Considerations and Practical Applications for Using Artificial Intelligence (AI) in Evaluations

Description

Tweet

Cekova, D., Corsetti L., Ferretti, S. and Vaca, S. (2025). Considerations and Practical Applications for Using Artificial Intelligence (AI) in Evaluations. Technical Note. CGIAR Independent Advisory and Evaluation Service (IAES). Rome: IAES Evaluation Function. https://iaes.cgiar.org/evaluation

Contents

Acknowledgments

Executive Summary

1. What is the AI Technical Note

1.1. How This Note Supports the CGIAR Evaluation Framework
1.2 What To Expect in this Technical Note

2. What is AI

2.1. AI in Evaluation: Concepts
2.2. How GenAI Works

3. How to use AI

3.1. GenAI Applications: Chatbots, Copilots, and Beyond
3.2. Responsible and Ethical AI Governance
3.3. Key Ethical Considerations for AI in Evaluation
3.4. How to Choose the Right AI Tools
3.4.1 Key Factors to Consider When Choosing an AI tool

4. Building AI Integration Competencies for Evaluators

4.1 Building the 'AI Muscles'
4.2 Remain Aware of Risks and Sensitivities
4.3 Building the Toolbox and Practices
4.4 Designing AI into the Whole Evaluation Process
4.5 Leverage AI for Knowledge Exchange
4.6 Negotiate and Agree Use
4.7 Supervision at Different Levels
4.8 Transparently Documenting AI Use

5. When to use AI in Evaluation Phases

5.1 AI Uses per Evaluation Phase
5.1.1. Research and Evidence Management
5.1.2. Evaluation Design, Research and Evidence Management
5.1.3. Evidence Collection
5.1.4. Data Analysis
5.1.5. Dissemination, Documentation, and Reporting
5.2. A Practical Guide to GenAI Conversations: Beyond Single Prompts to Meaningful Dialogue

Bibliography and Further Reading

Annex 1. Glossary of Terms

Annex 2. Non-Exhaustive List of AIs for Evaluation

Annex 3. CGIAR IAES GenAI Prompt Record by MERL Task

Executive Summary

The CGIAR 2030 Research and Innovation Strategy commits organizational change with seven ways of working, including â??Making the digital revolution central to our way of workingâ?•. In that context, Artificial Intelligence (AI), introduces both opportunities and risks to evaluation practice. Guided by the CGIAR-wide Evaluation Framework, integrating AI tools requires a governance approach to balance innovation with ethical responsibility, ensuring transparency, fairness, accountability, and inclusivity. This Technical Note encourages and guides CGIAR evaluators to ethically explore, negotiate, and experiment with AI tools:

Explore: Evaluators are invited to discover how AI, especially GenAI, can enhance evaluation efficiency, from scoping and data analysis to reporting. The Note provides practical guidance on AI applications and examples to support creative yet responsible exploration.

Negotiate: Integrating AI should be openly discussed with commissioners, stakeholders, and teams. The Note prioritizes jointly defining boundaries, expectations, and ethical parameters (transparency, accountability, and data sensitivity) at each phase of evaluation.

Use AI Responsibly: While AI tools are evolving, evaluators are encouraged to pilot and iterate their use. The document supports experimentation through practical tips, prompt examples, and tool selection criteria, all while emphasizing documentation and learning from each use case. Effective AI governance is grounded in core principles: Transparency requires clear documentation of AI tool usage, data sources, model limitations, and decision-making processes. Accountability involves assigning responsibility for AI decisions and outputs and establishing oversight and redress mechanisms. Fairness and inclusion must proactively mitigate bias and discrimination, with particular attention to underrepresented groups and data gaps. Data privacy and security must align with applicable data protection regulations and ensure secure handling practices. Human oversight ensures that evaluators retain control over processes and can intervene as needed.

In operationalizing ethical AI governance in CGIAR evaluations, due diligence is required in assessing AI tools for ethical alignment before deployment: reviewing the transparency of vendors, the documentation of models, and their intended use cases. Where relevant, components involving AIâ??especially those engaging human subjects or sensitive dataâ??should undergo ethics review. AI applications must be adapted to the local and cultural contexts in which evaluations are conducted, as what is suitable in one setting may be inappropriate in another. Additionally, participants should be informed about the use of AI systems and the implications of data collection or processing to ensure informed consent.

Ethical AI governance should be embedded in the entire evaluation lifecycle. During the design phase, evaluators should define AI tools to use, why they are selected, and assess risks. In data collection, AI tools should be used in ways that uphold data privacy and protection standards and avoid reinforcing harmful stereotypes or excluding groups. During the analysis phase, the role of AI in supporting interpretation should be documented, with an acknowledgment of limitations or biases. In dissemination, documentation, and reporting, AIâ??s contribution, limitations, and human validation should be disclosed. By rapid adaptation of content across formats, languages, and complexity levels, AI opens possibilities for broader, more inclusive communication of findings. Finally, the follow-up phases should include a reflection on the ethical implications observed and how these lessons can

improve future evaluations.

By embedding methodological flexibility into the evaluation processes, AI adoption would contribute to integrity, equity, and learning in an era of rapid technological advancement. This Technical Note is a conversation starterâ??as a â??Betaâ?• version, it will evolve based on responsible real-world experimentation and continuous reflection. Evaluators are encouraged to be responsive to stakeholder input throughout the evaluation processes, to ensure relevance, accuracy, and inclusivity.

Category

1. Uncategorized

Date
09/09/2026
Date Created
13/06/2025
Author
admin