Report on *Quality Assurance*

(Focus Group Discussion part of the 2022 Technical Report Learning & Optimization process)

June 2023

Table of Contents

1. Introduction

A quality assurance (QA) focus group discussion was held on June 20, 2023, for two and a half hours (10am–12:30pm). There were approximately 20 participants representing the QA assessors from the 2022 reporting process; and Initiatives; knowledge management; Monitoring, Evaluation, Learning and Impact Assessement (MELIA); Project Coordination Unit (PCU); and Portfolio Performance Unit (PPU).

The workshop was designed to learn from the past QA process realized for the 2022 Type 1 Report, completed in February 2023, and to implement the design of the next QA process. Feedback gathered from different stakeholders involved in this focus group discussion will feed into the Learning & Optimization Report.

During the focus group discussion, an overview of 2022 QA was presented, as well as some key issues and possible improvements for 2023, which were organized into three working sessions, each with its own question-and-answer session. Reflecting the pillars of the Learning & Optimization process, the three working sessions were:

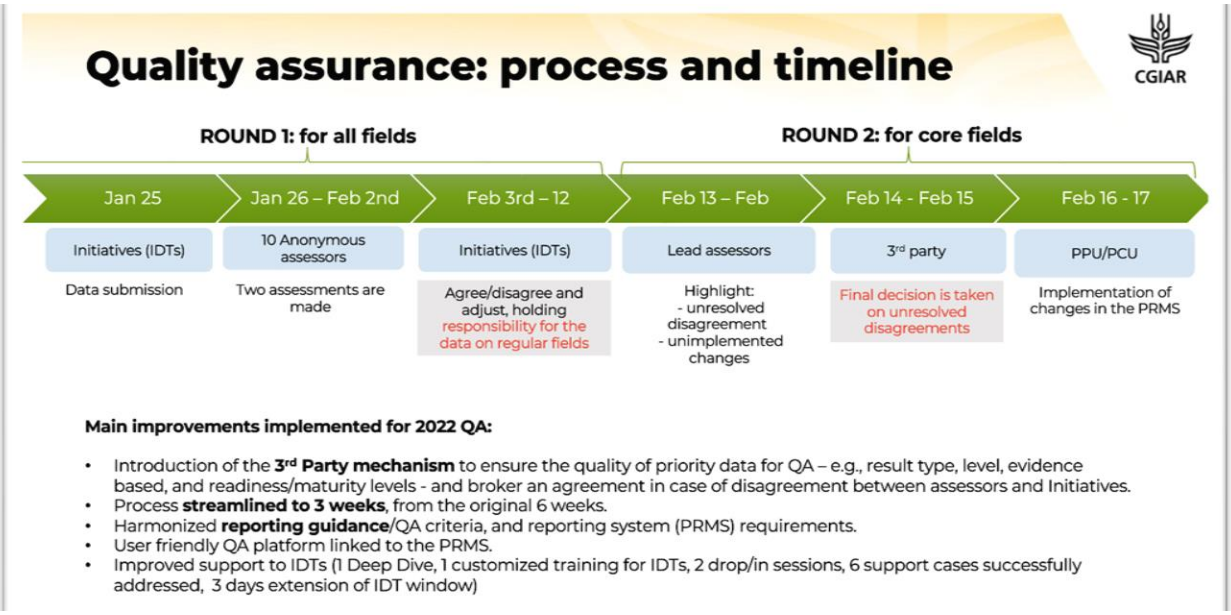
- QA product and system challenges: Quality and scope, timeline and needs of the Performance and Results Management Framework (**PRMS**)
- QA process challenges: Governance, third-party mechanism, Initiative Design Team involvement
- QA capacity and content challenges: Need for training, re-organize result indicators

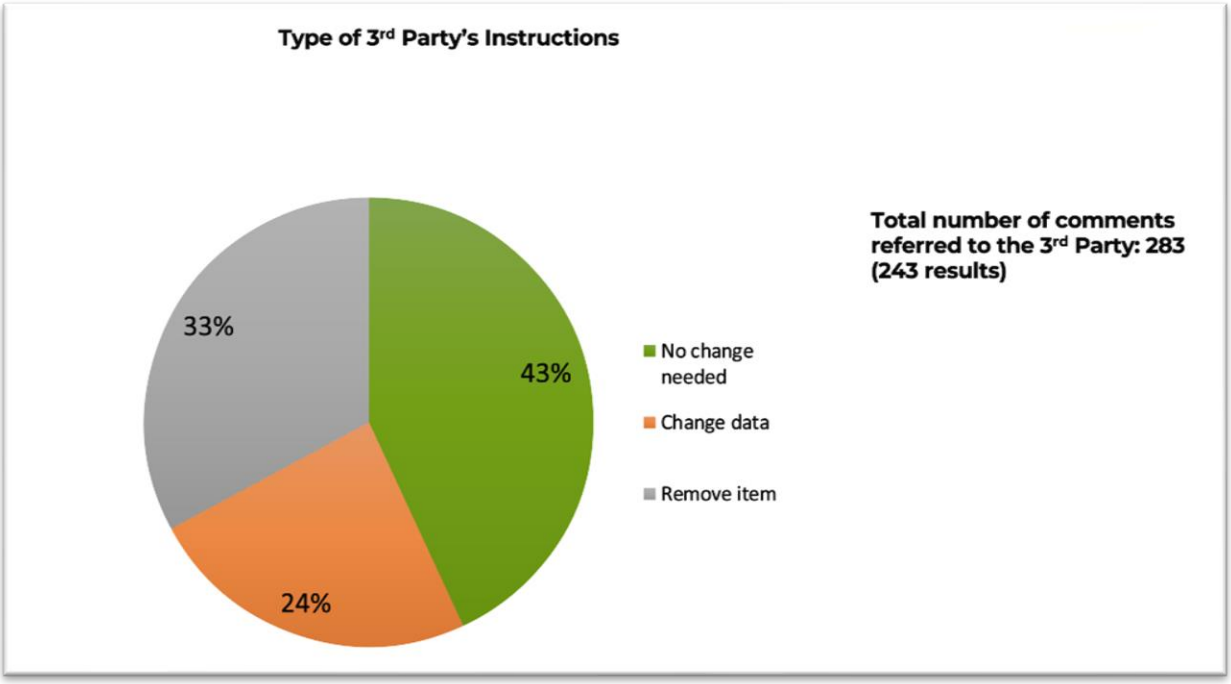
To ensure an inclusive dialogue, a Miro board was also used to collect feedback during the workshop and for the three following days. Application of Chatham House Rules allowed people to express themselves freely, as individuals, and a facilitator relatively external to the QA process enhanced this open approach.

The present report reflects the discussion conducted. It is based the recordings and notes left on the Miro board and was validated by participants. Action points and recommendations derived from the group discussion are found at the end of this document, as a conclusion.

2. Context: The 2022 QA process

The two slides below, presented during the focus group discussion, summarized the major achievements of the 2022 QA process that took place between January 25 and February 17, 2023. Most of the participants of the focus group discussion were directly involved in that QA process and offered different perspectives and represented different stakeholders. Challenges and learning from the 2022 QA experience were first introduced to the focus group discussion participants to refresh their memory and to feed into the discussion.





3. Summary of discussions

3.1. Working Session 1: QA Product and System Challenges

The first working session, on QA product and system challenges, discussed the scope and quality of the QA process, the timeline, the PRMS bottlenecks and needs, issues related to the QA Platform and other technicalities. The issues discussed are listed below.

3.1.1. Satisfaction with the product

Participants are **generally satisfied with the current QA scope and quality** and did not raise concerns about the relevance and current balance between the focal drivers of the QA process that were mentioned, i.e. accuracy, believability, objectivity, good reputation, completeness, accessibility, timeliness.

However, a concern was raised about the overall scope of the QA process regarding **knowledge products**. Participants proposed enhancing the QA process by expanding the scope to include a broader range of knowledge product types beyond papers published in peer-reviewed journals, such as gray literature, reports, and presentations. They questioned the focus on journals, as they are already well-defined sources. The roles of Centers or CGSpace librarians in QA were also discussed, and some participants felt that the QA of knowledge products other than peer-reviewed papers should not solely fall on them.

In this regard, the initiative guidance task group was mentioned as working with the metadata group to develop a clear typology of knowledge products. Efforts are under way to improve the quality of results uploaded into CGSpace, including the update of core metadata for knowledge products and ensuring standardized information.

On a broader level, participants highlighted the value of **linkages across results** to demonstrate impact and outcomes' attribution to CGIAR research, emphasizing the need for accurate and consistent reporting. Moreover, there are suggestions for **more detailed and meaningful reviews**.

3.1.2. QA frequency and timeline

Participants expressed concerns about the **last-minute rush of result submissions**, emphasizing that submitting a large volume of submissions close to the deadline compromises the quality of the data. They acknowledged the need for better planning and time management to ensure there was sufficient time for thorough reviews and the maintenance of data quality before the submission deadline.

Regarding **QA frequency**, there is a general preference for conducting **QA twice a year**, as this helps maintain control over key performance indicators and prevents them from getting out of control. Some participants, however, favored quarterly reviews while others suggested having a single reporting cycle, as it is easier to manage and provides a more streamlined process. Moreover, instead of increasing the number of reporting batches, there was a suggestion to conduct more in-depth reviews to ensure a thorough evaluation of results.

Participants expressed support for leaving **PRMS open throughout the year**, allowing results to be entered as they come up, which would make the process more efficient and effective. However, concerns were raised regarding the lowest granularity for including results, as attracting too many low-level results could pose challenges. The need to consider different levels of results, such as outputs and outcomes, when determining the reporting frequency, was highlighted. For outcomes, participants suggested to report once per year. Moreover, participants cautioned against uploading results in the PRMS that are then used and available to funders before they are quality assured, as they may undergo changes during the QA process. It is emphasized that data entered into PRMS should not be quality assured multiple times and should be completed and ready for QA. A suggestion was made to filter results through the Results Dashboard, where only quality assured results would appear, ensuring that non-quality assured results are not shared with funders.

3.1.3. PRMS (Performance Reporting and Monitoring System) and the QA Platform

Participants generally find the QA process and section in the PRMS useful and easy to navigate, suggesting a positive user experience. The following areas of improvements were discussed.

Participants expressed the need for an easier **data entry process** in PRMS, highlighting that the current method of entering results individually is laborious. They suggested considering improvements to streamline and simplify the data-entry process, potentially exploring options for bulk data entry or more efficient data input methods, e.g. based on Excel spreadsheets.

The option of adding a feature to change the "type" and “level” of result in PRMS should be an easy yet very needed fix for the coming reporting experience.

Participants raised concerns about the **lack of integration between PRMS and the theory of change (TOC) and Results Framework**. They noted that PRMS does not allow tracking progress against the TOC or Results Framework, indicating a need for better alignment and integration between these systems. Participants noted that PRMS currently does not order results by Work Package or theory of change (TOC), making it more of a repository than a tool for effective tracking against the TOC. There was a suggestion to improve the functionality of PRMS by incorporating features that align results with Work Packages or TOC, enabling better tracking and monitoring of progress.

Additionally, participants highlighted the importance of **filtering and downloading capabilities** within PRMS, particularly for Work Packages or Initiative Leads who require a better understanding of ongoing activities. A query function was proposed to enhance the search capabilities in PRMS, specifically allowing filtering based on Work Package outputs rather than just result titles. Participants emphasized the need for enhanced filtering and downloading capabilities in PRMS, specifically suggesting the ability to filter and download records, potentially in the form of Excel tables. This feature would facilitate data analysis and reporting for users.

The **overall scope of the PRMS** was also discussed. Some participants questioned whether PRMS can act as a repository, as they believe it may be built for a different purpose. There was also recognition of the PRMS not only as a reporting tool but also as a repository for storing and accessing results, enabling users to track progress over time.

3.1.4. Other technicalities

An idea was proposed to open an email account where individuals could send their QA-related doubts and potential issues to get these issues addressed before reporting on the QA platform.

Furthermore, participants noted the importance of linking the responsible person with the PRMS result in question during the QA and review process. This would enable effective communication and collaboration, particularly for follow-up actions or clarifications.

3.2. Working Session 2: QA Process Challenges

Themes from Working Session 2 concerned the QA process challenges: Governance, the third-party mechanism, Initiative Design Team involvement, strategies and tools for streamlining the process, ensuring access to evidence.

3.2.1. The third-party mechanism

The third-party mechanism, introduced in 2022 QA, has ensured the **quality of priority data** for QA and brokered an agreement in case of disagreement between assessors and Initiatives. The major issues faced were a lack of clarity in some of the decisions that were taken and Inaccessibility of evidence, which affected result acceptance in an inconsistent way.

The participants expressed concerns regarding the **clarity of decisions** made by third-party assessors. They highlighted the need for clearer and more transparent assessments, where the third-party's decisions can be more easily understood.

Additionally, issues related to **evidence accessibility** were raised, with some assessors unable to access the evidence provided by the initiatives. Broken links in the evidence further complicates the process, requiring additional efforts to retrieve and review the supporting information. To address these challenges, it is recommended to establish clear rules from the beginning, ensuring that evidence is submitted through accessible repositories and emphasizing that evidence not meeting these criteria will not be accepted — thus, results not meeting these requirements will be unsubmitted by the PRMS (but not deleted). By clarifying the expectations and rules, the assessment process can be streamlined, reducing the need for back-and-forth communication at the end of the process.

3.2.2. Streamlining processes and ensuring quality with AI and automation

The QA process was streamlined to three weeks in January/February in order to provide timely, quality-assured data to the reflect process/dashboard. However, it is suggested that it could be further streamiled by increasing the **use of AI** to automate certain tasks in the QA process, e.g. spotting links between evidence and results and accelerating assessments through AI-based keyword analysis. Participants proposed leveraging AI and automation to enhance the efficiency and effectiveness of the QA process. Specifically, they suggested using AI to help identify links between evidence and results. This would be particularly useful in cases where lengthy documents or reports serve as evidence, as it can be challenging for assessors to locate the relevant sections quickly. AI-powered algorithms could potentially analyze keywords or other markers to identify the relevant information, expediting the assessment process automatically. Moreover, AI could be used, for instance, to spot restricted access evidence before the QA process starts. By implementing such automation, QA tasks can be performed more swiftly and accurately, freeing up valuable time for both assessors and initiatives.

Moreover, it was suggested to reflect on how to **further automate and better focus the QA effort regarding knowledge management** and journals in particular. Currently, some 90% of peer-reviewed journals are automatically validated by the Manufacturer Quality Assurance Program. The participants discussed the need for streamlining the publication process. They suggested implementing a centralized process driven by staff affiliation and scope of data. This would involve extracting publications based on initiatives' scope and making them available in the PRMS. By automating this process, the burden on individuals involved would be reduced. Additionally, initiatives could select publications that align with their theory of change, ensuring relevance and consistency. This discussion raises the question of whether CGSpace should have quality levels to ensure that only high-quality results are displayed.

3.2.3. Management, access and quality of evidence

Participants expressed the need for improved management of evidence for the reported results throughout the year, particularly within Initiatives/Centers, to ensure their quality for public viewing and accessibility in repositories.

The existence of internal procedures for documents and evidence of **quality control** was questioned, with consideration that it might vary across Centers and Initiatives. Participants highlighted the importance of avoiding a last-minute rush to submit evidence and the associated burden on librarian time and CGSpace. Challenges in providing evidence for impacts further along the impact pathway are acknowledged. Different types of evidence are needed to capture progress along different stages of the pathway.

The **accessibility** and linking of evidence are crucial for ensuring the QA process and making it publicly available on the dashboard, with accessible links. There is need to explore ways to link evidence that cannot be made public, finding a solution to ensure the appropriate accessibility and sharing of restricted evidence. Differentiating between evidence suitable for public display and evidence that demonstrates policy influence or widespread technology uptake is important. Suggestions were made to publish certain types of evidence, such as knowledge products and key performance indicators, on platforms like CGSpace, while restricting access to other types of evidence on SharePoint based on access rights. It is emphasized that CGSpace may guarantee restricted type of access to entries (only to CGIAR users, only to the QA team) and, thus, be suitable also for storing evidence that should not be of puplic domain. These functionalities should be better explained to Initiatives.

Some other points were made regarding evidence management:

- Concerns were raised about the confidentiality of QA assessors when using center repositories; measures to ensure the anonymity of assessors need to be considered.
- Integration of evidence from Center-maintained repositories, such as the Monitoring, Evaluation and Learning platform, with PRMS was suggested as a way to streamline the process and avoid duplication of entries.
- The importance of clear guidelines and decisions regarding evidence quality and un-submission of results that do not pass QA was emphasized.
- A point was made on the evidence requirements for gender and climate scores: When tagging results with climate change and/or gender, participants noted that the request for precise details of where to find evidence was often redundant. They highlighted that the title of the result generally contains the necessary information, making the additional evidence request unnecessary.
- Evidence description guidelines should be included.

3.2.4. Alignment with standards and quality assurance

The possibility of aligning with international standards, specifically the International Organization for Standardization (ISO) quality assurance, was generally welcomed by participants. Exceeding expectations in terms of quality can be a significant selling point for funders. The need for a better understanding of the requirements and processes to align with international standards was highlighted.

3.3. Working Session 3: QA Capacity & Content Challenges

The themes discussed in Working Session 3, regarding QA capacity and content challenges, went from re-organizing result indicators, to training, to improving guidance.

3.3.1. Re-organize result indicators

In 2022, an unexpected number of "other output indicators" OT (approximately 600) were submitted and their QA was squeezed for time concerns. Concerns were raised by the focus group discussion participants about a lack of OT alignment in the PRMS. Participants expressed the need for clearer guidelines or definitions to ensure consistency in reporting this category of results, which should be definitely smaller or even removed.

Guidance should further be precise about when and **how to report events** (e.g., high-level meetings or conferences), including event inputs and outputs, to better link results across capacity development, knowledge products, innovation development, OT or other result type.

The development of a mechanism to **track stakeholder engagements and partnerships** was mentioned, which can support the Type 1 Narrative and potentially provide evidence for policy/strategy or any other outcome as soon as engagement outputs derive from non-CG stakeholders' involvement.

3.3.2. Improving the reporting guidance

Guidance was considered poor for **capacity development and capacity change**, with need for providing examples of good titles and descriptions for different types of capacity development activities such as lectures, workshops, short courses, and formal degrees (Ph.D., M.Sc., B.Sc.). Participants expressed the importance of clear guidance on the result, indicator, and evidence for each type and level of capacity development activity.

Some participants raised an issue about the attribution and assessment of **gender disaggregation**, especially in innovation use and virtual capacity-building activities. The breakdown of innovation use into male and female categories was found to be challenging and may not accurately reflect the gender distribution. Participants mentioned instances where a group of individuals (both male and female) joined a session using a single link, making it hard to verify their gender. The assumption of a 50:50 gender ratio included in the reporting guidance was deemed incorrect and undermining the credibility of the assessment.

3.3.3. Enhancing knowledge management

The knowledge management domain should be further improved and harmonized with the QA effort. Several challenges and possible solutions were discussed.

The **involvement of Center librarians in the QA** process was discussed. Participants suggested providing training and increasing awareness among Center librarians regarding quality assurance processes, specifically for journal articles. This would enable them to contribute effectively to the quality assurance efforts within the repository, potentially reducing the workload for other stakeholders involved.

A point of discussion concerned the **types of knowledge products** to be accepted in CGSpace across different Centers that, currently, have different policies and reporting rules. Participants expressed the need for clear guidance on what should be considered a knowledge product and what should not (yet to be considered as evidence for other results or a piece of research not to report). This aspect should be clarified, as well as the minimum quality for knowedge products to be reported as such (and uploaded to CGSpace) and whether reporting them under "other" should be further clarified. The initiative guidance task group was mentioned as working with the metadata group to develop a clear typology of knowledge products.

Finally, concerns were raised regarding the current **publication process**. Participants mentioned that there was a rush at the end of the year, and different systems were used by each Center, leading to inefficiencies. Data had to go through multiple platforms and Center librarians played a role in checking the data. The risk of individuals waiting for a handle and needing to delete and re-enter information, was discussed, especially if they had drafted a knowledge product but never received a handle. While the in-house quality assurance efforts were acknowledged, it was noted that they might not meet the requirements of an external quality assurance process. Participants emphasized the need for improved systems and training.

3.3.4. Need for training and capacity development

Several areas for improvement though training were discussed:

- Concerns were raised about the absorption of comments from assessors and **divergent responses among assessors**, indicating a need for addressing consistency and enhancing the acceptability of assessor comments. Participants proposed evaluating the QA process to improve feedback management and reduce disagreements.
- Concerns were raised regarding the expectations for QA and PRMS **data entry**. Participants felt that better guidance should be provided to data inputters to improve the quality of data entry, enhance understanding of the structure and results framework, and avoid considering day-to-day meetings as capacity development activities. Some participants also expressed the opinion that the expectation for QA should not solely rest with Center or CGSpace librarians.
- Participants discussed the importance of **linking results within the PRMS** to showcase connections and patterns. Some initiatives have already explored this feature, which has led to intriguing insights. The need for additional training on linking results and understanding the purpose and benefits of such linkages is emphasized.
- **Titles of results** are identified as an ongoing issue that needs attention to improve clarity and searchability.

3.3.4. Quality Assurance Community of Practice (QA CoP):

The idea of establishing a QA community of practice (CoP) was raised and its value was discussed. Some participants believed that it could add value if managed effectively, while others expressed doubts about its usefulness from the perspective of data inputters.

4. Major recommendations and action points

Working Session 1: QA product and system	
1	Improve the functionalities of PRMS/QA Platform to be able to conduct a deeper QA regarding matching results against ToCs.
2	Ensure the possibility of changing result type and level without unsubmitting and then submitting a new result.
Working Session 2: QA process	
1	Design the QA process timeline to reduce the final workload and improve the quality of the data submitted and their QA, e.g., by introducing a preliminary batch.
2	Review the QA process for knowledge products, including harmonized procedures to be followed at the Center level, in close collaboration with knowledge management team, what has to be QAed by external assessors and what relies on Centers responsibility.

3	Clearly define and communicate the governance mechanism underpinning the QA process and how the decision of un-submitting the results that did not comply with the evidence/quality requirements is taken.
4	Gather a better understanding of the requirements and processes to align the current QA process with international quality standards (ISO).
5	Leveraging AI and automation to enhance the efficiency and effectiveness of the QA process.
6	Defining a way to share evidence with assessors before the QA, especially by studying improved functionalities of CGSpace for restricted access.
Working Session 3: QA content and capacity	
1	Redefine the indicators to better capture and classify results that were reported under “Other” categories, e.g. design of a specific indicator for partnership/stakeholder engagement.
2	Studying the option of establishing a community of practice reflecting on QA across CGIAR (approach, training, cost/benefit analysis).
3	Update Standard Indicator Description Sheets (SIDS), reporting guidance and, accordingly, QA criteria and instructions, especially regarding capacity development and capacity change, how to report events that could alternatively fit capacity development or knowledge products or innovative capacity or capacity development type, or Other Outcome and the “Other output ” categories.