CGIAR PLATFORM FOR
BIG DATA IN AGRICULTURE
PLAN OF WORK AND BUDGET
2021

CGIAR, Platform for Big Data in Agriculture

The CGIAR Platform for Big Data in Agriculture is a cross-cutting program of the global CGIAR consortium of non-profit research institutes looking into virtually every aspect of food security spanning: genomics, breeding, agroecology, climate science, and the socioeconomic drivers and context of food systems change. The Platform tends to data standards and data sharing, digital innovation strategy and technology transfer, and research into the intersection of digital technologies and agricultural development in emerging regions.

CGIAR is a global research partnership for a food secure future dedicated to reducing poverty, enhancing food and nutrition security, and improving natural resources.

https://bigdata.cgiar.org/

TABLE OF CONTENTS

1

ADJUSTMENTS TO THEORIES OF CHANGE

The Platform for Big Data has over the course of four years developed a cross-cutting digital innovation function for CGIAR. The Platform set out to demonstrate that data standards, sharing, and analysis; partnerships and technical communities of practice; and applied innovation processes, strategy, and management can mutually reinforce each other to accelerate inclusive, impactful digitization in our sector and keep CGIAR abreast and evolving with digital disruption in how CGIAR research is done and the contexts in which it is delivered. The Platform Theory of Change (TOC) has both organizational and sector-level milestones, focused on helping CGIAR and partners increase their capacity to embrace big data, and information and communications technologies (ICTs) through:

1 enhanced collaboration across Centers, Programs and partners in using state-of-the-art data standards, analytics and ICTs;

2 enabling unrestricted discoverability of inter-linked datasets to understand and tackle multi-faceted food security challenges in new data and evidence-driven ways;

3 leveraging CGIAR expertise in the broader big data and ICT sphere, to establish CGIAR as an innovative digital thought leader; and

4 development of initiatives using proven big data innovations to drive agricultural growth in developing countries.

This Theory of Change has not changed, and the Platform can point to important progress under each of these goals. Implementation of the Platform, however, has highlighted the importance of change management within the System to achieve the full potential of digital tools and technologies within our organization and our sector. In 2021, the Platform aims to help launch a new One CGIAR digital strategy building on or helping to embed key learnings from the Platform, and will promote ways to tend to the change management needed to mainstream it across the new, more unified CGIAR organization.

Understanding the role of CGIAR in facilitating sector-wide digital transformation has deepened over the course of the past four years. As a result, the Platform now places more emphasis on strategic alliances to help our partners apply global data standards for making development data Findable, Accessible, Interoperable and Reusable (FAIR) and to demonstrate its value via analytic pipelines in support of commonly-encountered use-cases (e.g. climate adaptation options by location, analysis of return on investment of fertilizer use). The Platform will also synthesize evaluation and learnings from the Inspire Challenge--CGIAR's signature digital innovation process--regarding how innovation strategy and management can be applied to increase the uptake and impact of digital innovations in agricultural research for development.

2

PLANS AND EXPECTED PROGRESS TOWARDS OUTCOMES

The Platform for Big Data catalyzes data-driven innovation across CGIAR through data standards and sharing; analytics; partnerships and technical communities of practice. These approaches are reinforced by synergistic innovation processes to accelerate inclusive, impactful digitization in agriculture, helping CGIAR evolve in step with rapid changes in the research landscape, and establishing CGIAR technical leadership in the global digital agriculture sector. 2021 will be a year of transition to a unified CGIAR, and BIG DATA will contribute to the digital dimensions of this vision and seek to launch a new digital strategy in support of it.

Building momentum for digital innovation throughout the One CGIAR transition will be a central theme for the Platform in 2021. Applied, large-scale modeling leveraging CGIAR expertise with the computing power and data of industry partners will demonstrate the power of unified research infrastructure. Seven technical communities of practice--now totalling 5,000 members--will leverage domain expertise from across CGIAR and beyond to promote specific proofs of concept demonstrating the power of collaborative networks for One CGIAR. The Platform will seek to crowd in new investors in the Inspire Challenge, CGIAR's signature digital innovation process for fostering data-driven innovations, and integrate this process with CGIAR-wide research innovation strategy and management. The whole of the Platform team will be engaged in supporting the transition to a digital One CGIAR.

MODULE ONE: ORGANIZE

Under Module One ("Organize") the Platform will continue to build a global knowledge base and data ecosystem supporting agricultural research for development through the Global Agricultural Research Data and Innovation Network (GARDIAN), a pan-CGIAR and partnership-driven data ecosystem. The GARDIAN team will continue to enable search across all CGIAR repositories and connect more repositories from strategic partners, substantially building on GARDIAN's 170,000 publications and almost 28,000 datasets. Module One will work with the CGIAR Excellence in Agronomy (EiA) initiative, a key partner and client which will also contribute to the data standards, management, and analytic capabilities on offer for CGIAR and other researchers. The data ecosystem will be enriched by a data management toolkit with guidelines, services and tools, allowing users to more easily align with best practices towards making data assets open, FAIR-standards compliant (i.e. data that are findable, accessible, interoperable, and reusable), and ethically managed.

The FAIR data discoverable via GARDIAN will help researchers derive value from large amounts of data, aided by workshops to build CGIAR data science capacity. These efforts will be augmented by data processing and machine learning algorithms available through GARDIAN, and by an analytic environment, Collaborative GARDIAN Labs (CGLabs). In 2021 the team will prototype connectors between GARDIAN data and two commonly-used crop simulation models (WOFOST and DSSAT) leveraging data-to-model translators developed by the University of Florida as part of the Agricultural Model Intercomparison and Improvement Project (AgMIP).

CGIAR's open and FAIR knowledge base will be further enriched through GARDIAN's FAIR data workflow for helping researchers easily apply data management standards to high-value legacy data. The Platform will facilitate efforts for data to be "born FAIR" through a user-friendly Agronomy Ontology, the Socioeconomic Ontology (SEOnt) and an improved Agronomy Field Information Management System (AgroFIMS). AgroFIMS allows agronomists to generate FAIR standards-compliant field books for data collection via any one of three freely available applications.

Module One will also continue to improve the CGIAR Expert Finder that allows users to discover CGIAR's research by location, activity, subject area, funder and more.

MODULE TWO: CONVENE

The "Convene" module of the Platform will engage across CGIAR and our global networks of collaborators, beneficiaries, funders, and impact partners as we build collective actions to shape the future of digital agriculture, propelling the "One CGIAR" reform.

In 2021 the Platform will hold virtual events exploring the partnerships, data, and innovations that will support and accelerate CGIAR contributions to the Impact Areas of the CGIAR 2030 strategy. The Platform will build on the momentum of the CGIAR Convention on Big Data in Agriculture, which is becoming a reference in the digital agriculture space for innovation strategy, technical agenda-setting, collective action, and showcasing the depth and breadth of CGIAR's data science capabilities and its research and partner networks. The Convention has steadily increased fundraising, and 2021 will be a critical year for determining if this event can continue.

The Platform is a bridge to state-of-the-art innovation in industry, and exciting new partnerships will unfold in 2021. BIG DATA, Google X, Digital Green, Hewlett Packard Enterprise, and Yara will pilot solutions for responsible management and use of farm and farmer data for large-scale analytics and digital financial services. The Platform will deepen its research collaborations with the French Digital Sciences Institute (INRIA) and Cambridge University into emerging topics at the intersection of Artificial Intelligence and agricultural research for development to ensure that AI can be used to mitigate, rather than exacerbate systemic risks in global food security.

BIG DATA technical Communities of Practice (CoP) will engage with ongoing initiatives to offer their substantial intellectual and social capital as needed, and highlight community-specific data, methods, tools, and analytic products with an eye to mainstreaming these across CGIAR through proofs of concept and partner engagement:

The Data-Driven Agronomy CoP will facilitate cross-regional and transdisciplinary learning on digitally enabled extension. In collaboration with the Policies, Institutions, and Markets Program, the CoP will define and develop a roadmap for the digital "extension agent of the future". The effort will bridge elements of agricultural analytics, common digital agriculture tools and approaches, human-computer interaction, and socioeconomics, building on and contributing to digitally-enabled agricultural extension across CGIAR. The community plans to feature the effort in a series of webinars, discussion briefs, and research papers.

The Crop Modeling CoP will continue improving global coordination of crop modeling efforts. Through strategic collaborations with Purdue University, University of Florida and CGIAR centers, the CoP will close crop modeling knowledge and data gaps and make more data available for crop modeling purposes. To do this, the CoP will develop tools to classify environments, test remote-sensed data in crop simulations, and promote guidelines and recommendations for minimum datasets to effectively leverage crop models. The CoP will also coordinate big data analysis through the International Wheat Improvement Network (IWIN) with CIMMYT, and work towards increasing digital data collection capacity with the Alliance of Bioversity-CIAT (Alliance). Several webinars and training workshops are planned for knowledge sharing, capacity and partnership building, and to collectively address current challenges.

The Ontologies CoP enhances CGIAR ontologies and vocabularies as valid reference ontologies that are becoming more widely adopted for research data annotation (e.g., Crop Ontology, Small Fisheries and Aquaculture Ontology). With the active participation of members from the public and private sector, the CoP will continue increasing the robustness of its products and connecting the results to standard open public semantic frameworks for agri-food data (including AGROVOC and the Food Ontology, among others). Through linked data tagging, this effort supports the cross-domain research needed to drive impacts under the CGIAR Research Strategy and enables complex data queries through GARDIAN and other data discovery platforms. The CoP will produce recommendations regarding the governance of the ontologies developed by the working groups and align activities with the need of the One CGIAR Strategy.

The Socioeconomic Data CoP will implement a Socioeconomic Ontology with an Ontology Independent Metadata Schema to enhance the interoperability of messy socioeconomic data (together with Ontologies CoP), develop partnerships to validate digital trust and transparency, and continue to support the informal group of CGIAR Internal Review Boards on the ethics of human subjects data collection under One CGIAR.

The Geospatial Data CoP will build the community's technical capacity to develop crop analytics that contributes to in-season crop predictions (e.g., crop area planted, phenology, yield prediction) and build external alliances to accelerate their use. The community will assess and propose actions to address cross-cutting geospatial analysis capabilities needed to implement the CGIAR research strategy, further develop a pan-CGIAR tool for guiding investments in climate adaptation for small agricultural producers, and collaborate with the Excellence in Breeding Platform on defining key operating procedures and data annotation and management practices for drone data.

The **Information and Data Management CoP** will continue tackling issues relating to enhancing the CG Core metadata standard, CGIAR's data and publications repositories, and leveraging best practices towards Open, FAIR, and secure data assets. CoP working groups will collaborate to consistently implement a revised One CGIAR Policy tackling Open and FAIR data assets, and help address CGIAR-wide Data Management Maturity Assessment recommendations. These efforts will be supported through monthly meetings, by enhancing capacity to help researchers annotate and upload data assets to repositories, and via workshops to test and refine the Platform's data management tools and services.

MODULE THREE: INSPIRE

Module Three ("Inspire") is the innovation component of the Platform, with a portfolio of twenty-one digital innovation projects linking CGIAR researchers with external partners through the Inspire Challenge. In 2020, the Platform awarded COVID-19 'Rapid Response Grants' to established Inspire Challenge projects that proposed an additional, localized impact in terms of COVID response, recovery, and long-term resilience. The results of these projects will be known in early 2021. The 2020 Inspire Challenge selection process included additional selection criteria related to response, recovery, and resilience. In 2021 the final cohort of seven final new start-up projects, winners from the 2020 Challenge, will be implemented, and three scale-up grant awardee will enter their final year. The team will focus on showcasing existing projects to potential funders and strategic allies and on synthesis and learning from its portfolio and the five years of the Challenge. This synthesis will provide important learning for future digital innovation under One CGIAR.

Without additional funding, no new innovation grants will be awarded in 2021. In previous years an array of partners (seed companies, agribusiness, multilateral lending banks, and development agencies) have expressed interest in potentially funding the Challenge. In 2021 the Platform will assess ways to enable this vital new digital innovation process to continue promoting the breadth and depth of CGIAR research, linking it to new partnerships and avenues to impact at scale under One CGIAR.

Program Management Unit: Collaboration with the GENDER Platform

BIG DATA is working to enable gender data to be leveraged to its full potential to improve our understanding of relationships between gender, agriculture, and rapidly digitizing economies and societies where CGIAR works. Working with the Generating Evidence and New Directions for Equitable Results (GENDER) Platform, we will advance the visibility of sex-disaggregated and gender-sensitive research by building researchers' capacity for metadata tagging and searching. Gender researchers' individual and organizational risks will be reduced through support for responsible and ethical management of sensitive data. Further, the GENDER Platform resource hub will source publications and data from GARDIAN.

In 2020 the Platform included more gender-sensitive selection criteria in the Inspire Challenge, agricultural research and development projects. Eleven total projects (four Rapid Response, seven Inspire Challenge) will be reviewed for insights related to gender and big data in agricultural innovation in 2021 at their one year mark of maturation. We aim to capture and share relevant learnings with the wider agricultural innovation sector.

The Platforms will promote human-centered design (HCD) methodologies for the development of digital tools and services in CGIAR. In 2021, the Platform aims to synthesize the HCD research into an educational output and pilot it with real human-computer interface issues.

3

FINANCIAL PLAN FOR THE COMING YEAR, INCLUDING USE OF W1/2

Module 1 ORGANIZE: Almost half of the total W1/W2 funds allocated to the Organize Module in 2021 will be disbursed to partners across CGIAR, and collaborators outside the System. Just over 15% is earmarked for enhancements to the GARDIAN data ecosystem.

Module 2 CONVENE: Module Two will continue to be implemented between PMU and the Communities of Practice. The CoPs will drive pilot/proofs of concept that demonstrate their research innovations in practice and seek to embed them in the One CGIAR structure, and the PMU will continue to build alliances that leverage the array of BIG DATA Platform offerings, particularly the CGLabs analytic environment and the GARDIAN ecosystem.

Module 3 INSPIRE: The Inspire Challenge, CGIAR's signature digital innovation process, has begun to attract W3 funding (over US$500k to date) and has a growing portfolio of projects with an array of impactful stories. This presents an opportunity to build it into a new fundraising and partnership development mechanism in support of the evolving portfolio under One CGIAR. The Platform (similar to CRPs) was curtailed by one year as the One CGIAR process takes root, which means that the Challenge will need to shut down and make no new innovation grants unless it can be integrated with pan-CGIAR innovation efforts. All Inspire Challenge funds will be fully executed by the end of 2021.

Additional explanations for table 3 (optional):

TABLES

TABLE 2A. Planned Milestones

MODULE	MAPPED TO SUB-IDO	2022 FP OUTCOMES	MILESTONES	INDICATE OF THE FOLLOWING	MEANS OF VERIFICATION	CGIAR CROSS-CUTTING MARKERS FOR THE MILESTONE			
						FOR GENDER	FOR YOUTH	FOR CAPDEV	FOR CC
M1	• {primary} CC Enhanced institutional capacity of partner research organizations • CC Increased capacity for innovations in partner research organizations • CC Improved forecasting of impacts of climate change and targeted technology development • CC Increased capacity for innovation in partner development organizations and in poor and vulnerable communities	M1 Outcome:1.1. A demand-driven analytics environment is available.	2021 - 1.1.6. 2021: At least two high-value data products created through CoP engagement and by leveraging GARDIAN, with analytics made available through Collaborative GARDIAN (CG) Labs.	Reworded/ rephrased from proposal	Data cleaning and processing scripts, and/or augmented datasets of relevance to CGIAR's core mission openly available via GARDIAN data ecosystem.	1	N/A	2	1
		M1 Outcome: 1.1. A demand-driven analytics environment is available.	2021 - 1.1.7. 2021: At least one cloud-based machine learning application developed in conjunction with BIG DATA CoPs and scientists from and beyond CGIAR.	Reworded/ rephrased from proposal	Cloud-based machine learning application/s of relevance to CGIAR scientists openly available via GARDIAN data ecosystem.	1	N/A	2	1

MODULE	MAPPED TO SUB-IDO	2022 FP OUTCOMES	MILESTONES	INDICATE OF THE FOLLOWING	MEANS OF VERIFICATION	CGIAR CROSS-CUTTING MARKERS FOR THE MILESTONE			
						FOR GENDER	FOR YOUTH	FOR CAPDEV	FOR CC
		M1 Outcome: 1.1. A demand-driven analytics environment is available.	2021 - 1.1.8. 2021: At least two new services/scripts that can be used to clean and process GARDIAN data.	Reworded/ rephrased from proposal	Services/scripts to clean, process, analyze, and/or more easily derive insight from available data assets openly available via the GARDIAN data ecosystem.	1	N/A	2	1
		M1 Outcome: 1.1. A demand-driven analytics environment is available.	2021 - 1.1.9. 2021: At least one data-to-model pipeline for the generation of model-ready data, tested and available for wider use via CG Labs.	Reworded/ rephrased from proposal	Data-to-model translators available via CG Labs to enable decision support by easing the ability to leverage crop simulation models.	N/A	N/A	2	N/A
		M1 Outcome: 1.1. A demand-driven analytics environment is available.	2021 - 1.1.10. 2021: Data science approaches for innovation in agricultural research and wider use of CG Labs enabled.	New/ changed	Enhanced data science capacity and creation of high-value data products.	1	1	2	1
	• {primary} CC Enhanced institutional capacity of partner research organizations • CC Increased capacity for innovations in partner research organizations • CC Increase capacity of beneficiaries to adopt research outputs	M1 Outcome: 1.2. CGIAR resources are discoverable and reused.	2021 - 1.2.9. 2021: GARDIAN includes data assets from at least two new partner repositories beyond CGIAR.	New/ changed	At least 200 new datasets discoverable via GARDIAN	1	1	2	1

MODULE	MAPPED TO SUB-IDO	2022 FP OUTCOMES	MILESTONES	INDICATE OF THE FOLLOWING	MEANS OF VERIFICATION	CGIAR CROSS-CUTTING MARKERS FOR THE MILESTONE			
						FOR GENDER	FOR YOUTH	FOR CAPDEV	FOR CC
	• CC Increased capacity for innovation in partner development organizations and in poor and vulnerable communities • CC Improved capacity of women and young people to participate in decision-making	M1 Outcome: 1.2. CGIAR resources are discoverable and reused.	2021 - 1.2.10: 2021: At least 50 datasets scoring at least ▯ for all FAIR indicators annotated using GARDIAN's FAIR workflow and PII checker.	New/ changed	At least 50 datasets with high FAIR scores discoverable via GARDIAN.	1	1	2	1
		M1 Outcome: 1.2. CGIAR resources are discoverable and reused	2021 - 1.2.11: 2021: Early prototype of a smart data query and aggregation tool developed to mine GARDIAN's data pool and enable easy aggregation of relevant datasets.	New/ changed	Alpha version of a tool that allows easier dataset aggregation available via GARDIAN.	1	1	2	1
		M1 Outcome: 1.2. CGIAR resources are discoverable and reused.	2021 - 1.2.12: 2021: Early prototype of pan-CGIAR "Expert Finder" developed in consultation with early-adopter Centers.	New/ changed	Alpha version of a tool that allows users to discover CGIAR's research by location, activity, Center etc. made openly available.	1	1	2	1

MODULE	MAPPED TO SUB-IDO	2022 FP OUTCOMES	MILESTONES	INDICATE OF THE FOLLOWING	MEANS OF VERIFICATION	CGIAR CROSS-CUTTING MARKERS FOR THE MILESTONE			
						FOR GENDER	FOR YOUTH	FOR CAPDEV	FOR CC
		M1 Outcome: 1.2. CGIAR resources are discoverable and reused.	2021 - 1.2.13.: 2021: GARDIAN interface and data exploration enhancements to enable easier leveraging of data assets and data management tools and services.	New/ changed	New GARDIAN user interface, providing enhanced data exploration and download, and access to data management tools and services.	1	1	2	1
	• CC Increased capacity for innovations in partner research organizations • CC Increased capacity for innovation in partner development organizations and in poor and vulnerable communities • {primary} CC Increase capacity of beneficiaries to adopt research outputs • CC Improved forecasting of impacts of climate change and targeted technology development	M1 Outcome: 1.3. Standards and semantics are utilized to enable FAIR (Findable, Accessible, Interoperable and Reusable) agricultural data.	2021 - 1.3.11: 2021: User-friendly workflows and tools, training, and documentation for easy refining and implementation of the CG Core Metadata Schema v.2.0 and ontologies across Center publications and data repositories and GARDIAN.	Reworded/ rephrased from proposal	1. At least three trainings on best practices in data management, including data annotation with CG Core metadata elements and ontology concepts. 2. User-driven enhancements to data annotation workflows. 3. Data prioritization framework available via GARDIAN to help users identify high-value legacy datasets for FAIRification.	1	N/A	2	N/A

MODULE	MAPPED TO SUB-IDO	2022 FP OUTCOMES	MILESTONES	INDICATE OF THE FOLLOWING	MEANS OF VERIFICATION	CGIAR CROSS-CUTTING MARKERS FOR THE MILESTONE			
						FOR GENDER	FOR YOUTH	FOR CAPDEV	FOR CC
		M1 Outcome: 1.3. Standards and semantics are utilized to enable FAIR (Findable, Accessible, Interoperable and Reusable) agricultural data.	2021 - 1.3.12: 2021: Further development of AgroFIMS to pilot: Easy generation of field books and digital data collection from "non-traditional" multi-locational agronomic survey and demo trials; use of AgroFIMS with the Open Data Kit (ODK) platform, in addition to the KDSmart app currently enabled.	New/ changed	1. Beta version of AgroFIMS tested for data collection from agronomic surveys via KDSmart and other apps (e.g. ODK). 2. Further use-driven enhancements to the Agronomy Ontology (e.g. for data dictionary-type applications) openly available for wide use.	N/A	N/A	2	1
		M1 Outcome: 1.3. Standards and semantics are utilized to enable FAIR (Findable, Accessible, Interoperable and Reusable) agricultural data.	2021 - 1.3.13. 2021: Agronomy Ontology optimized for AgroFIMS trial and survey functionalities, and easier to interact with for novice users.	New/ changed	Enhanced ontology available via GitHub and indexed by Ontology Lookup Service	N/A	N/A	2	1

MODULE	MAPPED TO SUB-IDO	2022 FP OUTCOMES	MILESTONES	INDICATE OF THE FOLLOWING	MEANS OF VERIFICATION	CGIAR CROSS-CUTTING MARKERS FOR THE MILESTONE			
						FOR GENDER	FOR YOUTH	FOR CAPDEV	FOR CC
	• {primary} CC Enhanced institutional capacity of partner research organizations • CC Increased capacity for innovations in partner research organizations • CC Increased capacity for innovation in partner development organizations and in poor and vulnerable communities • CC Enhanced individual capacity in partner research organizations through training and exchange	M1 Outcome: 1.4. Enhance capacity, catalyze culture change to further CGIAR OA/OD compliance and public goods mandate.	2021 - 1.4.7: 2021: Course materials and webinars for researchers and data managers on best practices for data management and maximizing FAIRness of CGIAR resources, with a focus on gender aspects.	New/ changed	1. At least two trainings on best practices in data management towards open and FAIR outcomes, with a focus on enabling better gender disaggregated data. 2. Course on FAIR Data Management ("Best Practices for Open, FAIR, and Ethical Data") offered at least once.	2	N/A	2	N/A
		M1 Outcome: 1.4. Enhance capacity, catalyze culture change to further CGIAR OA/OD compliance and public goods mandate.	2021 - 1.4.8: 2021: At least two workshops and/or trainings for data and information managers and researchers on ways to render datasets FAIR, including through the use of standards-compliant field books for data collection.	New/ changed	1. Documentation and videos to enable collection of FAIR agronomic data, and to improve FAIRness of legacy data. 2. At least two workshops/trainings for researchers and data managers on the use of standards-compliant field books to generate FAIR data.	N/A	N/A	2	N/A

MODULE	MAPPED TO SUB-IDO	2022 FP OUTCOMES	MILESTONES	INDICATE OF THE FOLLOWING	MEANS OF VERIFICATION	CGIAR CROSS-CUTTING MARKERS FOR THE MILESTONE			
						FOR GENDER	FOR YOUTH	FOR CAPDEV	FOR CC
M2	• CC Enhanced institutional capacity of partner research organizations • CC Enhanced individual capacity in partner research organizations through training and exchange • {primary} CC Increased capacity for innovations in partner research organizations	M2 Outcome: 2.1. CGIAR is more broadly engaged in the big data community.	2021 - 2.1.8: 2021: CoPs around geospatial data, socioeconomic data, ontologies, data-driven agronomy, information and data management, livestock data, and crop modeling establish CoP networks across CGIAR and produce outputs addressing key constraints in data and analytics.	Reworded/ rephrased from proposal	2020 Flagship products from CoPs released and applied in 2021. 2021 Products from CoPs (shareable, harmonized data across disciplines, position papers, small pilots).	1	1	2	N/A
		M2 Outcome: 2.1. CGIAR is more broadly engaged in the big data community.	2021 - 2.1.9: 2021: New CoP on Information and Data Management incorporated into BIG DATA governance.	Reworded/ rephrased from proposal	Terms of reference for Information and Data Management CoP approved and adopted by BIG DATA Platform governance bodies. Updated CGIAR standards for disaggregation and discovery of gender and youth data.	1	1	2	N/A

MODULE	MAPPED TO SUB-IDO	2022 FP OUTCOMES	MILESTONES	INDICATE OF THE FOLLOWING	MEANS OF VERIFICATION	CGIAR CROSS-CUTTING MARKERS FOR THE MILESTONE			
						FOR GENDER	FOR YOUTH	FOR CAPDEV	FOR CC
		M2 Outcome: 2.1. CGIAR is more broadly engaged in the BIG DATA community.	2021 - 2.1.10: 2021: Hold high-level Convention on Big Data in Agriculture, with wide participation of CGIAR and non-CGIAR actors. Establishment of collaborative agreements.	Reworded/ rephrased from proposal	2021 Virtual convention events, convention reports, and communications materials.	1	1	1	1
		M2 Outcome: 2.1. CGIAR is more broadly engaged in the BIG DATA community.	2021 - 2.1.11: 2021: Develop a pre-competitive, pro-commercial, multi-stakeholder alliance for food security research, data sharing, and technology matchmaking and transfer.	New/ changed	Alliance Memorandum of Understanding.	N/A	N/A	1	N/A
	• CC Enhanced institutional capacity of partner research organizations • {primary} CC Enhanced individual capacity in partner research organizations through training and exchange • CC Increased capacity for innovations in partner research organizations	M2 Outcome: 2.2. CGIAR increases its capacity to work on priority topics more quickly, more effectively and at greater scale.	2021 - 2.2.4: 2021: New, high-frequency computational methods applied to sustainability analysis leveraging CGIAR and partner data.	New/ changed	New analytic products of high-frequency computational methods applied to sustainability analysis, leveraging CGIAR and partner data.	0	0	2	0

MODULE	MAPPED TO SUB-IDO	2022 FP OUTCOMES	MILESTONES	INDICATE OF THE FOLLOWING	MEANS OF VERIFICATION	CGIAR CROSS-CUTTING MARKERS FOR THE MILESTONE			
						FOR GENDER	FOR YOUTH	FOR CAPDEV	FOR CC
	• {primary} CC Enhanced institutional capacity of partner research organizations • CC Increased capacity for innovations in partner research organizations	M2 Outcome: 2.3. CGIAR develops as a learning organization.	2021 - 2.3.13: 2021: Implement an aligned, pan-CGIAR action plan for information technology tools, processes, and infrastructure.	New/ changed	Launch CGIAR Digital Strategy in Q2.	1	1	2	1
		M2 Outcome: 2.3. CGIAR develops as a learning organization.	2021 - 2.3.14: 2021: Establish or leverage shared big data services to further pan-CGIAR collective action on priority research themes (especially cloud services enhancing pan-CGIAR research efforts including on climate adaptation and ecosystem services).	Reworded/ rephrased from proposal	At least two high-value shared services investments that expand our capacity to work with large datasets.	N/A	N/A	1	N/A
		M2 Outcome: 2.3. CGIAR develops as a learning organization.	2021 - 2.3.15: 2021: Develop capacity building activities linked to high demand data science skills.	Reworded/ rephrased from proposal	At least two workshops (online and in person) to build CGIAR big data capacity. Pilot online data science academy for CGIAR researchers.	N/A	N/A	2	N/A

MODULE	MAPPED TO SUB-IDO	2022 FP OUTCOMES	MILESTONES	INDICATE OF THE FOLLOWING	MEANS OF VERIFICATION	CGIAR CROSS-CUTTING MARKERS FOR THE MILESTONE			
						FOR GENDER	FOR YOUTH	FOR CAPDEV	FOR CC
M3	• CC Improved forecasting of impacts of climate change and targeted technology development • Optimized consumption of diverse nutrient-rich foods • {primary} Increased resilience of agro-ecosystems and communities, especially those including smallholders	M3 Outcome: 3.1 CGIAR shows how data-driven approaches yield results in poverty reduction, enhanced nutrition or environmental benefits.	2020 extended to 2021 - 3.1.6. Synthesis of Inspire project successes and failures in 2019; best practice guidance provided.	Identical to proposal	Synthesis report examining the Inspire Challenge in light of digital innovation strategy.	0	0	1	0
		M3 Outcome: 3.1 CGIAR shows how data-driven approaches yield results in poverty reduction, enhanced nutrition or environmental benefits.	2021 - 3.1.11: 2021: Implement up to seven start up Inspire projects (awarded in 2020) and up to four scale up projects awarded in 2019.	Identical to proposal	2020/2021 Cohort project reports; completed selection process and awards for up to four new Inspire and four scale-up Inspire projects.	1	1	1	1
		M3 Outcome: 3.1 CGIAR shows how data-driven approaches yield results in poverty reduction, enhanced nutrition or environmental benefits.	2021 - 3.1.12: 2021: Synthesis of Inspire project successes and failures over the course of 2017-2021, policy documents, best-practice guidance.	Identical to proposal	Synthesis report examining the Inspire Challenge in light of digital innovation strategy.	N/A	N/A	2	N/A

TABLE 2B. Planned evaluations/reviews, impact assessments and learning exercises

PLATFORM	MODULE	STATUS	PLANNED STUDIES/LEARNING EXERCISES IN THE COMING YEAR FOR GENDER	GEOGRAPHIC SCOPE	WHO IS COMMISSIONING THIS STUDY
BigData	M3	On Going	End-Project Report consolidating project successes, failures and lessons learnt from the prototyping of Croppie	National, Regional, Peru	Alliance of Bioversity and CIAT
BigData	M1,M2	On Going	Meta-analysis of digital agriculture from clearinghouse	Global	Alliance of Bioversity and CIAT
BigData	M1,M2	On Going	Digital strategy research	Global	Alliance of Bioversity and CIAT

TABLE 2C. Planned major new collaborations

NAME OF PLATFORM/CRP OR NON-CGIAR COLLABORATOR	BRIEF DESCRIPTION OF COLLABORATION AND VALUE ADDED
USDA - U.S. Department of Agriculture	Collaborate to address issues relating to crop interoperability.
SCiO - Big Data in Food Systems	Collaborate to enhance the GARDIAN data ecosystem.
ICARDA	Collaboration focused on enabling best practices towards open and FAIR(ER) data through the TRANSFORM module of the Excellence in Agronomy initiative.
CIMMYT	Collaboration focused on enabling best practices towards open and FAIR(ER) data through the TRANSFORM module of the Excellence in Agronomy initiative.
IRRI	Collaboration focused on enabling best practices towards open and FAIR(ER) data through the TRANSFORM module of the Excellence in Agronomy initiative.
IITA	Collaboration focused on enabling best practices towards open and FAIR(ER) data through the TRANSFORM module of the Excellence in Agronomy initiative.
ICRISAT	Collaboration focused on enabling best practices towards open and FAIR(ER) data through the TRANSFORM module of the Excellence in Agronomy initiative.
CIAT	Collaboration focused on enabling best practices towards open and FAIR(ER) data through the TRANSFORM module of the Excellence in Agronomy initiative.
AfricaRice	Collaboration focused on enabling best practices towards open and FAIR(ER) data through the TRANSFORM module of the Excellence in Agronomy initiative.
FAO - Food and Agriculture Organization of the United Nations	Collaboration to: (1) enable GARDIAN data assets to be discoverable via FAO-AGRIS; (2) address issues relating to crop interoperability.
The World Bank	Work with the World Bank data management team to enable wider testing, refinement, and use of GARDIAN FAIR data workflow and associated data management resources.
UF - University of Florida	Collaboration on building model-ready pipelines into Collaborative GARDIAN (CG) Labs.
EMBRAPA - Empresa Brasileira de Pesquisa Agropecuária	Collaborate to address issues relating to crop interoperability.
Accenture	Assist in the finalization of an appropriate digital strategy for the CGIAR and identify key priorities for the BIG DATA Platform.

NAME OF PLATFORM/CRP OR NON-CGIAR COLLABORATOR	BRIEF DESCRIPTION OF COLLABORATION AND VALUE ADDED
CAM - University of Cambridge	The Centre for the Study of Existential Risk and the BIG DATA Platform are conducting horizon scanning research to stay abreast of emerging topics in digital agriculture.
CIMMYT	Leading two Big Data Platform Communities of Practice (Socioeconomic data and crop modeling) and participating actively in the Geospatial community of practice.
University of Edinburgh	Leading the Livestock Data for Decisionmaking Community of Practice, an externally hosted CoP of the BIG DATA Platform.
ILRI	BIG DATA will join an ILRI committee guiding information infrastructure and governance. ILRI is also a key partner in the Information and Data Managers Community of Practice that is joining Big Data.
X - the Moonshot Factory	Collaboration on model development for large-scale, dynamic crop modeling
UC Davis - University of California, Davis	Inspire winner 2020: "Citizen-H2D3: Pilot in Rwanda" UC-Davis will lead activities in data analysis.
CSIRO - Commonwealth Scientific and Industrial Research Organisation	Inspire winner 2020: "Citizen-H2D3: Pilot in Rwanda" CSIRO will lead activities in data collection and analysis. All partners will contribute to project and results dissemination.
RECONCILE - Resource Conflict Institute	Inspire winner 2020: "Big data in resilience of rangeland communities" Partners in Kenya and Kyrgyzstan will support the piloting.
GMV - GMV	Inspire winner 2020: "Big data in resilience of rangeland communities" GMV is responsible for the technical aspects of the data platform development.
KyRICH - Kyrgyz Research Institute of Crop Husbandry	Inspire winner 2020: "Big data in resilience of rangeland communities" Partners in Kenya and Kyrgyzstan will support the piloting.
WOCAT - World Overview of Conservation Approaches and Technologies	Inspire winner 2020: "Big data in resilience of rangeland communities" Partners in Kenya and Kyrgyzstan will support the piloting.
NOAA - National Oceanic and Atmospheric Administration (United States)	Inspire winner 2020: "The ClimaCell Locust Project" NOAA will serve as data partner
IFDC - International Fertilizer Development Center	Inspire winner 2020::N-ALLyzer: From Nitrogen to ALL other nutrients" IFDC will contribute with existing and new relevant data on yield and nutrient uptake under a wide range of environmental and management conditions; AI, and modeling.
OptionLine - Optionline LLC	Inspire winner 2020::N-ALLyzer: From Nitrogen to ALL other nutrients" Optionline: help with a combined approach of Machine Learning and Artificial Intelligence (AI), App development and greenhouse trials.
Producers Direct	Inspire winner 2020: " Croppie - the PhotoCropping app" Producers Direct will manage the project; facilitate pilot testing activities with smallholders through in-country teams and farmer / youth networks in Peru and Uganda; develop the prototype app; support for training data collection.

NAME OF PLATFORM/CRP OR NON-CGIAR COLLABORATOR	BRIEF DESCRIPTION OF COLLABORATION AND VALUE ADDED
tec - Tecnológico de Costa Rica	Inspire winner 2020: "Hola Talia"-Boosting extension service through AI" Instituto Tecnológico de Costa Rica will provide expertise on artificial intelligence.
ALIN Africa - Arid Landscape Initiative	Inspire winner 2020: "Big data in resilience of rangeland communities" Partners in Kenya and Kyrgyzstan will support the piloting.
ClimaCell.org	Inspire winner 2020: "The ClimaCell Locust Project" ClimaCell.org will be coordinating and project managing the activities, and ensuring the voice of the vulnerable are represented in the product design and strategy. ClimaCell is ClimaCell.org's technological partner.
HPE - Hewlett Packard Enterprise	Collaboration on data fabric solutions and supercomputing for agroclimatic modeling
PureScan AI - PureScan AI	Inspire winner 2020: "Rapid, Low-Cost Aflatoxin detection using AI" PureScan AI will support: -Product Development -Data transparent platform/application development -Onground implementation, Industry connects
IDEO.org	Inspire winner 2020: " Croppie - the PhotoCropping app" Ideo.org will support user-experience design - including gamification and use incentives.
RAB - Rwanda Agriculture and Animal Resources Development Board	Inspire winner 2020: "Citizen-H2D3: Pilot in Rwanda" RAB will support project management, data collection, and data analysis.
NCAR - National Center for Atmospheric Research	Inspire winner 2020: "The ClimaCell Locust Project" NCAR will serve as data partner.
CORBANA - Corporación Bananera Nacional	Inspire winner 2020: "Hola Talia"-Boosting extension service through AI" CORBANA and MAG will provide in-depth agronomic knowledge and data sources for the project, as well as facilitating the testing of a live prototype with Costa Rican banana farmers.
MAG - Ministerio de Agricultura y Ganadería (Costa Rica)	Inspire winner 2020: "Hola Talia"-Boosting extension service through AI" CORBANA and MAG will provide in-depth agronomic knowledge and data sources for the project, as well as facilitating the testing of a live prototype with Costa Rican banana farmers.
Viamo	Inspire winner 2020: "Citizen-H2D3: Pilot in Rwanda" VIAMO will lead activities in data collection.
Accenture	Assist in the finalization of an appropriate digital strategy for the CGIAR and identify key priorities for the Big Data Platform

TABLE 3. Planned Budget

	W1/W2		W3/Bilateral	Center Own fund	Total	COMMENTS ONMAJOR CHANGES
	2020 Carryover	**2021 Budget**				
M1	$730,827.00	$885,600.00	$90,000.00	$0.00	$1,706,427.00	
M2	$878,967.00	$1,054,400.00	$0.00	$0.00	$933,367.00	
M3	$519,169.00	$550,000.00	$400,000.00	$0.00	$1,469,169.00	
CRP Management & Support Cost	$83,495.00	$600,000.00	$80,000.00	$0.00	$763,495.00	
Strategic Competitive Research grant	$0.00	$0.00	$0.00	$0.00	$0.00	
Platform Total	**$2,212,458.00**	**$3,090,000.00**	**$570,000.00**	**$0.00**	**$5,872,458.00**	

The header "PLANNED BUDGET" spans the W1/W2, W3/Bilateral, Center Own fund, and Total columns.

TABLE 4. ESTIMATED 2020 CARRYOVER & 2021 BUDGET TABLE

	BUDGET 2021 W1/W2			
	Carryover from 2020	**2021 Budget**	**Total 2021 Budget**	**Comments on Major changes**
Personnel	$1.134.398	$1.639.458	$2.773.856	
Consultancy	$350.000	$231.909	$581.909	
Travel	$137.000	$128.988	$265.988	
Operational Expenses	$579.000	$840.498	$1.419.498	
Collaborators & Partnerships	$0	$0	$0	
Capital & Equipment	$0	$0	$0	
Indirect costs	$12.060	$249.147	$261.206	
Platform Total Budget	**$2.212.458**	**$3.090.000**	**$5.302.458**	