

AI in evaluation: from Principles to Practice

1. The real gap is not principles — it is operationalization

👉 Moving from **ethical guidance** → to **standard, consistent practice**

2. What needs to evolve?

A. Standards

From principles → to rules of practice

- Document AI use (prompts, workflows)
- Ensure traceability to evidence
- Maintain human judgment

B. Competencies

From experimentation → to AI literacy

- Understand limits & risks
- Critically interpret outputs
- Combine evaluation + digital skills

C. Institutional frameworks

From pilots → to governance

- QA protocols for AI-assisted work
- Clear safeguards & accountability
- Secure data environments

3. AI should not replace evaluators — it should augment them

→ Freeing time for **judgment, ethics, and context**

→ Trust depends on **consistent and transparent application**, not ad hoc use