

# Custom LLM Solution Architectures

---

A Practical Roadmap for Technology Professionals

---

Today's Presenter:



**STEVEN HOLT**

DATA SERVICES CONSULTANT  
@  
ONLINE BUSINESS SYSEMS

---

**Results. Guaranteed.**

**online**  
business systems

09/23/2024

# PRAIRIE DEV CON

WEB | DEV | CLOUD | AI



EXCELLENCE IN RECRUITMENT



# The “*A Little About Me*” Slide...

- Manitoba born and raised.
- ~15-years working in Data & Advanced Analytics
- Sr. Data Architect & Data Competency Lead for Data Services @ Online Business Systems
- Member of “The Innovation Lab” @ Online Business Systems (AI/GenAI & Web3)

# AGENDA

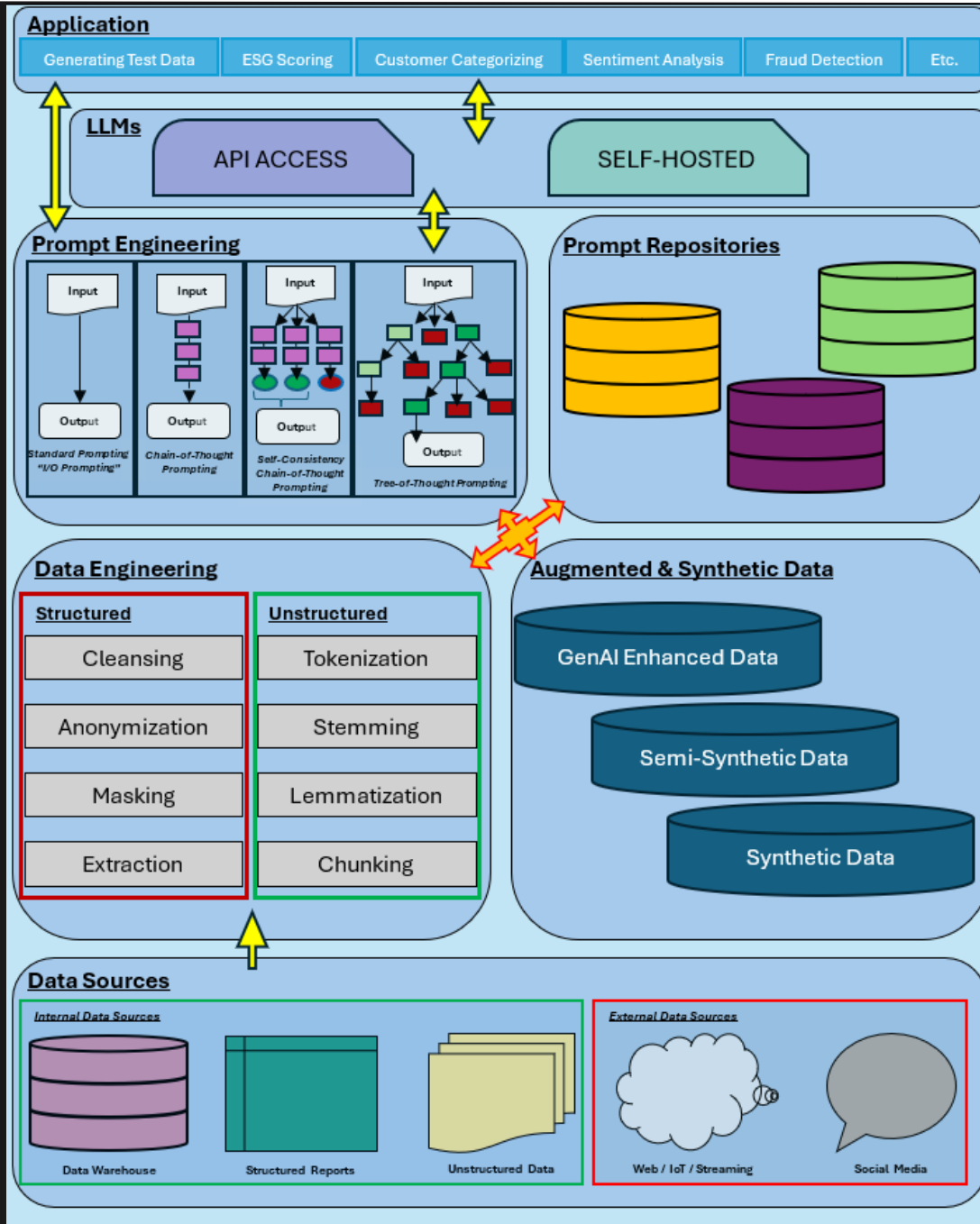
- 01 Introduction: The Full Picture
- 02 Layer 1: *Data Sources*
- 03 Layer 2: *Data Engineering / Augmented & Synthetic Data*
- 04 Layer 3: *Prompt Engineering / Prompt Repositories*
- 05 Layer 4: *LLMs*
- 06 Layer 5: *Applications*
- 07 Wrap-Up & Questions??

*But first ...*

Some terms to get familiar with.

- LLM (Large Language Model)
- GPT (Generative Pre-Trained Transformer)
- Prompt





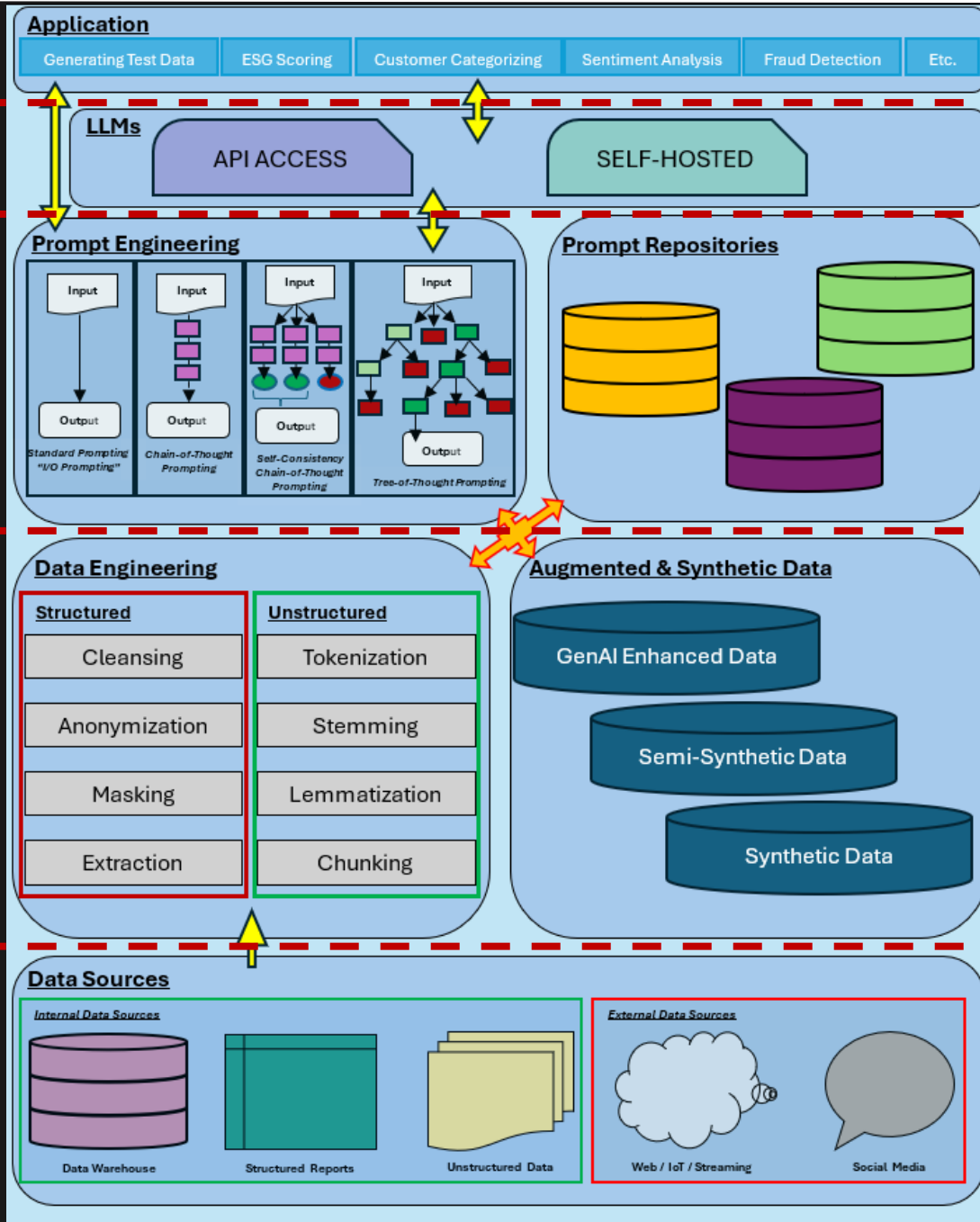
Layer 5

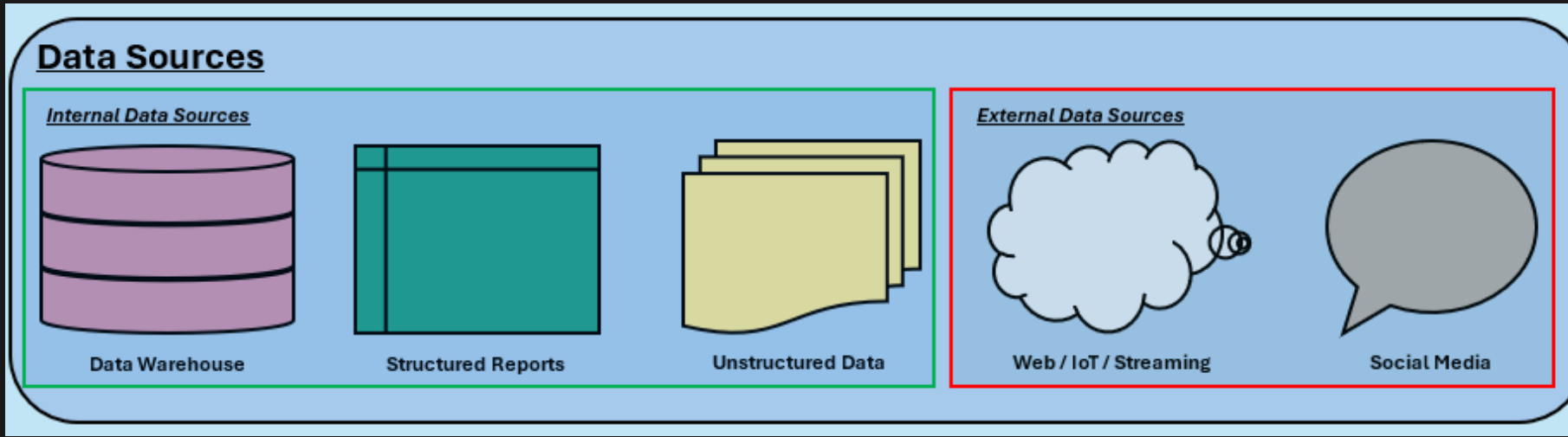
Layer 4

Layer 3

Layer 2

Layer 1





## Layer 1: Data Sources

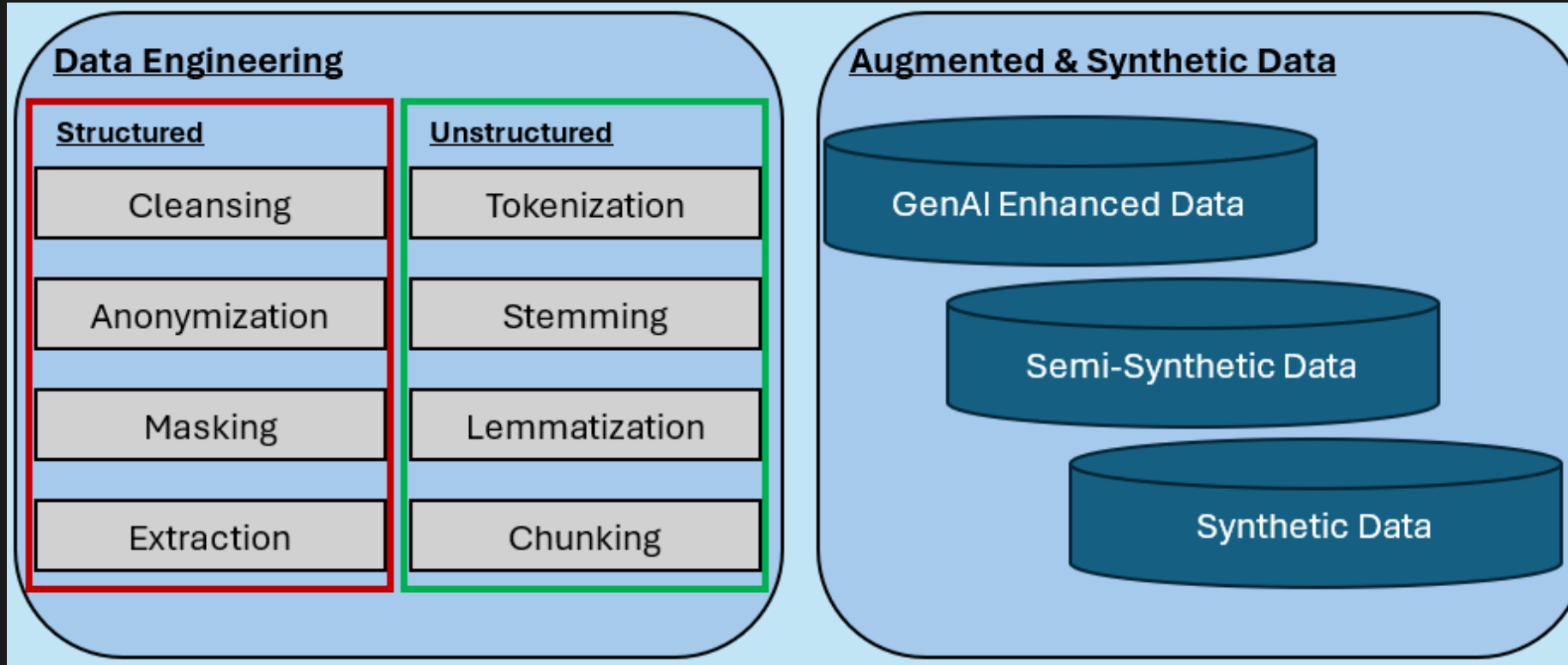
### Internal Data Sources:

- Data Warehouse(s)
- Structured Reports
- Unstructured Data

### External Data Sources:

- Web
- IoT
- Streaming Data
- Social Media





## Layer 2: Data Engineering / Augmented & Synthetic Data

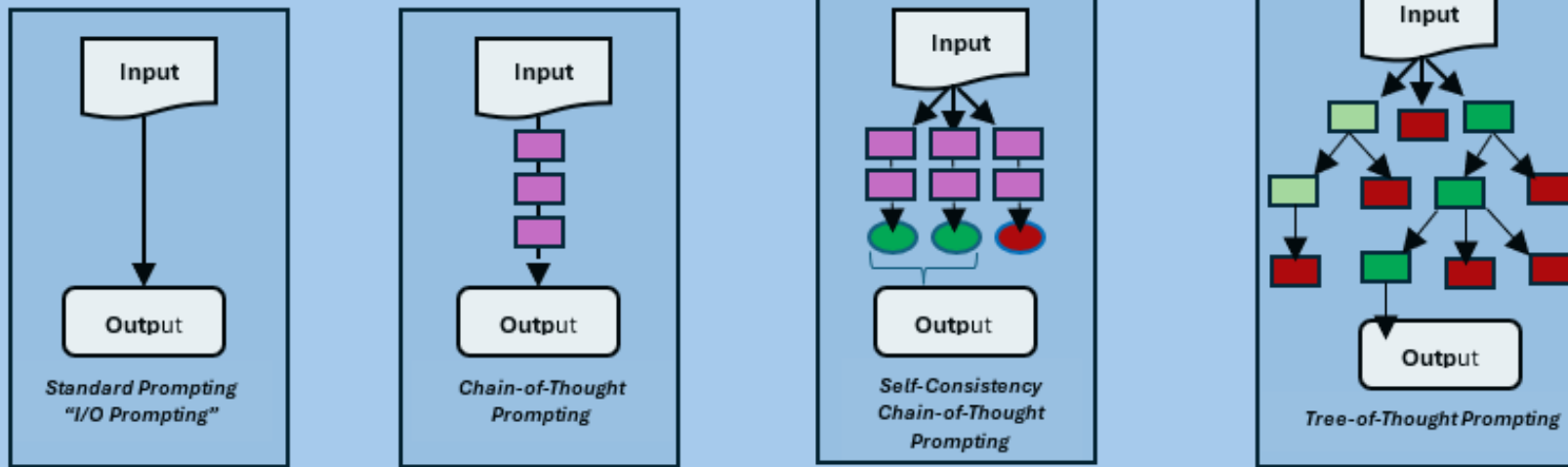
### Data Engineering:

- Data Extraction
- Data Masking/Anonymization
- Data Cleansing

### Augmented & Synthetic Data:

- Generative AI Enhanced Data
- Semi-Synthetic Data
- Synthetic Data

## Prompt Engineering



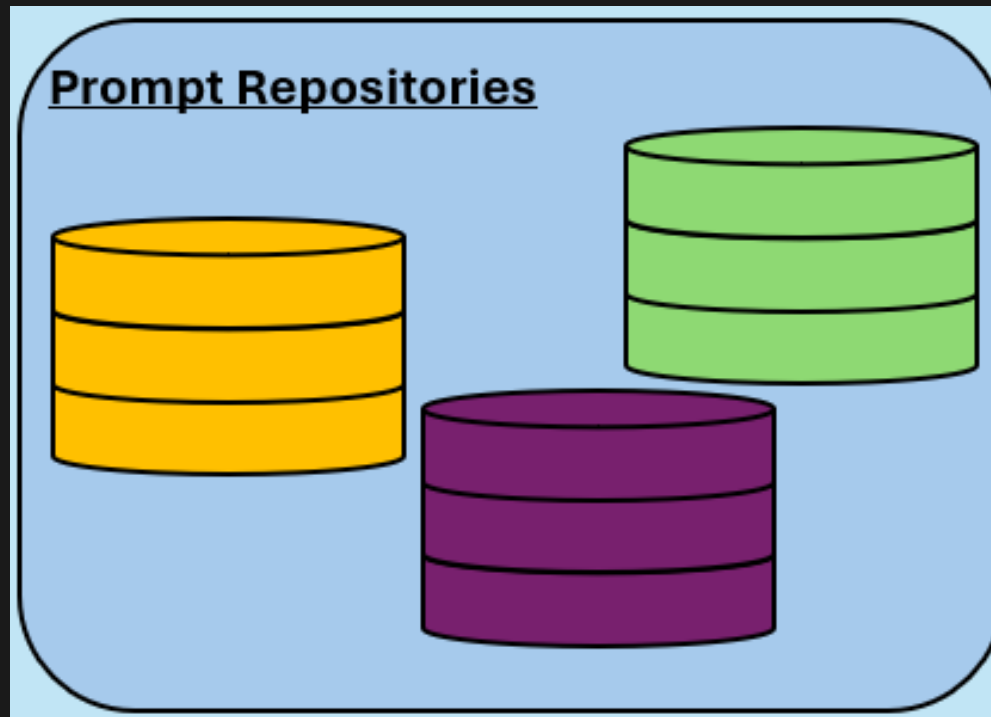
## Layer 3: Prompt Engineering

### Simple Prompt Engineering:

- I/O Prompting
- Chain-of-Thought Prompting
- Interactive Learning Through Dialog Prompting
- One-Shot/Few-Shot Prompting
- Counterfactual Prompting ("What if...?")

### Complex Prompt Engineering:

- Self-Consistency Chain-of-Thought Prompting
- Tree-of-Thought Prompting
- Graph-of-Thought Prompting
- Recursive Prompting
- Meta Prompting (Prompting for Prompts)
- Hybrid Model Prompting (e.g. Reflection)



## Layer 3: Prompt Repositories

### Simple Prompt Engineering:

- I/O Prompting
- Chain-of-Thought Prompting
- Interactive Learning Through Dialog Prompting
- One-Shot/Few-Shot Prompting
- Counterfactual Prompting ("What if...?")

### Complex Prompt Engineering:

- Self-Consistency Chain-of-Thought Prompting
- Tree-of-Thought Prompting
- Graph-of-Thought Prompting
- Recursive Prompting
- Meta Prompting (Prompting for Prompts)
- Hybrid Model Prompting



## Layer 4: LLMs

### API Access Models:

- GPT-3.5/GPT-4 (OpenAI)
- Gemini (Google)
- Copilot (Microsoft)

### Self-Hosted Models:

- Llama/Llama 2 (Facebook)
- XGen (Salesforce)
- BERT (Google)

### Application

Generating Test Data

ESG Scoring

Customer Categorizing

Sentiment Analysis

Fraud Detection

Etc.

## Layer 5: Application

### Operational Applications:

- Generating Synthetic Test Data
- Data Categorization
- Fraud Detection
- Forecasting

### Interactive Applications:

- Hypothetical Analysis
- A/B Testing & Evaluation
- Solution/Design Advisor
- Employee Training/Mentorship

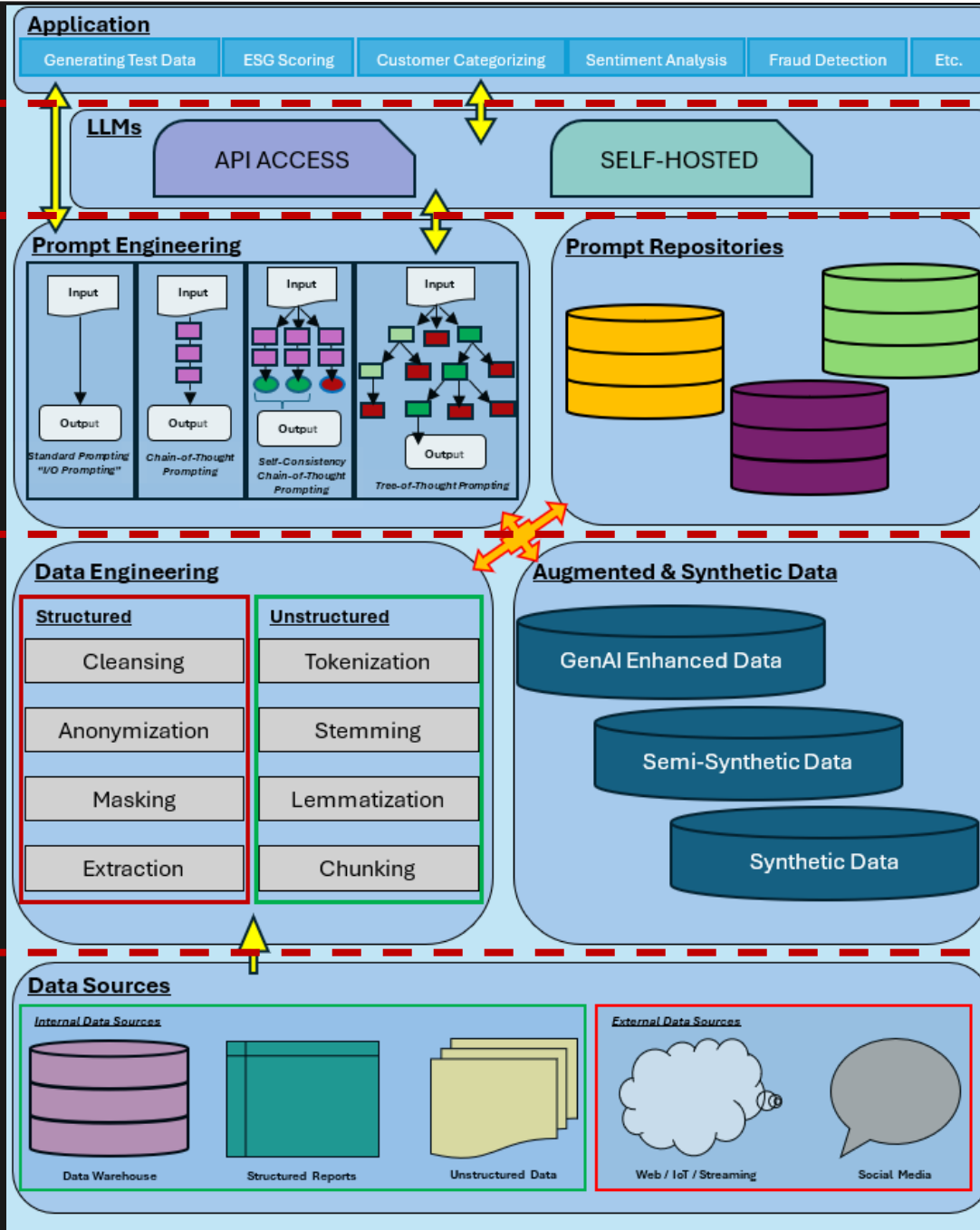
Layer 5

Layer 4

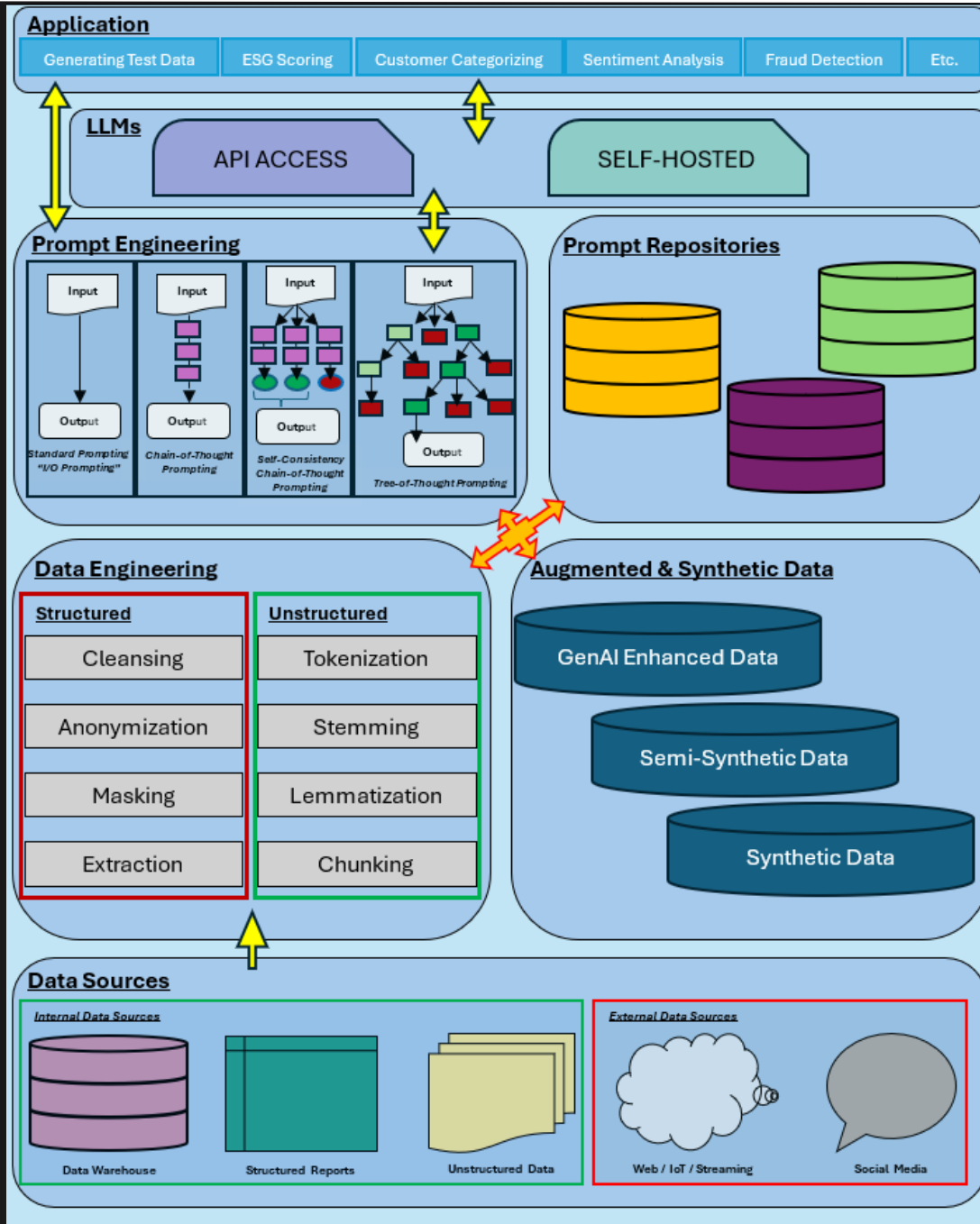
Layer 3

Layer 2

Layer 1







# Final Thoughts...

# Questions?



# Thank You!