programs (CRP) or Platforms do you work?	Multiple choice	a) A4NH b) GLDC c) WHEAT d) CCAFS e) Livestock f) PIM g) FISH h) MAIZE i) RTB j) FTA k) RICE l) WLE m) Genebanks Platform n) Excellence in Breeding (EiB) Platform o) GENDER Platform p) Big Data Platform q) Other

#	Question	Question Type	Answers
A1.3	Which Community of Practice, coordinated by the Big Data Platform, do you belong to?	Multiple choice	a) Data-driven agronomy b) Crop modeling c) Geospatial data d) Ontologies data e) Info and data management f) Livestock data g) Socio-economic data
A1.4	What is your role within the Community of Practice	Multiple choice	a) CoP member b) CoP leader c) CoP administrator d) Other _____
A1.5	Which type of external partner do you consider yourself affiliated with?	Multiple choice	a) Academia b) Private sector c) Policy maker d) International organizations/Foundation e) Research institute f) Other _____
A1.6	In which year did you participate in the Inspire challenge competition?	Multiple choice	a) 2017 b) 2018 c) 2019 d) 2020
A1.7	Were you selected as a finalist?	Binary	a) Yes b) No
A1.7.1	Have you received any feedback on your application?	Binary	
A1.7.2	Comment (optional)		
A1.7.3	How soon did you hear about the outcome?	Multiple choice	a) After less than a month b) Between 1 and 3 months c) Between 4 and 6 months d) Between 6 months and a year e) More than a year
A1.7.4	How fair/open would you rate the process? (Please rate 1 to 5, with 1=not fair; 5=highly fair)	Likert scale	
A1.8	When did you participate in the convention?	Multiple choice	e) 2017 f) 2018 g) 2019 h) 2020
A1.8.1	To which extent were your expectations met?	Likert scale	1 to 5
A1.8.2	What was the added value of	Open question	

#	Question	Question Type	Answers
	the Convention to your work? (Optional)		
A1.9	How often do you make use of the GARDIAN portal?	Multiple choice	a) Every day b) At least once a week c) At least once a month d) At least once every 6 months e) Less often than the above
A1.10	For what purpose did you first start to use the GARDIAN portal?	Multiple choice	a) Academic Thesis, (e.g. Bachelor Masters, Doctoral) b) Professional, non-academic research/work c) Interest not related to academic or professional work d) Other _____
A1.11	How did you learn about GARDIAN?	Multiple choice	a) Search Engine (Google, yahoo, etc.) b) Recommended by a colleague external to CGIAR network c) From my CGIAR network (Center, Research Program, Community of Practice, etc.) d) Social media e) Other _____
A1.12	Do you have your work accessible through GARDIAN	Binary	
A1.12.1	If yes, have you noticed an increased interest in it (i.e. number of consultations)?		A) Yes B) No f) I don´t know

Relevance

Please respond on the extent to which you agree or disagree with the following statements on a scale of 1-5 whereby 1 means "strongly disagree" and 5 means "strongly agree"

#	Statement	Question type	Answer	Audience
B1	The Big Data Platform's products, analytical tools and/or activities I engaged in add value to my work	Likert scale	Points 1-5 + "I don't know"	ALL
B2	The Platform's products and activities address my new and evolving needs.	Likert scale	Points 1-5 + "I don't know"	Centers & CRPS
B3	How did you use the GARDIAN portal? For what type of projects did you find it useful? (Optional)	Open Question		

Effectiveness

Please rate on a scale from 1 to 5, where 1 is "None" and 5 is "Very good"

#	Statement	Question type	Answer	Audience
C0	How would you rate your level of knowledge about the Platform mandate and results?	Likert scale	Points 1-5 +	ALL

Please respond on the extent to which you agree or disagree with the following statements on a scale of 1-5 whereby 1 means "strongly disagree" and 5 means "strongly agree"

#	Statement	Question type	Answer	Audience
C1	The Big Data Platform contributes to CGIAR efforts in fostering digital innovation for agricultural research.	Likert scale	Points 1-5 + "I don't know"	ALL
C2	When I engaged with the Big Data Platform, I felt that my specific needs at professional level were met.	Likert scale	Points 1-5 + "I don't know"	ALL
C3	The Platform has accelerated CGIAR's progress towards mapping data, methods, and tools to support cross analytics in AR4D (i.e. themes, sectors, regions)	Likert scale	Points 1-5 + "I don't know"	ALL
C4	The Platform has accelerated progress towards better data management and stewardship in CGIAR.	Likert scale	Points 1-5 + "I don't know"	CGIAR, CRPs, External Partners
C5	Through the Platform it is possible to engage with a wider agriculture data and innovation network	Likert scale	Points 1-5 + "I don't know"	CGIAR Centers, CRPs, Partners
C6	CGIAR Centers increased their ability to discover agricultural publications through the Platform.	Likert scale	Points 1-5 + "I don't know"	CGIAR Centers and CRPs
C7	CGIAR Centers increased their ability to make effective use of data for analytics thanks to the Platform's support and products.	Likert scale	Points 1-5 + "I don't know"	CGIAR Centers and CRPs
C8	Ethical practices in data collection, improved within CGIAR over the last four years.	Likert scale	Points 1-5 + "I don't know"	Centers and CRPs
C9	Ethical practices in data management and analysis improved within CGIAR over the last for years.	Likert scale	Points 1-5 + "I don't know"	Centers and CRPs
C10	Data search via GARDIAN is time-efficient compared	Likert scale	Points 1-5 + "I don't know"	ALL

#	Statement	Question type	Answer	Audience
	to other similar portals I used.			
C11	Using GARDIAN I can easily **find** Metadata and data I am looking for.	Likert scale	Points 1-5 + "I don't know"	ALL
C12	Using GARDIAN I can easily **access** the datasets I found (including authentication and authorization).	Likert scale	Points 1-5 + "I don't know"	ALL
C13	Using GARDIAN I can easily Integrate datasets I found with other data (including i**nteroperate** with applications or workflows for analysis, storage, and processing)	Likert scale	Points 1-5 + "I don't know"	ALL
C14	Using GARDIAN I can easily reuse the data I found (i.e. Data was well-described so it was possible to replicate and/or combine it in different settings).	Likert scale	Points 1-5 + "I don't know"	ALL
C15	My expectations were met with my engagement in the Community of Practice.	Likert scale	Points 1-5 + "I don't know"	Members of the CoP
C16	The Inspire challenge made significant contributions to digital innovations for research.	Likert scale	Points 1-5 + "I don't know"	CGIAR Centers, CRPs, Partners, participants to the challenge, CoPs.

Sustainability

Please rate the following questions using a scale from 1 to 5. How would you rate….?

#	Statement	Question type	Answer	Audience
D1	CGIAR's preparedness to take on a role of leadership in the international digital agriculture landscape?	Likert scale	Points 1-5 + "I don't know"	CGIAR, CRPs, External partners, Members of CoPs
D1.1	What makes CGIAR ready/not ready to take on such a leadership role?	Open question		CGIAR Centers, CRPs, Partners, CoP members
D2	Did you attend any courses or webinars provided by the Platform?	Binary	Yes/No	All

#	Statement	Question type	Answer	Audience
D2	How satisfied are you with online courses and webinars provided by the Platform	Likert scale	Points 1-5 + "I don't know"	All

Looking forward

#	Question	Type of question	Answer	Audience
D1	Are you aware of any initiatives that showcase the value of (big) data analytics in Agriculture Research for development (AR4D) developed by the Platform or CGIAR?	Multiple choice	Y/N	ALL
D1.1	If yes, can you share title, website, organization/institute...?	Open Question		
D2	What is your dream-scenario for One CGIAR e-research and data-drive impact?	Open Question	_____	CGIAR, CRPs, External partners, Members of CoPs
D3	In which directions or shape should the Platform for Big Data in Agriculture evolve?	Open Question	_____	CGIAR, CRPs, External partners, Members of CoPs

End of the Survey

If you would like an individual interview (remote), or if you are willing to be contacted for follow-up, please leave your E-mail address (optional).	Open question
Is there any person that you think would have insight on these question and to whom we should also send the survey?	Open question

Annex 6: List of Documents Consulted

#	Category / Folder	Document Title	Year	Module
1	CoP Annual Report	CoP Crop Modeling Module 2	2019	M2
2	CoP Annual Report	CoP Crop Modeling Module 2	2020	M2
3	CoP Annual Report	CoP Crop Modeling Module 2	2021	M2
4	CoP Annual Report	CoP CSI Module 2	2020	M2
5	CoP Annual Report	CoP CSI Module 2	2021	M2
6	CoP Annual Report	CoP CSI Module2	2019	M2
7	CoP Annual Report	CoP Data-Driven Agronomy Module	2020	M2
8	CoP Annual Report	CoP Data-Driven Agronomy Module 2	2021	M2
9	CoP Annual Report	CoP Data-Driven Agronomy Module 2	2019	M2
10	CoP Annual Report	CoP Ontology Module 2	2020	M2
11	CoP Annual Report	CoP Ontology Module 2	2021	M2
12	CoP Annual Report	CoP OntologyModule2	2019	M2
13	CoP Annual Report	CoP Socio-Economic Data Module 2	2020	M2
14	CoP Annual Report	CoP Socio-Economic Data Module 2	2021	M2
15	CoP Annual Report	CoP Socio-Economic Data Module2	2019	M2
16	Platform Annual report	CGIAR Platform for Big Data in Agriculture Annual Report 2017	2017	All
17	Platform Annual report	CGIAR Platform for Big Data in Agriculture Annual Report 2018	2018	All
18	Platform Annual report	CGIAR Platform for Big Data in Agriculture Annual Report 2019	2019	All
19	Platform Annual report	CGIAR Platform for Big Data in Agriculture Annual Report 2020	2020	All
20	Plans of Work and Budget - POWB 2020 Module 1	AgroFIMS Module 1	2019	M1
21	Plans of Work and Budget - POWB 2020 Module 1	Bioversity Cross-Cutting Module 1	2019	M1
22	Plans of Work and Budget - POWB 2020 Module 1	IFPRI Cross-Cutting Module 1	2019	M1
23	Plans of Work and Budget - POWB 2020 Module 2	Capacity Building Module 2	2019	M2
24	Plans of Work and Budget - POWB 2020 Module 2	Convention Module 2	2019	M2
25	Plans of Work and Budget - POWB 2020 Module 2	CoP Crop Modeling Module 2	2019	M2
26	Plans of Work and Budget - POWB 2020 Module 2	CoP CSI Module2	2019	M2
27	Plans of Work and Budget - POWB 2020 Module 2	CoP Data-Driven Agronomy Module 2	2019	M2
28	Plans of Work and Budget - POWB 2020 Module 2	CoP Ontology Module 2	2019	M2
29	Plans of Work and Budget - POWB 2020 Module 2	CoP Socio-Economic Data Module 2	2019	M2
30	Plans of Work and Budget - POWB 2020 Module 2	Developpan-CGIAR Digital Strategy	2019	M2
31	Plans of Work and Budget - POWB 2020 Module 2	Evidence Clearing House Module 2	2019	M2

#	Category / Folder	Document Title	Year	Module
32	Plans of Work and Budget - POWB 2020 Module 2	New Data Products Module 2	2019	M2
33	Plans of Work and Budget - POWB 2020 Module 2	Shared Big Data Services for CGIAR Module 2	2019	M2
34	Plans of Work and Budget - POWB 2020 Module 3	Inspire Challenge Module 3	2019	M3
35	Plans of Work and Budget - POWB 2020 Module 3	Inspire Challenge Scale Up Runner Up 2019: Revealing informal food flows through free WiFi	2019	M3
36	Plans of Work and Budget - POWB 2020 Module 3	Inspire Challenge Scale Up Runner Up 2019: Using Commercial Microwave Links (CML) to estimate rainfalls	2020	M3
37	Plans of Work and Budget - POWB 2020 Module 3	Inspire Challenge Scale Up Winner 2019: Pest and disease monitoring by using artificial intelligence	2019	M3
38	Plans of Work and Budget - POWB 2020 Module 3	Inspire Challenge Winner 2019: Hungry cities: Inclusive food markets in Africa	2019	M3
39	Plans of Work and Budget - POWB 2020 Module 3	Inspire Challenge Winner 2019: Real-time East Africa live ground water use database	2019	M3
40	Plans of Work and Budget - POWB 2020 Module 3	Inspire Winner Scale Up Runner Up 2019: An integrated data pipeline for small-scale fisheries	2019	M3
41	Plans of Work and Budget - POWB 2020 Module 3	InspireChallengeWinner2019: Rapid genomic detection of aquaculture pathogens	2019	M3
42	POWB - BigDataSecretariatManagerial	POWB – BigData Secretariat Managerial		All
43	Programs Descriptions and Reports - Proposal documents	1. Base Budget CRP and Flagship Big Data Module 1		M1
44	Programs Descriptions and Reports - Proposal documents	2. Base Budget CRP and Flagship Big Data Module 2		M2
45	Programs Descriptions and Reports - Proposal documents	3. Base Budget CRP and Flagship Big Data Module 3		M3
46	Programs Descriptions and Reports - Proposal documents	Big Data – Addendum: response to the Full proposal ISPC Commentary and other adjustments		All
47	Programs Descriptions and Reports - Proposal documents	CHAIR'S SUMMARY3rd System Council meeting, 23 November 2016	2016	All
48	Programs Descriptions and Reports - Proposal documents	Comments from the Consortium Office on IA Big Data Platform		All
49	Programs Descriptions and Reports - Proposal documents	ISPC Commentary on the full proposal for the "Leveraging CGIAR data: Bringing Big Data to Agriculture, and Agriculture to Big Data" Platform for Phase 2 (2017-2022)	2016	All
50	Programs Descriptions and Reports - Proposal documents	Leveraging CGIAR data: Bringing big data to agriculture, and agriculture to big data - Full Proposal and Resubmission	2016	All
51	Management Team minutes 2020 Google Drive	Management Team meeting minutes 2020	2020	All
52	Program Management and Governance	CGIAR Platform for Big Data in Agriculture International Advisory Board Terms of Reference – July 3, 2017	2017	All
53	Program Management and Governance	CGIAR Platform for Big Data in Agriculture Management Team Fact Sheet –Updated November 8, 2017	2017	All
54	Program Management and Governance	CGIAR Platform for Big Data in Agriculture Steering Committee Terms of Reference –August 3, 2017	2017	All
55	Steering Committee minutes 2020 Google Drive	Steering Committee meeting minutes 2020	2020	All
56	Program Financial Reporting and Management	CIAT Financial Statements December 31, 2019	2019	All
57	Program Financial Reporting and Management	IFPRI Financial Statements and Schedules	2020	All

#	Category / Folder	Document Title	Year	Module
58	List of projects	2020. List of projects 2020.xls	2020	All
59	List of projects	2021. List of projects POWB 2021 - Module 3 Inspire.xls	2021	M3
60	List of projects	2021. List of projects POWB 2021.xlsx	2021	All
61	Program Staffing and Partners	2020. Partners and projects they relate	2020	All
62	Program Staffing and Partners	2020. Program Staffing and Partners	2020	All
63	Program Staffing and Partners	Platform Partnerships report		All
64	Project documents and reports	List of innovations.xls		All
65	Project documents and reports	List of policies.xls		All
66	Project documents and reports	Outcome Case Studies - Reporting 2020 Evidences	2020	M3
67	Project documents and reports	Project Highlights Reported in 2020	2020	All
68	Digital Strategy OneCGIAR	ACCELERATING CGIAR'S DIGITAL TRANSFORMATION	2019	All
69	Digital Strategy OneCGIAR	CGIAR 2030 Research and Innovation Strategy	2020	All
70	Digital Strategy OneCGIAR	Chair's Summary113th System Management Board ('SMB') Meeting	2019	All
71	Digital Strategy OneCGIAR	DIGITAL INVESTMENTS OR INITIATIVES MATURITY AND PRIORITIZATION QUESTIONS		All
72	Digital Strategy OneCGIAR	DMMA MATURITY ASSESSMENT REPORT	2020	All
73	Digital Strategy OneCGIAR	ENTERPRISE ARCHITECTURE MAPPING	2019	All
74	Digital Strategy OneCGIAR - Focus group summaries	DIGITAL STRATEGY GENEBANKS PLATFORM FOCUS GROUP WORKSHOP		All
75	Digital Strategy OneCGIAR - Focus group summaries	DIGITAL STRATEGY COMMUNICATIONS FOCUS GROUP WORKSHOP SYNTHESIS/ COMPILATION		All
76	Digital Strategy OneCGIAR - Focus group summaries	DIGITAL STRATEGY CROP MODELING FOCUS GROUP WORKSHOP SYNTHESIS/ COMPILATION		All
77	Digital Strategy OneCGIAR - Focus group summaries	DIGITAL STRATEGY EiA FOCUS GROUP WORKSHOP		All
78	Digital Strategy OneCGIAR - Focus group summaries	DIGITAL STRATEGY GENDER PLATFORM FOCUS GROUP WORKSHOP		All
79	Digital Strategy OneCGIAR - Focus group summaries	DIGITAL STRATEGY GEOSPATIAL FOCUS GROUP WORKSHOP		All
80	Digital Strategy OneCGIAR - Focus group summaries	DIGITAL STRATEGY IDM FOCUS GROUP WORKSHOP		All
81	Digital Strategy OneCGIAR - Focus group summaries	DIGITAL STRATEGY IT FOCUS GROUP WORKSHOP SYNTHESIS		All
82	Digital Strategy OneCGIAR - Focus group summaries	DIGITAL STRATEGY MEL FOCUS GROUP WORKSHOP SYNTHESIS		All
83	Digital Strategy OneCGIAR - Focus group summaries	DIGITAL STRATEGY ONTOLOGY CoP FOCUS GROUP WORKSHOP SYNTHESIS		All
84	Digital Strategy OneCGIAR - Information Architecture Maps	INFORMATION ARCHITECTURE CIAT		All
85	Digital Strategy OneCGIAR - Information Architecture Maps	INFORMATION ARCHITECTURE RESEARCH IT CIMMYT		All
86	Digital Strategy OneCGIAR - Information Architecture Maps	INFORMATION ARCHITECTURE RESEARCH IT CIP		All
87	Digital Strategy OneCGIAR - Information Architecture Maps	INFORMATION ARCHITECTURE RESEARCH IT ICRISAT		All
88	Digital Strategy OneCGIAR - Information Architecture Maps	INFORMATION ARCHITECTURE RESEARCH IT IRRI		All

#	Category / Folder	Document Title	Year	Module
89	Digital Strategy OneCGIAR - Information Architecture Maps	INFORMATION ARCHITECTURE RESEARCH IT IWMI		All
90	Digital Strategy OneCGIAR - Information Architecture Maps	INFORMATION ARCHITECTURE RESEARCH IT WorldFish		All
91	Digital Strategy OneCGIAR - Information Architecture Maps	Toward a Digital One CGIAR Strategic research on digital transformation in food, land, and water systems in a climate crisis	2021	All
92	Inspire Challenge Evaluation	CGIAR Platform for Big Data in Agriculture Inspire Challenge Review (2017-2020)	2021	M3
93	Misc. June 24th Request	1. Big Data External Partners 2020		All
94	Misc. June 24th Request	2. Big Data 2017_Attendee List		M2
95	Misc. June 24th Request	2. Big Data 2018_Attendee List		M2
96	Misc. June 24th Request	2. Big Data 2019_Attendee List		M2
97	Misc. June 24th Request	2. Big Data 2020_Attendee List		M2
98	Misc. Program Participant Agreements	AfricaRice Program Participant Agreement	2017	All
99	Misc. Program Participant Agreements	Bioversity Program Participant Agreement	2017	All
100	Misc. Program Participant Agreements	CIFOR Program Participant Agreement	2017	All
101	Misc. Program Participant Agreements	CIMMYT Program Participant Agreement	2017	All
102	Misc. Program Participant Agreements	CIP Program Participant Agreement	2017	All
103	Misc. Program Participant Agreements	ICARDA Program Participant Agreement	2017	All
104	Misc. Program Participant Agreements	ICRAF Program Participant Agreement	2017	All
105	Misc. Program Participant Agreements	ICRISAT Program Participant Agreement	2017	All
106	Misc. Program Participant Agreements	IFPRI Program Participant Agreement	2017	All
107	Misc. Program Participant Agreements	IITA Program Participant Agreement	2017	All
108	Misc. Program Participant Agreements	ILRI Program Participant Agreement	2017	All
109	Misc. Program Participant Agreements	IRRI Program Participant Agreement	2017	All
110	Misc. Program Participant Agreements	IWMI Program Participant Agreement	2017	All
111	Misc. Program Participant Agreements	WorldFish Program Participant Agreement	2017	All
112	Assessments and Reviews	ISPC Assessment of the Platform on Big Data revised proposal (2017-2022)	2016	All
113	CAS Secretariat Meeting with The Platform	Big Data Platform_Information		All
114	CAS Secretariat Meeting with The Platform	CAS Secretariat-Big Data Meeting Highlights_Feb17.docx	2021	All
115	IEA_Expression of Interest for CGIAR Big Data Analytics Platform	CIAT Big Data and ICT EoI submission 14-8-2015	2015	All
116	IEA_Expression of Interest for CGIAR Big Data Analytics Platform	CIAT_Big Data and ICT Cross cutting platform EoI_ Cover letter.pdf		All
117	IEA_Expression of Interest for CGIAR Big Data Analytics Platform	ICARDA EOI_CoordinatingPlatforms_August11 FINAL		All

#	Category / Folder	Document Title	Year	Module
118	IEA_Expression of Interest for CGIAR Big Data Analytics Platform	ICRISAT CP Big Data ICT ICRISATvF		All
119	IEA_Expression of Interest for CGIAR Big Data Analytics Platform	IFPRI 2015 - EOI - Big Data ICT Coordination Platform final 081415		All
120	IEA_Expression of Interest for CGIAR Big Data Analytics Platform	ILRIICRAF EOIs_CoordPlat-BigData__17Aug final 1000 hrs		All
121	Initial Q&A and requests 27July	2018 Annual Report. OutcomesCaseStudiesSummary-BigData-Reporting		M3
122	Initial Q&A and requests 27July	2018 Convention Feedback		M2
123	Initial Q&A and requests 27July	2019 Annual Report. OutcomesCaseStudiesSummary-BigData-Reporting		M3
124	Initial Q&A and requests 27July	2019 Convention Feedback (Responses)		M2
125	Initial Q&A and requests 27July	2020 Annual Report. OutcomesCaseStudiesSummary-BigData-Reporting		M3
126	Initial Q&A and requests 27July	Annex 1 of Research Collaboration Agreement - Drone Tech	2020	All
127	Initial Q&A and requests 27July	CGIAR Platform for Big Data in Ag_2020 Convention Report		M2
128	Initial Q&A and requests 27July	CGIAR_OFDA_Policy_Approved_16April2021		M1
129	Initial Q&A and requests 27July	Consolidated list of projects from 2017 to mid-2021	2020	All
130	Initial Q&A and requests 27July	Final virtual meeting Agenda GR Policy WG March 24-26 2020 v4		All
131	Initial Q&A and requests 27July	International Advisory Board (IAB) composition		All
132	Initial Q&A and requests 27July	List of participants - fourth meeting of the scientific advisory committee on the global information system		All
133	Initial Q&A and requests 27July	Module 1. Focalpoints		M1
134	Initial Q&A and requests 27July	Research Collaboration Agreement - Drone Tech		All
135	Initial Q&A and requests 27July	Steering Committee (SC) Composition 2017-2021		All
136	Introductory Meeting- 28 July 2021	BDP_CAS_EvaluationDeck_BK_MD		All
137	Introductory Meeting- 28 July 2021	Notes Pages Big Data Presentation		All

Annex 7: Component Studies Executive Summaries

Annex 7.1 ORGANIZE

Didier L. et al

As part of CGIAR research activities, CGIAR generates datasets and publications but is CGIAR a Big Data provider or Big Metadata provider? The slightly misleading term of Big Data, more often seen from the angle of volume or velocity than the angle of variety more appropriate to datasets produced by CGIAR may have had some impacts internally but also externally on clarity of the mission of the CGIAR Platform for Big Data in Agriculture. Nevertheless, other messages like Open Access were well-received together with the willingness to contribute to the Platform, including from CGIAR Center's data managers.

Among the three modules supporting the Big Data Platform, ORGANIZE deals with knowledge management in its digital form, data collection tooling, data analysis, and facilities sharing services with metadata discovery (datasets and publications) to enable optimal reuse of data found relevant to a specific interest related to a scientific investigation. Optimizing data reusability is the goal of this module with publication coupling in a search query to associate scientific knowledge and context to a dataset or an integrated collection of datasets. A full report for the ORGANIZE module was drafted according to a validated methodology to evaluate the Big Data platform across the three modules. Interoperability of datasets and their metadata was then the driving motto of the activities of the module with the focus on Open Access and Open Data (OAOD) policy and the Findability, Accessibility, Interoperability, Reusability (FAIR) principles of the available resources with guidance for new resources. Each of the 15 CGIAR Centers had its own metadata service and the purpose of Big Data's GARDIAN platform was to harmonize the views on these repositories while evaluating their fitness for purpose using the developed FAIR metric and preparing for the metadata enrichment using ontologies adopted or developed in the Communities of Practices (CoPs) of the CONVENE module. GARDIAN successfully achieved a one-stop-shop for datasets including coupling with relevant publications.

The harvesting process from the existing Center's repositories is quantitatively a success, GARDIAN shows an increase of availability of data and publications throughout the project (2017 to 2021). Overall, the contributions and efforts from the Centers towards OAOD have been very positive with raised awareness of interoperability principles and an increase of +30% of OA for data and publications, comparing the three years before the launch of GARDIAN to the three years after (2018-2020). However, regarding the entire collection of metadata (in August 2021), only 49% of publications and 28% of datasets were available as OA with a change of +5.5% and +10% respectively since the launch of the Platform. The FAIR principles captivated minds and contributed to this very good result concerning new entries but the FAIR metrics used are misleading the level of interoperability achieved to reuse the datasets. For Open Science, open access is necessary but not sufficient. The FAIR metric itself may need improvements and refinements, e.g., the Reusability score calculated as the average of the other three scores (Findability, Accessibility, and Interoperability) is not informative or sufficient.

Contributing to FAIR is the richness of metadata as part of the ORGANIZE module and the use of the work from the CoP ontology for example. However, this is not visible in the metadata records and the metadata service (database) does not seem to allow annotations other than the AGROVOC vocabulary. Some Centers mention using them in their local repositories but a loss of richness of metadata occurs during harvesting from GARDIAN. Overall, no feedback from CoPs, CS2-CONVENE, or CS3-INSPIRE is sufficiently visible in GARDIAN to illustrate a synergy of work. The GARDIAN geospatial data exploration is very interesting but only a demo without links to the metadata search. A data collection mobile tool for in-field measurement, AgroFIMS, has been released to harmonize data and metadata right from their creation. Unfortunately, two years after the launch of AgroFIMS the upload of the data and metadata following the CG Core is not yet available.

On the computational facility to enable data analytics the proposed environment CG Labs appears powerful with cloud-based storage and shared spaces to initiate collaboration between data scientists and

other researchers. From its set-up, it is a useful but also empty shell and besides sharing several existing R and python libraries[6] it did not take on with visible demonstrations of its usefulness at this stage.

Conceptually, the GARDIAN proposal and the FAIR metrics were seen as useful but in practice, did not generate a high volume of traffic (10 times less than the Big Data Website alone in 2020). The interest appears to have grown from about 250 monthly users in April 2020 to 950 around December 2020 until May 2021 and decreased to 350 in July 2021. Comparing behaviors between 2020 and 2021, the small relative decrease in user's visits to all pages occurs in favor of browsing the publications (results from a search) is nonetheless encouraging in terms of usability. In relation to web traffic, the existing CGSpace may have been competing with GARDIAN, as both are providing metadata search services across CGIAR, however using CGSpace for publications and GARDIAN as a specific platform for data without losing the capacity to query for both at the same time from GARDIAN is possible using interoperability settings.

The focus on FAIR principles encapsulating interoperability settings was an efficient implementation strategy to capture minds and raise awareness of OAOD, with concepts and clear benefits across Centers and practical actions with common goals in relation to raising FAIR metrics scoring.

Achievements of the ORGANIZE module are going in the right direction but still have a long way to go to enable seamless data-driven analytics as leverage to innovation agriculture research. The online survey revealed that CGIAR's Platform for Big Data in Agriculture was fostering innovation in agriculture research but when asked about CGIAR as a leadership voice in digital agriculture, half of the sample skipped the question (n=55). As contributing to an increase in interdisciplinarity, progress has been made and a small but significant increase in Center collaboration (publications involving two or more Centers) was observed. Several outcomes of the Big Data platform should be useful to the design of the next open digital e-infrastructure initiative, but lessons need to be learned.

7.1.1 Recommendations:

Directions taken in the ORGANIZE module were compatible with One CGIAR's 2030 Strategy concerning interoperability aspects with the FAIR motto, but greater attention needs to be given to the interdisciplinary aspects of such initiative which has an impact on each facility the e-infrastructure has to offer. This means:

1. Change the wording Big Data so the name of the platform adapts to the CGIAR focus on the variety of datasets and needs for interoperability towards addressing 'Harnessing Digital Technologies for Decision-making across Food, Land and Water Systems'.
2. Demonstrate interoperability benefits with data integration.
3. Develop greater knowledge management from the great variety of datasets to meet interoperability requirements.
4. Give attention to metadata standardization (without replicating models already available) with appropriate semantic annotations, metadata on data quality, and metaquality (e.g., FAIR description and metrics).
5. Develop visual analytics that are appropriate to CGIAR domains, and primarily at the basis, the geographical and temporal aspects (date and reference period) for the Platform but also in combination with semantic queries results.
6. Add a semantic ontological knowledge base (semantic engine) in addition to semantic enrichment of metadata from harmonized vocabulary and ontological terms, i.e., providing semantic reasoning along the keyword searching discovery capacity
7. Share internal services in an interoperable manner (metadata and data services) using standards from international bodies, e.g., as Open Geospatial Consortium (OGC), World Wide Web Consortium (W3C).

---

[6] To facilitate uptake of data analytics in CG labs, the ORGANIZE Module team provided, in collaboration with the geospatial CoP, a series of existing data science and machine learning libraries in R and python, packed as a community Jupyter Docker Stack (available at https://hub.docker.com/r/scioquiver/notebooks) . Their use within CG labs has been limited despite a relatively high download (not necessarily to be used within CG labs).

8. Develop data analytics using interoperable services provided and be aware of the knowledge structuration.
9. Record end-user usage, and results from data analytics methods, with feedback to the knowledge structuration, i.e., usage and results as dynamic metadata and 3rd type of the Open Science aspects completing publications and datasets, e.g., scripts, models, and software (models as statistical or machine learning but also biophysical models, crop models).
10. Evaluate the role and integration of this specific e-infrastructure within the existing landscape of CGIAR services, for example, CGSpace being more known as the reference point to look for CGIAR publications metadata.
11. Track web analytics systematically from the beginning to evaluate impact in terms of usability for CGIAR researchers and outside for each new launch and facilities provided, then, be able to refine these facilities accordingly. This monitoring is also useful in relation to Quality of Science (e.g., view and download metrics).
12. Integrate efforts from the design and development of this e-infrastructure with efforts in demonstrating its usability (i.e., with CONVENE and INSPIRE).

Annex 7.2 CONVENE

Mathew K. et al

With the launch of a new research modality to advance the One CGIAR 2030 Research and Innovation Strategy, making the digital revolution central to the way of working is one of seven new implementation approaches prioritized in the strategy (the seventh way of working). Key elements of this approach include 'engagement with partners in developing cutting-edge, context-appropriate digital solutions, improved access to and use of data and digital innovations targeting small-scale farmers, pursuing new digital applications to accelerate learning and knowledge sharing among partners underpinned by principles of findability, accessibility, interoperability, and reusability (FAIR) for all CGIAR data'. Thus, leveraging the unprecedented opportunity provided by today's digital revolution is at the front burner of One CGIAR to accelerate progress towards the achievement of the Sustainable Development Goals (SDGs). The full proposal of the (as then called) Big Data Coordination Platform (now the Platform for Big Data in Agriculture – the Platform) expressed a rationale of elevating its visibility in the global community on data science. The proposal claims that CGIAR is "largely invisible in the broader field of big data analytics and ICT development". The CONVENE module of the Platform has an expressed desire to bring the big data community to agriculture to demonstrate the exciting science and business opportunities that exist in rural development and agriculture in developing countries.

This Component Study makes three assertions. The first assertion is that there is a distinction to be made between increasing visibility of CGIAR in the data science community and employing big data analytics to enhance the impact of agricultural research-for-development (AR4D); both for researchers with diverse disciplinary skill sets and the wider NARES/extension and Communities of Practice (CoPs). Second, CONVENE has a particularly important 'playmaker' role to undertake to enhance the role of the Platform by supporting the analysis of existing projects, models, and data. This could potentially enable the ORGANIZE module to take a forward-looking view of how FAIR processes may be mainstreamed in AR4D through the selection of dedicated sites (via the INSPIRE module) from where use cases of the applications of big data analytics in supporting evidence-based decision-making could emerge. Third, CONVENE can enable CGIAR to become a learning organization that supports research co-design, co-curation of data and models, and regional partnerships that foster a role for AR4D in the pursuit of the Sustainable Development Goals (SDGs). A robust learning organization has the potential to positively impact the Quality of Science (Qos) (CGIAR, 2020).

The Platform's Annual Convention and the CoP and Working Group mechanisms of the CONVENE module can leverage this strategic opportunity via follow-up proposals that articulate a ToC that supports the monitoring of both the performance of small teams of researchers and CGIAR projects and programs. Lessons emerging from the Agronomy Community of Practice (CoP), notably the co-design workshops and production of 10 use cases, leads us to conclude that are benefits to be derived if the Platform were to proceed in 'mission mode' to demonstrate the applications of big data analytics in transforming water, land, and food systems. For this purpose, we recommend that the CONVENE module should invest time and resources to arrive at a *modus operandi* of how big data analytics can support the seamless integration of biophysical models, decision models, and machine learning models for a given bundle of

CGIAR services constituting a particular domain (for example, biotechnology and life sciences). This will enable the ORGANIZE module to streamline the organization and use of CG Labs and GARDIAN to better amplify the benefits of big data analytics for the benefit of a broad constituency of stakeholders who could be the focus of strategic grants via the Inspire Challenge competition.

7.2.1 Main Findings of the CONVENE Component Study

Evaluation of the Platform found that approximately 80% of participants of CoP's under the CONVENE Module were from outside the CGIAR system. Further, several respondents to the online survey revealed that they were members of multiple CoPs. Out of a total of between 900-1,300 CoP members, membership of the Information and Data Management CoP was the lowest. Approximately 40% of the online survey respondents (n=106) pointed out that the Annual Convention met their expectations and increased their awareness of big data and open access, and helped them network. This suggests that the CONVENE module served as a window to raise awareness of big data and open access issues among a non-CGIAR audience. However, the bigger analytics function of the platform, which would have entailed a needs assessment of CGIAR scientists, was largely overlooked. It should be acknowledged, however, that the Platform undertook wide-ranging internal and external consultations, and an early survey to understand computing needs. Five incidental benefits of this process of engagement include a newsletter by the Ontologies CoP with 471 subscribers and 9,666 YouTube views in 2019. Incidentally, one of the top-line learnings from the CRP evaluation was that there were missed opportunities for leveraging modeling capabilities across programs and domains. This begs the question of why the Platform has not made much headway in developing those relationships considering that it came into effect five years after the CRPs were established.

Notwithstanding the attempt of the Platform to engage, its inability to make headway in forging relationships to advance the research needs for big data analytics can in large part be attributed to a weak theory of change (Toc). The Toc was influenced by: (a) the need to engage with the challenges of international agricultural development and (b) organizational changes due to the research strategy for One CGIAR in 2017. However, the Platform missed an opportunity to effectively articulate how it can support AR4D by establishing a stronger causal link between applications of big data analytics and adoption rates for CGIAR technical and management options. As a result, while some progress has been made to advance the findability and accessibility principles, crucially the interoperability and reuse components of the FAIR principles have remained largely nascent. But the increasing numbers of attendees to the annual conventions offer the hope that issues of interoperability can be addressed more explicitly in the future. One of the benefits of engaging with data interoperability is that the connections between three dimensions of modeling could be better articulated going forward to better exploit the full benefits of FAIR principles via AR4D. The three dimensions of modeling that the Platform can advance in support of AR4D are (a) bio-physical resource models (such as crop, soil, hydrological models), (b) empirical decision-making models (such as agent-based models, foresight, back casting models) and (c) machine learning models focused on data exploration and model prediction. While the slowness of organizational reform within CGIAR may have played a role in distracting from forging relationships necessary to sustain big data analytics, we could find no evidence that this was a major reason (McCalla, 2014).

This evaluation concludes that the selection of Inspire Challenge projects should not be guided by the interests of promoting blue sky thinking on big data applications in agriculture alone but also by the goal of developing use cases that demonstrate the applications of big data analytics via deeper engagement with NARES, local think tanks, and regional universities. In this regard, the CONVENE module has an important role to play in supporting Quality of Science (QoS) through Working Groups and CoP mechanisms. This has the potential to produce long-standing impact through downscaling of global environmental models and upscaling of local crop models and by addressing associated data harmonization challenges. This can undoubtedly impact positively upon adoption rates for CGIAR technical products (for example, seeds and soil conservation models) and build regional capacity to design management models that prepare for and respond to changes in demography, diets, climate, and agricultural terms of trade (Renkow, 2018). Lessons emerging from the Agronomy CoP, notably the co-design workshops and production of 10 use cases lead us to conclude that the Platform should proceed in 'mission mode' to demonstrate the applications of big data analytics to transform water, land, and food systems. For this purpose, we recommend that the CONVENE module focus on reviewing existing CGIAR data, projects, and models to arrive at a *modus operandi* of how it can proceed from bio-physical

models/decision models to machine learning models for a given bundle of CGIAR services constituting a particular domain (for example, biotechnology and life sciences). This will enable the ORGANIZE module to streamline the organization and use of CG Labs and GARDIAN to better amplify the benefits of big data analytics for the benefit of a broad constituency of stakeholders who can potentially become the strategic focus of Inspire Challenge grants.

7.2.2 Recommendations

The theory of change for 'Harnessing Digital Technologies for Timely Decision- making Across Food, Land and Water Systems' (Harnessing Digital Technologies) remains focused on first-generation issues of information technology and misses a crucial element: an understanding of the needs of users of big data analytics services within CGIAR and its network of partners. In other words, how can the Platform support the incubation of questions by users (or science questions) to drive the types of data analysis based on co-curation of meta-data? The big data analytics function of CGIAR can amplify the key role of governments to scale up benefits of AR4D by identifying elements of institutional reform that can lead to positive economy-wide outcomes for cross-cutting themes of youth and gender empowerment that are core priorities of AR4D under One CGIAR. The challenge for the Platform is therefore one of developing a framework for modeling that can help CGIAR make sense of the complexity of socio-ecological systems, predict the next landscape-level risk, and develop regional capacity to respond based on the systematic co-curation of previous CGIAR case studies and models. In line with the recommendation of the 'Synthesis of Learning from a Decade of CGIAR Research Programs', we, therefore, recommend a sharp departure from the ToC outlined by Harnessing Digital Technologies to instead focus on:

1. Developing a prioritization of CGIAR System-level/IDO outcomes with the greatest potential to impact upon SDGs 2 (Zero Hunger) 6 (Clean Water and Sanitation) and 13 (Climate Action). These three SDG goals in particular have the potential to demonstrate the role of big data analytics in advancing sustainability of water-energy-food interactions via AR4D.
2. Developing a ToC that articulates how big data analytics can enable CGIAR research to impact AR4D.
3. Developing a unified framework for modeling with a focus on a given geography to strengthen feedback loops between results of CGIAR field trials, and the design of policy instruments (guidelines, standards, notifications, circulars, and directives) through support to appropriate authorities.
4. Developing a computing workflow for how data will be organized, transformed, and visualized to support identification of a robust monitoring framework that would enable the attribution of policy changes to AR4D.
5. Designing a learning program tasked with identifying verifiable metrics to evaluate a big data pilot intervention at a dedicated site, preferably embedded within a regional network of NARES partners.

Annex 7.3 INSPIRE

Erik B. R. et al

The Inspire Challenge aimed to:

- Bridge the deep subject matter expertise of CGIAR researchers with the capabilities of external partners
- Challenge research organizations to collaborate with industry to leverage public good data

7.3.1. Main Findings of the INSPIRE Component Study

An analysis of the submission process to the INSPIRE Challenge shows that the Platform team did not allocate adequate resources to process and provide more productive feedback to the Inspire Challenge submitters that could have encouraged further development of the Innovation ideas proposed by many CGIAR researchers. In 2020 alone, 513 proposals were submitted and only seven were funded. The Inspire Challenge promoted a rich awareness among CGIAR Centers in exploring new technologies for the digitalization of agriculture, shown by the hundreds of innovation ideas proposed. These proposed innovation ideas were not optimally rechanneled into CGIAR Centers to facilitate exploring new technologies for the digitalization of agriculture.

The Inspire Challenge projects selected through the Inspire process were judged according to a criterion of Innovation, Data Usage, Scalability, Impact, and Sustainability. Therefore, the selected projects were not always tightly connected to CGIAR research priorities. The projects were not selected in relation to goals, related to the building of an internal structure, or making CGIAR data more accessible. There was no clear guidance that specified the objectives of the project and how it would benefit CGIAR research. Most projects were not regularly working with the other two Platform modules (ORGANIZE and CONVENE) to identify and harmonize the existing ontologies necessary for interoperability on the project. The projects did not get clear guidance on how to establish minimum standards for deployment into the Platform or CGIAR research. There was low coordination of the project's activities with the two other modules of the Platform, and with other Inspire Challenge projects.

The evaluation also found that there was a marked participation disbalance in the process and a resulting disbalance in the distribution of Inspire Challenge grants among CGIAR Centers.

One of the main conclusions of the [Inspire Challenge review](#) from March 2021 was that the Inspire Challenge was largely successful at creating partnerships between CGIAR Centers and other digital agricultural innovators and pointed out that most partnerships were with private-sector startups or firms.

This evaluation found that the INSPIRE module's contribution to digital innovations for research and delivery of research cannot be underestimated.

One researcher quoted:

> *"Personally, I think they (Inspire challenge) triggered some really nice collaborations. The challenge system with relatively very small money started some small startup projects. Brought to us some very smart people. That collaboration continues and number of projects, mining social media, signals of concerns, they brought to us great smart computer scientists, combination went very well."*

7.3.2 Recommendations

1. The selection process for new projects that aim to solve agricultural development challenges should develop proper feedback mechanisms and give advice and guidance to project applicants on how to improve the proposals.
2. More resources should be allocated to the team managing the selection process of new research projects to harness the power of digital technologies by giving more extensive feedback.
3. A guidance document should be provided with clear instructions on how the results of the project should be implemented and connect back to the CGIAR Centers, Initiatives, and the Platform, to maximize the impact of selected projects.
4. The winning projects should become test cases for testing new CGIAR capabilities and identifying weaknesses, and not be just stand-alone projects.
5. The Platform should create tighter governance mechanisms between the winning projects, the relevant CGIAR Centers, and the Platform.
6. Attention should be given to achieve a more even distribution of grants between CGIAR Centers.

Annex 8: CoP Case Study

Promoting data annotation for semantic interoperability - Spotlight on the Ontologies Community of Practice

John K. et al

Annex 8.1 Background

The CGIAR Platform for Big Data in Agriculture (the Platform) is intended to harness the capabilities of big data to accelerate and enhance the impact of international agricultural research. A key component of the Platform's theory of change was to establish new partnerships. This was envisaged through elevating CGIAR's visibility and status in the global community on data science as the 'go-to' institution for data analytics in agricultural development. The Platform's CONVENE module sought to enable interactions and collaborations among scientists across CGIAR Centers and CGIAR Research Programs (CRPs) and with external big data partners. Among strategies to achieve this, was to establish or work with existing Communities of Practice (CoPs) and Working Groups (WGs). This spotlight on the Ontologies CoP is part of the external evaluation of the Platform commissioned by the CGIAR Advisory Services (CAS) in 2021.

Several CoPs were created under CONVENE at the onset in 2017, and among them was the *Ontologies CoP* (the other CoPs were *Agronomy Data*, *Crop Modeling*, *Geospatial Data*, *Livestock Data*, and *Socio-Economic Data)*. The Ontologies CoP was an adoption and enhancement of the previously existing Crop Ontology Community Project. The CoP has aimed to support the other CoPs of the Platform to develop ontologies for data description and to identify innovative solutions that can support quality data annotation. The CoP taps into both CGIAR in-house and external expertise for its goals. Success for the Ontologies CoP is envisaged through the adoption of best practices, recommended ontologies, and guidelines in the selection, use, and application of semantics for data harmonization. These are to be applied at different stages of data collection and storage to ultimately enhance semantic interoperability, multi-disciplinary data platforms, and graph databases in the actualization of FAIR principles. The CoP has aimed to expand coverage of ontologies beyond crops to encompass data on fisheries and aquaculture, livestock, socio-economics, water management, and agroecology. The CoP has also sought to stimulate knowledge sharing on ontologies among stakeholders, such as researchers, data managers, domain experts, experts in ontology design, and platform development teams.

The CoP is led by Elizabeth Arnaud from the Alliance of Bioversity International and CIAT. She was the leader of the pre-existing Crops ontology community of practice. The CoP has four working groups named by the ontologies they aim to develop and maintain. These are, (a) *Plant Phenotypes Ontologies* led by the Alliance of Bioversity International and CIAT, (b) *Fish Ontology*; led by WorldFish with support from Bioversity International[7] (c) *Agronomy Ontology*; led by IFPRI and the Alliance of Bioversity international and CIAT, and (d) *Socio-Economic Ontology;* led by IFPRI. CoP membership is through a LinkedIn group with 225 members as of 20th October 2021, a YouTube channel with 300 subscribers, and a newsletter with 472 subscribers as of the same date. A core group of the CoP comprising 32 representatives from CGIAR Centers and external partners meets every month to advance the objectives of the community. The larger membership of the CoP has been meeting at the Annual Big Data Convention to interact in person and to advance the CoPs agenda including documenting accomplishments and designing work plans for the next year. In May 2021, the CoP started exploring governance frameworks for individual ontology products, starting with the Crops Ontology. This included organizing meetings as well as a webinar on the proposed Governance Operational Model for Ontologies (GOMO). However, no substantive efforts were observed to have gone into firming up a governance framework for the CoP.

Annex 8.2 Key Findings

---

[7] CIAT and Bioversity International merged in 2020 to become the Alliance of Bioversity International and CIAT

Our findings on the Ontology CoP are organized by the four main criteria of the external evaluation namely relevance, efficiency, effectiveness, and sustainability as follows:

8.2.1 Relevance

The Ontologies CoP has registered steady growth of individual membership to its LinkedIn group suggesting relevance to the Platform's CGIAR and external stakeholders. Steady growth has also been seen in subscriptions to its newsletter as shown in Figure 1 which covers the period from 2017 until October 2021.

Annex 8 - Figure 1: Membership and subscriptions to Ontologies CoP

The CoP's annual reports indicate the engagement of 21 partners in total throughout 2018, 2019, and 2020. CGIAR Centers and Universities were the most common partners as indicated in Figure 2. In 2018, the CoP registered 10 partnerships followed by an increase to 14 in 2019 after one partnership was dropped and five new ones were acquired. In 2020, the CoP registered 18 partnerships, having dropped two and registered six new ones. A desire to formalize a governance framework for the CoP was expressed by the CoP's leadership, toward an enhanced basis for strengthening partnerships.

Annex 8 - Figure 2: Partnerships registered in 2018, 2019, and 2020 (Source: CoP annual reports)

The steadily increasing number of individual membership and organizational partnerships is indicative of the momentum the CoP has gained throughout the implementation of the Platform. Of note is the likelihood for a large proportion of LinkedIn group members and newsletter subscribers to remain inactive as pertains to reading shared content and more so contributing to or responding to content.

Nonetheless, the richness, diversity, and value of membership and partnerships in the CoP were indicated by the key informants to be of value in the external evaluation. These included sentiments such as, "*(The) Ontologies CoP is very successful, engaging a lot of partners*" and "*I like a lot, the Ontologies CoP; because it tries its best to bring new people inside and bring new knowledge*". High levels of interactivity and knowledge sharing fostered through the CoP's organized online and in-person meetings were evident among the key informants.

8.2.2 Efficiency

The CoP uses webinars as a mode of capacity development among its constituents. Since 2019, the CoP has organized 17 webinars on topics related to ontologies and semantic interoperability. The 17 webinars each had over 40 live attendees; the highest attended webinar recorded 72 attendees[8]. Webinar recordings have been published on the CoP's YouTube Channel. The webinars were delivered by 39 speakers from diverse backgrounds, mostly external partners. This included stimulation of interactions with universities and corporations in the industry such as the BASF Group, Bayer, and the KWS Group. The 33 videos published on the channel have accumulated 10,172 views, about half of which were of the webinar titled "*Machine Learning and Ontology*". The attendance and viewership, diversity of speakers, novelty of the content, and interactivity of sessions suggest efficient use of resources given the low costs associated with organizing webinars and the CoP's fruitfulness in using them to foster capacity development among its stakeholders. The CoP's ability to tap into knowledge and expertise from external speakers of diverse backgrounds is a demonstration of building on the mutual interests and benefits among stakeholders to efficiently deliver on its mandate.

8.2.3 Effectiveness

The CoP was observed to be instrumental in persuading CGIAR colleagues to pay attention to ontologies, not as mere academic standards but as mechanisms for unlocking the potential of big data through semantic interoperability. The Crop Ontology (CO) was initiated in 2008 to address the need for valid lists of defined breeders' traits and variables by breeding data management systems and field books. Maintained by the plant phenotypes working group of the Ontologies CoP, the CO has over 10,000 harmonized and validated variables and covers 33 crops. This is up from the 20 crops it covered in 2016. Its adopters include the integrated breeding program, the Global Agricultural Trials Repository, International Cassava database, the Boyce Thompson Institute, and Bayer. The Agronomy Ontology (AgrO) maintained by the CoP was created in 2014. Adoption of AgrO was also pegged on the uptake of the Agronomy Field Information System (AgroFIMS) by agronomists observed to have been delayed, pending the release of a new version expected by the end of the year. The Socio-economic Ontology (SEONT) being developed by the CoP was tested and validated with IRRI survey data. While CO is accessible via its website, AgrO and SEONT are accessible as open-source repositories under the Agricultural Semantics account on GitHub. We did not find evidence to authoritatively ascertain the extent of adoption by researchers and science of data actors for the CO, AgrO, and SEONT. Although the ontologies' distribution is through web portals, the corresponding traffic statistics cannot reflect the actual usage of the ontologies. The external evaluation interviews captured frustration among stakeholders about the slow uptake of ontologies despite the visible investments in their development. One interviewee stated, "*The CG(IAR) has been spending forever trying to develop ontologies, and people can spend their entire careers developing ontologies, but nobody uses it".* Among future expectations of the CoP's efforts is an ontology of small-scale fisheries and aquaculture. This is following the signing of an MoU between Worldfish and the Alliance of Bioversity International and CIAT in 2020. The CoP published a descriptor article on its workings and results in the *Patterns* journal in 2020. In the same year, the CoP was one of the key stakeholder groups contributing to the Platform's paper on One CGIAR's digital strategy. Through this, the CoP elevated focus on trends such as digital agriculture, machine learning, the data deluge, and the maturity of semantic web technologies as imperatives to strategically address the digital strategy.

8.2.4 Sustainability

That not all the CoPs funding came from the Platform and that the CoP has its roots in an established crops ontology community pre-existing the Platform suggested independence from the Platform for sustainability. Some key informants appreciated the work of the CoP to the extent of appealing for sustaining its work in One CGIAR. Publication of the descriptor article in the *Patterns* journal is not only bound to give CGIAR an elevated profile on the global science of data landscape but to also attract new expertise and partnerships for the CoP. A challenge was made by interview participants to strengthen

---

[8] According to the CoP, 72 attendees participated in the webinar on the use of knowledge graphs at BASF and KWS,

CGIAR's leadership on ontology as they deemed it weak. An internal survey within the CoP underpinned the importance of funding in the identification and use of ontologies in agriculture (Arnaud et al., 2020). The sustainability of the CoP is yet to be demonstrable in One CGIAR especially through funding. A keyword analysis ascertained that all but one of the 19 candidate initiatives for One CGIAR envisaged ensuring data interoperability explicitly by annotating dataset variables with ontologies where possible and controlled vocabularies were not possible. This was invariably under the policy compliance and oversight section of the proposals. As well as indicating attainment of the CoP's and the broader Platform's effort to elevate Open and FAIR data assets (OFDA) to the CGIAR-wide policy level, this suggests sustained relevance of the ontologies CoP in the transition towards One CGIAR. However, the proposals' commitment to "annotating data variables with ontologies" could imply the use of externally curated ontologies or settle for the use of controlled vocabularies as an alternative to ontologies developed and maintained within CGIAR. This suggests that demand for the development and maintenance of ontologies, the primary source of relevance, hence sustainability for the Ontologies CoP is not guaranteed in the new One CGIAR initiatives. Moreover, ontology objectives about data discovery and knowledge modeling were deemed by a key informant to be fully understood by the Platform yet not adequately by the rest of CGIAR.

Annex 8.3 Conclusion and Recommendations

The Ontologies CoP has been a key mechanism for advancing semantic interoperability through data annotation thanks to the Platform's support. It carries a major part of the promise to actualize the Platform's goal of harnessing the capabilities of big data to accelerate and enhance the impact of international agricultural research. As a continuation of the previously existing Crops Ontology Community of Practice, the CoP has grown the crop ontology by the number of concepts and crops covered. It has also advanced the AgrO ontology and the new SEONT ontology. That it was not possible to establish evidence on the extent of uptake for the CO and AgrO ontologies within and outside CGIAR should be of concern for the evaluation of return on ontology-related investments. The CoP has not only identified and documented best practices for identification, selection, and usage of ontologies for data annotation. It has also pursued innovative mechanisms for efficiently harnessing and sharing expertise on ontologies among stakeholders within and outside CGIAR. This has included managing a growing number of institutional partnerships. The external evaluation found no clear indication on where the important role served by the CoP will be housed in One CGIAR.

Ontologies can be expected to play a major role in facilitating semantic interoperability to unlock the potential of big data in agriculture as powerful machine-readable representations of domain knowledge (Goldstein et al., 2021). More demonstrable effort is recommended to entrench the capacity and funding to develop and maintain ontologies for agricultural subdomains within One CGIAR. This is more so if CGIAR is to consolidate its position as a 'go-to' institution for thought leadership on the science of data in agriculture. In addition to instituting measures for rapid development and maintenance of CGIAR led ontologies, it is crucial to institute mechanisms for tracking and increasing the uptake of such ontologies. This is to reduce delays or opportunity costs incurred with the slow progression of ontologies in a digital technology landscape that is rapidly evolving. More crucially, a strategy on ontologies that aggressively prioritizes the growth of uptake will ensure a more rapid demonstration of the return on investment by CGIAR and its funding partners. Increasing the reach and depth of engagement by agriculture research and the science of data stakeholders on ontologies will continue to be relevant as it pertains to demonstrating CGIAR leadership in agricultural research and innovation. Efforts to sustain and enhance the CoP as a mechanism for such engagements are therefore recommended moving forward in One CGIAR.

8.3.1 References

Arnaud, Elizabeth, Marie-Angélique Laporte, Soonho Kim, Céline Aubert, Sabina Leonelli, Berta Miro, Laurel Cooper et al. "The ontologies community of practice: A CGIAR initiative for big data in agrifood systems." Patterns 1, no. 7 (2020): 100105.

Goldstein, A., Fink, L., & Ravid, G. (2021). A Framework for Evaluating Agricultural Ontologies. Sustainability, 13(11), 6387.

Annex 9: Overview of CGIAR Platform for Big Data in Agriculture

Annex 9.1 Platform Purpose and Objectives

According to the final July 2016 Proposal "*the Big Data Platform focuses on enhancing CGIAR and partner capacity to deliver big data management, analytics and ICT-focused solutions to CGIAR target geographies and communities through its ambitious partnerships with both upstream and downstream partners. In addition to developing new partnership models with big data leaders at the global level, the Platform seeks to promote CGIAR-wide collaboration across CRPs and Centers.*"

Box 1: Big Data Platform's tripartite objectives from the July 2016 proposal

1. ***Support and improve data generation, access, and management in CGIAR:*** *For CGIAR to embrace the power of big data analytics and be the leader in generating actionable data-driven insights for stakeholders, key requirements, enabling environment components, and critical gaps, which were identified during the scoping consultations. Through collaboration and co-creation with partners identified as the champions in bringing big data to agriculture, the Platform will provide support to CGIAR and partners to address the gaps, both organizational (i.e., Open Access/Open Data [OA/OD] compliance) and technical (e.g., providing useful datasets, tools, and services), and organize capacity building activities to sustain the efforts across the consortium.*

2. ***Collaborate and convene around big data and agricultural development:*** *CGIAR needs ambitious external partnerships to deliver the potential of big data to smallholder agriculture. Likewise, CGIAR is an attractive boundary partner for many private and public big data partners to engage in the context of agriculture in the developing world . This objective will set up system-level partnerships that Centers and CRPs can tap into and use to stimulate greater use of data analytics in CGIAR mission-critical research. Amongst other approaches, the Platform will provide opportunities and spaces for facilitated virtual collaboration and interaction among partners and stakeholders. A Big Data Convention will be organized to bring key actors to CGIAR and CGIAR to the key actors in a network that will be documented and nurtured. The Convention will focus on the generation of ideas and innovations. It will democratize big data opportunities, share progress amongst CRPs and Centers in promoting big data analytics. It will build capacity internally and externally on big data approaches in agriculture. Novel approaches to communications will increase exposure of CGIAR work on big data, and further engage a range of actors through novel approaches to partnerships.*

3. ***Lead by example and inspire how big data can deliver development outcomes:*** *Demonstrate the power of CGIAR big data analytics through "Inspire" projects that solve development challenges at the core of CGIAR SRF. These may include, but not be exclusive to, approaches that use big data analytics and ICTs to provide unprecedented multi-disciplinary data to researchers, deliver novel information to farmers, monitor the state of agriculture and food security in real time and inform critical national, regional and global policies and decisions. Venture capital (<$100k) will be provided to generate novel approaches, and some larger projects will be developed to deliver on the overall vision of the Platform: democratize big data to include smallholder farmers.*

Annex 9.2 Platform Initial Impact Pathways

In the 2016 Proposal, CGIAR defines *"big data as harmonized, interoperable, and contextually integrated datasets and publications from multiple disciplines relevant for CGIAR's research and development goals (CGIAR Consortium 2015a)".* The applications of insights gained from the study of such integrated datasets are greatly advancing knowledge and impact in several fields, including the -omics, biomedical, and ecological domains (Kitchin 2014; Madin et al. 2008; Jin et al. 2015). The technical definition of big data varies across disciplines but is generally characterized as having a high volume, velocity, variety, and variability (Laney 2001).

The Platform aims to increase the impact of agricultural development by embracing big data and ICT approaches to ***solve development problems faster, better, and at a greater scale.*** As outlined in the Strategy and Results Framework (SRF), this will initially be across CGIAR but will be extensible to agriculture at large. The Platform's theory of change focuses on increasing the capacity of CGIAR and partners to embrace big data and ICT approaches as shown in Figure 1.

Annex 9 - Figure 1: Initial impact pathways for the Platform for Big Data in Agriculture[9]

In 2016, the Platform was reported to be founded on the following **principles**:

1. *Process-oriented agile approach:* Establish processes and collaborative spaces needed to deliver goals

---

[9] Source: CIAT and IFPRI (2016). CGIAR Big data coordination platform. Proposal to the CGIAR Fund Council, 31 March 2016. International Center for Tropical Agriculture (CIAT), Cali, Colombia and International Food Policy Research Institute, Washington DC, United States of America.

in phases. Supported by agility, and iterative interactions with users to adapt emerging technologies to fulfill growing needs.

2. *Network approach through partnership:* Centered around how networks and Communities of Practice rather than single institutions leverage technology and new data resources as the basis for solving problems rather than single institutions. These Communities of Practice can leverage technology and new data resources to create a broader and deeper impact in programming.

3. *Iterative data needs assessment and technology landscape analysis:* To better understand Open Data initiatives and Big Data-based Information and Communications Technology for Development (ICT4D) initiatives, a regular data landscape analysis will be conducted for better alignment of the Platform with newly emerging agricultural research and development topics and big data technologies. This also involves the Platform working with its network partners to assess primary user needs through a multi-partner, multi-data stream, multi-country project in each region.

In the 2016 proposal, the authors acknowledge that the primary **assumption** of this theory of change is that data are a valued commodity for development that can be harnessed to deliver growth in agriculture in developing countries. It also assumes that CGIAR and its partners can identify business opportunities where rural institutions are weak, to deliver benefits to marginalized smallholder farmers. This assumption is central to the design of the Platform.

Annex 9.3 Platform Structure and Modules

The Platform operates a networked partnership model that is co-led by CIAT (with CIAT taking fiduciary and operational responsibility) and IFPRI; the Platform facilitates the convergence of CRPs, Centers, and external partners towards problem-solving. It comprises a nucleus secretariat whose primary objective is facilitating dialogue, collaboration, and communication across and between partners. The Platform Secretariat is driven by a knowledge-sharing approach in interacting and networking amongst partners.

The Platform operationalizes its tripartite objectives via three modules:

Annex 9 - Figure 2: Platform modules and their objectives [10]

Objective 1 Support and improve data generation, access, and management in CGIAR.	Objective 2 Collaborate and convene around big data and agricultural development.	Objective 3 Lead by example and inspire how big data can deliver development outcomes.
Module 1 ORGANIZE: organizes extant data and draws them together for unified and interlinked discoverability. It assesses their status and fitness for use identifies what and where gaps exist, and strengthens its analytical capacity for data-driven impact.	Module 2 CONVENE- Convenes the scientific resources across CGIAR with a X range of partners to generate new collaborative opportunities and bring big data to agriculture, and likewise, agriculture to big data.	Module 3 INSPIRE-Inspires work on big data by funding research by CGIAR scientists with partners to innovate new ways to handle big data relevant to agriculture for the benefit of poor smallholder farmers.

---

[10] Source: Adapted from CIAT and IFPRI (2016). Big Data Coordination Platform. Full Proposal (final version). July 2016.

9.3.1 The Platform's Linkages with CGIAR Centers, Platforms, and Research Programs

The 2016 proposal laid out the collaboration models between the Platform and other CGIAR Platforms and CRPs. The **ORGANIZE** module was conceptualized to actively work with Center and CRP researchers and CoPs (including the data, knowledge, IT, legal, and other relevant system-wide CoPs) in the inventory and management of datasets towards "open" and supporting tools for researchers' use. The goal is to support all Centers and CRPs to not only comply but overachieve open access and open data principles and CGIAR policy on these. The **CONVENE** module brings together big data practitioners from Centers and CRPs together with partners and other Platforms in spaces to encourage interaction and pursuit of common goals. The expectation was that the interactions will produce ideas that would encourage applications for funding under the **INSPIRE** module to develop them. The INSPIRE module creates opportunities for novel ideas to be realized through pilot projects with new partners to CGIAR under collaborative efforts embedded within CRP activities.

Besides, collaboration was envisioned with the Genetic Gains Platform (now Excellence in Breeding, also launched in 2017) through shared infrastructure for processing of genetic data, and the Genebank Platform (launched in TBC) to the extent to which environmental and socio-economic data can add value to passport information for germplasm collections. The Platform also envisaged collaborating with the CGIAR Gender Research Network, the Gender platform established in 2020.

The evaluation team conducted a desk review of the Platform's progress towards the integration of cross-cutting themes, namely: Gender, youth, capacity building, and climate change. The result of this exercise can be consulted in the Inception Report. As well, progress related to the issue of "Intellectual Assets" and "Open Access".

9.3.2 Platform Management and Governance

The Platform is co-led by CIAT and IFPRI. The CIAT leadership takes fiduciary responsibility, signing the performance contract with the system office, and consults IFPRI leadership as needed. Implementation is through a secretariat comprising a Big Data Coordinator (Platform coordinator), Platform Co-founders, Module One Leader, a Project coordinator, communication and engagement specialists, and modest administrative support. The Platform coordinator is responsible for intellectual leadership and representation, sign-off on deliverables, and has decision-making authority with respect to day-to-day operations of the Platform. Module 1 implementation is led by Dr. Medha Devare who was leading open access and open data work in the Consortium Office. The implementation of Module 2 and 3 is led by the Platform coordinator. An executive management team includes the two Module leaders and the two Platform co-founders, Dr. Andy Jarvis and Dr. Jawoo Koo. The executive team meets bi-monthly with Communities of Practice leaders (CoPs) leaders participating as observers.

The project has an eight-member steering committee reporting formally to the CIAT board, whose Chair and Director-General then report on the Platform as a whole to the System Office. The steering committee (SC) monitors the Platform's effectiveness and makes programmatic decisions. Its membership includes one permanent member each from CIAT, IFPRI, and the CGIAR System Office. Other members representing their constituents on a two-year rotating basis include one each from CRPs and Centers, and another three representing partners. The committee members assign a chairperson from among these three representatives of partners. The rotating membership is assigned to balance critical knowledge and experience on informatics, legal and intellectual assets, data management, data analytics. The committee was to meet in person at the Annual Convention and three to four times virtually. The workings of the SC were to be reviewed after 24 months. A five-member International Advisory Board (IAB) was to be set up, meeting face-to-face once per year, and virtually once per year. Its role was to explicitly examine how the Platform connects effectively with other global and regional efforts for continued relevance and novelty. The IAB membership was to represent related major initiatives such as Open Harvest and Global Open Data for Agriculture and Nutrition (GODAN) with expected two-year membership terms. Findings of the IAB were to be reported to the CIAT Board once per year.

In 2017, the steering committee and the management team including six CoP representatives, and the secretariat were instantiated. The three partner representatives to the steering committee were from IBM (private sector), the University of Florida (research), and the Bill and Melinda Gates Foundation (funder). With many management and governance-relevant decisions being made regularly, the steering

committee and the management team met monthly in 2018. The IAB was instantiated in 2019 and its members from Food and Agriculture Organization (FAO), GODAN, Google, Mars Inc, GIZ, and Ag Gateway met in person. In 2019, the steering committee was reported to have representation from all envisaged constituencies except for the CGIAR system office which was not represented. In that year, two external members rotated out and one extended their term. A new CRP representative was introduced in 2019. No changes were made in the Platform governance and management structures in 2020. The Steering committee held 11 meetings in 2020 and the executive management team held 12 meetings.

9.3.3 Platform Funding and Budget

According to the 2016 Platform's Proposal, Big Data Platform had a 6-year budget of US$30.2m primarily from Windows 1 & 2, representing an annual budget that ranges from US$3.9m to US$6.7m. In terms of the budget allocation per module, Module 1 received the largest budget share in 2017 (68% total) and 2018(58%) with the main cost driver being funding to Centers aimed at improving the effective management of CGIAR data and compliance with the Open Access, Open Data (M) Policy. Module 2's budget in 2017 was US$1.46 with a progressive growth by a standard 5% annually to maintain the fixed costs associated with creating an enabling environment. Similarly, Module 3's budget was projected to double by its fourth year from Year 1 (US$0.6m) to year 4 (US$1.31m). The budgeted cost for the Platform Secretariat was pegged at US$300k in the proposal and was covered under Module 2 CONVENE- with percentage allocations to cross-cutting themes such as Capacity Building (40%), gender, and youth-related activities (17%).

Annex 9 - Table 1: CGIAR Platform for Big Data in Agriculture - Funding and Budget (USD)

Module	2017	2018	2019	2020	2021	2022	Total
Module 1 ORGANIZE	4,336,320.51	3,172,574.31	2,261,673.74	1,159,962.09	1,125,489.82	1,192,411.46	13,248,431.93
Module 2 CONVENE	1,455,300	1,516,077	1,579,603.14	1,646,517.85	1,716,339.85	1,789,267.18	9,703,105.02
Module 3 INSPIRE	612,720	670,095.9	1,017,294.24	1,307,446.44	1,089,338.37	538,356.48	5,235,251.43
Management + Support Cost	300,000	315,000	330,750	347,288	364,652	382,885	2,040,575
Total	6,704,340.51	5,673,747.21	5,189,321.12	4,461,214.38	4,295,820.04	3,902,920.12	30,227,363.38

Source: Big Data in Agriculture re-submitted Proposal

Annex 10: Evaluation Team Background

Ibtissem Jouini, Team Leader

Ms. Jouini is a senior evaluator and researcher. She founded the EvalChange network in 2016: a group of independent consultants committed to making a lasting impact through their work giving special importance to the principles of gender equality, inclusiveness, and human rights. Over the last years, Ms. Jouini has contributed and led numerous independent evaluations where she designed rigorous and tailored methodologies applying several qualitative methods. Previous to that, Ms. Jouini worked for international development organizations (UNDP, GIZ, USAID) where she was involved in regional programs mainly related to the field of Governance. Ms. Jouini is a Tunisian national based in Spain.

Didier Leibovici, Subject Matter Expert

Didier Leibovici's expertise is in geospatial data analytics and after 15 years of research in leading UK universities (Oxford, Leeds, Nottingham, Sheffield), 5 years at IRD (France), 2 years at Sanofi-Recherche (France),4 years at INSERM (France) working within interdisciplinary and international contexts for European research programs with UK, France, LMIC (in Africa and South-Asia), he is setting up GeotRYcs, a geo-spatial-temporal data scientist consulting service. Didier has a PhD in Biostatistics and a Master's degree in computing-science; his scientific production in data analysis and geospatial science are on spatiotemporal data modelling and analysis within different contexts, such as epidemiology, agriculture, and agro-ecological monitoring, dynamics in population studies, location-based citizen crowdsourcing of environmental information within interdisciplinary projects. Didier's interests are in challenging the potential of interoperability developments to manage cross-domains scientific models involving geospatial data from heterogeneous sources.

Erik Bongcam-Rudloff, Subject Matter Expert

Erik Bongcam-Rudloff is a Professor of Bioinformatics at the Swedish University of Agricultural Sciences (SLU), Uppsala, Sweden. He received his doctorate in medical sciences from Uppsala University in 1994 and his Docentur in Bioinformatics in 2004. His main research deals with the development of bioinformatics solutions to solve biological research questions from the Life Sciences community. Bongcam-Rudloff has extensive experience in coordinating international research projects related to his group's research, including the COST Action BM1006 - "Next Generation Sequencing Data Analysis Network (SeqAhead), the EU FP7 "Broadening the Bioinformatics Infrastructure to unicellular, animal, and plant science" (ALLBIO) and the Horizon2020 "Bridging Biobanking and Biomedical Research across Europe and Africa", (B3Africa). He has also participated as WP leader on other EU projects as EMBRACE, Affinomics, and BBMRI, and acted as Grant holder for the COST Action CA15110 "Harmonising standardization strategies to increase efficiency and competitiveness of European life-science research" (CHARME). Professor Bongcam-Rudloff served as an executive board member in several international organizations relating to computational biology and bioinformatics as EMBnet, Goblet, and ISCB and worked in bioinformatics capacity building projects in Europe, Asia, Africa, and Latin America. Erik Bongcam-Rudloff is currently a member of the advisory board of the EU "Human Exposome Assessment Platform" (HEAP), the EU "Participatory Approaches to a New Ethical and Legal Framework for ICT" (PANELFIT) project, and "The Eastern Africa Network of Bioinformatics Training" (EANBiT).

Mathew Kurian, Subject Matter Expert

Mathew Kurian is Consortium Lead for the Belmont Forum project on cyber-enabled disaster resilience involving partners at Penn State University, Cranfield University, and the University of Sao Paulo besides UNHABITAT, Geneva, and Ministry of Water and Irrigation, Government of Tanzania. He previously led the establishment of the Capacity Development and Governance Unit at the United Nations University (UNU) in Dresden, Germany where he launched the Nexus Observatory (an online platform to support the monitoring of the SDGs) in collaboration with GIZ, Bonn. He has previously held staff positions at The World Bank and Consultative Group on International Agriculture Research (IWMI-CGIAR) and has served on the faculty of UNESCO-IHE, Delft, and University College London, UK. His most recent book *Boundary Science* (Elsevier 2021) inspired the launch of the climate panel- an online Platform that connects data with models and engages decision-makers at the level of local governments and communities in Sub-Saharan Africa: https://www.theclimatepanel.com.

John Kieti, Expert Information Communications Technology & Data Management

John is an exponent of social justice. He is passionate about digitalization and digital Platforms helping to solve social and economic problems. He has just over 20 years of experience in management information systems and building digital entrepreneurship ecosystems. He previously worked as a Data Manager, Analyst/Programmer, Head of Information Systems, Director of Programs, and Chief Operations Officer in various organizations. He has in the past designed and deployed information systems gathering vast data for aggregation and analysis at national levels. He led the teams organizing PIVOT East, Eastern Africa's premier conference and challenge for mobile technology startups between 2011 and 2014. He was part of the team conceptualizing the CTA's Pitch Agrihack challenge in 2013. John has served as a technology, innovation, entrepreneurship, and digital agriculture consultant for CTA, the World Bank Group, iHub, and the University of Nairobi, among other organizations. He holds an MBA and a BSc in Computer Science. He is a PhD candidate pursuing research on digital Platforms for agriculture. From this research, he recently published a peer-reviewed article on "*the sources of value creation in aggregator Platforms for digital services in agriculture*". John's vision is for technology, innovation, and entrepreneurship to unlock the economic potential of developing countries, starting with agriculture.

Stefania Sellitti, Monitoring, Evaluation and Learning Consultant

Stefania Sellitti is a development economist with a strong background in agriculture and rural development. She worked on several research projects with CIAT and CropTrust, focusing on the empowerment of workers in coffee estates in Latina America, on the knowledge about climate change in Nicaragua and Colombia, and the impact of CIATS's Genebank and bean collection. She is currently working as a teaching assistant at the NOVA School of Business and Economics in Lisbon. She has experience in Monitoring, Evaluation, and Impact Assessment, both within CGIAR, as an intern at the DG Agri of the European Commission, and as an external consultant for private companies, such as a Plan-Eval in Brazil and COATL in Portugal.

Annex 11: Conflict of Interest Statements

Original forms with signatures and additional detail are available upon request.

S/N	Conflict of Interest Statements	Ibtissem Jouini	Didier Leibovici	Mathew Kurian
1	Main employer and any other organization that provides you with remuneration (which may be named participants in the project/ program/ proposal you are being asked to review/evaluate	Eval Change Network	GeoTRYcs (cie)	The Climate Panel
2	Are you aware whether a relative, close friend, close colleague or someone with whom you have financial ties is receiving funding from or giving advice to a project/program/proposal you are being asked to review/evaluate?	Yes☐ Details: No☒	Yes☐ Details: No☒	Yes☐ Details: No☒
3	Does any project/program/proposal you are being asked to review/evaluate cite any of your own current research?	Yes☐ Details: No☒	Yes☐ Details: No☒	Yes☐ Details: No☒
4	Does any project/program/proposal you are being asked to review/evaluate name researchers with whom you have active collaborations, recently published joint papers or are in regular email correspondence?	Yes☐ Details: No☒	Yes☐ Details: No☒	Yes☐ Details: No☒
5	Does any project/program/proposal you are being asked to review/evaluate name any of your past PhD students are active participants?	Yes☐ Details: No☒	Yes☐ Details: No☒	Yes☐ Details: No☒
6	I declare that the information provided on this statement is true and complete	Dated: 20 July 2021	Dated: 14 July 2021	Dated 7 July 2021

S/N	Conflict of Interest Statements	Erik Bongcam Rudloff	John Kieti	Stefania Sellitti
1	Main employer and any other organization that provides you with remuneration (which may be named participants in the project/program/proposal you are being asked to review/evaluate	Swedish University of Agricultural Sciences, professors' position	International Finance Corporation/World Bank Group University of Nairobi	NA
2	Are you aware whether a relative, close friend, close colleague or someone with whom you have financial ties is receiving funding from or giving advice to a project/program/proposal you are being asked to review/evaluate?	Yes☐ Details: n/a No☒	Yes☐ Details: n/a No☒	Yes☐ Details: n/a No☒
3	Does any project/program/proposal you are being asked to review/evaluate cite any of your own current research?	Yes☐ Details: No☒	Yes☐ Details: No☒	Yes☐ Details: No☒
4	Does any project/program/proposal you are being asked to review/evaluate name researchers with whom you have active collaborations, recently published joint papers or are in regular email correspondence?	Yes☐ Details: No☒	Yes☐ Details: No☒	Yes☐ Details: No☒
5	Does any project/program/proposal you are being asked to review/evaluate name any of your past PhD students are active participants?	Yes☐ Details: No☒	Yes☐ Details: No☒	Yes☐ Details: No☒
6	I declare that the information provided on this statement is true and complete	Dated: 29 July 2021	Dated: 3 June 2021	Dated: 16 July 2021

Annex 12: Evaluation Terms of Reference (ToRs)

Evaluation of CGIAR Platform for Big Data in Agriculture
Draft, July 2, 2021.

12.1 Background

12.1.1 Rationale and Context of the Evaluation

In today's connected, data-rich world, big data presents tangible benefits and challenges revolutionizing the Agricultural Research for Development (AR4D) continuum as well as people's lives. The smart and effective use of data is key to unlocking and accelerating the achievement of the 2030 Agenda for Sustainable Development. Data innovations and digital tools bring critical capabilities for agile adaptation in food systems.

CGIAR's data and knowledge products should be, arguably, among its crown assets. To stay at the cutting-edge of the rapidly evolving digital world, the CGIAR invests in the curation and maintenance of these assets through a five-year (2017-2021) CGIAR Platform for Big Data in Agriculture (hereinafter, the Platform) approved by the System Council. The Platform is a coordinating mechanism to deliver a coherent data-driven and data-intensive strategy leveraging data capabilities and infrastructure. Its strategy focuses on collaboration among CGIAR Research Programs (CRPs) and Centers, leveraging external expertise to enable unrestricted discoverability of linked open datasets. "The ultimate goal of the Platform is to harness the capabilities of Big Data to accelerate and enhance the impact of international agricultural research. It will support CGIAR's mission by creating an enabling environment where data are expertly managed and used effectively to strengthen delivery on CGIAR SRF's System Level Outcome (SLO) targets."[11] An overview of the Platform is summarized in Annex 1.

The CGIAR Advisory Services Shared Secretariat (CAS Secretariat) supports and facilitates the CGIAR's independent advisory services, comprising the Independent Science for Development Council (ISDC), the Standing Panel on Impact Assessment (SPIA), and an independent Evaluation Function. CAS Secretariat's Evaluation Function supports the implementation of the CGIAR System multi-year evaluation plan to meet CGIAR System's needs for rigorous high-quality independent evaluations to inform decision making across the System. As part of its 2021 approved workplan and budget, the Evaluation Function is mandated to conduct a full-fledged external evaluation of CGIAR's Big Data in Agriculture Platform.

Earlier in the year, per its mandate and approved workplan, to meet the needs of System Council represented by Strategic Impact, Monitoring, and Evaluation Committee (SIMEC) On June 21, 2021, the Evaluation Function completed the Synthesis of Learning from a Decade of CGIAR Research Programs (CRPs). The high-level 2021 Synthesis pooled evidence from 43 CGIAR evaluations, reviews, syntheses, and assessments including the 2019 performance management standards pilot assessment for the Platform commissioned by CAS Secretariat and conducted by Dalberg Advisors on behalf of the CGIAR System. Another evaluative study related to the Big Data Platform was commissioned by CAS Secretariat's predecessor, CGIAR's Independent Evaluation Arrangement in 2018, a review of CGIAR's open access/open data policy and implementation support.

In addition to the aforementioned evaluative assessments, since its inception in 2017, the Platform has been the subject of several other reviews and studies, worthy of mention. In response to CGIAR System Management Board's request for a digital strategy that identified CGIAR's comparative advantage relating to big data,[12] a strategic research study was conducted on digital transformation in food, land, and water

---

[11] Big Data Coordination Platform: Full Proposal 2017-2020

[12] Chair's Summary, 13th CGIAR System Management Board ('SMB') Meeting. Approved May 3 2019

systems in a climate crisis[13] in support of the 2030 Research and Innovation Strategy. Related to this, a high-level assessment of digital strategy across CGIAR was conducted. Also, in 2021, a review of the Inspire Challenge, assessed the Platform's Inspire Challenge program (2017-2020) and its broader contributions to catalyze partnerships and digital agricultural innovations.

With the launch of a new research modality to advance the One CGIAR 2030 Research and Innovation Strategy, making the digital revolution central to the way of working is one of the seven new implementation approaches prioritized in the strategy (seventh way of working). The key elements of the 2030 strategy's seventh way of working include engagement with partners in developing cutting-edge, context-appropriate digital solutions, improved access to and use of data and digital innovations targeting small-scale farmers, pursuing new digital applications to accelerate learning and knowledge sharing among partners underpinned by principles of findability, accessibility, interoperability, and reusability (FAIR) for all CGIAR data. Thus, leveraging the unprecedented opportunity provided by today's digital revolution is at the front burner of One CGIAR to accelerate progress towards the achievement of the Sustainable Development Goals (SDGs).

Against this backdrop, the recent 2021 Synthesis of Learning from a Decade of CGIAR Research Programs brought to the fore thematic evidence gaps related to digital innovations revealing limitations on the evaluative evidence available on digital innovations. There has been hitherto no comprehensive independent process evaluation of the Platform in its entirety. Informed by the evaluative evidence needs identified in the synthesis, and in response to the request of CGIAR System Council, the Evaluation Function under CAS Secretariat seeks to commission an evaluation of the Platform. The evaluation would assess the Platform's effectiveness, design, and delivery and distil lessons and recommend actionable operational and strategic approaches for the future One CGIAR.

12.2 The Evaluation

12.2.1 Evaluation Purpose and Scope

This evaluation will serve the dual purposes of accountability and learning. It will be both summative and formative in nature and will assess the design, scope, implementation status and the capacity to achieve the Platform objectives. It will collate and analyze lessons learned, challenges faced, and best practices obtained during implementation as a guide for future planning. It will assess the performance of the project against planned results and the preliminary indications of potential sustainability of results. The evaluation will provide essential evaluative evidence for decision-making by the CGIAR System Council, Big Data Platform management, and its partners.

The evaluation will cover all the activities of the Platform from its initiation in 2017 through mid-2021 considering the need for timely evidence with the drivers, the transition to One CGIAR, and the COVID-19 pandemic. The evaluation will integrate cross-cutting themes of Gender, Diversity, and Inclusion (GDI), youth, climate change and capacity development as well open data and intellectual assets.

The main objectives of the evaluation of the Big Data in Agriculture Platform are to:
- Assess the relevance of the Platform design, theory of change (ToC) and the Platform's role in positioning CGIAR as a learning organization, its ability to cultivate new digital alliances, pursue data innovation in support of its mission;
- Identify the supporting factors and constraints behind achievement of the Platform and each of its modules and the validity of the ToC assumptions in light of the results achieved, including its response to COVID-19;
- Assess the Platform governance, management, and implementation processes;
- Provide recommendations relevant to the future development and implementation aligned with One CGIAR Way of Working 7 – Making the Digital Revolution Central to Our Way of Working and One CGIAR initiatives related to digital technologies, to include inter alia, "Harnessing Digital Technologies for Timely Decision-Making across Food, Land, and Water System" (Systems

---

[13] https://cgspace.CGIAR.org/handle/10568/113555

Transformation Action Area)[14] and, if applicable, other system-wide recommendations.

The formative and summative component will address both effectiveness of the Platform implementation strategy and the results. This includes the implementation modality, partnership arrangements, institutional strengthening, beneficiary participation, sustainability of the Platform. The evaluation will include review of the project design and assumptions made at the beginning of the project development process. It will assess the extent to which the project results have been achieved, partnerships established, capacities built, and cross cutting issues integrated. It will also assess whether the project implementation strategy has been optimum and recommend areas for improvement and learning.

12.2.2 Key Stakeholders

The key stakeholders of this evaluation with their particular interests are presented in Table 1.

Table 1: The Platform Evaluation key stakeholders

Type of Stakeholder	INTEREST	
	Accountability	Learning
CGIAR System Council & Funders	✓	
CGIAR System Board	✓	✓
The One CGIAR Portfolio Performance Management Team	✓	✓
MD, Institutional Strategy and Systems, Global Director, Digital Services.		✓
Initiative Design Teams (IDTs)		✓
Project Coordination, Monitoring and Performance Management Unit		✓
Big Data Platform Management		✓
Big Data Platform Steering Committee, International Advisory Board		✓
Big Data Focal Points in all CGIAR Centers		✓
CGIAR partners involved in generating and use of CGIAR knowledge products.	✓	✓
All the Big Data Communities of Practice (CoPs)		✓
End Users of Big Data Platform	✓	✓

To the extent feasible given the resource and time allocated to the evaluation, key stakeholders will be widely consulted and engaged throughout the evaluation process through relevant channels and using the appropriate engagement tools.

12.2.3 Evaluation Criteria and Questions

The evaluation will examine project implementation against the hereunder criteria by addressing the following (broad but not exhaustive) questions.

Table 2: Evaluation criteria and questions

Criteria	Key Evaluation Questions
Relevance	1. To what extent are the Platform's objectives relevant to the needs of its stakeholders and target groups?
Efficiency	2. Have resources (funds, human resources, time, expertise etc.) been allocated strategically and timely to achieve Platform outcomes?

---

[14] 13th CGIAR System Council Meeting, SC13-02 Pre-read: CGIAR 2022-2024 Investment Plan

Effectiveness	3. To what extent did the Platform achieve its intended and unintended outcomes? 4. How effective has the Platform been in building digital capabilities and partnerships supporting CGIAR research? 5. To what extent have Platform outputs and outcomes contributed to changes in the organization and its stakeholders as relates to their use of data and digital technologies?
Sustainability	6. To what extent are the Platform products and communities positioned to be effective in the future, seen from the perspectives of scientists and of the end users of digital agriculture products and innovations? 7. To what extent would the Platform outputs outlive the existence of the Platform in its current form?

The evaluation criteria and key questions are further detailed with sub-questions in Annex 2 and will be elaborated in consultation with relevant stakeholders at the inception phase towards the development of the evaluation matrix.

12.3 Approach and Methodology

The evaluation will be primarily desk-based and use a mixed-methods design. Methodological rigor in the evaluation design will be adhered to. The inception report will include a detailed evaluation matrix and a description of the proposed methodological approach. The inception report and other key deliverables will be peer-reviewed by evaluation and Subject Matter Experts (SMEs). CAS Secretariat's processes will guide, and quality assure the evaluation process.

Quantitative data will be collected via online survey instruments, data will be disaggregated (wherever possible) by age and gender. Quantitative analyses would also be performed to the extent possible on available quantitative indicators and metadata from the relevant data sets (including GARDIAN). Qualitative techniques would combine an extensive review of extant documentation on the Platform, content analysis of the evaluative evidence from the 2021 synthesis exercise, open and semi-structured interviews with internal and external stakeholders, and focus-group discussions. It is also recommended that case studies be presented for each Platform Module to understand the user perspectives and experiences. The use of data science techniques such as machine learning algorithms incorporating Artificial Intelligence and data mining where relevant to expand the data collection and analysis of data sets is also encouraged. Data sources will be triangulated to ensure transparency and independence of judgment and to minimize bias.

Stakeholder groups to be interviewed would be elaborated during the inception phase and include key Platform partners, the Platform's focal points at all Centers, data managers and information specialists at all Centers, and users of the Platform. The evaluation team shall determine whether to seek additional information and opinions from representatives of any of the external thought partners to the Platform. To increase credibility, the particular value will be placed on the triangulation of the data and solid argumentation of the conclusions drawn and recommendations made. The evaluation would be conducted in close collaboration with the Platform of Big Data in Agriculture.

CAS Secretariat will guide the evaluation process and ensure that the evaluation team uses appropriate tools and technology to enhance data access and, that data analysis is robust. CAS Secretariat will also ensure the effective communication of evaluation results with evaluation stakeholders.

12.3.1 Expected Limitations to the Evaluation

The evaluation's remit and its resources limit the extent to which it can collect primary information from the Platform's vast network of partners. Therefore, the evaluation will use reports and other documents, a representative sample of interviews, surveys, and limited ground-truthing to gather evidence on the evaluation questions and validate its findings.

12.4 Evaluation Timeline and Management

12.4.1 Evaluation Phases and Timing

The evaluation is scheduled to take place between July and December 2021, for transmission to the System Council in December 2021, after vetting with SIMEC. An indicative time frame for the evaluation and expected deliverables is provided in Table 3 (see Annex A3 for a detailed schedule), to be elaborated in the inception report.

Table 3: Indicative Evaluation Timeline, with Milestones and Selected Deliverables, 2021

Phase	July (Weeks)				August (Weeks)				September (Weeks)				October (Weeks)				November (Weeks)				December (Weeks)		
	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3
Inception		☐ Document Review ☐ Briefings ☐ Inception report																					
Inquiry						☐ Desk review ☐ Interviews ☐ Surveys ☐ Module case studies ☐ Data analysis																	
Reporting														☐ Preliminary findings ☐ Report development ☐ Draft report ☐ Report QA review ☐ Validation workshop ☐ Final report									
Management Response																				☐ Management Response			
Dissemination																					☐ Targeted webinars ☐ Knowledge products		

Preparatory phase: During the preparatory phase CAS Secretariat, in consultation with relevant stakeholders, will review key documents and define the scope and issues surrounding the evaluation, and carry out the following tasks:

- Develop the Terms of Reference (ToR);
- Consult the ToR with stakeholder groups (SIMEC, Global Science Group Director- Systems Transformation, evaluands);
- Select and contract the evaluation team leader and in consultation with her/him, the evaluation team.

Inception phase: The inception phase is the responsibility of the team leader. The inception report will focus on the following elements:

- Preliminary project theory model(s); refinement of the evaluation questions, elaboration of evaluation methodology including quantitative and qualitative approaches through an evaluation framework ("evaluation matrix");
- A stakeholder analysis identifying key stakeholders, networks and channels of communication. This information should be gathered from the Platform documents and discussion with the Platform team;
- A preliminary list of strategic issues of importance for emphasis during the inquiry phase;
- An indicative evaluation report outline and division of roles and responsibilities between the evaluation team leader and the external evaluation team; people to be interviewed and possible surveys to be conducted and a debriefing and reporting timetable.
-

These elements will be drawn together in an inception report to be agreed between the team and the CAS Secretariat, which will subsequently represent the contractual basis for the team's work and deliverables of the evaluation. As a requirement to finalize the inception report, a consultation will be arranged between CAS Secretariat, the evaluation team and peer-reviewers to interrogate the evaluation approach and methodology and enhance the evaluation matrix.

Inquiry phase: The evaluation team will collect the evidence according to the plan detailed in the inception report, complete its analysis, and prepare a preliminary list of findings and conclusions.

Reporting phase: In the reporting phase, the evaluation team will prepare a presentation of preliminary findings, to debrief the CAS Secretariat and Platform Management and to seek validation, factual corrections, and feedback.

The team would develop the draft evaluation report for the CAS Secretariat's comments and factual corrections. Under the CAS Secretariat's guidance, the report would be reviewed by a team of external peer-reviewers. With the feedback from relevant stakeholders, the evaluation team would finalize the evaluation report taking into account comments according to the team's judgement.

Management Response: During this phase, CAS Secretariat will liaise with the Project Coordination, Monitoring and Performance Unit through its relevant Tasks Units- Project Coordination Unit (PCU) and, Monitoring and Performance Management Unit (MPMU)[15] to coordinate the preparation of the management response with the Platform management. The management response will be published on the CAS Secretariat website.

Dissemination: The evaluation report, the executive summary and the evaluation brief and other knowledge products along with the management response, will be published on the CAS Secretariat's website. In line with the dissemination and knowledge management strategy to be developed at the inception phase, tailored presentations will be made to targeted stakeholders and learning events organized with internal and external stakeholders.

12.4.2 Evaluation Management and Responsibilities

The Evaluation Lead, Svetlana Negroustoueva, of the CAS Secretariat manages the evaluation process, under the overall direction of the CAS Secretariat Director, Allison Grove Smith. Questions or comments regarding this Terms of Reference should be directed to CGIAR Advisory Services Evaluation (CGIAR) CAS-Evaluation@CGIAR.org copying s.negroustoueva@CGIAR.org.

The evaluation will be conducted by an independent team of experts (the evaluation team). The team leader has final responsibility for the evaluation report and all findings and recommendations, subject to adherence to CGIAR Evaluation Standards. The primary responsibilities of the team leader will be:
- Setting out the methodology and approach in the inception report;
- Guiding and managing the evaluation team during the inception and evaluation phases;
- Overseeing the preparation of, and quality-assuring, data collection outputs by other members of

---

[15] If these entities do not yet exist, the interaction will be with the existing System Office Programs Unit.

the team;
- Consolidating team members' inputs to the evaluation products (inception report and the evaluation report);
- Where necessary, representing the evaluation team in meetings with stakeholders;
- Delivering the inception report, draft and final evaluation reports.

The evaluation team is responsible for submitting the deliverables highlighted in 3.3 and detailed in Annex 3 to CAS Secretariat, these include but are not limited to:
- An inception report;
- Three Module case studies;
- A brief presentation of preliminary findings, for the debrief with the Platform management and CAS Secretariat;
- Draft report of the Platform evaluation, N.B the CAS Secretariat will provide a template for the draft and final reports;
- A final evaluation report following the report template with a maximum of 25 pages, and written in plain English in line with CAS Secretariat's style guide;
- A two to three-page executive summary, and a set of annexes with additional information apart from the main body of the report;
- PowerPoint presentations covering the main points of the evaluation, including purpose, methods, findings, conclusions, recommendations, and additional notes relevant to the evaluation. The CAS Secretariat will provide the relevant templates.

The CAS Secretariat will be responsible for planning, initially designing, initiating, and managing the evaluation. It will also be responsible for the quality control of the evaluation process and outputs, and dissemination of the results. The Evaluation Function Lead supported by a Senior Evaluation Officer will provide support to the team throughout the evaluation.

The Platform's management, steering committee and focal persons will respond to the Evaluation team's needs for information throughout the evaluation: documentation and data, access to partners and staff for engagement with the evaluators, and information on partners and stakeholders. These actors will be also be responsible for giving factual feedback on the draft evaluation report.

To ensure the independence of the evaluation, the CAS Secretariat's staff will not participate in meetings where their presence could bias the responses of external stakeholders. Adequate consultations with evaluation stakeholders will be ensured by the evaluation team and the CAS Secretariat throughout the process, with debriefings on key findings held at various stages of the evaluation. The Evaluation Function Lead will ensure transparent and open communication with stakeholders during each of the key evaluation phases.

12.4.3 Evaluation Team

The evaluation team will comprise six (6) team members drawn from the vetted Subject Matter Expert (SME) and Evaluator roster maintained by CAS: (1) Evaluation team leader- Evaluator; (3) Senior SMEs in digital innovation and (1) Mid-level SME in Information, Communications Technology (ICT) and data management. They will be supported by (1) mid-level evaluation analyst (consultant) for data collection, analysis, and Knowledge Management (KM). The team would conduct the evaluation in conformity with international and CGIAR evaluation standards.

The team members will have a strong cumulative experience in conducting complex, global strategic evaluations with suitable background relating to big data in agriculture and working knowledge of CGIAR and its research. The multi-disciplinary evaluation team would combine competencies and expertise in the following areas:
- Data generation, analysis, management and governance
- Power relationships and politics around information (social science)
- ICT governance, risk management, and international and national regulatory frameworks
- Partnerships, in particular, with the private sector
- Research or development agencies on issues, programs and policies related to agriculture and natural resources and digital technologies

- A strong understanding of Gender, Diversity and Inclusion (GDI) issues
- High-level expertise in reviewing and processing a large number of documents, conducting one-on-one and group interviews using appropriate technology in data collection and analysis and communication of evaluation results.

Each evaluation team member will be carefully vetted for any present or future conflicts of interest (COI).

The **team leader** will have a minimum of 15 years' experience in evaluation, with extensive experience in regional or global strategic-level evaluations with working knowledge of the use of digital technologies and data science. The team leader must have experience in leading teams, excellent analytical, synthesis and communication skills (written and verbal) and demonstrated skills in mixed qualitative and quantitative data collection and analysis techniques. The team leader will manage the team of a subject-matter experts and two (2) team members as above with the following qualifications:
- At least a master's degree in Development Economics/Planning, Digital Systems, Computer Science / Engineering, Data Science, Economic, Public Administration, and Management and in any other related university degree
- Extensive expertise, knowledge, and experience in the field of evaluation of development programs
- At least 10 years of experience in working with international organizations and donors
- Experience of program formulation, monitoring and evaluation
- Experience in designing, developing, implementing, and evaluating technology-assisted projects
- Skills on high-quality analysis, reporting in English and time management for timely deliverables submission
- Proven experience coordinating program activities with governmental, nongovernmental, and private-sector partners.

Peer-reviewers with relevant subject -matter expertise will be called up at necessary stages of evaluation design and implementation for enhanced rigor and validity.

12.4.4 Deliverables and Dissemination of Findings

The inception report: the inception report, which builds on the terms of reference for the evaluation, outlines the evaluation team's proposed approach to the main phase of the evaluation as follows: (i) elaborating the scope and focus of the evaluation; (ii) developing the methodological tools for gathering evidence; (iii) providing a detailed evaluation matrix; (iv) clarifying the analytical frameworks to be used by the evaluation; and (v) providing a detailed work plan for the evaluation.

The evaluation report- the main output of this evaluation - will describe findings and conclusions, based on the evidence collected in the framework of the evaluation questions defined in the inception report, and recommendations logically following the conclusions. The recommendations will be evidence-based, relevant, focused, clearly formulated, and actionable. They will be prioritized and addressed to the different stakeholders responsible for their implementation. The main findings and recommendations will be summarized in an executive summary. The main report should be concise (no longer than 25 pages – excluding the Executive Summary and Annexes) and written in plain English. The evaluation team will be expected to produce a three-page brief of key findings and lessons, following a template provided by the CAS Secretariat.

Review of the draft evaluation report- The evaluation team will submit a zero-draft report to the CAS Secretariat as part of the quality assurance process. Upon the acceptance of a draft of adequate quality, CAS Secretariat will share this first draft report with a team of peer reviewers. The first draft will be shared with the Platform team for their review and comments- for any errors of fact and highlight the significance of any such errors in any conclusions. Subsequently, a discussion version of the report will be presented to SIMEC for feedback. With the feedback of SIMEC integrated, the discussion version of the report will be presented to System Council for their input which will guide the final evaluation report.

The final report shall be submitted by email to the Evaluation Function Lead in electronic editable form (MS Word) aligned with CAS Secretariat's style guide. The final report will follow a standardized structure and template to be provided by CAS Secretariat. CAS Secretariat will finalize the report by having it

professionally edited. The final evaluation report will be published on the CAS Secretariat's website.

Presentations: The team leader and evaluation team where necessary will present the evaluation results to key CGIAR stakeholders via various communication channels to targeted audiences.

12.4.5 Contract and Payment Schedule

The CAS Secretariat is hosted by CGIAR System Organization through an arrangement with the Alliance of Bioversity International and the International Center for Tropical Agriculture, at its offices in Rome, Italy. Contracting will be carried out by our hosting entities and under their name on behalf of CAS Secretariat. The members of the evaluation team are expected to abide by the Conflict of Interest policy of the CAS Secretariat and must maintain independence in fact and appearance from the Platform under review throughout the duration of the assignment. Each evaluation team member must sign and return statements indicating their understanding and compliance with the policies of the CAS Secretariat and its host institutions. All contracting fees and conditions will be administered in line with the approved policy for consultants. Confidentiality provisions are covered in these contracts. All collected data must be shared for the confidential records kept within the CAS Secretariat; informants should be duly notified to adhere to ethical evaluation principles.

1. Annex A1 to ToR: Background: CGIAR Big Data in Agriculture Platform

A1.1 Purpose and Objectives

According to the final July 2016 Proposal, the Platform focuses on enhancing CGIAR and partner capacity to deliver big data management, analytics and ICT-focused solutions to CGIAR target geographies and communities through its ambitious partnerships with both upstream and downstream partners. In addition to developing new partnership models with big data leaders at the global level, the Platform seeks to promote CGIAR-wide collaboration across CRPs and Centers. Big Data Platform's tripartite objectives culled from the July 2016 proposal are:

1. **Support and improve data generation, access, and management in CGIAR:** For CGIAR to embrace the power of big data analytics and be the leader in generating actionable data-driven insights for stakeholders, key requirements, enabling environment components, and critical gaps, which were identified during the scoping consultations. Through collaboration and co-creation with partners identified as the champions in bringing big data to agriculture, the Platform will provide support to CGIAR and partners to address the gaps, both organizational [i.e., Open Access/ Open Data (OA/OD) compliance] and technical (e.g., providing useful datasets, tools, and services), and organize capacity building activities to sustain the efforts across the consortium.

2. **Collaborate and convene around big data and agricultural development:** CGIAR needs ambitious external partnerships to deliver the potential of big data to smallholder agriculture. Likewise, CGIAR is an attractive boundary partner for many private and public big data partners to engage in the context of agriculture in the developing world. This objective will set up system-level partnerships that Centers and CRPs can tap into and use to stimulate greater use of data analytics in CGIAR mission-critical research. Amongst other approaches, the Platform will provide opportunities and spaces for facilitated virtual collaboration and interaction among partners and stakeholders. A Big Data Convention will be organized to bring key actors to CGIAR and CGIAR to the key actors in a network that will be documented and nurtured. The Convention will focus on the generation of ideas and innovations. It will democratize big data opportunities, share progress amongst CRPs and Centers in promoting big data analytics. It will build capacity internally and externally on big data approaches in agriculture. Novel approaches to communications will increase the exposure of CGIAR work on big data, and further engage a range of actors through novel approaches to partnerships.

3. **Lead by example and inspire how big data can deliver development outcomes:** Demonstrate the power of CGIAR big data analytics through 'Inspire' projects that solve development challenges at the core of CGIAR SRF (Strategy and Results Framework). These may include, but not be exclusive to, approaches that use big data analytics and ICTs to provide unprecedented multi-

disciplinary data to researchers, deliver novel information to farmers, monitor the state of agriculture and food security in real-time, and inform critical national, regional, and global policies and decisions. Venture capital (<$100k) will be provided to generate novel approaches, and some larger projects will be developed to deliver on the overall vision of the Platform: democratize big data to include smallholder farmers.

A1.2 CGIAR Platform for Big Data in Agriculture- Structure, and Modules

The Platform operates a networked partnership model that is co-led by CIAT (with CIAT taking fiduciary and operational responsibility) and IFPRI; the Platform facilitates the convergence of CRPs, Centers, and external partners towards problem-solving. It comprises a nucleus secretariat whose primary objective is facilitating dialogue, collaboration, and communication across and between partners. The Platform Secretariat is driven by a knowledge-sharing approach in interacting and networking amongst partners. The Platform operationalizes its tripartite objectives via three modules:

1. **ORGANIZE:** The ORGANIZE module aims to fully open access CGIAR's intellectual assets, address technical and organizational challenges, and provide CGIAR researchers with an enabling environment to strengthen data analytical capacity and develop practical, big-data-driven use cases in a coordinated way. As a minimum, this module seeks to align CGIAR Centers on open access and open data and ensure compliance with CGIAR's Open Access Policy, ratified in late 2013 by all 15 Centers.

   The Module works with Center and CRP researchers and Communities of Practice (CoPs) (including the data, knowledge, Information Technology (IT), legal, and other relevant system-wide CoPs) in the inventory and management of datasets towards "open" and supporting tools for researchers' use. The goal is to support all Centers and CRPs to not only comply but overachieve with regard to open access and open data principles and CGIAR policy on these.

2. **CONVENE:** Through the CONVENE module, the Platform aims to implement an annual CGIAR Big Data and Information and Communications Technology (ICT) Convention where representatives from Centers and CRPs will share information, develop joint initiatives, and collaborate with key external actors in the big data space. It established Communities of Practice (CoPs) across Centers for defining data standards and interoperability protocols, dovetailed with the OA/OD initiatives. It operates virtual collaboration spaces and sponsors quarterly webinars and capacity-building workshops and connects with existing initiatives, other GIAR CoPs, Conventions on ICT4D and Big Data.

   The CONVENE Module assembles big data practitioners from Centers and CRPs together with partners and other Platforms in spaces to encourage interaction to produce ideas that qualify to be funded for further development under the Inspire Module.

3. **INSPIRE:** The INSPIRE module seeks to generate high profile, collaborative applications of big data in agriculture through small- and medium- sized "Inspire" initiatives, embedded within larger CRP-related initiatives, that bring to bear new partnerships on core CGIAR challenges. These consist of one to three-year initiatives, a case in point is the INSPIRE Challenge initiative, to generate new analytical approaches, scientific outputs, and high-profile examples of how big data analytics can deliver agricultural development in CGIAR target geographies.

   The Inspire Module seeks to create opportunities for novel ideas to be realized through pilot projects with new partners to CGIAR under collaborative efforts embedded within CRP activities.

Figure A 1: Big Data PLATFORM Objectives, Minimum Success Factors (MSF) and Modules

Objective 1: Support and improve data generation, access, and management in CGIAR.	**MSF: Compliance with the open access and open data policy of CGIAR, ensuring donors and investors in CGIAR can be confident that data are being managed and shared effectively across all CGIAR operations.** • Module 1-ORGANIZE: organizes extant data and draws them together for unified and interlinked discoverability. It assesses their status and fitness for use identifies what and where gaps exist, and strengthens its analytical capacity for data-driven impact.
Objective 2: Collaborate and convene around big data and agricultural development.	**MSF: New partnership models developed with upstream and downstream partners, from public and private sectors, to deepen and widen CGIAR's capacity on big data analytics and use** • Module 2: CONVENE- Convenes the scientific resources across CGIAR with a X range of partners to generate new collaborative opportunities and bring big data to agriculture, and likewise, agriculture to big data.
Objective 3: Lead by example and inspire how big data can deliver development outcomes.	**MSF: Established models for innovation and other approaches will be used to achieve this, built on the foundation of collaboration from Objective II.** • Module 3: INSPIRE-Inspires work on big data by funding research by CGIAR scientists with partners to innovate new ways to handle big data relevant to agriculture for the benefit of poor smallholder farmers.

Adapted from the Big Data Coordination Platform- Full Proposal (final version) July 2016.

A1.3 Management and Governance

Leadership of the Platform is provided through a secretariat, which consists of a Big Data coordinator, Platform co-founders, a project coordinator, Module One leader, communications specialists, and administrative support. The Platform coordinator doubles as the leader for both Modules Two and Three respectively. In addition, the Platform also has focal points in all 15 Centers through which it liaises with Centers as needed.

As per governance, the Platform relies on its steering committee led by a chair and comprising five other representatives; permanent members both from CIAT and IFPRI respectively. Other members are partners, Centers and Research Programs respectively. Another permanent member is a representative of the CGIAR System Office. The three other steering committee members are on a 2-year rotating basis and representatives of the CRPs, Center and Partner respectively.

A1.4 Platform Principles

1. Process-oriented agile approach: Establish processes and collaborative spaces needed to deliver goals in phases. Supported by agility, and iterative interactions with users to adapt emerging technologies to fulfill growing needs.
2. Network approach through partnership: Centered around how networks and communities of practice rather than single institutions leverage technology and new data resources as the basis for solving problems rather than single institutions. These communities of practice can leverage technology and new data resources to create broader and deeper impact in programming.
3. Iterative data needs assessment and technology landscape analysis: To better understand Open Data initiatives and Big Data based Information and Communications Technology for Development (ICT4D) initiatives, a regular data landscape analysis will be conducted for better alignment of the Platform with newly emerging agricultural research and development topics and big data technologies. This also involves the Platform working with its network partners to assess primary user needs through a multi-partner, multi-datastream, multi-country project in each region.

A1.5 Partnership Ecosystem

Big Data Platform relies on a network of diverse partners comprising All CGIAR Centers and 12 CRPs as well as 70 external thought partners[16] such as international organizations, academia, research institutes, private companies including global players on big data analytics. Big Data's partnership ecosystem spans upstream knowledge generators, through downstream knowledge users.

A1.6 Funding and Budget

According to the Platform's Proposal, Big Data Platform had a six-year budget of US$30.2m primarily from Windows 1 & 2, representing an annual budget which ranges from US$3.9m to US$6.7m. In terms of the budget allocation per module, Module One received the largest budget share in 2017 (68% total) and 2018(58%) with the main cost driver being funding to Centers aimed improving the effective management of CGIAR data and compliance with the Open Access, Open Data (M) Policy. Module Two's budget in 2017 was US$1.46 with a progressive growth by a standard 5% annually to maintain the fixed costs associated with creating an enabling environment. Similarly, Module Three's budget was projected to double by its fourth year from year one (US$0.6m) to year four (US$1.31m). Budgeted cost for the Platform Secretariat was pegged at US$300k in the proposal and was covered under Module Two CONVENE- with percentage allocations to cross-cutting themes such as capacity building (40%), gender and youth-related activities (17%).

Table A 1: CGIAR Big Data in Agriculture Platform- Funding and Budget (USD)

Module	2017	2018	2019	2020	2021	2022	Total
Module 1 – ORGANIZE	4,336,320.51	3,172, 574.31	2,261,673.74	1,159, 962.09	1,125, 489.82	1,192,411.46	13,248,431.93
Module 2 – CONVENE	1,455,300	1,516,077	1,579, 603.14	v1,646, 517.85	1,716,339.85	1,789,267.18	9,703,105.02
Module 3 – Inspire	612, 720	670, 095.9	1,017, 294.24	1,307, 446.44	1,089,338.37	538,356.48	5,235,251.43
Management + Support Cost	300,000	315,000	330,750	347,288	364,652	382,885	2,040,575
Total	6,704,340.51	5,673,747.21	5,189, 321.12	4,461,214.38	4,295,820.04	3,902,920.12	30,227,363.38

Source: Big Data in Agriculture resubmitted Proposal, 2016

2. Annex A2 to ToR: Evaluation Criteria, Key questions and Sub-questions

Table A 2: Evaluation Criteria, Key questions and Sub-questions

Criteria	Key Evaluation Questions	Sub-Questions
Relevance and Strategic Fit	a) To what extent are the Platform's objectives relevant to the needs of its stakeholders and target groups?	a) Were the Platform mechanisms and approaches aligned with Center and key partners' priorities, capacities, and expectations? b) To what extent have cross-cutting themes GDI, youth, climate change, Open Data and Intellectual Asset issues been considered in project design and implementation? c) How appropriate are the Platform's outputs in the light of its operating environment and to what extent are these properly used, resilient and adaptable to local and evolving constraints including COVID-19 Pandemic?

---

[16] Source: https://bigdata.CGIAR.org/about-the-platform/

Criteria	Key Evaluation Questions	Sub-Questions
Efficiency	b) Have resources (funds, human resources, time, expertise etc.) been allocated strategically and timely to achieve Platform outcomes?	d) How adequate has been the high-level, technical, institutional, and administrative support from the Platform's partners? e) How efficient was the implementation: use of resources, timeliness?
Effectiveness	c) To what extent did the Platform achieve its intended and unintended outcomes?	a) To what extent has the Platform enabled Centers to comply with CGIAR's Open Access and Data Management (OADM) Policy? b) How has the Platform contributed to change in organizational data policy and its implementation? *(potential data sources: OA/OD Policy in effect, Information and Data Managers' CoP, CGIAR Core metadata schema, System organization stakeholders).* c) To what extent did the Platform achieve the planned outputs and outcomes noted in the proposal? *(Source: Big data reports and associated evidence).* d) What have been the main contributing or limiting factors for the results achieved? e) How has the Monitoring, Evaluation and Learning (MEL) system facilitated or inhibited achievement? f) To what extent has the Platform's governance and institutional mechanisms helped to create ownership among key stakeholders?

	d) How effective has the Platform been in building digital capabilities and partnerships supporting CGIAR research?	g) Partnerships
		  i. To what extent has the Platform accelerated Centers' progress towards making their data Findable, Accessible, Interoperable and Reusable (FAIR)?
		  ii. To what extent has the Platform enabled CGIAR to engage with the wider Big Data community? *(data sources: Community of Practice participation, reported partnerships, Convention participation, Inspire Challenge evaluation).*
		  iii. Technical communities of practice: how has the Platform enabled (or not) the expansion of external engagement in technical communities of practice? *(data sources, CoPs and communications team).*
		  iv. How has the Platform enabled external partners to have their data searchable alongside CGIAR data via GARDIAN? *(data source: https://gardian.bigdata.CGIAR.org/).*
		  v. Incentives to publish process data: To what extent has the Platform promoted incentives to publish scientific meta data among its partners?
		  vi. Partnership ecosystem & capabilities: What capabilities and constraints do the internal and external partnership typologies in different regions contribute to the Platform's outputs and outcomes?
		  vii. To what extent has the Platform strengthened capacities to promote demand for and use of FAIR data?
		  viii. To what extent did the Platform conventions open the way to build/deepen novel partnerships that leverage CGIAR expertise to shape the future of digital agriculture?
		h) Identification of digital research and innovation challenges
		  i. To what extent has the Platform's technical communities of practice (CoPs) engaged with partners to stay abreast of digital innovations related to CGIAR research domains, and sought to develop new digital research capabilities in CGIAR? *(data source: CoP sessions at Conventions, annual work plans and reports by CoP).*
		  ii. How effective has the Platform been in identifying challenge areas for targeting digital innovation linked to the CGIAR research portfolio? *(data source: Inspire Challenge evaluation).*

Criteria	Key Evaluation Questions	Sub-Questions
		i) New/Improved Methods. To what extent has the Platform catalyzed the development of new digital methods for research or delivery of research? *(Potential data sources: CGIAR publications focused on digital innovations before and after 2017, in grey literature and peer-reviewed publications; evaluation of the Inspire Challenge.*
		j) Data collection and analysis questions and potential sources i. What Platform generated outputs would improve CGIAR data collection and analysis (Data s*ources: GARDIAN, AgroFIMS, CG Labs, Community of Practice products*) and who is using or co-developing them? ii. What outputs from the Platform target enabling CGIAR to manage potential legal or reputational risk regarding data privacy and security? *(data sources: responsible data guidelines, 'managing agricultural privacy' report for system engineers, GARDIAN PII (personally identifiable information) detector.* iii. To which extent are the Platform's quality control mechanisms improving (or not) the shared data quality?
		k) New knowledge and innovations: To what extent has the Platform contributed to digital innovations for research and delivery of research in the CGIAR? *(Potential data sources: "digital" outputs by CRPs before and after 2017, Inspire Challenge evaluation, CoP outputs and reports).*
	e) To what extent have Platform outputs and outcomes contributed to changes in the organization and its stakeholders as relates to their use of data and digital technologies?	a) To what extent have Platform investments resulted in digitally-enabled research innovation in CGIAR? *(Potential data sources: uptake of Platform tools by internal and external stakeholders: GARDIAN, Expert Finder, AgfroFims by Alliance Bioversity-CIAT, IITA/Excellence in Agronomy, FAO; uptake and use of CoP outputs and reports by CGIAR and CoP members).* b) To what extent have Platform investments equipped CGIAR with new cross-cutting capabilities to use its data to address commonly posed research questions regarding agriculture, climate, and food systems? *(data sources: analytic products on climate risk and fertilization generated using Big Data analytic environment).* c) To what extent do Platform investments make more data available for agricultural analytics, and facilitate the use of these data? *(data sources: growth of data and publications discoverable over time via GARDIAN, growth in GARDIAN partners over time).* d) To what extent has the Platform helped change CGIAR culture and practice regarding responsible, ethical data collection, management, and analysis? *(Potential data sources: interviews with data managers, performance management team at System Office, datasets and publications discoverable over time via GARDIAN, user traffic of GARDIAN, System Management Board minutes from 13th meeting).*
Sustainability	f) To what extent are the Platform products and communities positioned to be effective in the future, seen from the perspectives of scientists and of the end users of digital agriculture products and innovations?	g) To what extent do the internal and external stakeholders engage with the Platform (e.g. the Convention, CoPs, innovation grant process, data processes and tools) value it? (Data sources: evidence from the Inspire Challenge interviews. Thematic analysis of semi-structured interviews from the digital strategy research). h) What Platform-generated insights, products, and communities have contributed to the One CGIAR reform/reorganization? i) What Platform-generated artefacts, policies, products, communities and approaches are being integrated into One CGIAR? j) What are the lessons learned for future design of similar initiatives? k) To what extent does the Platform position CGIAR with a leadership voice in digital agriculture in the eyes of its international partners?

Criteria	Key Evaluation Questions	Sub-Questions
	l) To what extent would the Platform outputs outlive the existence of the Platform in its current form?	a) To what extent do the internal and external stakeholders own and seek continuity of its programmatic elements? If so, which and why? If not, why not? *(Data sources: evidence from the Inspire Challenge interviews. Thematic analysis from semi-structured interviews from the digital strategy research).* b) What are the lessons learned to facilitate the translation of Platform's outputs and outcomes to CGIAR's way of working 7-Making the Digital Revolution Central? c) How would capacities built in partners ensure sustainability of results? d) What are the key factors in management and governance structured to ensure success and sustainability of the Platform?

3.   Annex A3 to ToR: Indicative Evaluation Schedule

Table A 3: Indicative Evaluation Schedule

Evaluation Phase	Tasks	Outputs	Responsible	Dates
Preparatory	Draft evaluation ToR /ToR Revisions	Final evaluation ToR	CAS Secretariat	9 July
	Selection of consultants from the vetted roster	Evaluation team contracts.		
Inception	Onboarding and briefing of the external evaluation team Development of the Inception report with the evaluation matrix	Draft inception report with evaluation matrix	Evaluation team	23 July
	Consultation with peer reviewers on the methodology and approach.	Final inception report and evaluation matrix	Evaluation	30 July
Inquiry	Desk review		Evaluation Team	1-8 October
	Survey	Survey instrument		
	Interviews	Interview notes		
	Module case studies	Case study notes		
Reporting	Analysis and report development	Detailed report outline for feedback to CAS		
	Validation workshop	PPT	Evaluation Team and CAS Secretariat	11-15 October
	Submission of draft Platform evaluation report	Draft Platform evaluation report	Evaluation Team	25 October
	Report review by CAS, peer-reviewers and key stakeholders as needed.	Compiled feedback by peer-reviewers and key stakeholder groups.		5 November
	Drafting of the final report integrating the feedback	Draft final report		5 November – December 6
	Presentation of Draft final Report to SIMEC for feedback	Draft final report, PPT	CAS Secretariat	

Evaluation Phase	Tasks	Outputs	Responsible	Dates
	Revision of the draft final report integrating SIMEC's feedback	Revised draft final Report	Evaluation Team	
	Presentation of draft final Report to System Council	Draft final report. PPT	CAS Secretariat/Evaluation Team	
		Final report	Evaluation Team	10 December
Management Response	Management Response coordinated by Project Coordination, Monitoring, and Performance Unit.	Management response	CAS Secretariat liaising with Project Coordination, Monitoring, and Performance Unit.	December 2021
Dissemination	Development of knowledge products and knowledge management in line with the Dissemination strategy for the Evaluation.	Evaluation briefs and knowledge products.	CAS Secretariat/Evaluation team where necessary.	December Onwards