

Multimodal AI Assistant Test Document

This PDF demonstrates the multimodal processing capabilities of the Echoes AI Assistant.

Key Features:

- Text extraction from PDFs
- Image processing with OCR
- Audio transcription
- JSON structured data handling
- Enhanced RAG with multimodal embeddings

The system can now process various file types to build comprehensive knowledge bases and provide r

Technical Implementation:

- Uses PyMuPDF for PDF processing
- Tesseract OCR for image text extraction
- Speech recognition for audio files
- JSON parsing for structured data
- Unified multimodal processor interface