WORKSHOP REPORT

Accelerating Food-Land-Water Systems Research in CGIAR through Responsible AI Integration

Colombo, Sri Lanka | September 25-27, 2025

Mariangel Garcia Andarcia, Surajit Ghosh, Carolina Iglésias Martins, José Berenguer, Paulo Silva, and Daniel Jimenez

December 2025

Authors

Mariangel Garcia Andarcia, Research Group Leader - Water Futures Data and Analytics, International Water Management Institute (IWMI), Colombo, Sri Lanka
Surajit Ghosh, Regional Researcher - Integrated Water Data Science, IWMI, Colombo, Sri Lanka
Carolina Iglésias Martins, Consultant, IWMI, Colombo, Sri Lanka
José Berenguer, Consultant, CGIAR System Organization, Madrid, Spain
Paulo Silva, Consultant, IWMI, Colombo, Sri Lanka
Daniel Jimenez, Senior Scientist, Alliance Bioversity-CIAT, Montpellier, France

Acknowledgments

This work was conducted as part of the CGIAR Accelerator for Digital Transformation. We would like to thank all funders who supported this research through their contributions to the CGIAR Trust Fund (www.cgiar.org/funders). We sincerely acknowledge all CGIAR Centers who represented the workshop and the CGIAR System office for their support.

CGIAR Accelerator for Digital Transformation

The CGIAR Accelerator for Digital Transformation aims to be a thought leader in digital and data science research across food, land, water, and climate systems, accelerating the delivery of research and digital services responsibly using data and tools such as AI, and improving decision-making based on evidence and insights into complex systems.

Citation

Garcia Andarcia, M.; Ghosh, S.; Martins, C. I.; Berenguer, J.; Silva, P.; Jimenez, D. 2025. *Accelerating food-land-water systems research in CGIAR through responsible AI integration.* Report of the Workshop on Accelerating Food-Land-Water Systems Research in CGIAR through Responsible AI Integration, Colombo, Sri Lanka, 25-27 September 2025. Colombo, Sri Lanka: International Water Management Institute (IWMI). CGIAR Accelerator for Digital Transformation. 40p.

Front and back cover photo: IWMI

Disclaimer

Contents

Executive Summary

The Workshop, titled “Accelerating Food–Land–Water Systems Research in CGIAR through Responsible AI Integration,” was held from 25 to 27 September 2025 at IWMI, under the CGIAR Digital Innovation Initiative. Organized by the IWMI Digital Transformation Accelerator (DTA), the workshop brought together over 26 participants from 10 CGIAR centers and partner organizations to co-develop a shared vision for responsible, human-centered, and interoperable artificial intelligence (AI) in research. The event aimed to advance CGIAR’s broader digital transformation agenda by defining how AI can act as a collaborative partner in scientific discovery by enhancing data-driven insights, analytical capacity, and cross-center integration.

The purpose of the workshop was fourfold: (i) to assess CGIAR’s AI maturity and existing digital assets, (ii) to co-create a vision and operational roadmap for AI adoption, (iii) to define organizational Objectives and Key Results (OKRs) for responsible AI integration, and (iv) to develop conceptual prototypes showing the AI Co-Scientist vision in action. Across three intensive days, the sessions combined diagnostics, design-thinking approaches, and collaborative prototyping to identify pathways for embedding AI ethically and effectively within CGIAR’s Food–Land–Water (FLW) research systems.

On Day 1, participants conducted a comprehensive mapping exercise of the AI ecosystem, identifying over 40 active AI-related tools, models, and datasets across CGIAR. This diagnostic provided the first consolidated baseline of AI maturity and readiness across centers. Discussions revealed significant opportunities for applying AI in data harmonization, predictive analytics, and decision support, while highlighting ongoing challenges such as fragmented data management, uneven governance structures, and limited institutional interoperability. Day 2 emphasized human-centered and ethical AI design. Using empathy mapping and scenario development, participants explored how AI can best support researchers while maintaining transparency, inclusivity, and equity. The conversations reinforced the importance of aligning technological progress with ethical safeguards, ensuring that AI tools amplify human creativity and serve the needs of diverse users, from scientists to smallholder communities. Participants also reflected CGIAR’s institutional readiness, noting that while the emerging CGIAR 365 Data Lake is a strong foundation, it must evolve toward a federated, AI-ready architecture with standardized metadata and shared governance protocols. The focus of Day 3 was on integration and prototyping. Multidisciplinary teams developed five prototype concepts that illustrated AI’s potential to enhance research and operations. These included applications in donor matchmaking, geospatial intelligence, automated reporting, and conversational data analysis through natural language interfaces. The discussions also explored Model Context Protocols (MCPs) and federated architectures as mechanisms for connecting AI systems and datasets across centers, promoting efficiency and collaboration.

By the end of the workshop, participants had co-created a shared CGIAR-wide vision and set of OKRs, established the groundwork for an AI asset registry, and identified key priorities for governance, capacity development, and infrastructure enhancement. The outcomes align directly with CGIAR’s Digital Transformation and Innovation Strategy, reinforcing the system’s commitment to responsible AI that is transparent, interoperable, and impact oriented.

The workshop concluded with a collective call to action to move from vision to implementation by operationalizing the AI Co-Scientist framework through sustained collaboration, ethical design principles, and open innovation. Participants agreed that harnessing AI responsibly across CGIAR’s Food–Land–Water research domains can accelerate progress toward sustainable, equitable, and climate-resilient food futures for all.

Introduction

Agricultural and environmental systems today face mounting and interconnected challenges driven by climate change, resource scarcity, and population growth. Rising global temperatures, shifting rainfall patterns, and increasing frequency of extreme weather events threaten both productivity and livelihoods, especially in climate-vulnerable regions. At the same time, water scarcity, soil degradation, and biodiversity loss are accelerating, while global food demand is projected to rise by over 50 percent by 2050. These trends place enormous pressure on natural resources and demand an urgent transformation in how food, land, and water (FLW) systems are managed.

Traditional research and policy approaches—while rich in scientific rigor—are increasingly constrained by the scale, speed, and complexity of these interrelated crises. Decisions that once relied on seasonal or local data must now consider global interdependencies and nonlinear feedback between agriculture, hydrology, and climate systems. This calls for a new generation of tools and methods capable of integrating massive, multi-source datasets and translating them into actionable insights. In this context, Artificial Intelligence (AI) emerges as a powerful enabler of transformation, offering new ways to understand, predict, and manage change across the FLW continuum.

“Given CGIAR’s mission and the increasing role of AI in agricultural innovation, these advancements raise important strategic questions for us.”

"How can we leverage AI to accelerate breeding programs, optimize decision-making in soil health and pest control, or improve water management?"

Dr. Ismahane Elouafi

AI holds transformative potential across every stage of agricultural and environmental research. In data-rich ecosystems, AI can automate data collection, harmonization, and pattern recognition from diverse Earth Observation (EO), sensor, and socio-economic datasets. In data-scarce contexts, it can extrapolate insights from partial information, helping identify vulnerabilities and anticipate risks in near real time. For instance, machine learning (ML) models can improve crop yield prediction, optimize irrigation scheduling, assess land degradation, or detect flood and drought patterns long before they become visible to the human eye. Natural language processing (NLP) and knowledge graph models can help synthesize scientific literature and grey data, accelerating evidence-based policy and innovation.

More importantly, the rise of generative AI and agentic systems marks a shift from automation to collaboration between humans and machines. AI no longer merely processes data; it now participates in scientific reasoning, hypothesis generation, and decision support. The emerging concept of an “AI Co-Scientist”, a system capable of assisting researchers in designing experiments, analyzing results, and translating insights into solutions, captures this evolution. When embedded within CGIAR’s mission, such systems can enhance research productivity, foster innovation across disciplines, and democratize access to advanced analytical capabilities among scientists and partners.

Recognizing the urgency of system-level digital integration, CGIAR established the Digital Transformation Accelerator (DTA) as a cross-cutting initiative to leverage digital innovations, including AI, big data, and cloud computing, for scientific and operational excellence. The DTA aims to build a cohesive digital ecosystem across CGIAR centers, enabling shared infrastructures, interoperable data systems, and responsible innovation frameworks.

Within this agenda, the workshop on Accelerating Food–Land–Water Systems Research through Responsible AI Integration served as a milestone event to advance the vision of responsible AI adoption in CGIAR’s FLW research systems. The workshop directly supported DTA’s goals by convening experts, researchers, and digital practitioners to collectively assess current AI capacities, identify gaps, and co-design pathways toward ethical, transparent, and high-impact AI applications. It also provided an opportunity to align AI research priorities with CGIAR’s institutional strategy on digital transformation and its commitment to responsible, human-centered innovation.

As part of the CGIAR Digital Transformation Accelerator (DTA), this initiative aims to assess the real opportunity for AI applications within CGIAR’s research ecosystem by engaging with Science Programs (SPs), accelerators, and research teams. The findings will guide the development of AI-driven methodologies, governance frameworks, and training programs to accelerate scientific discovery in food, land, and water (FLW) systems.

The initiative will approach three key areas (Figure 1):

1. Needs and recommendations for CGIAR to adopt a responsible AI-forward approach in pursuit of its mission.
2. The current scope and potential of AI being used to accelerate CGIAR research (AI Co-Scientist).
3. The current scope and potential of AI applications for farmers and other food, land and water (FLW) system actors.

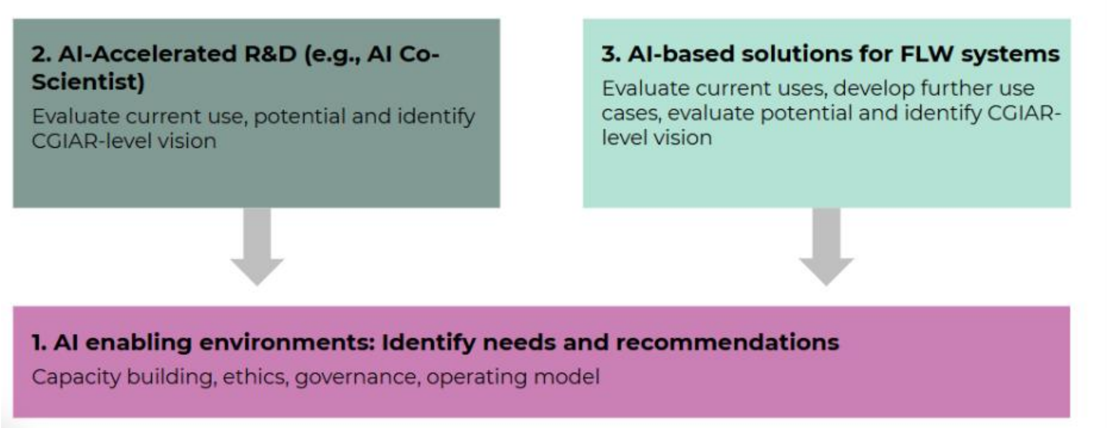


Figure 1*. The relation between the proposed activities as part of the initiative*

Workshop Structure

Part 1: Mapping the Present

The process began with an assessment of the current state ("As-Is"), carried out through center-specific breakout groups when possible. These groups applied a partial ORID framework (Objective, Reflective, Interpretive, Decisional) to explore how each center is engaging with AI across the four research stages: prospecting, planning, executing, and closing. The aim was to surface what is currently being done, how it feels, and participants' views on AI's role at each stage. This exercise generated not only insight into technical and institutional capacities but also revealed emotional and cognitive barriers that may hinder broader or more effective AI adoption. These findings laid the groundwork for Part 2 (i.e., "To-Be"), where participants envisioned a future in which AI is more intentionally embedded across CGIAR's research workflows.

Part 2: Envisioning the Future

Building on this, participants reassembled into five thematic breakout groups, each focusing on a specific pillar of responsible AI integration, such as Strategic Intent, Organizational Readiness, Governance and Ethics, Evaluation and Monitoring, and Adoption and Use. These groups carried forward insights from the ORID sessions into a series of future-oriented activities (as depicted on Figure 2):

- A modified 2x2 matrix regarding Four Strategic Patterns for AI Transformation (Wolfe, Choe & Kidd, 2025) was used to map emerging visions for AI integration, asking participants to consider whether proposed changes would be incremental or transformational, and whether they would lead to increased or decreased reliance on human agency. This exercise was intentionally designed to highlight cognitive dissonance and prompt reflection on trade-offs, assumptions and tensions across the system.
- A hands-on session through LEGO® SERIOUS PLAY® methodology allowed participants to externalize and visualize abstract concepts through physical modelling. This hands-on process facilitated cross-disciplinary dialogue, making complexity more tangible and inclusive.
- To translate the collective vision into measurable outcomes, the final section introduced the Objectives and Key Results (OKRs) framework1 as a practical tool for defining institutional priorities, tracking progress, and promoting shared accountability, and as an alternative to traditional performance metrics.

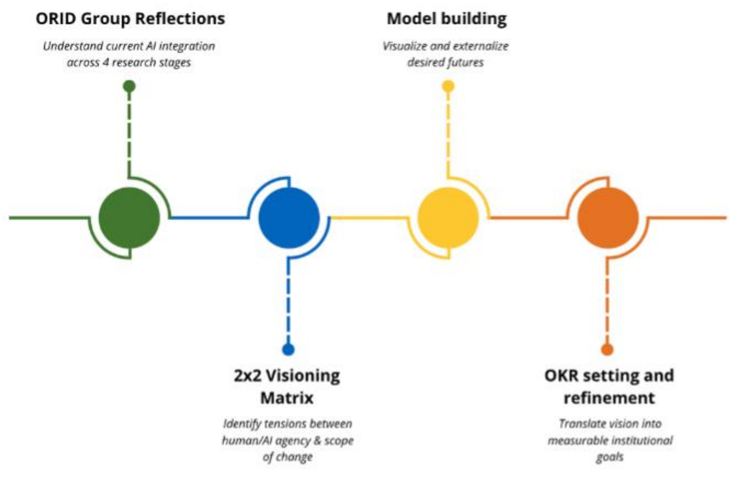


Figure 2. Workshop exercises for parts 1 and 2. (*Source:* Author's creation)

Part 3: Planning the Journey

Under the theme "Planning the Journey," Day 3 centered on the practical realization of CGIAR's AI roadmap and how to operationalize responsible, human-centric AI across FLW systems research through design, prototyping, and institutional alignment. The session moved participants from ideas to concrete AI prototype concepts through collaborative design. Multidisciplinary teams developed use cases aligned with CGIAR's FLW mission, ethical and human-centred AI, and cross-center scalability. Using templates and design-thinking tools, they framed problems, user needs, and outcomes, then created mock-ups and workflows. Five key concepts emerged: "Ask the Data" conversational data access, a GeoAI Advisor for flood and drought monitoring, AI-driven donor matchmaking, an automated research reporting assistant, and an ethical AI governance dashboard.

A technical segment introduced MCPs and federated AI architectures to link diverse models and datasets while preserving data sovereignty, positioning CGIAR 365 as an AI-ready data lake. In the afternoon "demo day," groups dramatized their solutions and received structured feedback on relevance, scalability, interoperability, and ethics. Discussions focused on simple, user-centred, open-source designs and proposed advancing selected prototypes into pilots through the Digital Transformation Accelerator.

Part 1 - Mapping the Present

Centers' Efforts on AI

The first major session invited five-minute presentations from CGIAR centers, each summarizing their ongoing AI-related work. These concise pitches revealed the breadth of current AI use cases across the system. Examples included machine learning applications for crop yield prediction, EO-based flood and drought monitoring, automated land-use classification, and predictive modelling for pest and disease surveillance. Participants also presented AI-driven initiatives in genomic selection, agricultural value chain analytics, and natural language processing for literature synthesis.

This mapping exercise demonstrated that CGIAR collectively possesses a diverse and rapidly expanding portfolio of AI-related assets (Table 1). However, it also underscored the need for coordination and interoperability. Many centers reported that data fragmentation, limited computational infrastructure, and inconsistent metadata standards remain significant challenges. Several participants noted that while AI pilots exist within individual projects, there is still no unified framework for scaling, validating, or ethically governing these tools across the One CGIAR system.

Table 1. Summary uses cases presented by the Centers.

Center	Key AI Techniques Used	Main Application Domains	Regions of Work
Alliance of Bioversity & CIAT	Computer Vision (YOLO), Explainable ML, LLMs	Pest & disease detection, digital phenotyping, policy synthesis	Latin America, Africa
CIP (International Potato Center)	CNNs, LLMs (Amazon Bedrock), Bayesian ML, XGBoost	Genebank management, biodiversity monitoring, socio-environmental modeling	Andes–Amazon, Peru
IFPRI	LLMs (RAG agents), ML for agricultural stats, web-scraping	Food policy, market monitoring, advisory research	Global
ICRISAT	Computer Vision, ML (Soil health mapping), LLMs (Knowledge extraction)	Carbon credit MRV, genebank data exploration, digital agro-advisory	India
IRRI	Deep Learning, Predictive ML, Drone CV, Breeding Data Pipelines	Predictive breeding, phenotyping, remote sensing	Asia (Philippines, Vietnam, India)
ILRI	ML, LLMs (RAG), ABM (HEURISTICS)	Rangeland mapping, livestock health, socio-economic modeling	East Africa
IITA	Hybrid ML + LLM, CV, Gene-editing AI	Agronomy chatbots, breeding research, knowledge systems	West & Central Africa
WorldFish	CV, ML, LLMs	Fisheries monitoring, advisory, sustainability modeling	Africa, Asia
IWMI	GeoAI, CNN, LLMs, Computer Vision, Digital Twins	Hydrology, drought/flood forecasting, water advisory	South Asia, Africa
ICARDA	ML, RS, CV, Cloud analytics	Breeding data pipelines, pest management, climate adaptation	CWANA, MENA

Next, the groups start exploring how AI is currently used across the CGIAR research cycle. Drawing on both pre-workshop survey responses and group discussions, this exercise surfaced with a consistent pattern: while interest in AI is high, its current use is fragmented, exploratory, and largely unsupported by institutional structures. Importantly, the same cognitive and emotional pressures that shape research workflows, such as competitive funding dynamics or siloed expertise, also shape how researchers engage with AI tools. Table 1 summarizes how AI is being used across each research stage and the associated barriers. Aisha Walcott-Bryant, head of Google Research Africa, delivered the keynote address, offering a compelling overview of Google's work in applying AI to address global challenges, with particular relevance to agri-food systems in the Global South. She outlined the strategic focus of her team within Google Research, which targets high-impact, scalable AI solutions in areas such as climate resilience, disaster response, and food security. Their approach centers on identifying problems where Google's technical capabilities, particularly in AI, computing power, and global infrastructure, can deliver measurable social benefits (Table 2).

Table 2. Insights per stage.

Research Stage	Observed use of AI	Barriers/Gaps
Prospecting[2]	Exploratory use for literature review, trend scanning, ideation (e.g., ChatGPT for brainstorming or scanning outputs). AI use is task-specific but not systematic.	Lack of awareness of others' work; low trust in AI-generated outputs; feelings of isolation and fear of missing out; limited strategic coordination.
Planning[3]	Some integration into protocol drafting, identifying research gaps, or team scoping via agile tools. AI used to map expertise, synthesize materials, and support early-stage design thinking.	Siloed planning practices; uneven adoption across centers; limited support or buy-in; cognitive overload and uncertainty around tools and data quality.
Executing[4]	Hands-on AI use for data standardization, NLP/ASR pipelines, or visualization; coexists with manual workflows.	Lack of internal validation processes; stress, time pressure, and concerns about accuracy and accountability; absence of clear pipelines and safeguards.
Closing[5]	Early stage use for reporting, presentations, or manuscript support. Some refine AI outputs; others express low trust.	Low confidence in AI quality; attribution concerns; fears of job displacement; lack of guidance on interpretation and validation.

Key Challenges Identified

The discussion highlighted that the primary barriers to impact are not purely technical but are related to data foundations, real-world implementation, and ensuring model reliability and validation. Figure 3 compiles the main patterns and insights discussed during the session.

- **Exploratory and Fragmented Use**: AI use across CGIAR's research stages remains largely exploratory and fragmented. It is task-specific, concentrated in early phases like prospecting and planning, and typically driven by individuals rather than embedded in institutional workflows. This signals a gap between experimentation and strategic integration.
- **Lack of Institutional Support and Safeguards**: Participants highlighted the absence of formal policies, approval processes, and institutional safeguards for AI use. Particularly in the execution and closing stages, concerns emerged around reliance on third-party tools without adequate vetting—raising red flags about privacy, intellectual property, and accountability.
- **Emotional and Cognitive Strain**: Emotional states, ranging from curiosity and optimism to anxiety and frustration, strongly shaped participants' engagement with AI. Pressure to keep up with innovation, combined with limited internal support, left many researchers feeling uncertain about how to use AI tools responsibly or effectively.
- **Uneven Adoption Across Centers**: The use of AI varied significantly across CGIAR centers. While some teams reported active testing of AI tools, others were unaware of what was available or lacked the capacity to apply them. This inconsistency reflects deeper challenges in institutional learning and shared infrastructure.
- **Low Trust in AI Outputs**: Doubts around the quality, contextual fit, and reliability of AI-generated outputs emerged across multiple stages. These concerns were especially strong in communication and reporting phases, where researchers felt unsure how much to trust, interpret, or validate AI-assisted content.
- **Unclear Standards and Accountability**: A recurring theme was the absence of shared language, validation mechanisms, or agreed criteria to guide AI usage. Participants raised important questions about attribution, transparency, and the broader implications of AI for early-career researchers, including fears around job displacement.

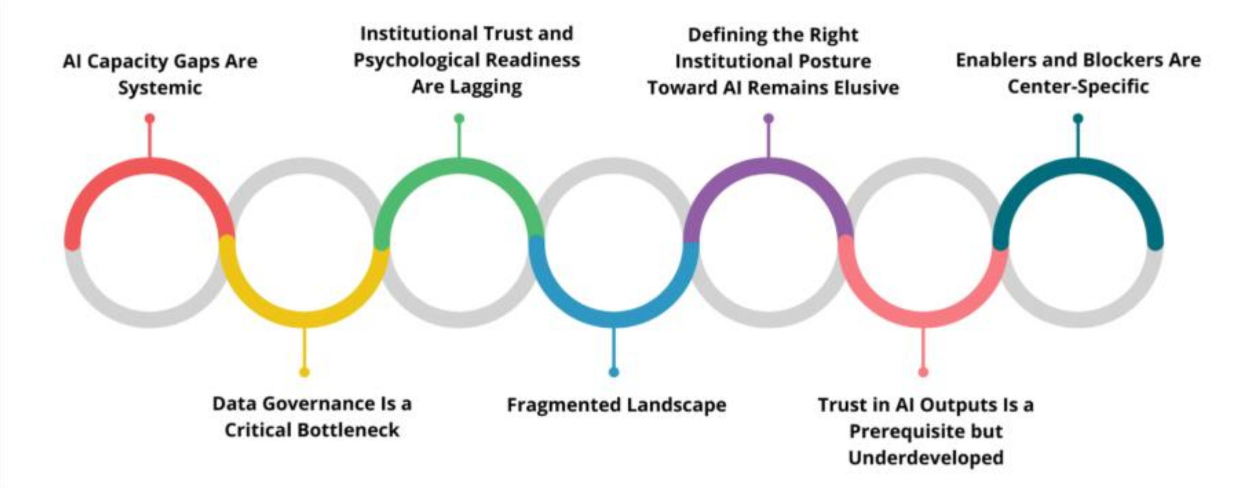


Figure 3. Compilation of major findings from the workshop and pre-workshop survey. (*Source:* Author's creation)

Potential Opportunities & Synergies

There was strong, immediate interest in leveraging the group's collective assets to produce joint outputs. Participants consistently highlighted the need to move from scattered, individual experimentation to more deliberate, cross-center learning. This included interest in piloting use cases jointly, pooling feedback on tools, and accelerating collective capacity. Many emphasized the potential to co-develop AI resources, such as shared datasets, toolkits, or practice notes, that reflect real-world priorities rather than abstract guidelines. Participants also noted the value of aligning efforts based on complementary strengths across CGIAR centers (for example, pairing expertise in data analysis with participatory research approaches). As showed in Table 3, one recurring theme was the need for shared infrastructure and documentation to reduce duplication and increase adoption. Lightweight but centralized mechanisms were proposed to provide practical support.

Figure 4. Participants mapping CGIAR's AI landscape *(photo*: IWMI)

Table 3. Detailed summary of findings on AI usage across R&D stages

<table>
<tr><th></th><th>Doing</th><th>Feeling</th><th>Thinking (direct quotes)</th></tr>
<tr><td>Overall Findings</td><td>AI is being used but not in consistent or coordinated ways. Usage is often task-specific, driven by individuals, and introduced informally into workflows rather than embedded into team practices or institutional standards.</td><td>Researchers expressed readiness to engage with AI, mixed with concern and frustration. Many participants want to push boundaries but feel unsure of the rules. This reveals a skill–confidence–governance gap.</td><td>Researchers are asking critical, thoughtful questions but usually in isolation. Their reflections point to a lack of collective strategy, shared language, or clear points of reference to guide their decisions.</td></tr>
<tr><td>Prospecting</td><td>Most still rely on manual workflows to identify funding calls, conduct landscape reviews or generate concept notes. AI is use is exploratory but not systematics (e.g., using ChatGPT to brainstorm or scan research outputs).</td><td rowspan="4">Excitement and optimism were widespread, especially around the idea that AI could speed up repetitive tasks, accelerate literature reviews, or enable more efficient planning.

Several participants said they were “fired up” or eager to finally test their tools or ideas if they had support and buy-in.

At the same time, people repeatedly expressed stress, time pressure, and cognitive overload. They were often unsure if they were using AI well or if they had the right data or skills to validate results.

There’s also a hidden fear of getting it wrong: being responsible for bugs, missteps, or missed opportunities because the tools or support aren’t there yet. Others feel alone in decision-making, unsure who is responsible for guiding use or interpreting results.</td><td>“Are we missing opportunities?”
“Are we competitive?”
“What’s everyone else doing?”</td></tr>
<tr><td>Planning</td><td>A few groups are starting to incorporate AI into protocol drafting, trend monitoring, or identifying research gaps. Agile tools like ClickUp or JIRA are being used in tandem with AI for initial design scoping. Still, much planning remains manual and siloed.</td><td>“Do we have enough data?”
“What are the right methods for this idea?”
“Should we build a GenAI chatbot, and how?”</td></tr>
<tr><td>Executing</td><td>More hands-on AI usage: for data standardization, generating code, building plots or running NLP/ASR pipelines. However, practices coexist with largely manual data collection and cleaning. Workflows are being stitched together without clear pipelines or validation checkpoints.</td><td>“Are these results valid?”
“How do I validate synthetic datasets?”
“Will this generate more bugs if I rely on AI?”</td></tr>
<tr><td>Closing</td><td>AI is used to support reporting and assist with presentations and donor outputs. Some are monitoring and refining outputs continuously, while others feel unsure how much to trust what AI generates.</td><td>“How can we ensure inclusive impact?”
“Can we trust what AI gives us?”
“What happens to jobs if this scales?”</td></tr>
</table>

Part 2 - Envisioning the Future

Participants identified that enabling responsible AI within CGIAR requires more than technical improvements. It depends on strengthening collaboration practices, establishing institutional safeguards, and embedding ethical and operational standards throughout the research cycle.

Where Researchers Want AI to Focus

Exercises during this part surfaced shared ambitions for what CGIAR should have in place within 3–5 years across all 5 main pillars: Strategic Intent, Organizational Readiness, Learning and Monitoring, Adoption and Use, and

Governance and Ethics. Based on interpretative analysis, these are the patterns found during the workshop, from most widely identified priorities to more emerging or niche priorities:

- **Cross-Center Collaboration and Knowledge Sharing**: There is a strong desire to move beyond isolated experimentation and toward system-wide collaboration. Participants envision a future where CGIAR uses AI to map internal expertise, identify research overlaps, and foster tool-sharing across centers. Large Language Models (LLMs) were seen as enablers to answer questions like "who has worked on this?" or "who has run a similar experiment?"
- **Deliberate Alignment on AI Strategy:** One of the main takeaways is that alignment around AI strategy must be built deliberately. Strategic readiness for AI is not only technical or procedural but also cognitive and cultural.
- **Responsible AI Governance and Use:** There was a recurring call for CGIAR to define where AI should not be used, adopt system-wide guidelines for responsible use, and establish accountability for ethical breaches. Participants emphasized the need for clear guidance on liability, risk, attribution, and quality assurance mechanisms (e.g., using AI for manuscript checking).
- **Empowering Human Talent and Safeguarding Learning:** Participants highlighted that CGIAR must invest as much in people as in tools. This includes protecting space for critical thinking (especially for early-career researchers) and avoiding over-reliance on AI that risks eroding human judgment. At the same time, building talent pipelines and capacity for ethical, context-aware AI use is essential, pairing adoption with training and domain expertise, so AI complements human insight.
- **Human-Centered Automation:** Many participants noted the opportunity to reduce administrative burden through automation. Tasks like filling out forms or handling routine reports should be offloaded to AI so researchers can focus on value-added work. This also includes AI's potential to support institutional learning at a scale by automating onboarding, capturing institutional knowledge, and generating learning content based on user feedback and organizational insights. However, care must be taken to ensure this does not lead to over-reliance or reduced human judgment.
- **Monitoring, Learning, and Feedforward**: Suggestions included refining Machine Learning Operations (MLOps) pipelines, embedding continuous feedback from users, and integrating AI into Results Framework Assessments. Participants emphasized the need to monitor outputs, alongside Responsible AI.
- **Legal, Cybersecurity, and Infrastructure Readiness:** Beyond research use, participants emphasized the need for strong foundations, including legal regulation, cybersecurity protocols, shared intranet systems, and auto-updated repositories to safeguard and standardize AI implementation across CGIAR.
- **Accelerating Discovery of New Research Frontiers:** Finally, participants noted the promise of AI in identifying novel research opportunities and gaps, surfacing emerging questions that cut across disciplines, and expanding the boundaries of CGIAR's research agenda.

To visualize these priorities, the figures below (Figure 5 and Figure 6) present a consolidated view of the insights shared across the five AI readiness pillars, highlighting the patterns that consistently surfaced during the workshop discussions.

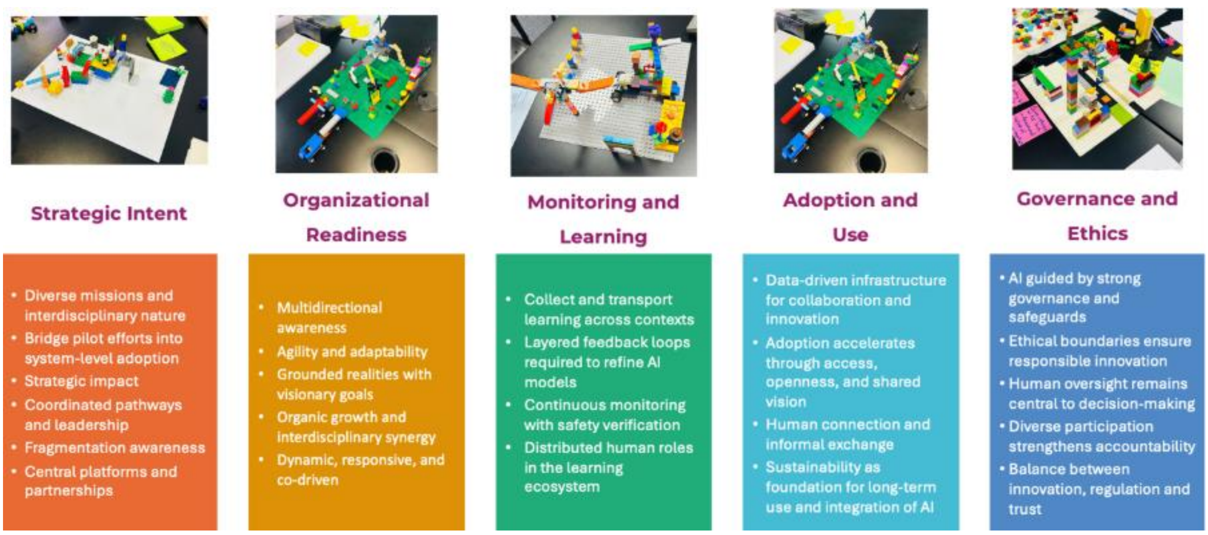


Figure 5. Workshop insights Across the Five Pillars. (*Source:* Author's creation)

Across all groups, participants prioritized system-level enablers that fostering cross-center collaboration, reducing cognitive and administrative burden, and embedding responsible use principles.

Figure 6. Lego model for Evaluation and Learning (*photo:* IWMI)

Table 3. Summarization of required components for R&D at CGIAR.

Pillar	Streamlining Research Workflows	Connecting People and Knowledge	Discovery and Insight	Automation of Admin or Routine Tasks	Talent and Capacity Development
Strategic Intent	**Focus on optimizing research speed, processes, and project management through targeted AI use**, such as: R&D optimization; Increase research project speed (e.g., breeding cycles); Streamline the process; Better project allocation, AI-powered capacity management across centers (HR vs Research).	**Use AI to strengthen collaboration, policy influence, and resource alignment across CGIAR**, such as: Process for convening and advising policy; Research policy; Prioritization of challenges and assignment of project tools; Resource mobilization and entering new markets; New and fully supportive network of research partners (tech).	**Use AI to identify emerging opportunities and new directions for research and innovation**, such as: Identifying new research frontiers.	**Use AI to reduce manual administrative work and support institutional learning**, such as: MELIA automation.	**Invest in human capital and strategic growth of internal capabilities**, such as: Investing in internal talent development.
Organizational Readiness	**Focus on formalizing and standardizing operational processes through regulation and frameworks**, such as: Legal regulation; Policy frameworks; Standardized manuals and guidelines; Shared intranet; Auto update of repositories and systems.	**Use AI to foster a shared organizational culture and collaborative knowledge networks**, such as: Vision-setting across centers; Shared conventions and recommendations; Participatory "show & tell" practices; Interdisciplinary ways of working; Normative intent to guide ethical collaboration; Greater institutional capacity for researching AI use and ethics.	**Leverage AI tools to enhance access, aggregation, and analysis of distributed knowledge**, such as: Retrieval tools for finding and aggregating data; Generated reports/literature reviews (e.g., across open access/non-OA, different languages, quality of output).	**Delegate routine tasks and internal processes to automated systems that support scale, learning, and protection**, such as: Cybersecurity policy implementation; Scaling expertise and learning through automation; Auto-generated insights and documentation.	**Build internal preparedness through human-centered investments in knowledge and skills**, such as: Developing domain expertise; Investing in soft skills.
Learning and Evaluation	**Focus on improving operational efficiency and feedback cycles through targeted AI interventions**, such as: Refining the MLOps pipeline based on feedback; Assessing CGIAR's Results Framework.	**Using AI to identify friction points and collaboratively solve them across centers**, such as: Applying AI to correct bottlenecks in shared research or operations.	**Use AI to support continuous monitoring, adaptive learning, and safeguarding**, such as: Enabling real-time monitoring with safety verification mechanisms.	**Delegating routine and reporting tasks to AI to free up researcher time**, such as: Automating user feedback reporting; Generating learning content through AI-assisted content creation.	–

Adoption and Use	**Focus on improving operational efficiency and reducing repetitive tasks through targeted AI interventions**, such as: AI/ML intervention on data processing; Annotation; Information on existing groups and research on the topic one wants to research; Who has received the same donor funding.	**Using AI to connect expertise, tools, and data across CGIAR to improve collaboration and avoid duplication**, such as: Using LLMs to identify who is working on what expertise; Findings who across CGIAR has conducted similar experiments; Sharing tools across centers; Leveraging platforms for joint data sharing, algorithm use, and analysis.	**Leveraging AI to enhance research**, such as: Enriching interactions, mechanisms to accelerate research and discovery across CGIAR.	**Delegating routine and administrative tasks to AI to let researchers focus on core science**, such as automatization of admin-related tasks that take time from research.	–
Governance and Ethics	**Applying automation to improve efficiency and reduce administrative burdens**, such as: Filling out admin forms using AI/automation.	**Using AI to enable shared governance and institutional transparency**, such as: Detecting liabilities and risks; Using AI for research quality assurance (e.g., ChatGPT manuscript review); Defining operational areas where AI should be excluded; Fighting undisclosed use of AI.	**Strengthening the role of AI in promoting research integrity, accountability, and contribution to the scientific community**, such as: Adopting and enforcing guidelines for responsible use; Upholding the mission to contribute to the open-source AI community; Mindful time allocation to original thinking and learning (especially for junior staff).	**Acknowledging the risks of unchecked automation**, such as: Entry-level jobs disappearing; Reduced value of human capital due to over-reliance on AI tools; Ensuring accountability in AI-supported decision-making.	–

Defining and Measuring Success

The last part of the workshop invited participants to define what meaningful progress looks like for responsible and inclusive AI across CGIAR. Groups engaged in a collaborative exercise to articulate OKRs that reflect CGIAR's aspirations across five strategic pillars. These OKRs, while aspirational, provide direction for future alignment, monitoring, and shared ownership of AI integration efforts.

Although Key Results are still in a formative stage, they offer valuable insight into how progress could be measured from new partnerships and data infrastructure to institutional policies and training (please see Figure 7).

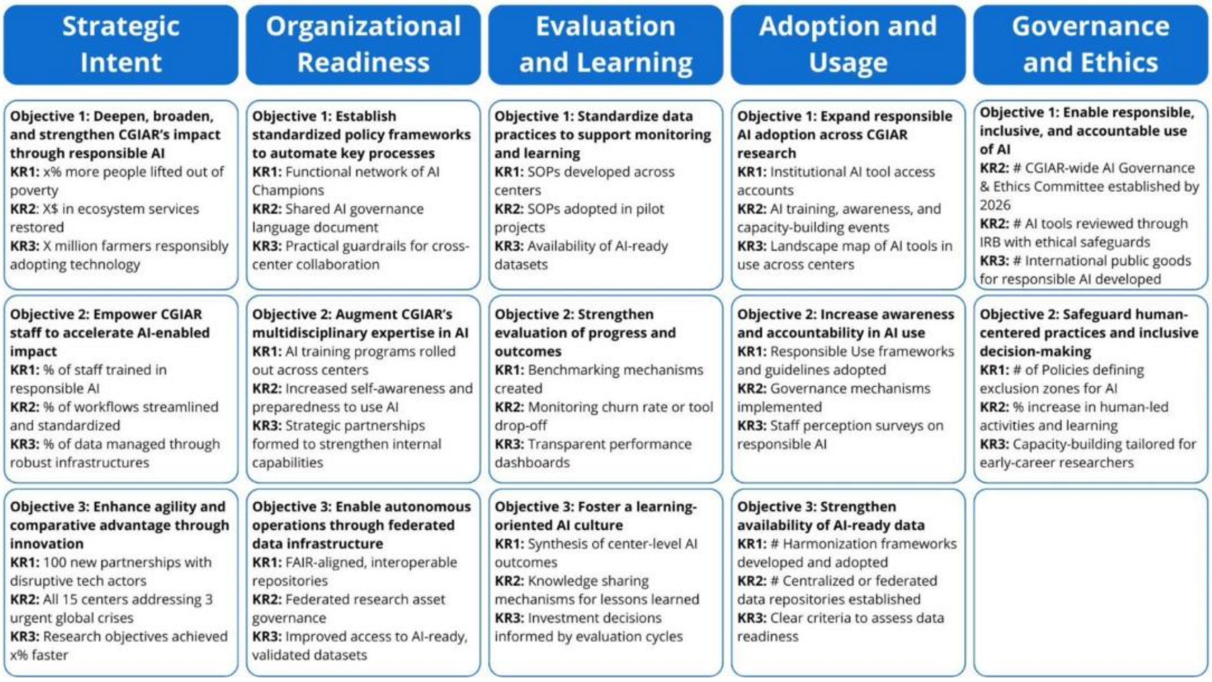


Figure 7. OKRs per pillar

The recommendations (Figure 8) based on the above figure (Figure 7) aim to support the groundwork for an AI roadmap for CGIAR.

Figure 8. Recommendations for an adaptive AI roadmap. (*Source*: Author's creation)

Part 3 - Planning the Journey

Human Centered Design and the Evolving AI Landscape

Leveraging the agreed upon future state where AI will increasingly play a role both in the disruptive innovations being delivered by CGIAR, but also in the way scientists can be assisted in optimizing their research work to amplify their impact on communities and ecosystems, the core team assigned to design and plan the AI Co-Scientist framework and platform presented to the audience their research findings on evolving trends around AI usage, advances in AI architectural patterns and tools, followed by a session explaining how Design Thinking can be applied in the context of designing and delivering user-centered AI experiences (Figure 9), to maximize the desired outcomes, and aligned with the strategic intent envisioned by the organization.

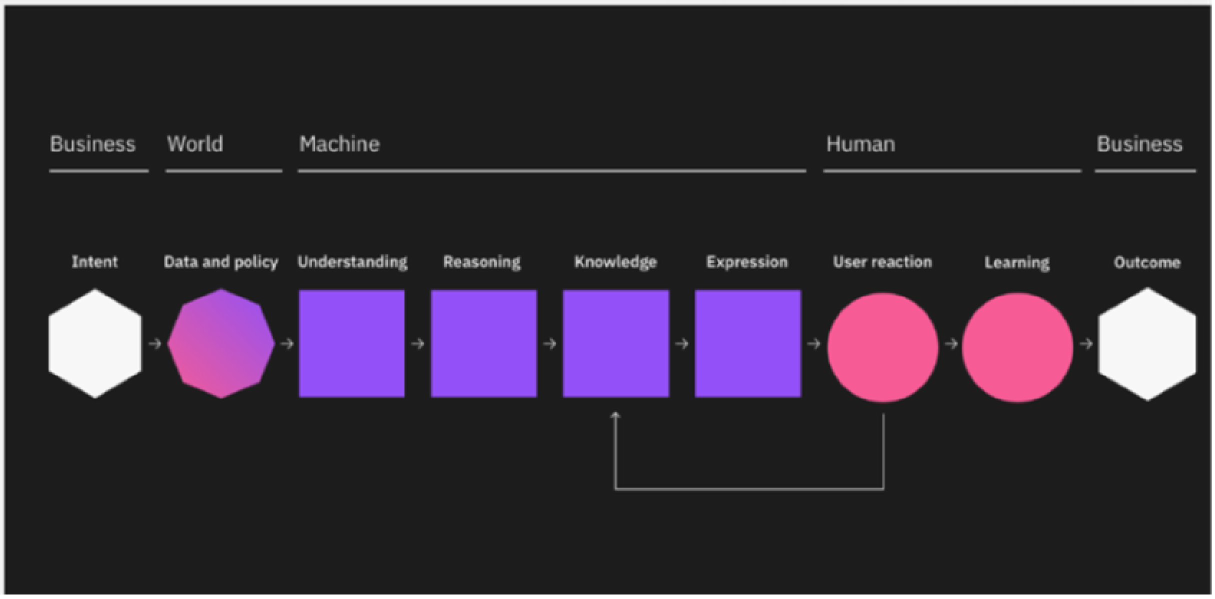


Figure 9. AI/Human context model. (*Source:* https://www.ibm.com/training/enterprise-design-thinking)

Setting the AI Platform Foundations and Domain Contextualization

This second day of the workshop ended with brainstorming sessions that allowed participants to submit and discuss proposals for the envisioned development of an AI framework (processes, guidelines, guardrails, enabling environment) and an AI platform (infrastructures, technology stacks, tools), expressing both desired capabilities across the value chain, but also concerns that must guide the design of the new framework and platform, in order to align with ethical principles and CGIAR's vision.Key takeaways from the brainstorming sessions pointed out as

very relevant aspects to take into consideration, such as:
- Current (existing) data platform(s) at CGIAR, while seemingly ideal to underpin the additional capabilities desired for implementing AI-based solutions, still need to gain maturity and broader utilization by the centers.
- CGIAR and each center must collaborate further in evolving current data architectures, data governance and data management practices, to better support AI.
- DTA's effort in developing collaboration hubs and communities of practice is perceived as having an increasingly important role to ensure capacity building across centers, on AI Tools, AI Models, and AI systems design.
- CGIAR is uniquely positioned to develop domain-specific models (on FLW domains) and promote multidisciplinary efforts across centers, suggesting there should be a significant effort into developing accelerators, models and use cases that can foster innovations and promote more impact.

From Ideas to Product Concepts and Prototypes

Overview of the design approach

The final day of the workshop marked the transition from ideation to concept development, introducing participants to the Digital Transformation Accelerator (DTA)'s approach for designing and implementing an AI Co-Scientist multi-agent system. This approach builds on leading global practices in AI-for-science—most notably Google's AI Co-Scientist initiative—while grounding these concepts in CGIAR's specific research, institutional, and ethical context.

Participants were briefed on how multi-agent AI systems can support different stages of the research lifecycle, augmenting human expertise rather than replacing it. This framing provided a shared conceptual foundation (Figure 10) before participants engaged in hands-on design activities.

Design Thinking–based prototyping process

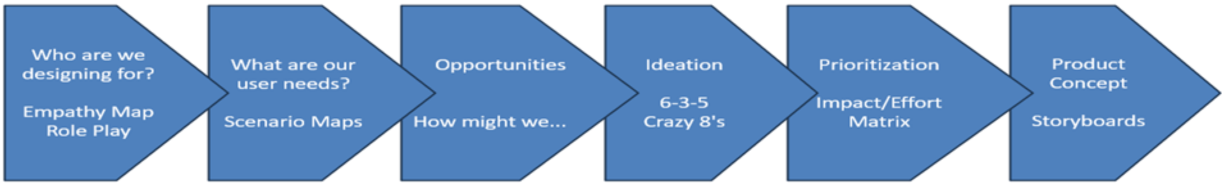


Figure 10. Design Thinking steps.

Participants then took part in a sequence of **Design Thinking activities**, leveraging insights, challenges, and opportunity areas identified earlier in the workshop. They were organised into **multidisciplinary design teams**, bringing together scientific, technical, and operational perspectives.

Each team was tasked with developing a **conceptual prototype** addressing a priority institutional or research challenge within one of the four stages of the research lifecycle:

- Prospecting
- Planning
- Executing
- Closing

To guide the design process and ensure strategic coherence, all teams applied three common criteria:

1. Alignment with CGIAR's food, land, and water (FLW) mission areas
2. Adherence to ethical, human-centred AI principles, including transparency, accountability, and inclusivity.
3. Potential for cross-center scalability and reuse, consistent with the One CGIAR approach

This framing ensured that the resulting concepts were not isolated ideas, but plausible candidates for system-wide implementation.

Idea Consolidation and Prioritization

As part of the prototyping process, groups were reorganized by research lifecycle stage (prospect, plan, execute, close). Each group revisited the ideas generated during earlier workshop sessions and mapped them explicitly to their assigned stage.

This step enabled participants to:

- Reconnect ideas to the original user needs and pain points that generated them.
- Assess ideas against impact, feasibility, and scalability considerations; and
- Conduct a structured prioritization exercise to select a single idea for prototyping.

The prioritization process surfaced both convergence (shared challenges across Research Centers) and divergence (stage-specific needs), reinforcing the value of a lifecycle-based approach to AI design.

Emergent Prototype Themes

Five prototype themes emerged, collectively spanning the full research lifecycle:

1. **AI for Donor Matchmaking** (Prospecting)
A system that uses AI to align funding opportunities with ongoing and proposed research projects based on thematic focus, impact areas, and geographic scope. The concept supports researchers in identifying relevant funding opportunities earlier and more efficiently.
2. **Ask the Data** (Planning)
An interactive natural-language interface that enables researchers to query datasets directly via AI-powered chat, integrated with CGIAR data repositories (e.g., Dataverse). This prototype addresses time-intensive data exploration and preparation tasks during research planning.
3. **GeoAI Advisor** (Executing)
A spatial analytics concept leveraging Sentinel-1 and other Earth Observation data to improve flood and drought monitoring. The prototype illustrates how AI can augment real-time analysis and decision-making. during research execution.
4. **Automated Research Reporting** Assistant (Closing)
A generative AI tool capable of synthesizing project updates, progress summaries, and publication insights. This concept targets reporting and knowledge-sharing bottlenecks at the closing stage of research.
5. **Ethical AI Governance Dashboard** (Cross-cutting)
A monitoring framework designed to track compliance with ethical AI standards and data governance protocols across centers and across all stages of the research lifecycle.

Prototyping Outputs and Design Principles

Each team developed **storyboards and low-fidelity mockups** to visualize their concepts. These artefacts focused on:

- user journeys and interaction points,
- high-level data flows and integrations,
- roles of human users versus AI agents, and
- governance and oversight considerations.

Participants were explicitly encouraged to apply the **design principles discussed on Day 2**, particularly:

- **transparency** (clarity on how AI supports decisions),
- **explainability** (ability to understand and interrogate outputs), and
- **inclusivity** (designing for diverse users and contexts).

As a result, the prototypes reflected CGIAR's commitment to **responsible innovation** and provided a credible bridge between conceptual AI systems and real research workflows.

Key Insights and Findings

- **Lifecycle framing was highly effective**: Structuring ideation and prototyping around the research lifecycle helped participants move from abstract ideas to concrete, actionable concepts.
- **Cross-cutting needs emerged alongside stage-specific ones**: While some challenges were unique to specific research stages, others—such as ethical governance and data access—cut across the entire lifecycle.
- **Human-AI collaboration was consistently prioritised**: None of the teams proposed fully autonomous AI systems; instead, all concepts positioned AI as an assistant or co-scientist.

Institutional Architecture and Interoperability

Following the prototype development, a technical session addressed interoperability and institutional architecture. Facilitators introduced the concept of Model Context Protocols (MCPs)—frameworks designed to connect diverse AI models and datasets through shared context layers. The session illustrated how MCPs can support distributed AI systems, allowing each center to maintain autonomy over its data while enabling cross-institutional collaboration.

Participants discussed the importance of a federated AI architecture, where each CGIAR center contributes data, models, or APIs to a shared ecosystem governed by common ethical and technical standards. This approach, they agreed, balances innovation with accountability, supporting open collaboration without compromising data sovereignty or privacy. The group also explored the linkage between MCPs and CGIAR 365, emphasizing the need to evolve the existing data lake into an AI-ready infrastructure. Proposed actions included establishing unified metadata standards, integrating model registries, and implementing automated quality control pipelines. Participants highlighted that such developments would not only accelerate AI applications but also reinforce transparency and reproducibility in research.

Prototype Demonstrations and Feedback

In the afternoon, each design team presented its prototype concept to the full group, simulating a mini "demo day." Participants then engaged in group storytelling exercises, developing AI-based prototypes aligned with research challenges.

- **Group 1** presented "Tinder Donor," an AI app that matches researchers with potential donors by analyzing proposal ideas and generating tailored pitches (Figure 11).

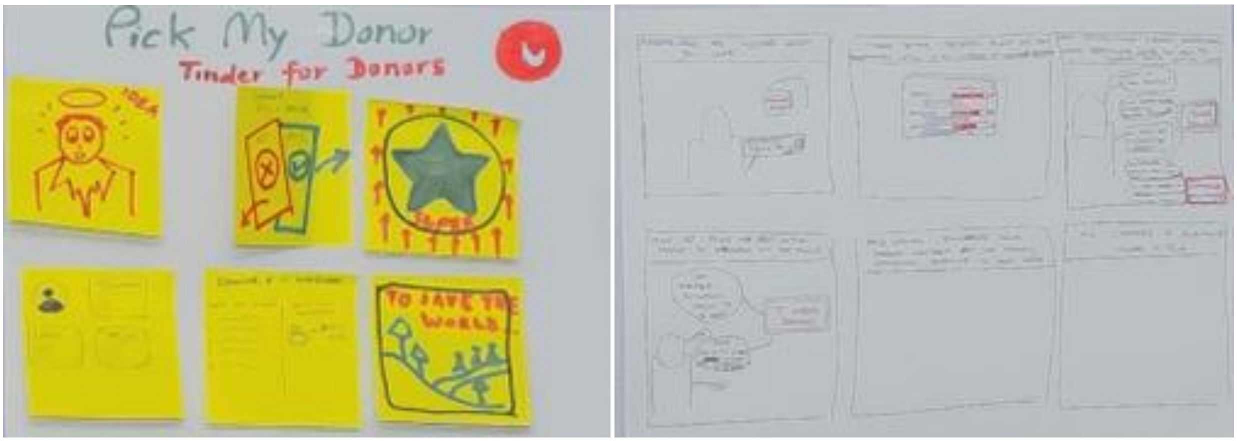


Figure 11. Group 1 demonstration (*photo:* IWMI)

- **Group 2**'s story, "David and Goliath," depicted the struggle of researchers against massive datasets, resolved by AI-assisted data access and harmonization (Figure 12).

Figure 12. Group 2 demonstration (*photo:* IWMI)

- **Group 3** introduced a GeoAI agent integrating spatial, socioeconomic, and biophysical data to support project design and decision-making, ensuring cross-sectoral learning (Figure 13).

Figure 13. Group 3 demonstration (*photo:* IWMI)

- **Group 4** dramatized the chaos of report deadlines and proposed an integrated AI system that automates data collation, visualization, and synthesis, freeing time for innovation and reflection (Figure 14).

Figure 14. Group 4 demonstration (*photo*: IWMI)

The presentations showcased both the creativity and technical acumen of participants. For instance, the "Ask the Data" concept illustrated how conversational interfaces could make data access more intuitive, while the "GeoAI Advisor" demonstrated the potential for real-time decision support in climate and disaster resilience. After each presentation, facilitators and peers provided structured feedback, assessing the relevance, scalability, and ethical compliance of each idea. Common recommendations included strengthening interoperability with existing CGIAR systems, enhancing user-centered design, and prioritizing open-source development models to maximize impact across centers. The discussion emphasized the importance of maintaining simplicity ("Keep It Simple") while adopting advanced architectures. The role of Model Context Protocols (MCPs) was highlighted as a means to connect diverse AI tools and support interoperability. Participants also explored how federated architectures, microservices, and open-source systems could facilitate collaborative multi-agent environments across centers.
The plenary discussion emphasized that these prototypes represented functional starting points for long-term AI initiatives under the DTA. Participants proposed that selected prototypes undergo further development through pilot testing and integration into existing research workflows.

Potential Use Cases Presentation

Three use case presentations at the workshop showcased concrete advances towards the AI Co-Scientist vision, each illustrating responsible AI integration in CGIAR research. CIP's presentation introduced an "AI Research Assistant" prototype enabling researchers to query datasets in plain English. The "Ask the Data" tool uses OpenAI and LangChain to translate natural-language questions into data queries and return clear results, lowering technical barriers to insight. In another presentation, the Portfolio Performance Unit (PPU), CGIAR introduced key AI tools developed by the organisation to enhance data-driven decision-making. Central to the session was the Semantic NLP Aggregator Platform (SNAP), which now features automatic summarization with embedded references, semantic search, clustering, and integration with OpenAI's ChatGPT for interactive analysis. SNAP supports both automatic and "Power User" modes, enabling users to explore PRMS datasets by meaning rather than keywords, generate thematic summaries, and validate findings via clickable links. The workshop also showcased the Prompt Batching Tool, which streamlines the generation of multiple AI prompts and outputs, enhancing efficiency in data processing. Additionally, the session covered the foundational role of Transformer models in modern NLP, highlighting their self-attention mechanism for understanding language context. A dashboard was presented to visualise SNAP outputs alongside performance data, supporting narrative development. These innovations aim to embed responsible AI into IWMI's research and reporting workflows. IFPRI's presentation highlighted the development of agile AI agents for data and policy analysis. One example was a retrieval-augmented search agent for IFPRI's Food Security Portal that combines semantic search with LLMs to deliver context-rich answers linked to original sources. Another pilot automated the tracking of global food-export restrictions by scraping multilingual news and using LLMs to extract policy changes, with cross-verification against official datasets to ensure accuracy. IFPRI also demonstrated multi-agent systems that interface with complex economic models to facilitate scenario exploration for researchers. Collectively, these use cases bring the AI Co-Scientist concept to life – AI tools acting as collaborative partners to accelerate data discovery, analysis, and reporting – in line with the workshop's vision of human-centred, responsible AI driving innovation in food, land, and water research.

Conclusion

Participants recognized that the workshop had successfully moved from diagnosis to design and from vision to initial implementation pathways. Several cross-cutting outcomes were reaffirmed:

- Agreement on the AI Co-Scientist vision as CGIAR's guiding framework for responsible AI integration.
- Consensus on establishing a CGIAR AI Asset Registry and Responsible AI Governance Framework.
- Identification of pilot prototypes for continued development under the DTA.
- Recognition of capacity-building needs and the importance of developing an AI ethics curriculum for CGIAR staff.

The workshop concluded with a renewed sense of purpose—affirming that responsible AI integration requires continuous dialogue, cross-center alignment, and a shared ethical foundation.

Appendix A. Participant statistics presented below (Figures 15, 16 and 17)

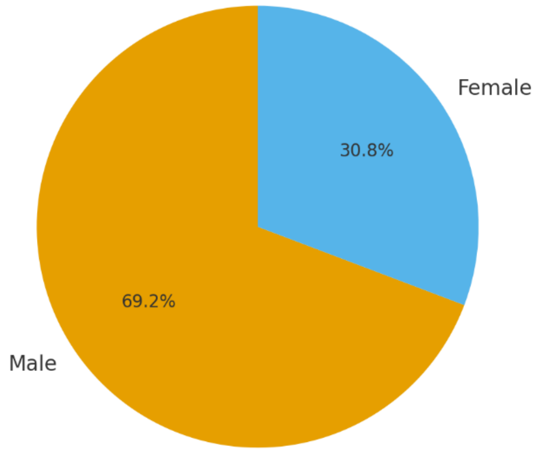


Figure 15. Gender diversity among the participants

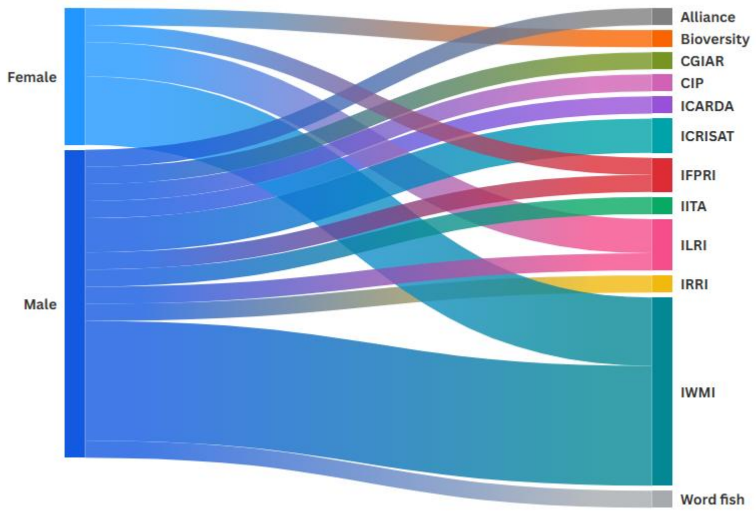


Figure 16. Participants' organization with gender diversity

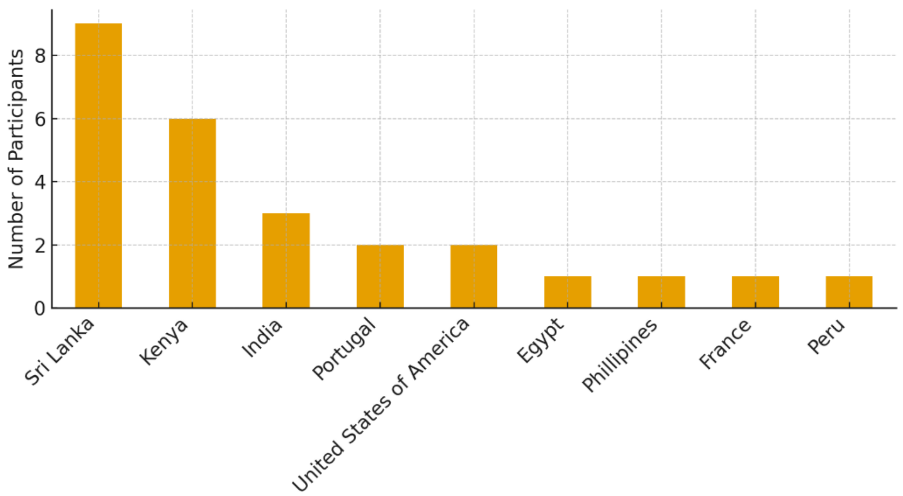


Figure 17. Participants' countries

Appendix B. Exercise 1: CGIAR's R&D Lifecycle + Enablers/Blockers

Guided breakout using ORID method across four research stages (Prospecting, Planning, Executing, Closing). Each group contributed to a shared matrix.

Table B.1. Group 1

	Doing	**Feeling**	**Thinking**
Prospecting	What problems can we solve with AI? Testing available/commercial/open source tools for literature review, ideation, brainstorming	I feel relieved because literature reviews can be done much faster	AI can do lit review better than researchers Does AI find/write any relevant literature?
Planning	Identifying research groupsby using AI tools Identifying suitable research methods and drafting research strategies + workplans		
Executing	Policy on AI tools Researchers are using external AI tools which are not complliant with data privacy	We feel anxious about shaing our private data I feel anxious because my AI model could lead to wrong recommentations → environmental/social impact	Need to test tools properly and explore risks → risks assessment tools, methods, etc. Mitigation measures
Closing	AI-generated manuscripts Using AI for improving/checking research quality	Worried for younger scientists Good for senior researchers to review reports/papers Bias for research literature review	Lack of deep knowledge
Enablers	AI governance with existing tools Partnership endorsement from governments		
Blockers	Better usage of AI → prompt engineering Tests for AI tools Reluctant to share data and resources Accountability – who is in charge of wrong decisions? Procurement challenges on AI resources – ambiguity in how to procure 'risky AI tools' Environmental constraints, water energy Regulating data from partners (governments) Data misingormation Data privating concerns		

Table B.2. Group 2

	Doing	Feeling	Thinking
Prospecting	Chatbots (e.g. ChatGPT or ILRI's Animal Health Chatbot or Grantable) Used to explore research ideas from core proposition formulated by a scientist	Lack of trust in scientific referencing Reduced creative thinking and critical	Low attention span Learning from others mistakes
Planning	Meeting AI recording transcription Moozwell for project management	Changed narratives Biases	
Executing	AI for generating Q/A for fine tunning LLMs Geospatial for cloud Developing questionnary OpenAI's Whis AI or Gemini for automatic transcription and optional translation of focus group discussion	Makes one lazy Saves time	
Closing	Generating progress report		We need more tools to detect undisclosed use of AI
Enablers	Open Access to certain extent Large database of expert knowledge - CGSpace		
Blockers	Lack of training Lack of GPU		

Table B.3. Group 3

	Doing	Feeling	Thinking
Prospecting	Resource unit sends an email with calls No AI for RFP's → manual Search engines + newsletters ChatPT • Putting concept note into proposal to initiate piloting of products/devices developed for pest detection/monitoring • Sourcing for donors to fund/approve the proposal	Frustrated to look for funding Pressure Frustrated by the centers competition	Are we missing opportunities? Competition
Planning	No AI - IRB – Internal Review Board (critical); word; portal (ABC) Research protocols with LLMs AI Assist for Brainstorming Inception Workshop Manual processing for resources plan	Angry because silos Optimistic Excited Fired-up / Excited to get the ball rolling to finally test the idea/revised concept, if and many will adopt it.	We don't take advantage of existing information. *Silos* Who else is doing the same? Collected already the data? *Silos* You have to share the pie, yet you have big plans.
Executing	Automation of some field work (phenotyping) Data collection processes e.g. training, workshop is manual / electronic AI Machine learning methods for analysis AI for data annotation Data cleaning, formatting, etc. manual - No AI Audio data collection in local language LLM / ASR / NLP	Motivated to actually see results and impact Curious and excited Stressed Time crunch Anxious	Hesitant / doubtful to implement variations / changes in protocol Are we in the right path to reach the outcome Time constraints for implementation Do I have the right data? Are we using the right methods?
Closing	Writing a lot – reports; papers Workshops and presentations to show results Use of AI for dissemination and natural language	Relief Optimistic	Impact Is it understandable? Do my results make sense?
Enablers	Data publishing not using AI Digital and data team Groups Evidences		
Blockers	Cost of AI tools No AI Policy IT Teams Silos		

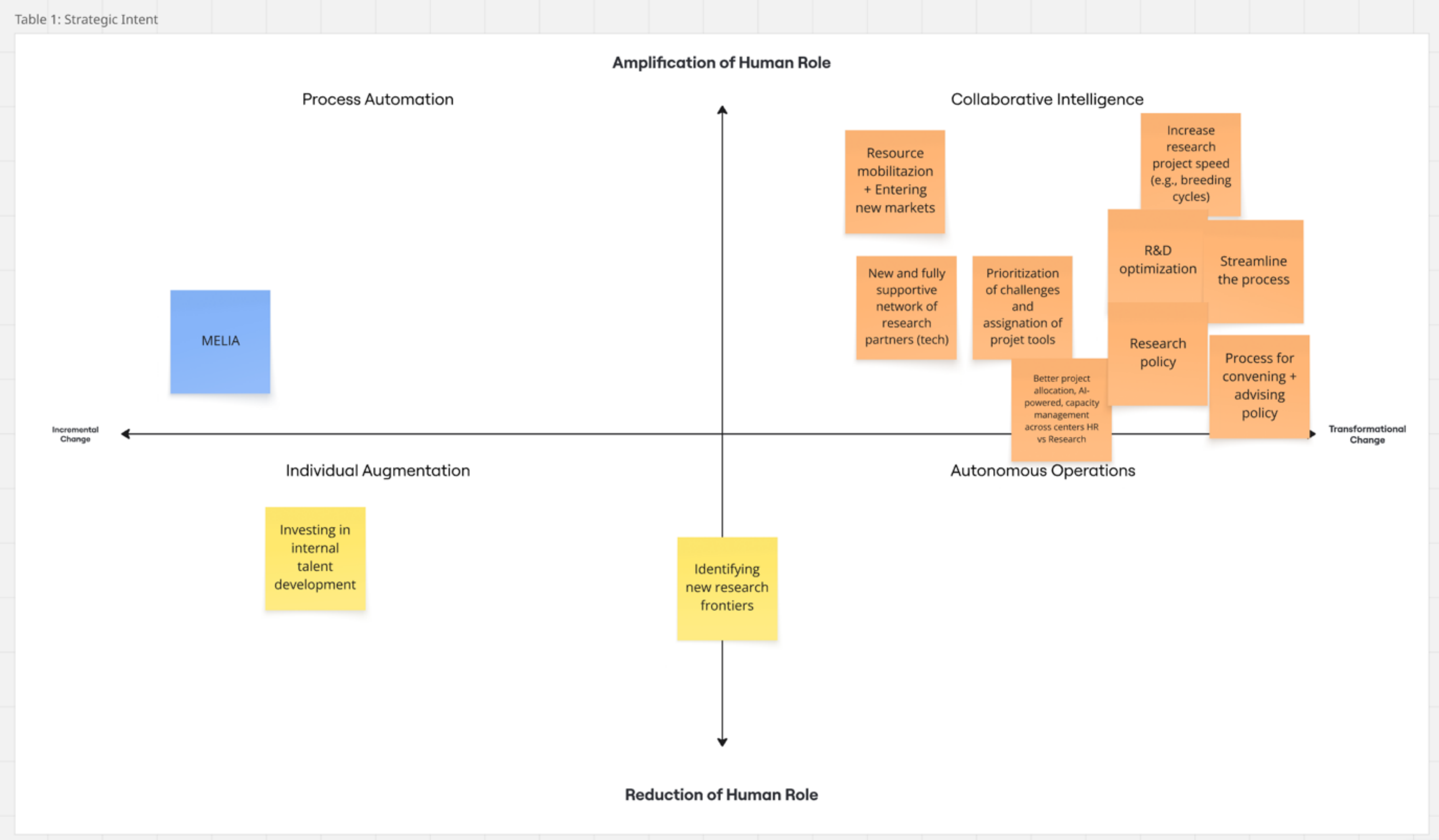
Table B.4. Group 4

	Doing	Feeling	Thinking
Prospecting	Using AI to identify current trends in specific areas Using AI to review CGIAR research to understand what's needed Setting up research questions with AI Exploring how AI can accelerate the development of research papers	Overwhelmed (due to stigma or pressure around AI tools) Noted that AI makes things feel "clear and fast" for writing proposals Cautious — need to be careful around ethical issues	"Now we can be up to speed" Acknowledgment that challenges lie ahead
Planning	Gathering AI-related knowledge and current trends Planning to execute with human-centered design Using AI to identify research gaps Identifying people who could fund research Using tools like JIRA and ClickUp aligned with agile practices Using AI tools for chat functionalities Exploring initial ideas and understanding how to solve a problem	General acceptance of AI among team members Questioning whether human validation is needed to mitigate risks in data and outputs Curious and exploratory — trying to extract something interesting and useful from trends or future-proof methodologies	AI tools help make planning more concise Questioning whether there is enough data or resources Finalizing literature critiques related to the idea Investigating existing methodologies for building a GenAI chatbot
Executing	Using AI tools to perform tasks Identifying the most time-consuming steps and understanding why Generating code for data analytics Building final plots Creating tools that can run through a structured pipeline	Perception that people are not ready for AI tools Lack of experience and platform knowledge is a barrier Uncertainty: "Is it up to me to decide every time?" "Will it cause more bugs?"; "Can I express visually what's in my mind?"	Observing current issues in the research context Questioning result validity Concerned with initial validation of generated or synthetic datasets Belief that AI can help accelerate workflows and lead to more effective results Recognizing need to improve prompting for expected analytics
Closing	Being more careful with AI-generated outputs Applying Agile frameworks Debugging using GPT Continuously monitoring and adjusting features based on requirements Creating PowerPoint presentations and donor reports	Amazed: "Oh my god, how did we live without AI?" Anxious — feeling under-capacity or overwhelmed Acknowledging the need to hire skills and better understand how to use AI wisely Frustrated by debugging inefficiencies and how AI changes existing code unpredictably	The inclusiveness and impact of AI should be more of a focus Concern that many jobs may be lost due to automation Recognizing how AI has shifted the way they work: "I couldn't have done debugging alone without AI"
Enablers	Collaboration centers focused on AI (internal or external partnerships supporting AI integration)		
Blockers	Skills vs. expectation tradeoff —> mismatch between what people know and what AI use requires AI hallucination —> reliability concerns AI inaccuracy in coding —> debugging and unexpected behavior create friction		

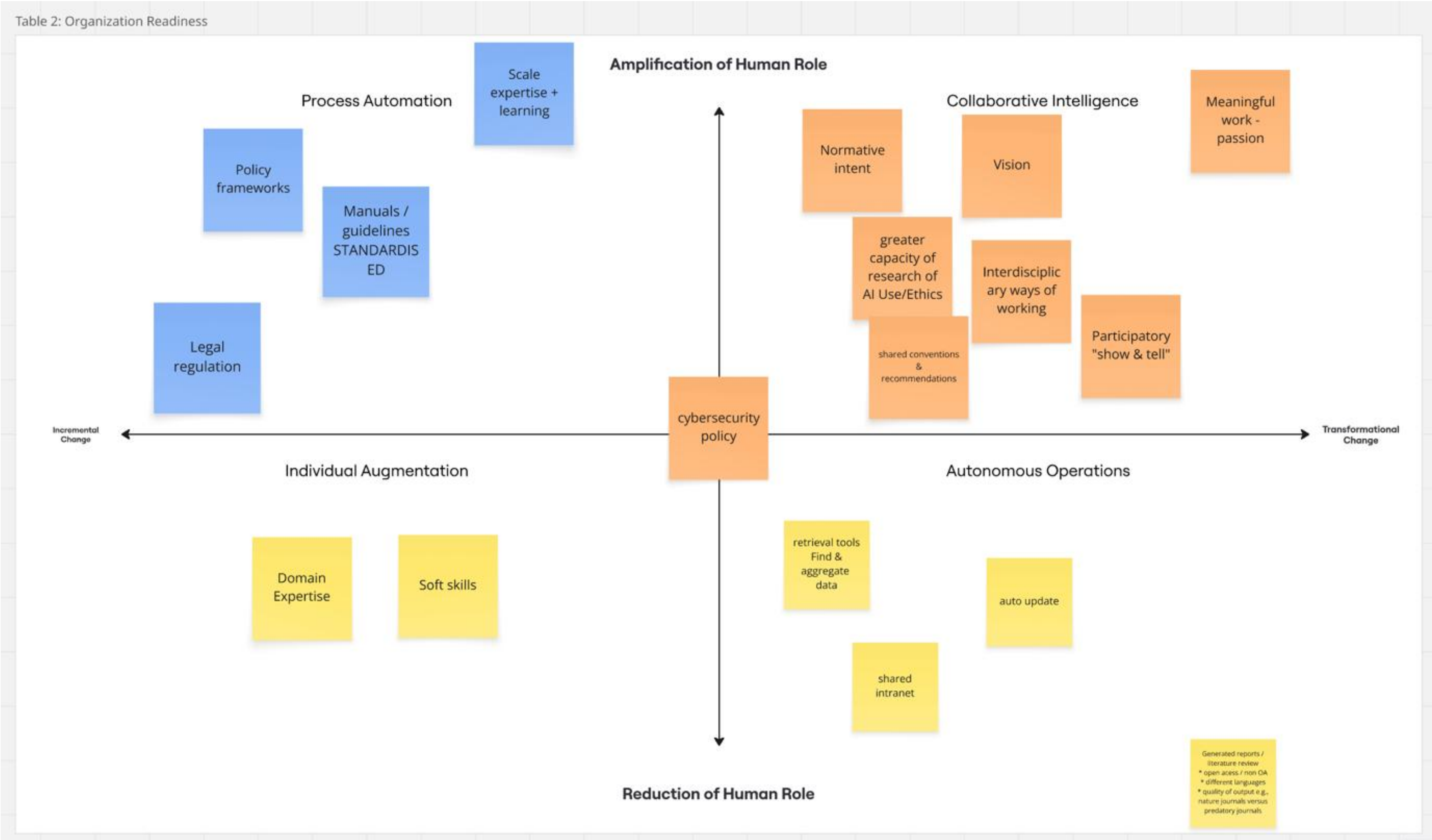
Appendix C. Exercise 2: To-Be (Strategic Intent, Organizational Readiness, Adoption & Use, Governance & Ethics, MELIA)

Breakouts worked on incremental vs. transformational change using the 2x2 + AI Intents framework.

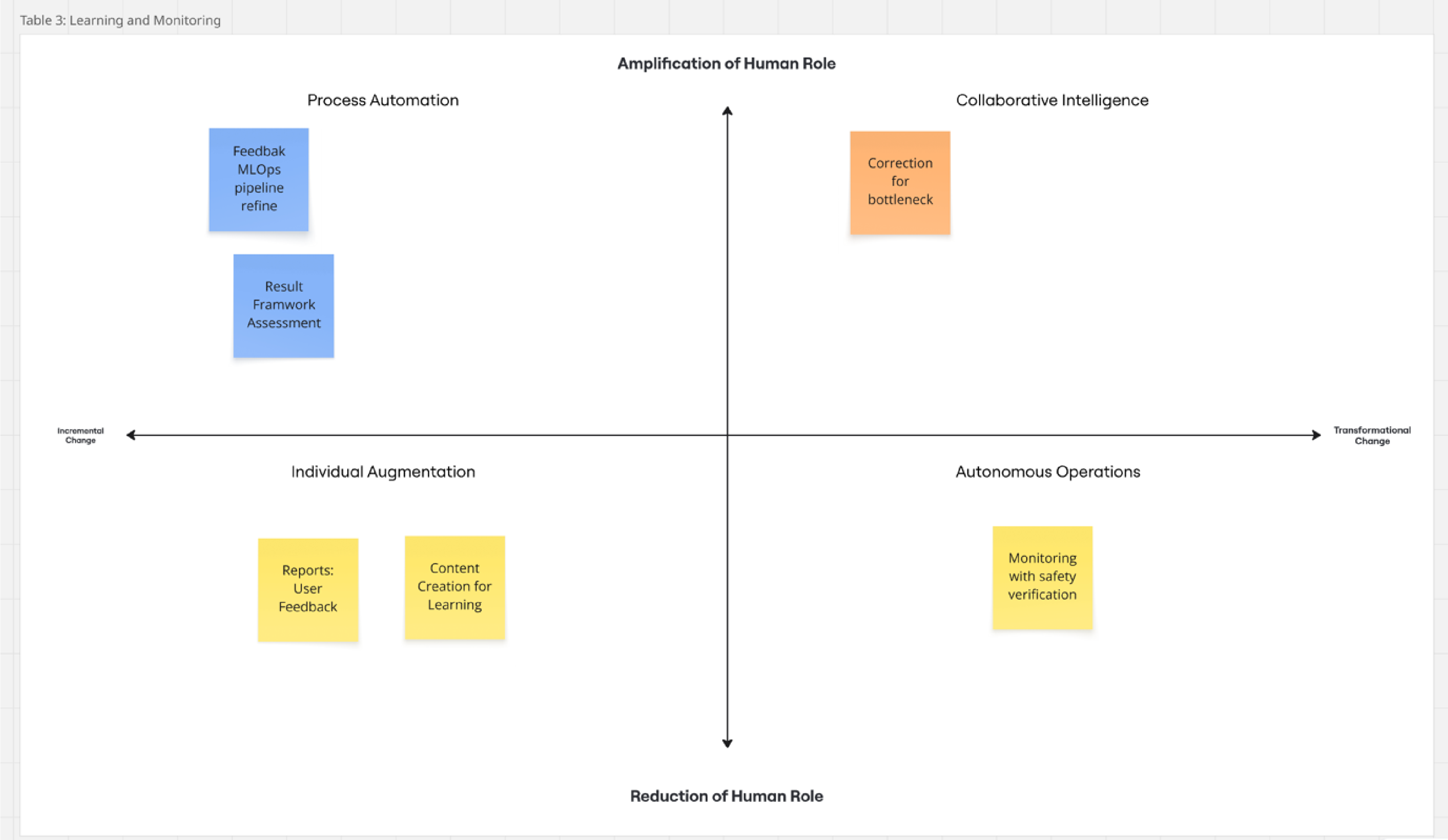
Group 1 – Strategic Intent



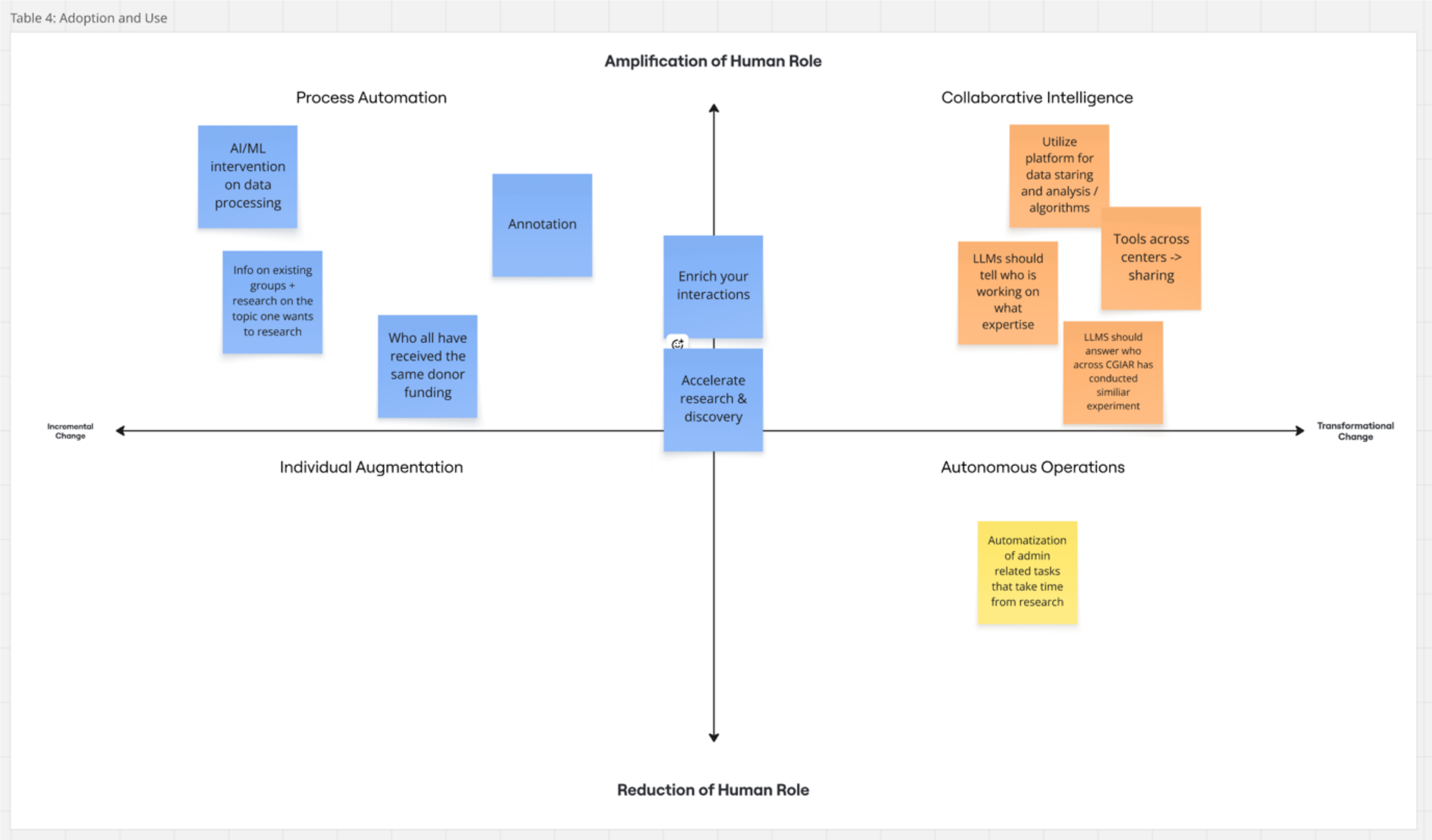
Group 2 – Organizational readiness



Group 3 – MELIA

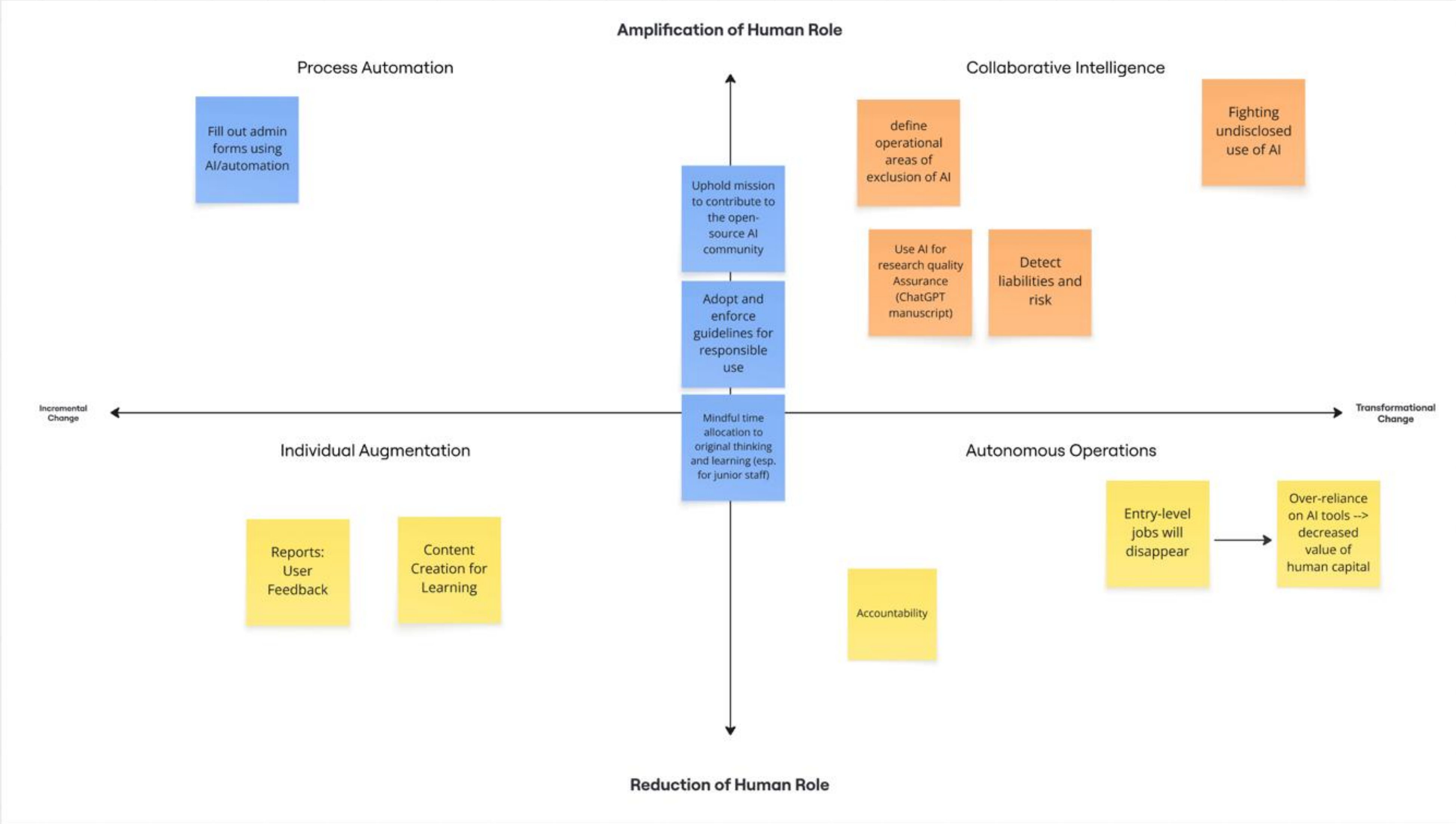


Group 4 – Adoption and use



Group 5 – Governance & ethics

Table 5: Governance and Ethics



Appendix D. Day 02 – OKRs

Pillar: Strategic Intent – 3-5 years

Stick notes area
Enhancing our convening abilities → accelerating policy change
Making science tools, knowledge and data more accessible to all
Open up new areas of research to catalyze innovation and remain competitive
Agility and adaptability to changing needs and demand
Creating new synergies across CGIAR
Ensuring safe, wise, inclusive and sustainable use of AI (maintaining results)
Creating new partnerships with tech partners
Diversify and improve our funding sources
Objective: Improving efficiency of research + from operations → reduces overheads
Serving the ORG's strategic vision

Objective #1: **Deepened, broadened, strengthen impact**

Why is this important for CGIAR? --

KR1: 50% more people out of poverty
KR2: 10 billion $ of ecosystem services restored
KR3: 7 million farmers have adopted tech responsibly

Objective #2: **Staff is empowered and enable to accelerate impact**

Why is this important for CGIAR? --

KR1: 50% processes + workflows streamlined + standardized
KR2: 70% staff trained on responsible use of AI
KR3: 100% data organized in strong infrastructure

Objective 3: **Becoming agile + adaptable to changing needs and demand → strengthening our comparative advantage**

Why is this important for CGIAR? --

KR1: 100 new strong partnerships with disrupting partners
KR2: All 15 centers wording on 3 urgent global crises
KR3: Research objectives reached 20% faster

Feedback

1. Are the objectives clear and meaningful?
- Not very relative to AI. More into technology
- No clear connection with AI.
- To high-level
- Very high-level → need to sharpen
- Objectives should be more defined
- You need very different funders for Objective #3
- The statement is not clear, e.g., “to strengthen”

2. Are the key results measurable and ambitious enough?
- How to measure role of AI in indicated KRs?
- Where is the AI?
- Slightly unrealistic. More clarity on how to measure.
- Too ambitious. May not related only AI. Lot of human and biophysical science required.
- KRs are measurable but not aligned with objectives.
- Right target but very inspirational.
- Objective 2 – Key results very measurable – good
- Objective 1 – Results not clear what role of AI is
- Overly ambitious – second that - +1; +1

3. Is the “why” compelling and aligned with CGIAR or researcher impact?
- Why? What is impact?
- Objective 2/3 seems much clearer
- It is aligned but why? → not there
- Lacking CGIAR (FWL) specificity
- Not compelling because it’s vague

4. What is missing or unclear?
- Objective 1
- AI element
- Better statement of objectives
- Are our staff ready for agile and adaptive work?
- Need of agile office spaces
- KR #3.1 – Need of agile partner for agile working mode
- Data security and privacy not defined
- Alignment with SDGs? Global South?
- Partnerships climate change

Pillar: Organization Readiness *Stick notes area* Enhance Collective Intelligence Ensure CGIAR – DTA research is meaningful, impactful, ethical, funded (?) + inclusive
Objective #1: **Process automation: Policy Framework** **Why is this important for CGIAR?** Standardized, practical guardrails + accountability to structure collaboration across centers + SPs KR1: a network of AI champions KR2: a shared language document KR3:
Objective #2: **??? augmentation – multidisciplinary expertise** **Why is this important for CGIAR?** So that we KR1: AI training program KR2: Self awareness of ….. KR3: No. of strategic partnerships (AI)
Objective 3: **Autonomous operations – common RAG federated repository** **Why is this important for CGIAR?** Augment FAIR ….. KR1: KR2: KR3:
Feedback ***1. Are the objectives clear and meaningful?*** - Love them :) - It's unclear what you want to achieve with the objective – e.g. more or less process automation - (About #1) too standardized methods might hinder innovation - Direction/action is missing - Rather generic, details are missing - Objective can be articulated more ***2. Are the key results measurable and ambitious enough?*** - Lacking quantitative targets - measurable, a bit generic - KRs nor measurable ***3. Is the "why" compelling and aligned with CGIAR or researcher impact?*** - Good whys

- Clear, comprehensible and compelling whys
- workflows for operating AI across country

4. What is missing or unclear?
- Availability of digital public infrastructure (DPI)
- organization personnel skills
- Objective 2 – Key Result Consultation for validating the AI tools. Required assessment.
- Unified AI (???) a single platform
- Quantitative measures

Pillar: Evaluation and Learning *Stick notes area* SOPs Monitoring
Objective #1: **Why is this important for CGIAR?** Standardization of data across centers KR1: Create SOPs KR2: Centers use SOPs KR3: AI-ready Data
Objective #2: Evaluate progress / process **Why is this important for CGIAR? ---** KR1: Use case benchmark KR2: Churn Rate KR3:
Objective #3: Learning **Why is this important for CGIAR?** Helps decide investments // Impact assessment KR1: Synthesis of the O/C or O/???'s KR2: KR3:
Feedback ***1. Are the objectives clear and meaningful?*** - They are very broad, only one word - Very concise - Objectives not clear - Too broad - No. sometimes, one single word is just not enough. ***2. Are the key results measurable and ambitious enough?*** - KRs lacy specificity - Wording is generic, impossible to know what it is measured - Not really. The goals should be SMART goals. - "Evaluation" → randomized control?? - Not really → try to be more specific - Results not clear - Key results are not measurable - I don't think so. Too broad, too unclear. - Yes, the (????) objective gives clear method for implementation but this is very high level. ***3. Is the "why" compelling and aligned with CGIAR or researcher impact?*** - To do a proper impact quality of the data is what matters most - Short of strong rationale - There is no why – why does M/E achieve ???? (investment is okay) is better

- So-so

4. What is missing or unclear?
- More words are missing!
- Focus should change to how to eval AI usage
- # of impact monitoring
- monitoring, evaluating + learning are all the same thing
- **Please add** explainability, bias, accuracy, monitoring
- **Process** – co-develop solution
- Almost everything is unclear here
- “Standardization” of data → what does it mean in terms of CG context?
- Business side is missing
- Could have explained with example
- Write objective on the scale expertise and learning AI intent

Pillar: Adoption and Usage

Stick notes area

People having time
Collective intelligence
Super computer
Land rover

Objective #1: **Reach a wider scale of users of AI tools (???) CG for high quality research results.**
Why is this important for CGIAR? Widespread use of AI tools for high quality research outputs. To reach the highest potential of everyone using AI tools.
KR1: Institutional accounts of AI tools for research access
KR2: # of training, capacity develop + awareness
KR3: Landscaping of the AI tools that are being used
[“Free time” – nr. Of hours per week that a researcher has to do what they want (quality time) - survey

Objective #2: **Increase awareness on responsible use of AI to accelerate research.**
Why is this important for CGIAR? Accountable use of AI
KR1: Frameworks and guidelines developed for the responsible use of AI.
KR2:
KR3:

Objective 3: **Make AI-ready data**
Why is this important for CGIAR?
KR1: Provide harmonization framework
KR2: Centralized data base
KR3:

Feedback

1. Are the objectives clear and meaningful?
- Yes
- Yes. #1 impact; #2 awareness + capacity; #3 data
- Objective 1: is use “good” use? What is “good” use?
- Good objective 2 :)
- Very subjective objectives. Not clear

2. Are the key results measurable and ambitious enough?
- Objective 3: Key results clear and overlapping with other groups.
- For adoption and usage specific KR could be added: KR → MEL studies on Adoption among beneficiaries
- Maybe add to Objective 2: KR #2 strict implementation of guidelines on responsible use of AI
- Key results: everyone has to do uncomfortable things or things we don’t like. It is called work.
- AI-ready data → define what “AI ready data” is and provide a tool to measure the readiness of data
- Guarantee – Quality of the work and processes.
- Some KR are difficult to translate into quantitative, measurable results
- Objective 1 & 2 are similar
- KR should be measurable. Good add the measure.
- Accountability measurable? Maybe researchers subscribing their AI assets.

3. Is the “why” compelling and aligned with CGIAR or researcher impact?
- NO. What’s the value for CGIAR?

- Obj 3 Great!!
- Fewer meetings → AI summaries
- As for our table, people here didn't explicitly answer the question on "why?"

4. What is missing or unclear?
- Conduct adoption study
- KR – Assess impact of AI use on human wellness. Work life balance
- Is getting a supercomputer realistic in terms of costs? (I have no idea how much $ it costs)
- More similar to governance
- Do we really need <u>more</u> people to have access? Or do we need the <u>right</u> people to have access to the <u>product</u> + <u>knowledge</u> (not necessarily AI itself)
- Staff training
- Free time – 0# Overlapping with governance
- Bit high level no practical step
- Objective 1, Key resylt 1 – 80% of CG researchers have AI enabled institutional accounts? Would this work?
- Centralized DB → maybe "federated" would be an option too
- How to avoid organization firing people because too much free time?

Pillar: Governance and Ethics *Stick notes area* - Define operational areas of exclusion of AI - Uphold mission to contribute to the open-source AI community - Mindful time allocation to original thinking and learning (esp for junior staff) - Setup a CGIAR AI governance structure and processes ← accountability - Sometimes AI leads to risks such as AI hallucinations by giving false positive answers - AI Intent 6 – Detect liabilities and risk
Objective #1: **Enable & ensure responsible use of AI in the CG** **Why is this important for CGIAR? --** KR1: Set up one dedicated CGIAR AI Governance and Ethics Committee by 2026 KR2: Make AI RB – includes AI – Ethics review before releasing the AI tools & provide technical expertise to IRB KR3: Produce international good for AI products [Ensure inclusiveness (?) along the process] [Ensure all research knowledge products disclose the use of AI use]
Objective #2: **Ensure Human Expertise and Human-Led interactions remain core.** **Why is this important for CGIAR? --** KR1: define critical activities where CGIAR policies mandate exclusion of AI KR2: Mindfully increase the %age of staff time spent on human interaction KR3: Provide capacity building activities to junior staff for AI [Increase time for in-person human to human interaction by xx%]
Objective 3: **--** **Why is this important for CGIAR? --** KR1: -- KR2: -- KR3: --
Feedback ***1. Are the objectives clear and meaningful?*** - Obj 3 can be a KR for Obj 1 - Objectives are inspirational <3 them - Nice statement of objectives - The objective are concise - Not specifying numbers but is going on the right direction ***2. Are the key results measurable and ambitious enough?*** - Yes - Many key results not measurable. Need to be SMART. - No measuring for 3rd obj - How to measure? #, at least… - Nothing measurable - Does % staff time interaction translate in better research?

3. Is the "why" compelling and aligned with CGIAR or researcher impact?
-N/A
- Sounds paranoid of AI domination where humans developed AI tools, platforms, etc.

4. What is missing or unclear?
- IRB needs example ethical/non-ethical AI
- AI will be missing
- Goods for AI products?
- Two KR's are overlapping.

CGIAR is a global research partnership for a food-secure future. CGIAR science is dedicated to transforming food, land, and water systems in a climate crisis. Its research is carried out by 13 CGIAR Centres/Alliances in close collaboration with hundreds of partners, including national and regional research institutes, civil society organisations, academia, development organisations and the private sector. www.cgiar.org

To learn more about this Science Program, please visit:
www.cgiar.org/cgiar-research-portfolio-2025-2030/digital-transformation

Contact
Mariangel Garcia Andarcia, Research Group Leader, Water Futures Data & Analytics (WFDA), IWMI (M.GarciaAndarcia@cgiar.org)

CGIAR DIGITAL TRANSFORMATION

International Water Management Institute

Alianza Bioversity & CIAT