

1. SOP Checklist for RAG Implementation within Enterprises

Phase 1: Preliminary Preparation

1. Needs Analysis Meeting

- Determine use cases and target users.
- Identify key business processes and pain points.
- Define expected outcomes and success metrics.

2. Data Collection and Evaluation

- Identify relevant data sources (PDFs, Word, Excel, databases, internal websites, etc.).
- Evaluate data quality and completeness.
- Determine sensitive data handling solutions.

3. Technical Environment Assessment

- Evaluate existing IT infrastructure.
- Determine deployment environment (cloud/on-premises).
- Assess security and compliance requirements.

Phase 2: System Development

4. Data Processing Workflow Establishment

- Data cleaning and formatting.
- Establish data indexing pipelines.
- Implement data update mechanisms.

5. RAG Core Development

- Select appropriate embedding models.
- Design retrieval strategies.
- Fine-tune prompt templates.

6. UI/UX Design

- Design user interface.
- Implement interaction flows.
- Conduct usability testing.

Phase 3: Testing and Deployment

7. System Testing

- Functional testing.
- Performance testing.
- Security testing.

8. User Training

- Write user manuals.
- Conduct training sessions.
- Establish feedback mechanisms.

9. Deployment and Launch

- Phased deployment.
- Monitor system performance.
- Collect user feedback.

Phase 4: Maintenance and Optimization

10. Continuous Improvement

- Regularly update the knowledge base.
- Optimize based on user feedback.
- Expand new features.

2. Complete Operation Flow and File Descriptions

├─ app.py

├─ private_gpt.py

├─ ingest.py

├─ requirements.txt

├─ .env.example (Not deployed, but convenient for recording variables)

├─ logo.png (Optional)

└─ source_documents/ (Can be empty)

3. Streamlit Cloud Deployment Solution

Prepare a professional, enterprise-grade Streamlit UI with the following features:

- Intuitive search interface.
- Conversational interaction.
- Display of result sources.
- Enterprise brand customization.
- Responsive design.

Deployment Steps

1. Register for a Streamlit Cloud account (free version).
2. Connect to the GitHub repository.
3. Select `src/app/main.py` as the entry file.
4. Set up environment variables (API keys, etc.).
5. Click deploy.

4. Enterprise Brand Customized UI

Design the UI to be easily customizable for enterprise brand styles:

- Replaceable logo.
- Customizable color schemes.
- Professional enterprise-style layout.
- Responsive design adapting to different devices.