Evaluability Assessment of CGIAR's Portfolio 2025-2030: Synthesis Report

A. Jersild, J. Lenne, G. Gullotta, M. M. Molinari,
IAES: S. Negroustoueva

March 2026

Acknowledgments

Contents

List of Figures

List of Tables

List of Boxes

Acronyms

BDN	Better Diets and Nutrition	RII	Regional Integrated Initiatives
Capsha	Capacity Sharing	S4I	Scaling for Impact Program
CoP	Community of Practice	SAAF	Sustainable Animal and Aquatic
EA	Evaluability Assessment		Foods
ERG	Evaluation Reference Group	SDG	Sustainable Development Goal
GEI	Gender, Equality and Inclusion	SIMEC	Standing Impact Monitoring and
GEYSI	Gender, Equality, Youth and Social		Evaluation Committee
	Inclusion	SME	Subject Matter Experts
GDI	Gender, Diversity and Inclusion	SFP	Sustainable Farming Program
ISDC	Independent Science for Development	SP	Science Program
	Council	ToC	Theory of change
MEL	Monitoring, Evaluation and Learning	ToR	Terms of Reference
MELIA-F	Monitoring, Evaluation, Learning and	TRA	Technical Reporting Arrangement
	Impact Assessment-Foresight	VfI	Value for Investment
MFL	Multi-Functional Landscapes	VfM	Value for Money
MR	Management Response		
MYEP	Multi-Year Evaluation Plan		
PCU	Program Coordination Unit		
PORB	Plan of Results and Budget		
PPU	Portfolio Performance Unit		
QA	Quality Assurance		
QoR4D	Quality of Research-for-Development		
QoS	Quality of Science		
R4D	Research-for-Development		

Glossary of Terms

Term	Definition	Source
Accelerator	Accelerators will undertake strategic research in their topical areas and, through working with all other Programs, bolster CGIAR's ability to reach and support targeted end-users while furthering collaboration, coherence, and integration across the entire Portfolio.	CGIAR Tech Reporting MELIA Glossary v4.1 - September 2025[1]
CGIAR Results Framework	The program logic explains how the development objective is to be achieved, including causal relationships and underlying assumptions. Related terms include results chain and logical framework.	PRMF Glossary
Evaluability Assessment	An Evaluability Assessment (EA) is a systematic examination of a program's readiness to be evaluated. An EA is the first step to ensuring that an evaluation will be cost-effective, will result in a useful learning process, and will produce findings that inform decisions and improve program performance. It helps clarify an evaluation's feasibility, scope, and value for money. Within CGIAR, 'evaluability' is a standard of the Evaluation Framework, and in the EA Guidelines. "Evaluability refers to the extent to which an intervention can be evaluated in a reliable and credible fashion; the concept is central to a culture of results. A strong focus on evaluability at the design stage facilitates overall 'measurability', monitoring and subsequent evaluation."	EA Guidelines.
Evaluand	The subject or object being evaluated-in other words, *what* the evaluation is about, or the intervention in scope.	
Evaluation	The systematic and objective assessment of an on-going or completed intervention (project, Program, or policy), its design, implementation, and results. In CGIAR, evaluation refers to an external, completely (commissioned by the Independent Evaluation Function) or largely (Initiative) independent and systematic study of an in-depth nature that uses evaluation criteria in line with Evaluation Policy. In addition to research, it also applies to central CGIAR organizations, support Programs and themes, and the system as a whole.	Monitoring, Evaluation, Learning and Impact Assessment (MELIA) Glossary Version 5: November 2021
Impact Area	CGIAR Impact Areas are: (1) Nutrition, Health and Food Security; (2) Poverty Reduction, Livelihoods and Jobs; (3) Gender Equality, Youth and Inclusion; (4) Climate Adaptation and Mitigation; (5) Environmental Health and Biodiversity. CGIAR is targeting multiple benefits across five Impact Areas to contribute to collective global targets for the transformation of food, land, and water systems across local, regional, and global levels.	CGIAR Tech Reporting MELIA Glossary v4.1- September 2025
Management Response	Management Response (MR) is a formal mechanism that helps ensure that evaluations are used to contribute to organizational effectiveness, learning, and accountability. MRs facilitate strategic engagement on evaluation findings and recommendations for appropriate follow-up actions through a formal process.	CGIAR 2022 Evaluation Policy

---

[1] This version of the MELIA Glossary builds on the MELIA glossary v3 used for 2024 Technical Reporting, which incorporated the most recent approved changes to the CGIAR Results Framework. It serves as an interim solution to ensure consistent taxonomy across guidance documents, tools, and platforms for Technical Reporting, until the CGIAR taxonomy, being co-developed by centers and the SO under the MELIA-F Project, is finalized (expected mid-2026).

Term	Definition	Source
Monitoring	A process of continuous or periodic collection and analysis of data to compare how well an intervention/policy is being implemented against expected progress and results, to track performance against plans and targets, to identify reasons for under or over achievement, and to take actions to improve performance.	PRMF Glossary
Performance and Results Management Framework	The Performance and Results Management Framework (PRMF) describes the processes, systems and measures for managing CGIAR's performance and results to support delivery of the CGIAR 2030 Research and Innovation Strategy.	CGIAR Tech Reporting MELIA Glossary v4.1- September 2025
Performance and Results Management System	The Performance and Results Management System (PRMS) underpins the PRMF, encompassing planning, monitoring, and reporting. It provides information to support informed decision-making, enabling real-time data collection and the day-to-day management of the Portfolio and Science Programs/Accelerators.	CGIAR Tech Reporting MELIA Glossary v4.1- September 2025
Portfolio	The CGIAR Portfolio is made up of the research programs and/or platforms carried out by the centers and the CGIAR System Partners in support of the CGIAR Strategy and Results Framework and which are supported by (1) the CGIAR Trust Fund, and/or (2) bilateral sources contractually aligned to such programs and/or platforms.	CGIAR Tech Reporting MELIA Glossary v4.1 - September 2025
Program	CGIAR Programs are the main vehicle for delivery of research and innovation from 2025-30. Programs serve as entry points to describe CGIAR's offer on a key topic, elevating CGIAR's visibility in global agendas and facilitating the continuation and formation of inclusive alliances and partnerships.	CGIAR Tech Reporting MELIA Glossary v4.1- September 2025
Sustainable Development Goal	Seventeen (17) internationally accepted, publicly broadcast objectives related to achieving Sustainable Development Goals (SDGs) adopted between 2015-30.	CGIAR Tech Reporting MELIA Glossary v4.1- September 2025
Value for Investment	Value for Investment (VfI) is an evaluation system underpinned by four principles: interdisciplinary, mixed methods, evaluative reasoning, and participatory. It is designed to bring clarity to answering evaluative questions about how well resources are used, whether enough value is created, and how more value could be created from the resources invested in a policy or Program. VfI is a re-framing of Value for Money (VfM) as good use of resources. This framing aims to shift our lens from 'the money' to a broader perspective that views policies and Programs as investments in value propositions with the potential to create significant social, cultural, environmental or economic value.	Better Evaluation

Executive Summary

The Evaluability Assessment (EA) of 13 Programs and Accelerators (P/As) in CGIAR's 2025-30 Research Portfolio under the 2030 Research and Innovation Strategy aligns to the System Council endorsed Multi-Year Evaluation Plan (MYEP) per the 2025-27 Workplan for CGIAR's Independent Advisory and Evaluation Service (IAES) (SC/M21/DP5).

Box 1. Definition of evaluability [2]

> **Evaluability** refers to the extent to which an intervention can be evaluated in a reliable and credible fashion and is central to a culture of results. A strong focus on evaluability at the design stage facilitates overall measurability, monitoring, and subsequent evaluation.

Aligned to the Terms of Reference (ToR) EAs of CGIAR's Portfolio 2025-30**,** the EA Framework from the Guidelines for Conducting and Using EA in CGIAR (2022), was applied to the 13 SP/As in the 2025-30 Portfolio. As a learning, advisory and quality assurance (QA)-focused exercise, it provided recommendations. The EA process followed four phases: (1) inception; (2) partial, light-touch EAs towards Independent Science for Development Council (ISDC) Review (SC22 update); (3) full implementation of EAs; and (4) a synthesis phase-all to meet the following three objectives:

1. Identify Portfolio-wide patterns in evaluability, e.g., strengths, recurrent constraints, and emerging system-level tendencies across the seven EA domains.

2. Provide actionable, system-level recommendations to strengthen evaluability across the Portfolio, with indicative near-term priority actions.

3. Compare findings with the 2023 Regional Integrated Initiatives (RII) EA Synthesis results to understand the status of CGIAR's evaluability conditions, persistent constraints, and what implications there are for future evaluation planning and system-level learning.

Method

To develop the Synthesis, the EA team followed a four-step analytical process: (1) systematic **evidence compilation** from each individual EA, based on review of core documents obtained between June–August 2025, such as proposals, inception reports, MELIA Plan and Results Framework, prioritization documents, and ISDC reviews; (2) **cross-Program synthesis by EA domain**; (3) **cross-domain analysis** to identify linkages across design, evidence, learning and resource systems; and (4) Research Portfolio level synthesis **of cross-cutting patterns**.

Figure 1. Seven domains of the EA Framework

A. Intervention Logic

B. MEL Systems and Resources

C. Gender diversity and inclusion

D. Long-term evaluability

E. Context and environment

F. Management and Stakeholder engagement and response

G. Cost Effectiveness and Efficiency
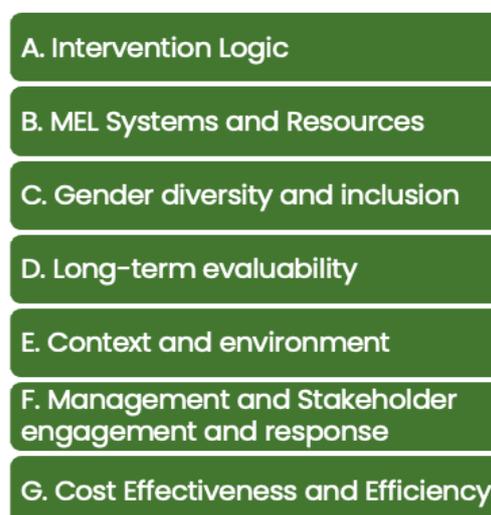
Source: Guideline for EAs in 2022

Analytical rigor was ensured through iterative team calibration (to align interpretation of domain criteria), expert peer review, and validation workshops with Monitoring, Evaluation, Learning, and Impact Assessment (MELIA) and Program teams. While the seven-domain EA Framework provided a unified structure, evaluability expectations were calibrated to intervention type–assessing causal pathways for Science Programs (SPs), enabling contributions for Accelerators, and scaling logic for Scaling for Impact Program (S4I). This preserved comparability while ensuring interpretive fairness and credibility across interventions with different mandates.

---

[2] CGIAR IAES. 2022. CGIAR Evaluation Framework. Rome: CGIAR Independent Advisory and Evaluation Service (IAES), Evaluation Function

The Synthesis interprets evaluability as a system property shaped by CGIAR's organizational norms, learning culture, and governance environment, rather than as a static feature of individual interventions.

Results

The Synthesis identified seven system-level themes that consistently recur across the 13 P/As. These themes cut across the EA domains and reflect organizational conditions that shape evaluability across the Portfolio. While the Portfolio demonstrates strong conceptual design and improved MELIA institutionalization, persistent constraints limit evaluability in practice across the seven themes: (1) ambitious but weakly operationalized theories of change (ToCs); (2) fragmented MELIA systems and uneven maturity; (3) strong GEI commitments but weakly operationalized; (4) short-term outcome horizons limiting long-term evaluability; (5) evaluation readiness assumed rather than planned; (6) learning practices present but not institutionalized; and (7) transparent finances but weak cost-performance linkages. Together, these seven themes converge into three cross-cutting evaluability patterns:

(1) **Strong design, weak operational grounding** (Domains A, B, D): ToCs and MELIA structures are in place, but contribution pathways, assumptions, and evidence roles require clearer operationalization.

(2) **Fragmented evidence and learning systems** (Domains A, B, D): Data are generated, but evidence use, GEI integration, and financial linkages remain inconsistent and siloed.

(3) **Weak linkage between resources and performance** (domains A, B, G): Financial systems track inputs well, but connections to outcomes and proportionality remain partial.

These recurring patterns informed four system-level recommendations.

Recommendations

Four recommendations featured in table 1 below address the structural drivers of evaluability constraints and provide a feasible 24-month agenda to strengthen evaluability across CGIAR. The table summarizes recommendations with indicative actors involved; see more detail in section five and table 4. Recommendations are subject to MR as per CGIAR-wide evaluation policy and MER evaluation guidelines, and the CGIAR Process Note on MR.

Table 1. Recommendations

Recommendations	Suggested involved actors
1. **Strengthen operational pathways for intervention logic and evidence generation.**	Office of the Chief Scientist, P/A Directors, MELIA focal points.
2. **Build an integrated MELIA-F & Finance architecture:** Create coherence across indicators, baselines, finance, learning and gender, equality and inclusion (GEI) metrics.	Office of the Chief Scientist, *via* PPU and Integrated Partnership Business Operations and Finance Unit, P/A Directors, MELIA focal points.
3. **Formalize evaluation readiness as a system requirement:** Integrate context scanning, risk analysis, and long-term evidence continuity into annual monitoring and reporting cycles.	Office of the Chief Scientist, *via* PPU, P/A teams, including Internal Audit experts; CGIAR Integrated Partnership Risk & Internal Controls Executive.
4. **Establish Portfolio-level evaluability governance:** provide coherence, accountability, and forward planning across all units.	CGIAR Chief Scientist, *via* PPU and PCU. Advisory by IAES, towards mid-line evaluations.

1 Introduction

1.1 Background

The Evaluability Assessment (EA) of 13 Programs and Accelerators (P/As) of the CGIAR's 2025-30 Research Portfolio under the 2030 Research and Innovation Strategy aligns to the System Council -endorsed Multi-Year Evaluation Plan (MYEP) per the 2025-27 Workplan or CGIAR's Independent Advisory and Evaluation Service (IAES) (SC/M21/DP5).

Box 1. Definition of evaluability and evaluability assessment

Evaluability refers to the extent to which an intervention can be evaluated in a reliable and credible fashion, and the concept is central to a culture of results. a strong focus on evaluability at the design stage facilitates overall measurability, monitoring, and subsequent evaluation.

An **Evaluability Assessment** is an early review of a proposed activity to ascertain whether its objectives are adequately defined, results verifiable, and evaluation questions answerable. An Evaluability Assessment establishes whether and how an intervention can be evaluated reliably and credibly.

Source: CGIAR IAES. Evaluation Policy, and 2022. CGIAR Evaluation Framework. Rome: CGIAR Independent Advisory and Evaluation Service (IAES), Evaluation Function.

Aligned to the ToR: EAs of CGIAR's Portfolio 2025-30, the EA Framework from the Guidelines for Conducting and Using EAs in CGIAR (2022), was applied to the 13 P/As in the 2025-30 Portfolio to respond to the following objectives:

- Support increased evaluability of P/As (interventions) across CGIAR's Portfolio 2025-27, and Portfolio evaluability.

- Support the development of a flexible and utility- focused monitoring, evaluation, and learning (MEL) approach at intervention level.

- Coordinate with the Independent Science for Development Council (ISDC), CGIAR's Program Coordination Unit (PCU) and the Portfolio Performance Unit (PPU) to maximize synergies and the use of results and learning from the EA process.

- Synthesize results across the 13 P/As EAs, towards an early assurance for the System Council of the Portfolio's evaluability.

Box 2. Outputs of the EA process, by Management Response (MR) status requirement

EA Reports (13)- Results from each EA exercise are presented in 13 separate reports, per P/A assessed. Each report outlines findings by assessment domain and provides actionable recommendations. EA reports are available to CGIAR to support **adaptive management,** including during the 2026 Pause and Reflect Stage or its equivalent.

Synthesis EA Report-The themes of the 13 EA reports, with concluding findings and recommendations, subject to a **MR**, and inclusion of accepted/partially accepted recommendations in the CGIAR's MR Tracker.

The EA process was implemented from January to November 2025 in close coordination with CGIAR stakeholders, PPU, PCU and aligned with ISDC. As a learning, advisory and Quality Assurance (QA) exercise, it aimed to provide specific actionable recommendations.

The EA process used four phases at programmatic level: (1) inception, (2) partial, light-touch EAs towards ISDC review; (3) full implementation of EAs; and (4) a final Synthesis phase. Table 2 maps the EA phase to the body of evidence and activities. Many documents were reiterated within the scope of the EA exercise.[3] This Synthesis Report an output of Phase 4- the culmination

---

[3] Key program documents included: the P/A proposal, Inception Reports, Results Framework, MELIA plans, comparative advantage, prioritization documents, Plan of Results and Budget (PORB), and ISDC reviews of Inception Reports. Individual EA reports contain full lists of documents reviewed.

of the 2025-30 Portfolio EAs, it integrates findings from the 13 interventions to surface system-level evaluability patterns and recommendations. Beyond summarizing results, it examines the conditions that shape evaluability across CGIAR's research, delivery, and enabling mechanisms–key insights to guide future design, governance, and investment decisions for the Portfolio.[4]

Table 2. Documents/resources reviewed across the EA process by phase

Phase, 2025	Activities/Documents/resources reviewed
1. **Inception** (February-July)	• P/As full design documents. • MELIA Needs Assessment Survey (February 2025). • P/As documents shared from April to June, e.g., theory of change (ToC), Monitoring, Evaluation, Learning and • Impact Assessment (MELIA Plan), and Results Framework.
2. **Light-touch EAs for ISDC review** (May-July)	• P/A Inception Reports & other documents from June 2025 (ToC, MELIA Plan, Results Frameworks). • Session and Bilateral meetings during CGIAR Science Week (7-12 April 2025).
3. **Full EA implementation** (August-October)	• P/As Inception Reports (June); documents submitted through October 2025 (ToC, MELIA Plan, Results Framework, Comparative Advantage, Prioritization, Plan of Results and Budget (PORB)); and their additional materials (e.g. presentations, MELIA representatives, outputs tracker). • P/As comments on the EA Inception Notes (July-August). • ISDC Review of 2025-30 Research & Innovation Inception Reports (August) and response from P/As to ISDC reviews (September). • Online EA Self-Assessments per P/A by EA domains (August). • Interviews with P/A teams. • Sessions for co-development validation of recommendations.
4. **Synthesis** (November-December)	• 13 P/As EA reports, feedback from P/As. • Sessions for co-development validation of recommendations.

The MR process–CGIAR's formal, time-bound commitment to implementing evaluation recommendations that are tracked and reported annually-applies to the EA Synthesis Report only, as explained above. The 13 individual EA reports are produced as guidance to the P/As and will be made public on a case-by-case basis, with their agreements.

This EA Synthesis marks the third application of CGIAR's EA Framework, following its use in the Gender Platform Evaluation, and of the four Regional Integrated Initiatives (RIIs), as part of IAES' 2023 Workplan.

1.1 Assessing Evaluability of CGIAR's 2025-30 Portfolio

Aligned with CGIAR's 2030 Research and Innovation Strategy, the 2025-30 Portfolio brings together 13 interventions to address interconnected global challenges such as climate change, gender and social inequalities, poor-quality diets, rural poverty, environmental degradation, fragility, and conflict.[5] The Portfolio is structured around a set of Impact Areas linked to multiple Sustainable Development Goals (SDGs), and comprises eight Science Programs (SPs), the Scaling for Impact Program (S4I), and four Accelerators (Figure 1). Collectively, these interventions aggregate expertise, partnerships, and organizational mechanisms intended to deliver CGIAR's ToC.

---

[4] Consult 2025 Evaluability Assessments: Everything You Need to Know–developed post SC22 update with key stakeholders during the EA implementation process.
[5] Please see discussion on emergent megatrends identified by ISDC in the Review of 2025-30 Research and Innovation Inception Reports.

Figure 1. CGIAR's 2025-30 Science and Innovation Portfolio
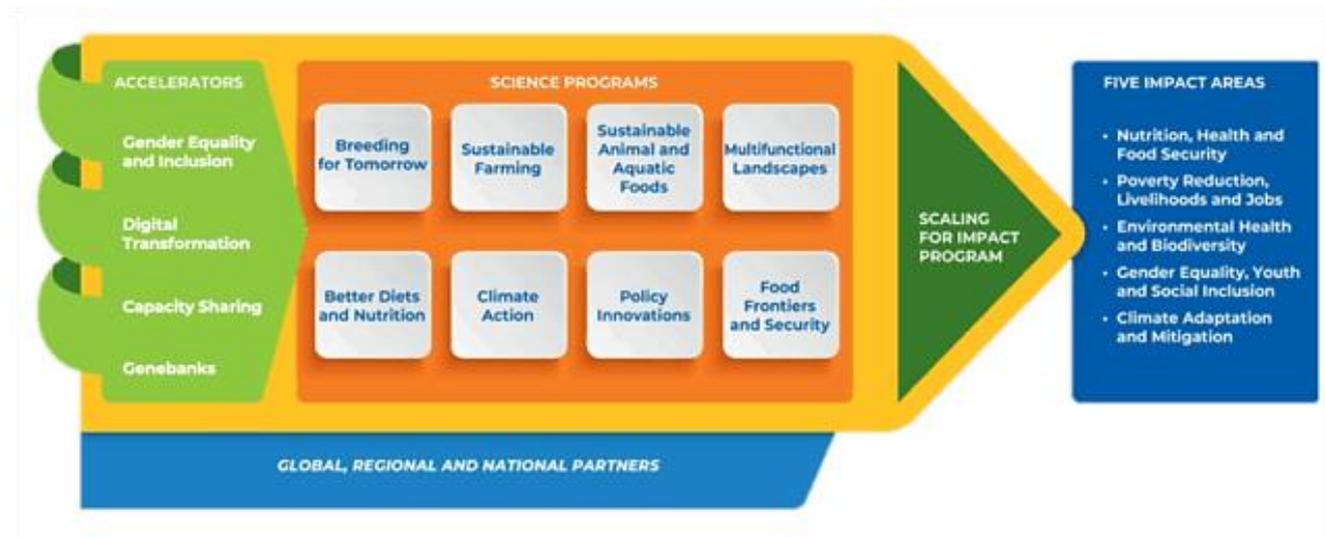
Source: CGIAR Research Portfolio 2025–30, Accessed November 2025

SPs generate thematic scientific evidence and innovation. Accelerators, by contrast, provide cross-cutting system capabilities–equity integration, digital infrastructure, capacity sharing, and genebanks. The S4I Program is the Portfolio's dedicated mechanism for scaling, supporting the SP/As in translating and scaling innovations and strengthening responsiveness to partner demand. These structural distinctions matter for evaluability *in* the following ways:

- **Program level: Distinct intervention roles create differentiated evaluability expectations.** SPs must articulate and operationalize complex causal pathways and multi-actor research trajectories. Accelerators must demonstrate how enabling functions contribute to measurable change across SPs; and the S4I Program must provide evidence of translation, uptake, and influence at system scale. Since these roles differ, evaluability requires differentiated expectations within a unified framework, allowing each intervention type to demonstrate contribution according to its role in the Portfolio.

- **Portfolio level: The 2025-30 Portfolio reflects a shift toward a systems-oriented research culture.** The ISDC Review of 2025-30 Research and Innovation Inception Reports states that the new Portfolio marks a movement away from a "traditional linear, technologically focused research agenda" to a "contemporary, systems approach, and impact-oriented research culture."[6] Systems change work requires clearer articulation of contribution, stronger MELIA integration, and more deliberate attention to learning, context, and organizational behaviors. These shifts make evaluability more dependent on understanding how interventions interact within a single, integrated portfolio architecture, rather than assessing them in isolation.

- **Operational level: Implementation realities shape what credible evaluative judgment looks like.** The Portfolio's design and delivery occur within operational realities that have direct implications for evaluability: multi-center implementation, blended and uncertain inter-annual funding (W1/W2, W3, bilateral), variable geographic environments, and diverse partner ecosystems. These evolving design and governance conditions shape what credible evaluative judgments look like across the Portfolio. They also underscore the necessity of a system-level synthesis, to understand how evaluability manifests within CGIAR's emerging Portfolio architecture. These conditions affect feasibility, evidence availability, and contribution visibility.

The Portfolio operates as an interdependent, matrix system in which research pathways, enabling functions, scaling mechanisms, funding structures, and governance arrangements jointly shape what constitutes credible

---

[6] CGIAR Independent Science for Development Council. 2025. ISDC Review of 2025-2030 Research and Innovation Inception Reports. Rome: CGIAR Independent Advisory and Evaluation Service.

evidence, contribution and learning. These dynamics underscore the need for a unified yet differentiated evaluability approach: one that respects the distinct roles of SPs, Accelerators, and S4I while assessing how evaluability conditions arise from CGIAR's broader institutional architecture.

This synthesis of the Portfolio's 13 P/As serves as a diagnostic of evaluability at a point in time to strengthen evaluation readiness, and as a foundation for system-level improvements that can support adaptive management, and steer strategic decision-making across CGIAR. In this EA synthesis, evaluability is used in a broad, systems-oriented sense, encompassing: (1) core evaluability conditions related to testability (e.g., intervention logic, data, indicators); (2) articulation of values-based objectives and pathways; and (3) system-level enabling or constraining factors that shape the feasibility and utility of evaluation. The intent of including these domains is diagnostic rather than judgmental–to surface conditions that affect evaluability feasibility and use, and to inform adaptation and decision-making at appropriate levels (Program, Portfolio, or system), rather than to assess Program performance or compliance. See **Annex 1** for the executive summaries of the 13 P/As-specific EA reports with references.

2 Purpose and Objectives of the Synthesis

Per the ToRs, this Synthesis Report consolidates the results of 13 EAs conducted between July and November 2025, to obtain a system-level view of evaluability across the Portfolio and to identify the institutional enablers or constraints of a credible, useful, and timely evaluation. The primary synthesis objectives include:

1. **Identify Portfolio-wide patterns in evaluability,** including strengths, recurrent constraints, and emerging system-level tendencies across the seven EA domains.

2. **Provide actionable, system-level recommendations** to strengthen evaluability across the Portfolio, with indicative near-term priority actions.

3. **Compare findings with earlier RII EA Synthesis results** to understand where CGIAR's evaluability conditions evolved, where constraints persist, and what this implies for future evaluation planning and system-level learning.

The four recommendations from this Synthesis are subject to MR as per CGIAR Evaluation Policy and MER Guidelines and the endorsed ToRs, to follow the **PPU Process Note.** Selected EA reports from P/As may be publicly available, based on the willingness of the P/A teams; however, their recommendations are not subject to the formal MR process.

3 Method

3.1 Analytical Framework

This Synthesis builds on the EA Framework (Figure 2), a structured approach to assess the readiness of Programs and Accelerators for credible evaluation. The Framework examines seven domains (A–G), which together capture the essential conditions for evaluability across a research-for-development (R4D) context. **Annex 2** presents the full set of domain questions.

Applying the seven domains across all 13 EAs (eight Science Programs, the S4I Program, and four Accelerators) provided a common analytical structure and enabled cross-Portfolio comparison. EA reports for each of 13 interventions formed the primary evidence base for this synthesis. The EA Framework demonstrates flexibility and robustness.

However, due to fundamentally different mandates of the three Program types, the EA team calibrated an interpretive lens in the following manner:

Figure 2*. Seven domains of the EA framework*

A. Intervention Logic

B. MEL Systems and Resources

C. Gender diversity and inclusion

D. Long-term evaluability

E. Context and environment

F. Management and Stakeholder engagement and response

G. Cost Effectiveness and Efficiency

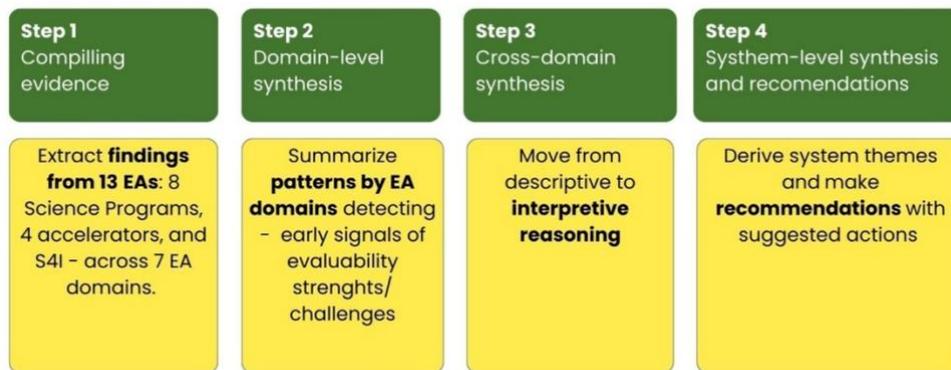Source: Guideline for Evaluability Assessments (2022)

- **SPs** were assessed for causal clarity, research-to-outcome pathways.
- **Accelerators** were assessed for enabling contributions, influence, and the changes they catalyze in partner systems.
- **S4I** was assessed for scaling logic, demand responsiveness, adaptation, and uptake conditions.

To maintain comparability while ensuring fair judgments, the EA team applied the same seven-domain Framework while adapting the criteria used to interpret evidence within each domain. In practice, this meant that all interventions were assessed against the same evaluability domains, however evaluative judgments reflected the function of each intervention type in the Portfolio (research, enabling, or scaling). Thus, comparability was achieved through a shared evaluative approach, and credibility was ensured by calibrating domain expectations to the distinct roles of SPs, Accelerators, and S4I. This approach allowed for methodological consistency without imposing unrealistic uniformity across intervention types.

3.2 Data Sources and Analytical Process

The Synthesis followed a four-step, iterative, analytical process designed to move from descriptive comparison to interpretive reasoning and system-level inference (Figure 3). The first step compiled evidence from the 13 EAs, extracting findings and recommendations by domain into a structured synthesis matrix to create a common evidence base. This was followed by horizontal analysis across interventions to identify recurring strengths and constraints within each domain. Next, cross-Domain reasoning examined relational patterns (e.g., how MELIA design influences GDI evaluability or cost-effectiveness), revealing systemic linkages and tensions. Finally, a system-level synthesis interpreted recurrent patterns across domains as reflections of CGIAR's organizational norms, incentives, and design logics, generating seven system-level themes and recommendations.

Figure 3. Four-step process to the EA Synthesis

Step 1 Compilling evidence

Step 2 Domain-level synthesis

Step 3 Cross-domain synthesis

Step 4 Systhem-level synthesis and recommendations

Extract **findings from 13 EAs**: 8 Science Programs, 4 accelerators, and S4I – across 7 EA domains.

Summarize **patterns by EA domains** detecting – early signals of evaluability strenghts/challenges

Move from descriptive to **interpretive reasoning**

Derive system themes and make **recommendations** with suggested actions

Analytical rigor of the EA process was strengthened through:

Team calibration: An iterative calibration process during regular EA team meetings to compare interpretations of evidence, clarify domain criteria, and harmonize judgments across all 13 EAs.

Rather than imposing a numerical rating scale, the team focused on developing shared evaluative reasoning, ensuring consistent application of the EA Framework and strengthening credibility of cross-portfolio inferences.

Peer review: Draft synthesis outputs were reviewed by subject-matter experts (SMEs) from a joint ISDC/Evaluation Function roster. Authors of individual EA reports cross-checked each other's work, providing additional triangulation and refinement (see **Annex 6** for biographies).

Validation measures: Preliminary findings and draft recommendations were reviewed and co-developed with stakeholders in three validation events (see Box 3). Feedback from these processes informed revisions to both analysis and narrative.

3.3 Analytical Orientation of the Synthesis

The Synthesis treats evaluability as a property shaped by the interaction between intervention design and CGIAR's organizational environment. Evaluability depends not only on technical quality of ToCs, indicators, baselines, or MELIA systems, but also on broader organizational factors such as governance arrangements,

documentation practices, funding structures, and the maturity of learning and accountability mechanisms across CGIAR.

Variation in documentation depth, timing, and MELIA capacity across the 13 interventions creates limits to comparability. Hence, the synthesis interprets evaluability as a dynamic property rather than a fixed condition-one that reflects both design quality and organizational context. The use of a shared seven-domain framework mitigated comparability limitations by providing a consistent set of evaluability anchors, even where the nature of evidence varied.

Comprehensive methodological details, i.e., data sources, analytical procedures, calibration processes, and validation steps are presented in **Annex 2**, with technical definitions in the Glossary of Terms above.

4 Findings

4.1 Analytical Transition from EA domains to Cross-Cutting Themes

The findings in **Annex 3** discuss how the 13 Portfolio interventions perform against the seven evaluability domains. Figure 4 below provides a Portfolio-wide snapshot of these conditions across domains A–G. The 'yes,' 'partial,' and 'no' judgments reflected in the Figure are directly derived from the guiding questions for each domain in the EA Framework and were assessed through systematic document review, interviews, and team calibration during each individual EA.

Figure 4 offers an indicative view of evaluability conditions across the Portfolio. However, because the guiding questions vary in importance, and because a 'yes' response does not always imply full evaluability, the Figure should be interpreted as an approximation of patterns rather than a numeric score. The distribution of yellow, dark gray, and light gray highlights relative strengths and gaps, but it does not constitute a weighted index or imply equivalence across domains or questions.

Figure 4. Portfolio-wide evaluability landscape across domains A-G
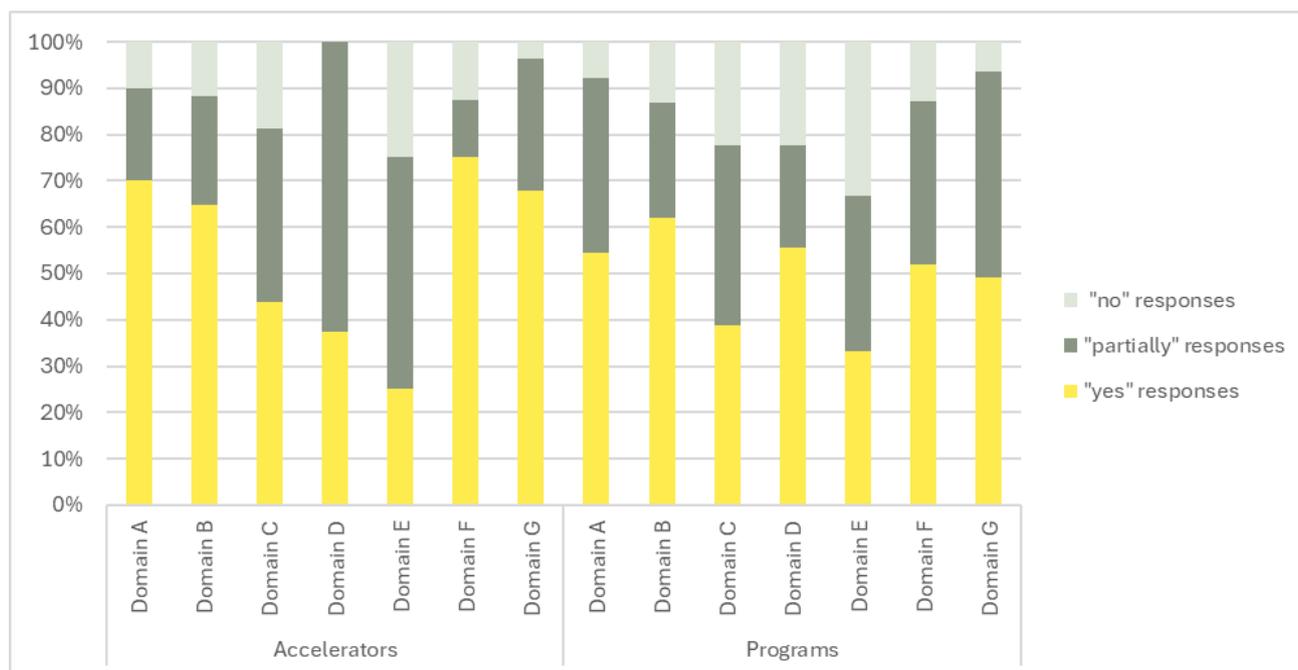
The clustering evident in the graph shows evaluability conditions present across all EA domains, with some domains considered more robust than others: among Accelerators, management and stakeholder engagement, and among Programs and MEL systems. The Synthesis shifts from assessing isolated conditions by domain to

analyzing the cross-domain interdependencies linking intervention logic, MELIA practice, inclusion, contextual readiness, and financial structures.

Using the cross-domain reasoning approach in Methods (and further detailed in Annex 2), the analysis identified seven system-level themes that consistently occur across the Portfolio. They reflect evaluability challenges and strengths that are not domain-specific. Technical issues (e.g., weak causal assumptions, incomplete baselines, limited stakeholder mapping) often reflect deeper institutional dynamics such as incentive misalignment, fragmented workflows, or uneven coupling between evidence systems and resource decisions. For this reason, the themes presented below should be understood not as re-statements of the EA domains, but as cross-domain expressions of an evaluability system in practice.

4.2 Thematic Summative Findings with Implications

4.2.1 Ambitious but Weakly-Operationalized Theories of Change

Across CGIAR's 13 interventions, the ToCs are conceptually strong, being well aligned with the 2030 systems-transformation framing and scientifically credible, but they remain only partly evaluable when viewed through a wider system lens. The logic is internally coherent (domain A) yet insufficiently linked to the evidence systems (domain B), time horizons (domain D), and financial architecture (domain G) needed to make it evaluable in operational terms.

As a result, design and delivery are often decoupled: Programs optimize for narrative coherence rather than evaluability (i.e., clarity about how and under what conditions results can credibly be demonstrated). Interventions tend to be theory-rich but evidence-thin. Many articulate plausible pathways but stop short of specifying who is responsible for generating, verifying, and resourcing the data that would make those pathways evaluable in practice.

Interpretive Insight: The recurrent cross-domain pattern suggests that CGIAR's evaluability challenge is not conceptual weakness but a failure to integrate across systems. Strong logic frameworks largely operate in isolation from MELIA systems that could validate them, from financial models that could cost them, and from management processes that could adapt them. This points to an organizational design culture that prizes ambition in research breadth and depth and coherence at the expense of realistic operationalization, perhaps a reflection of, in part, donor demand for greater delivery and impact, but within an environment of uncertain and declining funding. **Illustrative Evidence** includes:

- Breeding for Tomorrow (B4T) and Sustainable Farming Program (SFP) exemplify high-quality causal design but face evaluability limits due to under-specified feedback loops and multi-year biological lags.

- Despite its complexity, Better Diets and Nutrition (BDN) has institutionalized MELIA, which has enabled greater integration between domains A and B.

- GEI and Capacity Sharing (CapSha) articulate enabling pathways credibly but cannot yet trace contribution effects across Programs.

- Sustainable Animal and Aquatic Foods (SAAF) and Multi-Functional Landscapes (MFL) demonstrate how scientific and contextual complexity challenge linear causal assumptions.

System-Level Implication focuses on ToCs, which must evolve from static narrative instruments into integrated management tools that explicitly connect logic, evidence, and resources. Strengthening these cross-domain linkages between ToCs, MELIA, finance, and adaptive management will shift CGIAR from designing for compliance to designing for evaluability, towards improved coherence across the Portfolio.

4.2.2 Fragmented MELIA Systems and Uneven Maturity

Across the Portfolio, MELIA architectures are present and aligned in principle with CGIAR's Performance and Results Management Framework (PRMF) 2022-30. However, the planning documentation and MELIA arrangements show uneven readiness for evaluation. While most interventions outline monitoring systems capable of generating routine data, fewer provide sufficient information to assess whether these systems will

support adaptive decision-making, integrate qualitative and behavioral indicators, or consistently establish baselines. Several interventions propose internal indicators that are not linked to a cross-Portfolio registry, which may limit future comparability and synthesis.

Interpretive insight: This pattern suggests that, at the design and planning stage, MELIA systems are more fully articulated for compliance-oriented reporting than for evaluative learning. The documentation reveals stronger clarity on what will be measured than on how evidence will be interpreted and used for adaptation or how MELIA cycles will link to financial or managerial decision processes. When ToCs remain broad or under-specified (domain A), MELIA plans default toward output monitoring rather than outlining the evidence needed for contribution analysis. These conditions may constrain long-term evaluability (domain D), as current plans provide limited detail on baseline strategies, qualitative protocols, or tracer approaches that would allow progress to be meaningfully assessed over time. **Illustrative evidence** includes:

- B4T and SFP show stronger practice linking reflection activities, dashboards, and planned learning events to potential management adjustments.

- MFL, BDN, Climate Action, and Policy and Innovation outline quantitative indicators aligned to PRMF requirements but provide limited detail on how qualitative, behavioral, or QoS-related evidence will be generated or used. Baseline plans in their documentation vary in completeness.

- GEI and CapSha outline participatory learning intentions without their consistent mapping to standard PRMF/QoR4D structures; actor typologies and contribution pathways are often implicit in documentation. Several interventions propose internal indicators without specifying how these will feed into portfolio-level systems.

System-level implication points to a MELIA ecosystem that is structured but not yet coherently integrated. The necessary elements for monitoring exist, but the plans do not consistently demonstrate how evidence will flow into coordinated learning or resource decisions across CGIAR. Strengthening evaluability therefore requires moving from distributed documentation to a more collectively aligned MELIA architecture so that planned evidence systems support both management needs and strategic evaluation readiness.

4.2.3 GEI Commitments Strong but Weakly Operationalized

Across the Portfolio, GEI are institutionally visible but unevenly embedded in the evaluative fabric of P/As. Most Initiatives identify women, youth, marginalized groups and partner organizations as target groups, yet only a minority articulate how inclusion influences research design, uptake pathways, or outcome measurement. The pattern interacts closely with domain A (Intervention Logic) and domain B (MELIA Systems): inclusion is often presented as a strategic principle within ToCs but not supported by indicators, disaggregated data, baselines, or MELIA processes capable of capturing differential benefit.

Interpretive insight: This inconsistency reflects a broader system tension between normative commitments and evaluative practice. CGIAR's strategic approach positions GEI as central to transformation, but operational systems still treat it as a cross-cutting value rather than a measurable domain of performance. Consequently, evidence on equity outcomes remains anecdotal or activity-based instead of demonstrating empowerment, redistribution of influence, or institutional change. The finding also highlights an epistemic gap: inclusion data tend to be quantitative and descriptive, while qualitative or behavioral measures of agency, decision-making, or social legitimacy are under-developed. This limits the system's ability to substantiate gender-transformative claims or trace equity effects across levels of influence. **Illustrative evidence** includes:

- BDN and MFL programs integrate inclusion through participatory platforms and just-transition frameworks, but feedback loops linking these activities to higher-level outcomes remain incomplete.

- SFP and B4T programs show early integration of GEI within multi-actor partnerships, yet inclusion is often inferred from partner composition rather than from evaluative metrics.

- GEI Accelerator leads organizational discourse and policy guidance on inclusion but relies on broad stakeholder categories ('partners,' ''users,' 'beneficiaries') that mask differentiated accountability.

- CapSha and S4I began tracking youth and early-career engagement, illustrating progress toward social inclusion, though reporting is fragmented.

System-level implication of the uneven operationalization of inclusion points to a wider challenge in translating CGIAR's values into evaluable practice. Inclusion is widely affirmed but rarely structured as an evidence domain, leaving the system unable to demonstrate equity outcomes credibly. The implication is that GEI should evolve from a moral and narrative commitment to an evaluative function—one that connects representation, resource distribution, and outcomes.

4.2.4 Short-Term Horizons Limit Long-Term Evaluability

Across the Portfolio, sustainability and long-term evaluability are recognized in concept but rarely designed for practice. Nearly all 13 interventions articulate ambitions extending to 2030 and beyond, yet few include mechanisms for tracing outcomes, maintaining datasets, or following partners after Program closure. This pattern connects domain A (Intervention Logic), domain B (MELIA Systems), and domain D (Long-Term Evaluability). While logics describe transformation trajectories, MELIA systems seldom provide the temporal infrastructure—baselines, continuity plans, or *ex-post* provisions needed to substantiate lasting change.

Interpretive insight: The weakness lies not in intent but in planning horizons. Program cycles and funding models are short relative to the temporal dynamics of R4D. Many Initiatives equate durability with institutional uptake, assuming partner ownership ensures persistence. However, few define what sustained change looks like or how it will be evidenced. The result is a form of temporal myopia: evaluations capture progress within funding windows, while system-level transformation unfolds far beyond them. Long-term evaluability therefore requires treating evidence of continuity as part of design, not as a retrospective exercise. **Illustrative evidence** includes:

- BDN, MFL, and FFS programs link durability to behavioral and organizational resilience but lack indicators for persistence after 2030.

- SFP program introduces intermediate indicators and long-term partnerships that could underpin future *ex-post* assessments.

- GEI and CapSha assume sustained change through partner integration but provide no framework to verify post-program uptake.

- Genebanks is an outlier, and one that is funded in perpetuity, supported by Crop Trust's dedicated long-term funding and monitoring structures.

At system level, this finding exposes the relatively short planning horizons embedded in CGIAR's design and funding model. Programs aspire to long-term transformation but are bounded by short-term accountability cycles. Evaluability thus depends on building institutional memory and continuity mechanisms that extend beyond project timelines—turning sustainability from an assumption into a demonstrable system property.

4.2.5 Evaluation Readiness is Assumed, Not Designed

Across the Portfolio, evaluation readiness is largely assumed rather than planned. Most interventions possess sound MELIA systems and accessible contexts, yet few have institutionalized procedures for assessing their preparedness for credible, timely evaluation. This pattern spans domain B (MELIA Systems) and domain E (Context and Environment), intersecting with domain A (Intervention Logic), where feasibility is embedded in design. Evaluation feasibility often depends on tacit organizational knowledge, such as experienced staff or partner familiarity rather than explicit readiness plans, risk assessments, or contextual intelligence protocols.

Interpretive insight: CGIAR's evaluation culture favors optimism and aspiration over foresight. Program teams tend to treat evaluability as a default property of well-designed interventions, rather than as an operational condition requiring proactive evidence-based adaptive management. External shocks such as political transitions, climate variability, or data-access constraints, are typically addressed reactively, not anticipated. As a result, evaluation timing, data availability, and stakeholder engagement hinge more on circumstance than design. Strengthening readiness thus involves shifting from assumed feasibility to planned evaluability: integrating contextual scanning, risk analysis, and adaptive scheduling into the standard management cycle. Among P/As **illustrative evidence** includes:

- MFL and FFS operate in dynamic field environments and have experimented with flexible data collection modalities, yet few conduct formal environmental or political scans.

- GEI and CapSha depend on alignment with partners' MELIA calendars and organizational sensitivities, leaving evaluation scheduling and data access vulnerable to misalignment.

- Only SFP has recommended routine annual context scans as a standard evaluability practice.

Across the Portfolio, contextual awareness is informal, captured in discussions or reports but rarely documented as part of MELIA planning and implementation.

System-level implication is the recurrent assumption of evaluation readiness reveals an organizational blind spot: readiness is treated as given rather than planned. This pattern reflects a procedural culture focused on implementation rather than evaluative foresight. The system-level implication is that preparedness must be institutionalized, anticipating contextual risks, planning evidence windows, and ensuring accessibility for credible evaluation.

4.2.6 Learning Valued but Weakly Governed, Inhibiting Evidence-Based Decisions

A strong learning ethos exists across the Portfolio, yet institutionalization of learning systems remains largely weak. Nearly all P/As value reflection, adaptive management, and co-creation, but few have formal mechanisms to document, track, or operationalize lessons learned. This theme draws primarily from domain F (Management and Stakeholder Engagement) but also connects to domain B (MELIA Systems) and domain G (Cost-Effectiveness), where learning is a necessary link between evidence and decision-making. Across the 13 interventions, learning is frequent, but follow-through is inconsistent.

Interpretive insight: CGIAR's culture prizes reflection with limited investment in learning governance. The outputs of knowledge exchange through 'pause-and-reflect' sessions, workshops, and Communities of Practice, rarely feed into Portfolio learning or management decisions. Learning is event-based and person-dependent rather than systematized through feedback loops, accountability processes, or response tracking. This suggests a learning-performance gap: evidence is generated but not institutionally cycled back into adaptive planning, limiting the system's capacity to learn at scale. I**llustrative evidence** includes**:**

- GEI and CapSha cultivated active learning communities and participatory reflection spaces but lack systematic tracking of what insights are used and by whom.

- BDN and FFS programs show emerging good practices–embedding reflection into regular review cycles and linking evaluation findings to course corrections–yet these remain the exceptions rather than the norms.

- In MFL and Policy Innovation programs, learning sessions are frequent but undocumented; feedback loops depend heavily on individual leadership commitment.

At system level, this theme underscores that CGIAR values reflection without institutionalized learning as governance. Learning occurs through goodwill and local Initiative rather than structured accountability. The implication is that system maturity will depend on transforming learning from episodic practice to an embedded function—linking reflection, evidence, and management action in traceable ways.

4.2.7 Transparent Finances, Opaque Value

Financial systems across Portfolio 2025–30 demonstrate high transparency and sound fiscal management, yet the evaluability of efficiency and value-for-investment remains limited. In some but not all interventions, budgets and expenditures are mapped to Areas of Work and High-Level Outcomes, with weak linkages between financial data and performance evidence[7]. This theme draws primarily from domain G (Cost-Effectiveness and Efficiency) and intersects with domain A (Intervention Logic), domain B (MELIA Systems), and domain D (Long-term Evaluability), where performance information should inform, and be informed by, resource allocation.

Interpretive insight: The analysis reveals a structural disconnect between financial accountability and performance evidence. CGIAR's budgeting frameworks emphasize transparency and compliance, but they are not yet designed to demonstrate value creation. Legacy funding, overlapping projects, uncertain funding in outyears and mixed pooled–bilateral financing obscure causal linkages between investments and outcomes. The result is a

---

[7] Related study on domain G is available on IAES publications and EA portal.

system that can show where money is spent but not what value is generated. Integrating MELIA and finance systems—so that cost, performance, and learning data flow together—would enable evidence-informed resource decisions and enhance organizational credibility.

Beyond the issue of cost–performance linkage, the analysis also highlights a temporal misalignment between financial accounting cycles and the realization of outcomes and impacts typical of R4D. While expenditures are transparently recorded and attributed within defined budget periods, many outcomes—particularly those related to systems change, capacity development, policy influence, and behavioral or organizational transformation—materialize several years later, after costs were incurred. In the absence of explicit mechanisms to link historical investments to results not yet evident, financial data cannot readily support *ex-post*, longitudinal, or contribution-based evaluation. This constrains not only assessments of efficiency but also domain D (Long-Term Evaluability) domain G (Cost-efficiency; VfI), as current financial and MELIA arrangements do not consistently provide the temporal traceability needed to assess whether, how, and under what conditions earlier investments contribute to sustained change over time.

Addressing these issues would enhance effective use of Technical Reporting Arrangement (TRA) requirements within CGIAR, and its connection to MELIA frameworks and evaluative processes for greater evaluability, particularly for assessing value for investment and longer-term contribution. **Illustrative evidence** includes:

- SFP and SAAF programs began mapping bilateral funding contributions to specific outcomes and HLOs, demonstrating partial progress toward financial-performance integration.

- Genebanks and CapSha maintain robust fiscal oversight, but face challenges in translating enabling functions (e.g., capacity building, convening, policy support) into measurable efficiency metrics.

Across programming, legacy funding complicates cost-effectiveness analysis, historical investments continue to yield benefits. However, these are untagged and untraceable in current reporting cycles. Standardized costing guidance and shared Value for Investment (VfI) rubrics-currently under development for 2026–could offer a key entry point for addressing this evaluability gap.

System-level implication shows a financially transparent but evaluatively opaque system. The disjunction between expenditure and outcome data limits the capacity to articulate value for investment or adaptive efficiency. Evaluability depends on connecting financial stewardship with evidence of performance, creating a shared evaluative language that bridges finance, science, and strategy.

4.3 Cross-Cutting Patterns and Interactions

The conceptual diagram in Figure 5 shows evaluability as a system property, shaped not by performance in any single domain, but by the quality of connections linking design, evidence, inclusion, long-term planning, engagement, and financial systems across the CGIAR Portfolio.

Figure 5. Evaluability within CGIAR as an Integrated System, by Seven Themes

The seven themes identified through the synthesis are interdependent: each represents one dimension of evaluability, yet all emerge from the same underlying system dynamics. While the seven themes originate from different domains of the EA Framework, their interactions reveal that evaluability challenges arise from how the system connects–or fails to connect–to its own evaluative functions rather than technical gaps. Strong upstream capacities (e.g., well-articulated ToCs, routine monitoring) often lose evaluative momentum when not linked to downstream functions such as adaptive decision-making, long-term evidence strategies, or resource allocation.

To summarize these interactions, table 3 highlights three system-wide evaluability dynamics visible across the Portfolio.

Table 3. System-wide cross-domain patterns influencing evaluability

Cross-cutting pattern	Domains most affected	System-level implication
Strong design, weak operational grounding	A, B, D	ToCs and MELIA structures are in place, but evaluability depends on clearer operational pathways, explicit assumptions, and defined evidence roles.
Fragmented evidence and learning systems	B, C, F, G	Data are generated, but evidence use is inconsistent; learning remains siloed and episodic.
Weak linkage between resource use and performance	A, B, G	Financial systems track inputs well, but connections to outcomes are partial, limiting efficiency and VfI assessment.

Strengthening evaluability requires more deliberate links between intervention logic, evidence generation, resource allocation, and managerial decision-making. In this sense, evaluability emerges not simply as a technical property of individual P/As, but as an organizational capability rooted in how CGIAR governs for evidence use, coherence, and learning.

5 Recommendations

The four recommendations, presented below and summarized in table 4, translate the seven themes and three cross-cutting system patterns into system-level recommendations with suggested actions, for **MR**, and subsequent implementation subject to tracking in the MR Tracker. They do not target individual P/As; rather, they address the structural and cultural constraints that shape how evaluability is conceived**,** generated, and used across CGIAR. As the analysis above indicates, a persistent cultural assumption–that evaluability is an inherent property of well-designed interventions rather than a capability requiring foresight, planning, and coordination– limits the organization's ability to embed evaluation into routine decision-making. Taken together, these recommendations are designed to shift that culture from optimism to operational realism, from fragmented practices to coordinated systems, and from compliance-oriented reporting to evidence-informed management. Each **recommendation** specifies: (1) a system action–what must change to strengthen evaluability; and (2) a set of non-binding practical actions-for *how* this could be operationalized given varying capacities and resources. Collectively, the four recommendations constitute a feasible 24-month agenda for strengthening evaluability as a shared organizational capability.

Table 4. Recommendations, priority actions, and estimated responsible parties

Recommendation, with core purpose (what must be done)	Suggestive non-binding actions (how it could be done)	Suggested lead actors and suggested involved actors:
1. **Strengthen operational pathways for intervention logic and evidence generation.**	- Develop simplified operational ToC template. - Conduct ToC stress-test exercises. - Provide capacity strengthening on causal reasoning (PPU). - Formal linkage established/communicated on operationalizing ToC with planning & reporting. - Use Adaptive Management guidance to implement recommendations at P/A level.	Office of the Chief Scientist, P/A Directors, MELIA focal points.

Recommendation, with core purpose (what must be done)	Suggestive non-binding actions (how it could be done)	Suggested lead actors and suggested involved actors:
2. Build an integrated MELIA-F-Finance-Learning architecture: Create coherence across indicators, baselines, finance, learning and GEI metrics.	- MELIA-F and Finance Division Integration Taskforce. - Annual Learning and Decision Briefs for Adaptive Management. - Indicator harmonization. - Minimum MELIA Standards Guide. - Use Adaptive Management guidance to support MELIA–Finance–Learning architecture.	Office of the Chief Scientist, via PPU and Integrated Partnership Business Operations and Finance Unit, P/A Directors, MELIA focal points.
3. Formalize evaluation readiness as a system requirement: Integrate context scanning, risk analysis, and long-term evidence continuity into annual cycles.	- Develop Evaluability Readiness Checklist based on EA domains for Adaptive Management dashboard. - Semi-annual context-environment scans by domain (based on MR Tracker and other evidence). - Readiness simulations (2 SPs + 2 Accelerators). - Long-term evidence continuity plans.	Office of the Chief Scientist, via PPU, P/A teams, including IA experts; CGIAR Integrated Partnership Risk & Internal Controls Executive.
4. Establish Portfolio-level evaluability governance: provide coherence, accountability, and forward planning across all units.	- Include a CGIAR Evaluability Review and Planning Session at the annual MELIA CoP. - Develop State of the Portfolio Brief. - Conduct Stock take of implementing EA recommendations in MLE (2027). - Technical Guidance Bursts. - Quarterly learning exchanges under MEL CoP.	CGIAR Chief Scientist, via PPU and PCU. Advisory by IAES, towards midline evaluations.[8]

Behind co-development of recommendations, Box 3 provides insights. The EA activity was dynamically managed by IAES/EF cognizant of the multiple demands and evolving priorities of key internal stakeholders during the EA implementation period. The evidence points to multiple areas in which collaborative, early action would be necessary to enhance evaluability, effectiveness, and Portfolio cohesion, including through the CGIAR-wide MELIA-F effort. Taken together, the four recommendations and their corresponding suggested actions constitute a realistic, resource-aligned 24-month agenda for strengthening evaluability as a system capability across CGIAR.

---

[8] System Council approved Multi-Year Evaluation Plan (MYEP) (2025-27 work plan for CGIAR's Independent Advisory and Evaluation Service (IAES) (SC/M21/DP5). Midline evaluations of the programs and accelerators are scheduled for 2027.

Box 3. Recommendation co-development

Following the standard processes, the EA team used the UN Evaluation Group's Improved Quality on Evaluation Recommendations Checklist and aligned to the 2025 Review of CGIAR Management Response System to Independent Evaluations in finalizing the synthesis recommendations. [SC decision: M21-EDP9].

In line with Recommendation 1 of the MR System Review, the non-binding practical actions in this chapter are intended as **non-exhaustive, illustrative suggestions** to support MR development. These suggestions do **not** prescribe mandatory steps; rather, they offer possible ways recommendations could be operationalized and may be adapted—or not used—depending on context, capacity, and stakeholder relevance. Their purpose is to facilitate clearer and more actionable MR development while preserving managerial discretion. The following steps were taken:

- Presenting preliminary recommendations at a designated workshop on the side-line of the MELIA-F project launch in Nairobi, Kenya on October 14 2025.

- Online validation meeting with Portfolio leads and MEL focal points on November 5, 2025.

- Discussion with the *ad hoc* consultative group for the EA activity on November 12. 2025.

- Presentation of recommendations to the CGIAR's MEL Community of Practice on November 21, 2025.

The range of participants across meetings included representatives of PPU, PCU, MELIA leads from the P/As, funders and IAES (see Annex 5 for participants). Workshop discussions and comments were considered in finalizing the recommendations presented below (consult a related blog).

RECOMMENDATION 1: STRENGTHEN OPERATIONAL PATHWAYS FOR INTERVENTION LOGIC AND EVIDENCE GENERATION.

(Domains A, B, D): Strengthening causal clarity and operational feasibility will improve contribution analysis, enable judicious resourcing, and create a coherent pathway from research outputs to measurable change. To address this, CGIAR should adopt simplified, operational ToC templates, require explicit articulation of realistic intermediate outcomes, and embed assumptions and contribution logic directly into design and MELIA cycles. The non-binding suggestive measures include:

1.1 Develop a formal link between ToC implementation and planning and reporting functions: a simplified, operational ToC template would help integrating contribution pathways, detailed and plausible assumptions, evidence roles, and feedback loops. PPU should initiate this participatory process with P/A leads and MELIA focal points, and in consultation with PCU and IAES. The emphasis should be on the development of a template for P/A use, with guidelines focused on the development of evaluable ToCs.

- Timeline: A Q1-Q2 development timeline is suggested.

- Cost implication: Low to moderate cost with minimal new budget complementing existing reflection and planning cycles. The process requires workshop facilitation and MELIA staff time.

1.2 Develop and conduct a ToC 'stress-test exercise' for all P/As to identify unrealistic assumptions, missing actors, and under-specified outcome pathways. This exercise can be used on a periodic basis for ensuring ToC relevance and evaluability. PPU should create a user-friendly checklist with accompanying notes to guide the process and consult on its development with IAES, PCU and MELIA focal points.

- Timeline: A Q2 development timeline is suggested, along with development of the ToC template.

- Cost implication: Low cost.

1.3 Provide capacity building on causal reasoning, contribution logic, and the development of intermediate outcome indicators. PPU should convene to provide guidance and resources, on approaching causal reasoning and contribution logic in the context of complex systems change and evaluation, possibly as thought pieces or blogs to be published, topics for discussion in a MELIA Community of Practice (CoP), or a session featured in a MELIA workshop.

- Timeline: Ongoing, 2026-27.

- Cost implication: Staff costing, and occasional consultant costing.

1.4 Use CGIAR's Adaptive Management guidance to steer and embed a continuous, evidence-based learning and decision-making process supported. Portfolio Adaptive Management Guidance[9] can support implementing recommendations at P/A level for promote flexible and responsive approach across the portfolio.

- Timeline: Ongoing, 2026-2027.
- Cost implication: Staff costing

RECOMMENDATION 2: BUILD IN AN INTEGRATED MELIA-F AND FINANCE MECHANISM THAT SUPPORTS PORTFOLIO-LEVEL COHERENCE AND INCLUSIVE EVIDENCE USE TO OPERATIONALIZE TRA.

(Domains B, C, F, G): CGIAR could establish an integrated MELIA-F[10] and Finance mechanism to harmonize indicators, strengthen qualitative and GEI metrics, link MELIA with financial systems, and support systematic evidence use across the Portfolio. A unified system would demonstrate proportionality between investments and outcomes, track differentiated benefits, to help ensure consistent use of evidence in management decisions. Stronger MELIA and finance integration is a foundational step toward enabling future long-term cost-benefit and VfI analysis. Non-binding suggested measures include:

2.1 Develop and formalize clear guidance on how TRA 2025-30 reporting will interface with MELIA and financial systems to enable aligned cost/result pathways. The outline can consider: (1) minimum financial-performance linkage expectations; (2) reporting workflows between Programs, Accelerators, Finance, and MELIA focal points; and (3) guidance for integrating cost tagging, GEI tagging and performance evidence. The guidance should be co-developed by PPU and Integrated Partnership Business Operations and Finance Unit, in consultation with IAES and MELIA focal points.

- Timeline: Q1-Q4 2026
- Cost implication: Low to moderate; requires coordination and technical drafting.

2.2 Establish a MELIA-Finance Integration Taskforce to ensure PRMS, PORB and MELIA datasets speak to each other and to enable cost-performance visibility in CGIAR and integrate gender, equality, youth and social inclusion (GEYSI) tagging where feasible. The taskforce should develop shared definitions for outputs, outcomes, and cost lines; design dashboards linking performance indicators to expenditure categories; and pilot GEYSI-tagged resource tracking for at least two P/As. PPU and the Integrated Partnership Business Operations and Finance Unit could lead setting up the Taskforce and its workplan. It is important to note that the EA team is not proposing universal GEYSI tagging, nor is it implied that all expenditures should be classified through a GEYSI lens. Rather, the tagging is a selective learning-oriented mechanism to support evaluability where inclusion is material to an intervention's intent or risk profile. Thus, tagging refers to: (1) identifying when inclusion is an explicit objective, design feature, or risk consideration within an intervention; and (2) the designs ability to enable those activities or budget lines to be linked to relevant evidence streams (e.g., participation, influence, differentiated benefit).

- Timeline: Q1 2026–Q4 2027
- Cost implication: Moderate. Requires technical integration and data engineering; potential for earmarked donor support.

2.3 Require annual Learning and Decision Brief for each P/A for transparency on use of evidence in management decisions, Incorporating the brief into performance reviews would support institutionalizing evidence use as a leadership expectation. A 1–2-page format developed by management (office of Chief Scientist with Human Resources) can prompt reporting on: decisions taken; evidence sources used (PRMS, PORB[11], GEI disaggregation, and qualitative learning), changes adopted; and limitations.

- Timeline: Q2 2026

---

[9] Finalized by PCU at the time of finalizing this EA Synthesis report.
[10] Ongoing CGIAR-wide effort managed by PPU.
[11] Plan of Results and Budget

- Cost implication: Low. Fits within existing MELIA
  roles and management cycles.

2.4 Harmonize qualitative, behavioral, and GEYS indicators to better capture systems-change processes and differentiated benefits. Strengthen measurement of change processes central to systems transformation and inclusivity. Develop an indicator library including standardized behavioral indicators (both quantitative and qualitative changes); qualitative evidence protocols; and GEYSI indicators beyond sex disaggregation (intersectionality, participation influence, equitable benefit). Harmonization would support cross-portfolio synthesis and improve the evaluability of contribution pathways. PPU should provide leadership in this harmonization process, supporting MELIA focal points to engage with their colleagues in the programs on developing an initial list for cross-Portfolio comparison.

This suggested action must be a Portfolio-wide, multi-actor effort. Behavioral and qualitative indicators are not GEI-specific; rather they are central to understanding system change, contribution, and uptake across the P/As (e.g., capacity strengthening, policy, influence, digital transformation, scaling, and partnership change). GEYSI indicators constitute one critical subset of this broader indicator family. Their role is to ensure that behavioral and qualitative change processes are examined through an equity and inclusion lens-addressing who participates, who influences decisions, and who benefits from CGIAR's interventions. They are not intended to operate as a stand-alone indicator stream, nor to imply that responsibility for indicator articulation or implementation sits solely with the GEI Accelerator. The intent of Suggested Action 2.3 is therefore co-development: SPs and relevant Accelerators articulate behavioral and qualitative indicators aligned with their causal pathways, while the GEI Accelerator provides technical stewardship to help embed GEYSI dimensions consistently across these indicators. Portfolio-level leadership and coordination for this process sits with PPU, in collaboration with MELIA focal points and the GEI Accelerator.

- Timeline: Q2-Q3 2026
- Cost implication: Low

2.5 Develop a Minimum MELIA Standards Guide to provide consistent expectations for baselines, qualitative evidence, behavioral indicators, GEI integration, documentation, and timing. The guide could include minimum baseline requirements, expectations for GEI-disaggregated data, how to collect behavioral/qualitative evidence, and standards for linking MELIA cycles to financial reviews to enable consistency while preserving program flexibility. MELIA focal points or MEL CoP participants can brainstorm content in consultation with PPU and IAES, and develop a user-friendly document for Program use.

- Timeline: Q1-Q3 2026.
- Cost implication: Low. Primarily technical drafting and consultations.

2.6 Strengthen MELIA–Finance–Learning integration through Adaptive Management: Apply CGIAR's Portfolio Adaptive Management guidance to integrate adaptive cycles into MELIA–Finance–Learning processes, reinforcing coherence between financial, performance, and learning systems and supporting responsive, evidence-based management across the Portfolio.

- Timeline: Q1–Q4 2026.
- Cost implication: Low to moderate.

RECOMMENDATION 3: FORMALIZE EVALUATION READINESS AS A SYSTEM REQUIREMENT, NOT AN ASSUMED CONDITION.

(Domains D, E): CGIAR should embed anticipatory, context-responsive planning into annual cycles and MELIA architecture, requiring SPs, Accelerators and S4I to conduct structured environment scans, risk assessments, and long-term evidence planning, with minimum standards for baseline design, time-bound data continuity strategies, and explicit evaluability criteria informing scheduling and resource allocation. Systematizing readiness will reduce evaluation disruption and costs, improve credibility with evidence, and meet CGIAR's ambition for organizational and behavioral change. The non-binding suggestive measures include:

3.1 Introduce a CGIAR-wide Evaluability Readiness Checklist that P/As can use annually to clarify minimum evaluability conditions across domains C, D and E (Context, Long-Term Evidence, and Readiness). Content may include existence/quality of baselines, planned data access points, known political and environmental risks,

partner availability windows, long-term data continuity arrangements, capacity or infrastructure gaps, and alignment of MELIA timing with Program cycles. IAES should suggest a checklist format draft for adaptation by the Chief Scientist Office. The checklist should be socialized and used through the Adaptive Management dashboard to support accountability and Portfolio monitoring.

- Timeline: Q1-Q3 2026 (development and piloting during annual planning).

- Cost implication: Low; Checklists leverage existing MELIA cycles; minimal new funding.

3.2 Conduct semi-annual context and environment scans by domain, using the MR Tracker and other evidence sources, to anticipate contextual volatility and inform evaluation readiness. Scans should assess political and policy developments, climate and seasonal trends, regulatory and market shifts, partner stability, and other risks or opportunities affecting implementation and evaluability. PPU should lead the process, in collaboration with the Office of the Chief Scientist, Program teams, and country coordinators.

- Timeline: Semi-annually starting Q4 2026.

- Cost implication: Low–Moderate; primarily staff time with occasional consultant support.

3.3 Commission readiness simulations in two SPs and two Accelerators. A "stress-test" exercise will model evaluation feasibility, timing, risk and resource needs. Simulation components may include scenario-based access constraints, data availability checks, partner readiness assessment, timing conflicts, and resource bottlenecks. PPU to lead in the design and facilitation with IAES consultation and support and participation of Programs and Accelerators.

- Timeline: Q3 2026 – Q2 2027

- Cost implication: Low – process based; can be facilitated using existing staff

3.4 Develop long-term evidence continuity plans for each P/A to ensure that data needed to assess outcomes beyond 2030 do not disappear when funding cycles shift. Plans may include data repository commitments, agreements with centers or partners to maintain dataset, arrangements for post-project follow-up (e.g., tracer studies, lean longitudinal methods); and identification of critical indicators requiring long-term tracking. Action on this measure would be taken at the Program or Accelerator management level.

- Timeline: Q1 2026–Q4 2027.

- Cost implication: Moderate–some continuity arrangements may require small, dedicated funding; donor co-financing possible.

RECOMMENDATION 4: ESTABLISH PORTFOLIO-LEVEL EVALUABILITY GOVERNANCE TO ENSURE COHERENCE, ACCOUNTABILITY AND FORWARD PLANNING.

(Domains A-G): CGIAR should establish a system-wide evaluability integration mechanism to synchronize progress on evaluability reforms, maintain coherence across MELIA, Finance, Program design, and learning governance, and support collective decision-making on evaluability standards for the 2027-30 period. This mechanism is not a new structure but an organizing function that brings together existing capabilities–IAES, PPU, MELIA focal points, Finance, and P/A leadership–to ensure that evaluability remains visible, coordinated, and actionable at system scale. Because evaluability is a collective responsibility, not an attribute of individual interventions, an annual joint process is needed to monitor progress, recalibrate standards, and ensure evaluability maturity across the Portfolio. The non-binding suggested measures include:

4.1 Include a CGIAR Evaluability Review and Planning Session at the annual MELIA CoP to bring together IAES, PPU, MELIA focal points, Finance, and P/A leadership for shared review, alignment and forward-planning. PPU should plan the meeting in collaboration with other departments and chair the first meeting. The chair's responsibility may rotate to the other actors on an annual basis to foster joint accountability. Suggested agenda items may include: (1) a progress review of system-level actions under Recommendations 1-3; (2) re-calibration of minimum evaluability standards using a shared maturity matrix across domains A-G; (3) agreement on collective priorities for 2027-30; and (4) identification of gaps requiring guidance, capacity support, or policy adjustments.

- Timeline: Annual.
- Cost implications: Low to moderate, primarily for coordination and facilitation; venue or hybrid hosting costs; minimal new funding.

4.2 Produce an annual Evaluability State of the Portfolio Brief to synthesize evidence on evaluability progress, challenges, and emerging risks. IAES should lead this effort, based on a review of available data. Potential content may include progress against each major recommendation; comparative analytics (e.g., baselines completed, MELIA-Finance integration milestones, durability and GEI readiness); and emerging risks requiring system attention.

- Timeline: Starting Q1 2027, annually.
- Cost implications: Low, leveraging existing reporting cycles and MELIA data.

4.3 Update and issue Portfolio-Level Evaluability Standards (2027-30) to maintain consistency as Programs evolve, funding changes, and MELIA systems mature. Focus areas may include ToC quality and contribution logic, baseline and indicator requirements, qualitative/behavioral data expectations, minimum GEI evaluability indicators, and requirements for data/resource continuity.

- Timeline: Standards update in Q3 2027.
- Cost implications: Low, primarily staff time.

4.4 Coordinate Technical Guidance Rapid 'Bursts' twice annually to respond quickly to emerging evaluability needs, e.g., contribution analysis templates, GEI indicator harmonization, readiness checklist updates considering recommendation implementation status in the MR Tracker. Office of Chief Scientist in cooperation with the MELIA Community of Practice and centers, in consultation with IAES, can conduct semi-annual scans by domain (based on MR Tracker and other evidence).

- Timeline: Twice each year, 2026-30.
- Cost implication: Low, using virtual formats.

4.5 Facilitate cross-unit learning exchanges focused on evaluability to maintain methodological coherence while allowing for innovation and diversity across the P/As. IAES should lead, in cooperation with MELIA focal points from each of the P/As. These could include sessions on good practice in baseline design, behavioral/qualitative indicator development, tracer study protocols, cost-contribution mapping, or ensuring evaluability in politically constrained environments.

- Timeline: Quarterly from mid-2026.
- Cost implications: Low.

4.6 Stock take of implementing EA recommendations to inform design of the 2027 midline evaluations (MLE) for clear evaluability priorities. PPU and IAES should co-lead on this effort.

- Timeline: Q3-Q4 2027.
- Cost: Moderate, depending on depth; could be integrated into existing midline review plans, adoption and scaling trajectories after project closure. These actions represent targeted strategic investments that will strengthen institutional memory and reduce future evaluation costs.

Collectively, these four recommendations frame a feasible, sequenced plan for improving evaluability over 2026-27 and onwards, anchored in existing capacities, aligned with institutional reforming, MELIA-F projects, and designed to enhance the credibility, usefulness, and efficiency of CGIAR's evidence ecosystem.

6 Comparison with RII EA Synthesis

A comparison between 2025 EAs of the 13 P/As and the 2023 Synthesis EA of 4 RIIs)-while not an exact comparison–does show that while many evaluability challenges persist, there are signs of iorganizational maturation. The current Program-specific anomalies but system characteristics that have persisted across design cycles and funding arrangements. The recurrence is also reflective of a research culture where evaluation is viewed as an 'add-on' and not part of core business. Whereas the RII synthesis highlighted decentralization, incomplete MEL systems, and unclear evaluand boundaries, the Portfolio Synthesis extends these concerns to the level of institutional design and governance. With 13 interventions assessed using a common evaluability

framework, fragmentation between program design, evidence systems, and financial architecture becomes visible as a system property rather than a by-product of individual Initiatives.

Across both EA syntheses, the main conclusion is constant: evaluability is constrained less by technical gaps and more by the underlying operating logic and culture of the system. This underscores the importance of the Portfolio-level recommendations in this report, not as discrete improvements for individual interventions, but as the institutional reforms needed to address long-standing, cross-Portfolio, evaluability limitations.

7 Conclusion

Across the 13 EAs, a clear picture emerges: **CGIAR's Portfolio is conceptually strong, strategically aligned in its orientation toward achieving CGIAR's transformative goals, and institutionally committed to learning — yet evaluability remains uneven because the connective systems that translate evidence into decision and value are still developing.**

The seven cross-cutting themes (figure 1) reveal that evaluability is not a technical feature of P/A design but an organizational property of how CGIAR governs, learns, and allocates resources. Patterns of ambition without proportional operational clarity (Theme 1) are mirrored in fragmented MELIA architectures (Theme 2) and symbolic rather than evidenced approaches to inclusion (Theme 3). Short-term funding horizons (Theme 4) and *ad-hoc* readiness planning (Theme 5) limit the ability to sustain and verify results, while learning processes (Theme 6) and financial systems (Theme 7) still operate largely in parallel rather than as part of a unified evaluative ecosystem.

Conclusion from findings shows that he challenge for CGIAR is less about individual P/A improvement and more about culture and system integration—connecting the organization's evaluative, managerial, and financial functions around shared definitions of value, evidence, and accountability. The recommendations that follow therefore emphasize institution-building over compliance: strengthening causal reasoning, harmonizing MELIA and finance systems, operationalizing inclusion, and embedding feedback loops that make evaluability a normal way of working. In responding to these recommendations, it is likely that Portfolio evaluability will not only be strengthened, but it will also contribute to improved Portfolio cohesion which was a critical issue highlighted by the ISDC Review of the 2025 Research and Innovation Inception Reports.

In this sense, the synthesis points to a transition: from viewing evaluability as an assessment of readiness to understanding it as a governing logic—one that links science quality, learning, and performance management in service of a credible, adaptive, and equitable CGIAR.

7.1 Implications for 2027 Midline Evaluations by IAES

The EA synthesis informs how IAES can scope, sequence, and design the 2027 MLEs. The MYEP budget will require strategic prioritization, as not all P/As can or should undergo full independent evaluations. The synthesis suggests **a clustered, question-driven evaluation approach,** to conduct a small number of thematically coherent, cross-program evaluations, complemented by lighter evaluability checks or evidence-readiness reviews for the remaining interventions. A clustered approach offers several advantages:

- **Efficiency:** Pooling data collection, travel, and analysis across related interventions.

- **Coherence:** Allowing MLEs to address system-level issues identified in the Synthesis rather than producing fragmented Program-specific reports.

- **Strategic learning:** Focusing evaluative effort where it will unlock the greatest value for Portfolio-level adaptation in 2027-30.

- **Methodological appropriateness:** Acknowledging that interventions differ in maturity, purpose, and evaluability.

The table below proposes a Portfolio-aligned clustering, each with a suitable evaluation approach and illustrative MLE questions that IAES may refine in consultation with PPU, MEL CoP and MELIA focal points, P/A leadership, and centers.

Table 5. Proposed clusters and priority lines of inquiry for the 2027 MLEs

Cluster	Included interventions	Rationale	Suggested evaluation approach	Illustrative MLE questions
Genetic Innovation and Breeding Systems	B4T, SAAF, Genebanks	Shared biological pathways, long time horizons, strong legacy investments, complexity in causal chains.	Theory-based evaluation with contribution analysis and evaluability review of biological evidence systems.	How plausible and feasible are the contribution pathways linking genetic innovation to adoption and system outcomes? How are biological, temporal and regulatory lags being managed, monitored, and communicated? To what extent are MELIA systems capturing quality, performance, and VfI across breeding and genebank functions?
Climate, Nutrition and Production and Environmental Systems	Comparative Advantage, SFP, MFL, BDN, Food Frontiers	Large SPs with diverse field realities; share context dependent outcomes, long-term trajectories, and complex multi-actor settings.	Mixed-method formative evaluation with attention to context-responsiveness and early signals of behavioral and organizational change.	Are Programs demonstrating progress toward behavioral or organizational change (vs. only technical outputs)? Are MELIA systems adequately capturing differentiated effects (GEI, geography, partnerships)? How effectively are Programs adapting to contextual shocks and political cycles?
Equity, Gender, Policy and Systems Enablers	GEI-A, PI	Both seek organizational and behavioral change; evaluability hinges on qualitative evidence, GEI metrics, legitimacy, and influence pathways.	Contribution-focused evaluation emphasizing qualitative evidence, organizational influence, and GEI outcomes.	How credible is the evidence linking enabling work (gender, equity, policy brokering) to observed partner behavior change? Are GEI and policy influence outcomes sufficiently resourced, measured, and integrated in Program decision-making? What early signs exist of institutional uptake or shifts in norms/decision processes?
Digital and Capacity Enablers	DTA, CapSha	Both support system-wide capacities; causal pathways operate through partner adoption and improved capability.	Developmental and utilization-focused evaluation, assessing learning loops, partner responsiveness, and adaptation.	How are digital and capacity interventions influencing Program and partner capability and practice? Are learning loops functioning (Are insights acted upon? At what speed?) What elements of the enabling functions appear most (or least) catalytic?
Scaling and Innovation Pathways	S4I (treated as its own cluster)	S4I is structurally distinct: system-level scaling mechanism; works through SPs; requires developmental evaluation focus.	Developmental Evaluation (DE) oriented toward learning, adaptation, and real-time usefulness.	How effectively is S4I brokering pathways from research to use? What adaptations were made based on evidence generated in 2025-27? How is S4I shaping SPs' readiness for scaling, demand alignment, and partnership patterns?
Cross-Portfolio Evaluability and System Learning (Meta-evaluation)	All P/As	Addresses system-level patterns identified in the synthesis.	Meta-evaluation and synthesis	How have evaluability conditions changed across domains A-G since 2025? What is the maturity level of MELIA-Finance integration, GEI indicator harmonization, and learning governance? What cross-Portfolio gaps require corrective action before 2030?

The clustering approach enables IAES to adapt evaluation design to the maturity, purpose, and evaluability status of different interventions, ensuring that limited resources are deployed where they produce the greatest strategic value. It also ensures that the MLEs do not become 13 disconnected exercises but instead contribute to a coherent Portfolio narrative about progress, challenges, and readiness for the 2030 horizon.

The proposed questions intentionally align with the seven system-level themes from the synthesis–for MLE insights to directly address the structural factors shaping evaluability across CGIAR. This approach will allow **the 2027 MLEs to support:** (1) strategic mid-course correction where it is most needed; (2) evidence-informed design adjustments ahead of the 2027-30 latter half of Portfolio implementation; (3) improved comparability and coherence across the interventions; and a (4) stronger foundation for the eventual end-of-Portfolio evaluation.