

Technical Specification Document

TECHNICAL SPECIFICATION - AI DOCUMENT PROCESSOR

1. System Overview:

The AI Document Processor is a cloud-based solution for intelligent document analysis. It uses advanced machine learning algorithms to extract, classify, and analyze text data.

2. Architecture:

- Frontend: React-based web application
- Backend: FastAPI microservices
- ML Engine: Transformer-based models
- Database: PostgreSQL with vector storage
- Cache: Redis for performance optimization

3. Key Features:

- Multi-format support (PDF, DOCX, images)
- Real-time text extraction
- Natural language question answering
- Entity recognition and classification
- Sentiment analysis
- Multi-language support

4. Performance Requirements:

- Processing speed: < 5 seconds per document
- Accuracy: > 95% for text extraction
- Concurrent users: 1000+
- Uptime: 99.9% SLA

5. Security:

- End-to-end encryption
- GDPR compliant
- SOC 2 certified
- Regular security audits