

Analyse critique du projet QwenPDF Insight

Ce rapport présente une évaluation honnête et complète du projet QwenPDF Insight, avec les axes atteints et les améliorations nécessaires pour le rendre parfaitement abouti.

■ Ce qui est déjà solide :

- Pipeline RAG multimodal fonctionnel (texte, tableaux, images)
- Modules séparés : extraction, embeddings, génération, reranking
- Intégration FastAPI, modèle Qwen3 Embedding et Qwen2-VL-7B
- Résumé, QA, extraction d'entités, détection automatique des intentions
- Projet bien structuré, clean, modulaire

■ Ce qui peut être amélioré pour viser l'excellence :

1. Structure du projet :

- Ajouter un fichier `core_pipeline.py` ultra propre et orchestrateur
- Ajouter des tests unitaires (dossier `tests/`)
- Améliorer le README (architecture, démonstration, cas d'usage)

2. Intelligence du système :

- Ajouter un classifieur de prompt pour router dynamiquement
- Détection automatique des KPI, entités, noms, montants, graphiques
- Fusion texte + image dans la génération (prompting enrichi)

3. Interface utilisateur :

- Créer une interface web moderne (Streamlit avancé, React ou autre)
- Ajout d'un historique, affichage des chunks, visualisation interactive
- Logique de sessions multi-documents

4. Évaluation :

- Feedback utilisateur manuel
- Ajout de métriques de qualité (temps, pertinence des chunks)