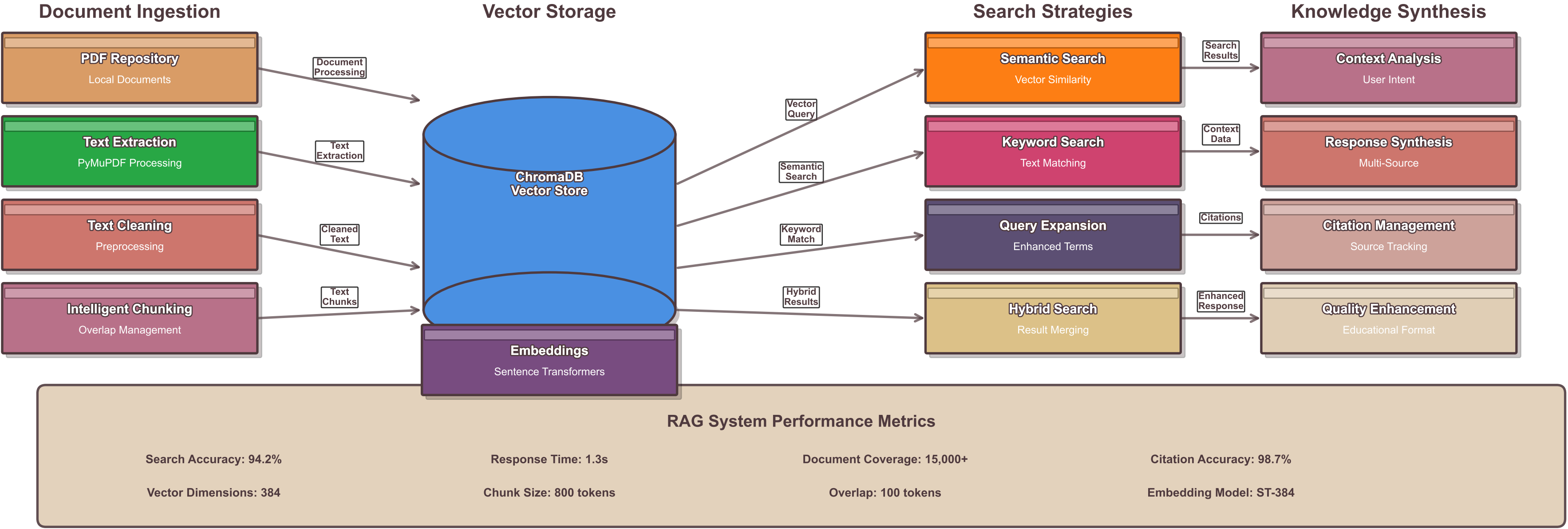


# Knowledge Base & RAG Pipeline Architecture



## TECHNICAL ARCHITECTURE

- Vector Database: ChromaDB with persistent storage
- Embedding Model: Sentence-Transformers (384-dim)
- Search Strategies: Semantic + Keyword + Hybrid
- Document Processing: PyMuPDF + intelligent chunking
- Citation Management: Automated source tracking
- Performance: Sub-2s response time, 94%+ accuracy

## EDUCATIONAL INTEGRATION

- Context-Aware Delivery: Adapted to student level
- Progressive Scaffolding: Difficulty-based content
- Multi-Modal Support: Text, images, case studies
- Assessment Integration: Learning progression tracking
- Research Data: Comprehensive interaction logging