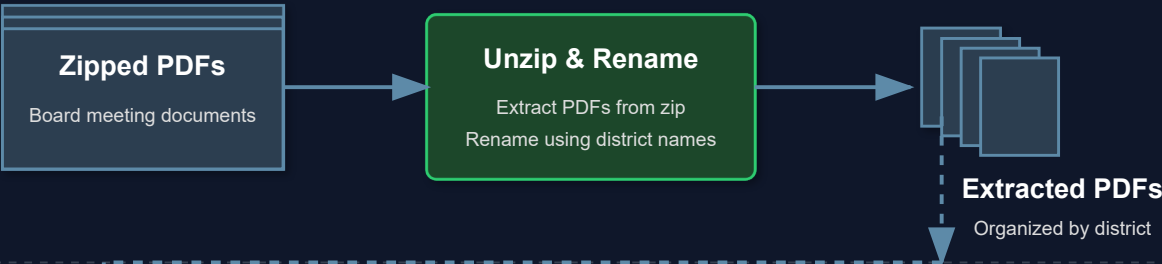
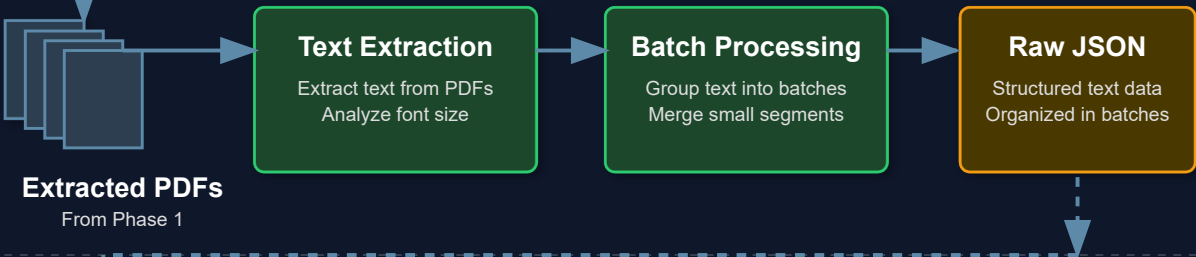


PDF to Batched JSON ETL Pipeline

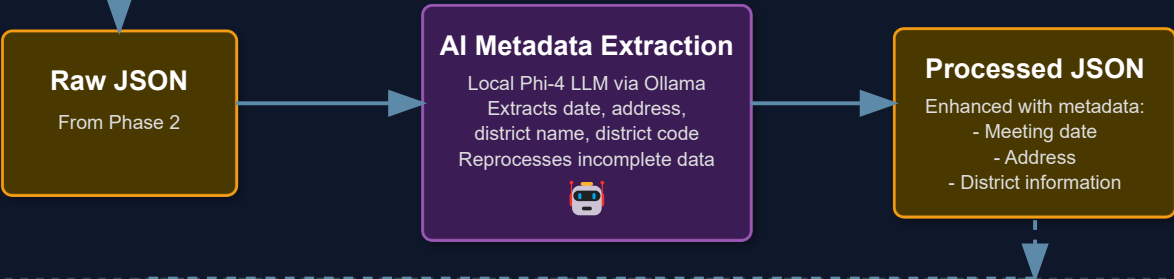
Phase 1: Data Extraction



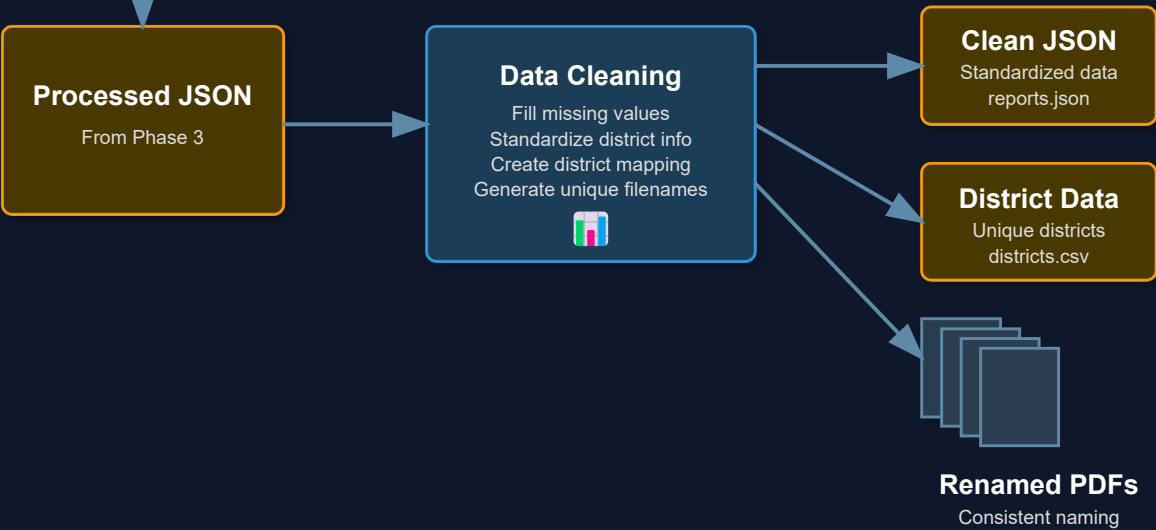
Phase 2: Text Extraction & Batch Processing



Phase 3: Metadata Extraction with LLM



Phase 4: Data Cleaning & Organization



Complete ETL Pipeline for School District Board Meeting Documents