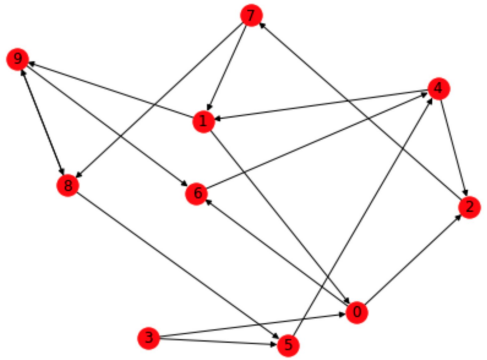
Transforming Structured Finance Compliance with AI

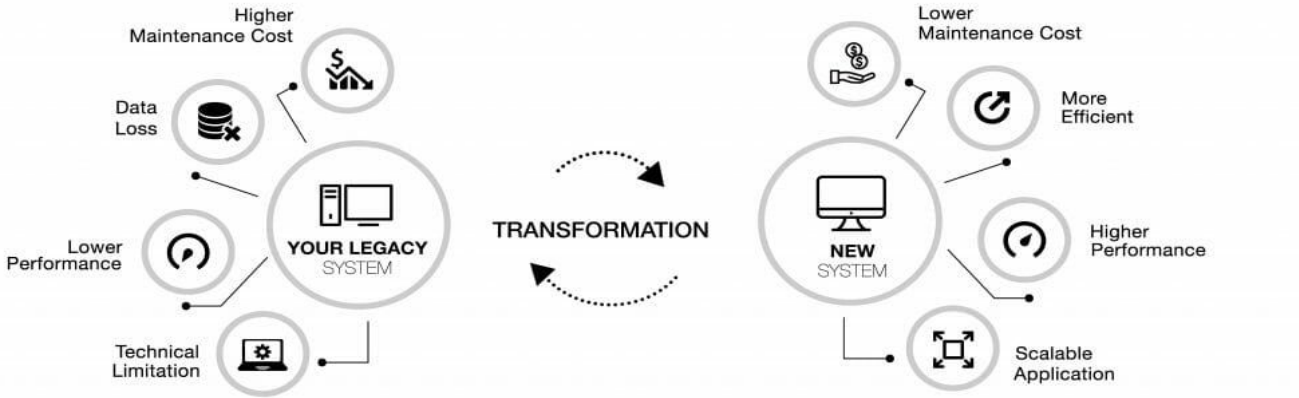
Automating the Conversion of Unstructured Financial Data into Structured, Regulatory-Compliant Assets

Hackathon Name & Date: Structured Finance Hackathon 2025
Team/Project Name: TEAM TONIC



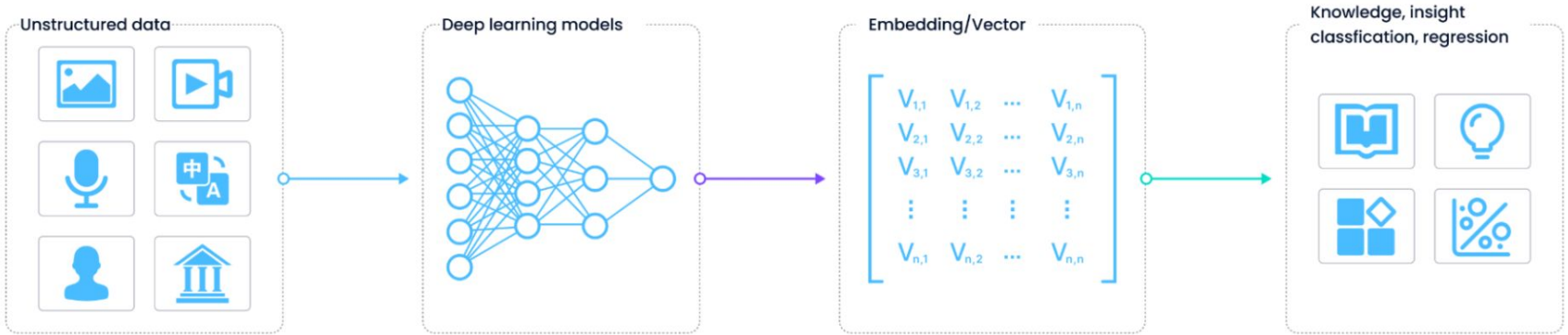
The Challenge in Structured Finance

- Manual processes in complex financial deal structuring, compliance, and risk management
- Legacy systems using Java & Rosetta complicating modern automation
- A 350-page ESMA compliance report that must be translated into machine-readable format



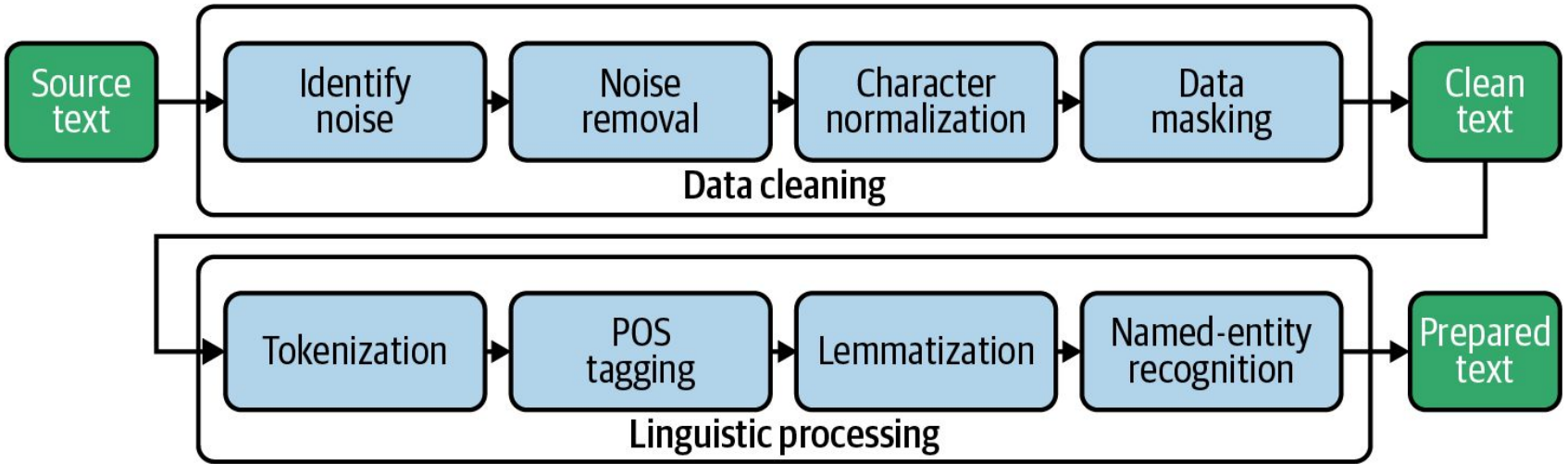
Our Vision

- Develop an AI model that reads raw, unstructured financial data and generates structured, compliant outputs
- Bypass legacy “monolithic” systems by leveraging modern frameworks (e.g., FastAPI, Python)
- Bridge the gap for regulatory compliance through automated data extraction and transformation



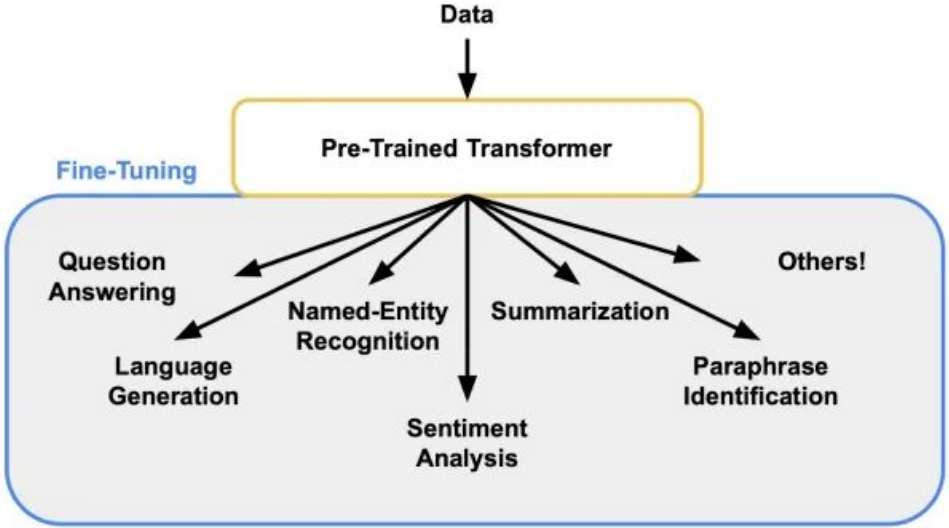
Phase 1: Converting Raw Data to an Annotated Dataset

- **PDF Ingestion & Conversion:**
  - Use PDF-to-LaTeX/Markdown tool to convert the ESMA report and other research sources
- **Data Chunking:**
  - Segment converted text into manageable parts and load into a Pandas DataFrame
- **Annotation & Analysis:**
  - Apply NLP (e.g., NER using spaCy or Hugging Face) to label key entities and regulatory markers
- **Benchmarking:**
- Establish baseline performance metrics for compliance extraction (precision, recall, etc.)



Phase 2: From Annotated Data to a Fine-Tuned Model

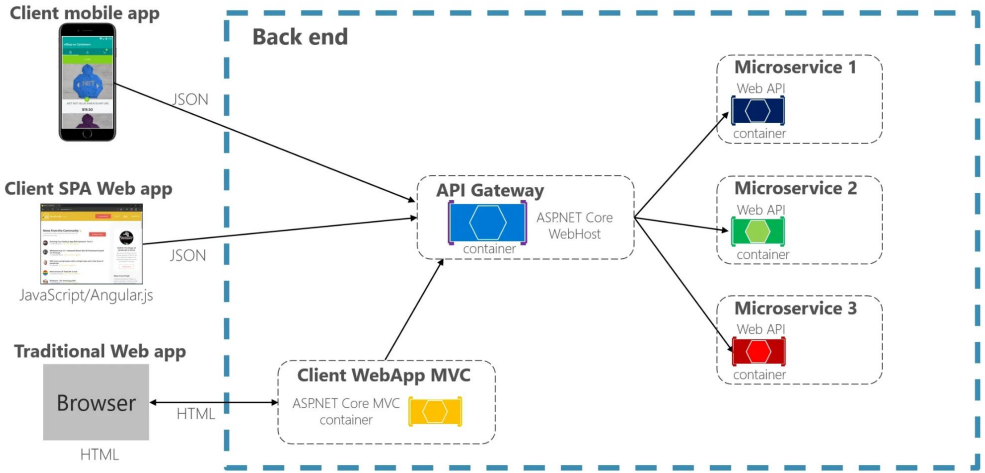
- **Dataset Transformation:**
  - Convert the enriched, annotated dataset into a ChatML-compatible training format
- **Model Selection & Fine-Tuning:**
  - Utilize the Osmosis-Structure-0.6B model for structured data tasks
  - Fine-tune the model on your specialized dataset for extracting structured financial events
- **Evaluation Benchmark:**
- Develop an evaluation benchmark to test the model's performance in generating compliant outputs



System Integration & API Development

- **FastAPI Server:**
  - Create a lightweight, modern API to expose model functionalities
  - Allow seamless integration with external systems and agentic models
- **Demo Application:**
- Build a demo app that pulls raw data from the Open Bank Project Sandbox
- Showcase how the API transforms this data into structured financial asset

Using a single custom **API Gateway service**



Expected Impact & Future Vision

- Rapid transformation from unstructured regulatory text to structured financial data
- Elimination of dependence on legacy systems (Java/Rosetta)
- A scalable pipeline that can ultimately integrate seamlessly with full DRR workflows and regulatory system