FAQ

Diagnosing Incorrect Answers

Colette may occasionally return incorrect answers. Follow these troubleshooting steps:

1. **Verify retrieval of correct source documents.**

   - If the correct answer *is not* present in any retrieved document, the failure is in retrieval. To fix:
     (a) Confirm that the relevant documents are included in your corpus.
     (b) Create a separate RAG index containing only those documents.
     (c) Query this index:
        - If the correct answer appears, but was missing in the full corpus, improve your retrieval model (e.g., use a larger or different embedding model).
        - If the correct answer still fails to appear, refine your indexing model.
   - If the correct answer *is* present in the retrieved documents, proceed to step 2.

2. **Assess the inference LLM.**

   - If relevant sources are retrieved but the generated answer is wrong, the inference LLM likely underperforms. Try a larger or alternative LLM.

3. **Escalate if unresolved.**

   - If the issue persists after improving retrieval and inference, report the problem with a non-confidential example for further investigation.

For a comprehensive list of RAG pipeline error sources, see the Colette technical note, p. 7:
COLETTEv2 Restitution 2025-03-07 (v0.3)

Reporting Issues

If you can't resolve a problem:

1. Check existing issues for similar reports on GitHub: https://github.com/jolibrain/colette/issues.

2. If none match, create a new issue including:

   - Description of the failure.
   - Example input and output.
   - Relevant logs or environment details.
   - Non-confidential document illustrating the error.

Installation and Runtime Errors

For errors during installation or execution:

1. Review the known issues list.

2. If missing, submit a new issue with:

   - Installation steps and environment (OS, Python version, etc.).
   - Full error logs and stack traces.