AutoPipelineAI v0.3.0 - System Architecture

LLM-Driven Agentic Framework for Autonomous ETL and DataOps

 100% Local Execution | Multi-Agent System | Natural Language Interface | PDF Intelligence | High-Performance DB | ⚡ Safe Code Execution

User Interface Layer

Streamlit Web Application | Interactive Dashboard | Real-time Updates

Mode Router & Dispatcher

Manual Mode	LLM Mode	Agent Mode	PDF Mode	Database Mode

ETL Module
Data Loading
Transformation
Schema Validation

Ollama Client
llama3, mistral, phi3
Prompt Templates
Code Executor

Agent Orchestrator
• ETL Agent
• Query Agent
• Profiling Agent

PDF Extractor
Text & Tables
NER Processor
Entity Extraction

DB Handlers
DuckDB
SQLite
SQL Interface

Cache Manager
 Memory Cache
 Disk Cache
TTL Management

Memory Store
 Conversation History
 Session State
Context Management

Configuration
⚙ Environment Variables
 Directory Setup
Settings Management

Security Layer
 Code Sandbox
 Safety Checks
Module Whitelist

DuckDB Analytics

SQLite Storage

Parquet Files

CSV Files

Legend
- Direct Data Flow
- Request Flow
- Bidirectional
- User Interface
- Modes
- Core Modules
- LLM Services
- Agents
- Storage

AutoPipelineAI v0.3.0 | Built with Python, Streamlit, LangChain, CrewAI, Ollama | 100% Local & Private | Open Source (MIT License)