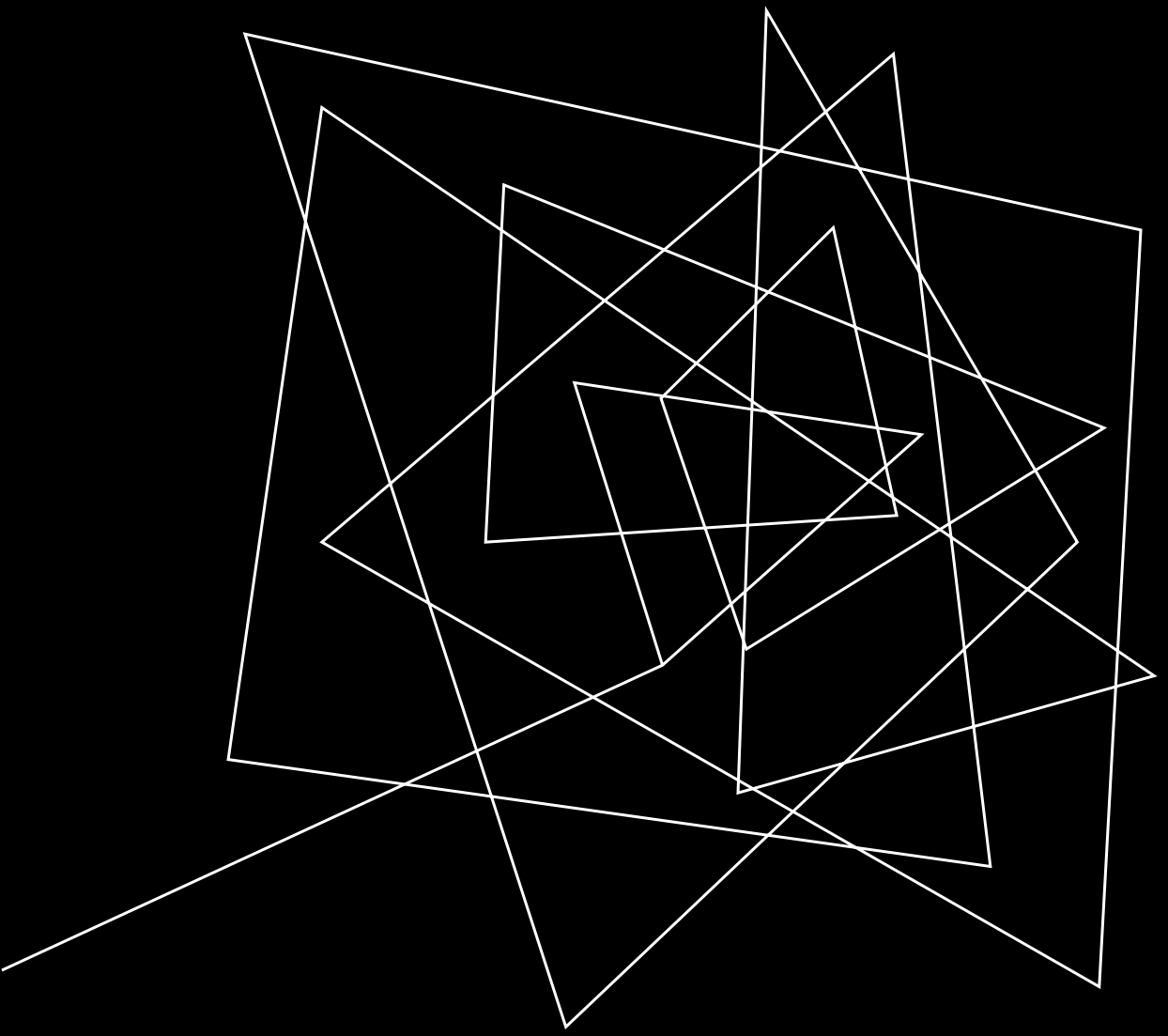
UNLEASHING THE POWER OF LLMS: REVOLUTIONIZING DRUG SAFETY AND BUSINESS AUTOMATION WITH MULTI-AGENT AI

#ATO - OpenAI 2025

Jeffery Painter, Sr. Director, Safety Innovation and Analytics, Quantitative leader for Global Safety at GSK. 18+ years in drug safety and statistical genetics. Background in software development, math and statistics. BS in CS and applied math (NCSU), JD (NCCU School of Law), MS CompSci (Georgia Tech). ASF member and contributor since 2002.

Venkateswarao Chalamalasetti, Sr. Manager and SME in Drug Safety Data & Systems. 10+ plus years of experience in PV systems and Exploring AI and ML in Pharmacovigilance. An Employee of Tech Mahindra and a contract of the GSK. Master of Computer Application (MCA).

Disclaimer: The views expressed in this presentation reflect the personal views of the authors and do not necessarily reflect the views of the authors' employers, nor any other institutions the author may otherwise be collaborating with. JP holds equity shares in GSK.



DRUG SAFETY

Pharmacovigilance

DRUG DEVELOPMENT PROCESS

Figure 1 – Drug development process (illustration by Jennifer Harris)

*Figure Referenced from https://sites.utexas.edu/melamed-lab/drug-development-cycle/

PHARMACOVIGILANCE: A HIGHLY REGULATED FIELD

- **Pharmacovigilance (PV)**
  - The science and activities relating to the detection, assessment, understanding, and prevention of adverse effects or any other drug-related problems.
    - FDA Definition: https://www.fda.gov/media/152487/download#:~:text=The%20science%20and%20activities%20relating,2002
- **Key Components of Individual Case Safety Report (ICSR):**
  1. a patient
  2. a reporter
  3. a suspect drug and
  4. an adverse event
- **Importance of Drug Safety**
  - Ensures public health, regulatory compliance and safety of medicines in the global market

Pharmacovigilance is primarily concerned with ***Adverse Events***

- An unwanted or unintended side effect resulting from taking a medication for its intended use, not necessarily harmful or serious

X

DRUG SAFETY TIMELINE

1960s

Thalidomide: tragic birth defects from anti-nausea medication leads to systematic testing of pharmaceutical products for developmental toxicity prior to marketing

Thalidomide: The Tragedy of Birth Defects and the Effective Treatment of Disease, 2011, J. Kim and A. Scialli

1970s

Birth of electronic medical records (EMR) begins at the Mayo Clinic in Rochester, Minnesota

1980s

EMR systems became standard in large hospital systems. Desktop computing enables private practice to begin capturing patient data

1990s

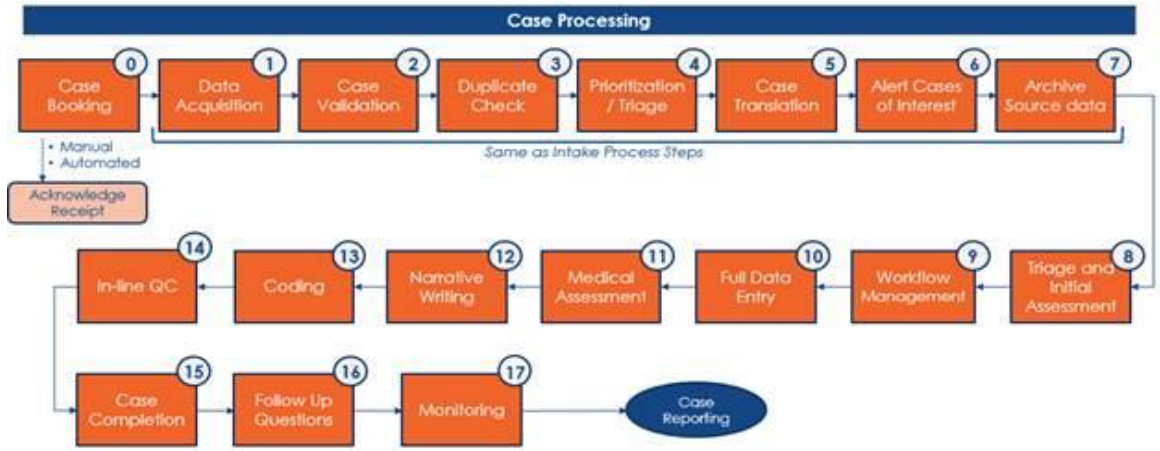
Rise of “big data” leads to development of advanced statistical methods and first use of ML for drug safety

Global and national safety report databases begin in earnest (FAERS, Vigibase, WHO-AERS)

A Bayesian neural network method for adverse drug reaction signal generation, 1998, A. Bate, et al.

DRUG SAFETY CENTERS ON DATA PROCESSING AND MONITORING

- Multiple Intake Channels:
  - Reports originate from email, call centers, social media, clinical trials, etc.
- Case processing and compliance
  - Case Intake and Duplication Detection
  - Prioritization and Triage
  - Alerting Cases of Interest
  - Medical Coding
  - Follow Up
  - Monitoring
- Goals:
  - Increase efficiency, reduce manual effort, reduce errors in data processing



High Level Overview of Case Intake

FDA (FAERS) PUBLIC DASHBOARD

Year
10 of 55

FDA Adverse Events Reporting System (FAERS) Public Dashboard

U.S. FOOD & DRUG ADMINISTRATION

Home | Search | COVID-19 EUA | Disclaimer | Report a Problem | FAQ | Site Feedback

Total Reports
18,065,129

Serious Reports (excluding death)
9,551,210

Death Reports
1,574,026

Reports by Report Type

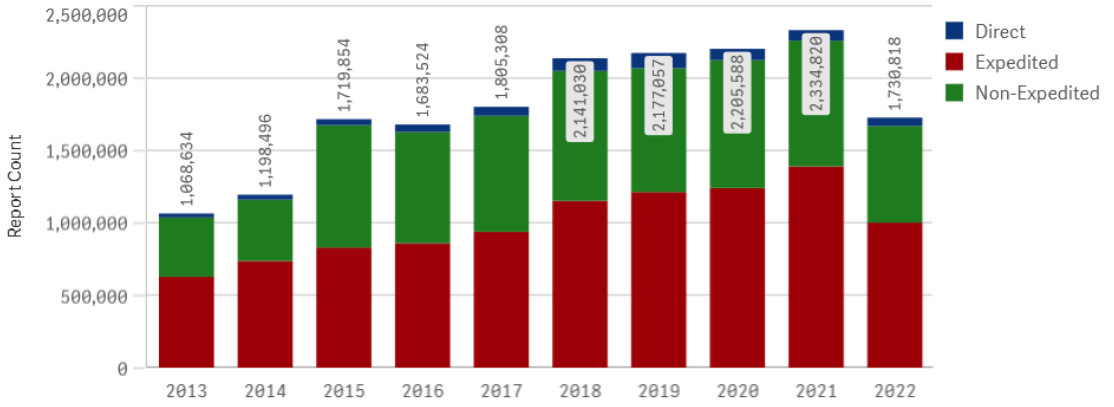
Since 1968 | Last 10 Years

Reports received by Report Type

Year | Report Type

	Total Reports	Expedited	Non-Expedited	Direct
Total Reports	**18,065,129**	**10,028,239**	**7,417,683**	**619,207**
2022	**1,730,818**	1,005,937	667,022	57,859
2021	**2,334,820**	1,393,711	868,560	72,549
2020	**2,205,588**	1,244,499	882,529	78,560
2019	**2,177,057**	1,216,611	855,058	105,388
2018	**2,141,030**	1,156,110	897,369	87,551

Reports received by Report Type



Data as of September 30, 2022

Vulnerability Disclosure Policy

This page displays the number of adverse event reports received by FDA for drugs and therapeutic biologic products by the following Report Types.

- Direct Reports are voluntarily submitted directly to FDA through the MedWatch program by consumers and healthcare professionals.
- Mandatory Reports are submitted by manufacturers and are categorized as:

ISSUES FACING PHARMACOVIGILANCE TODAY

More data than ever...

The COVID-19 pandemic and subsequent vaccines and medicines development led to more than 2.3 million new safety reports in 2021 alone.

Limited human resource...

Require automation to enable safety organizations to quickly process, evaluate and report safety findings to regulatory bodies.

Leveraging LLMs versus traditional ML

Traditional ML is time consuming and requires specific model building, continuous refinement and improvement.

LLM technology promises to enable faster capabilities to leverage AI across a variety of tasks without the need for custom building and fine tuning

LLM RESEARCH IN DRUG SAFETY

Building Trustworthy AI for Drug Safety

Safety Innovation and Analytics research publications

- Focus on Trustworthy AI: Ensuring reliability, reducing hallucinations and meeting regulatory requirements
  - Hakim, J. B., Painter, J. L., Ramcharran, D., Kara, V., Powell, G., Sobczak, P., ... & Beam, A. (2024). The Need for Guardrails with Large Language Models in Medical Safety-Critical Settings: An Artificial Intelligence Application in the Pharmacovigilance Ecosystem. *arXiv preprint arXiv:2407.18322*.
  - Stegmann, J. U., Littlebury, R., Trengove, M., Goetz, L., Bate, A., & Branson, K. M. (2023). Trustworthy AI for safe medicines. *Nature Reviews Drug Discovery*, *22*(10), 855-856.
- Enhancing Drug Safety User's Documentation Search Capabilities
  - Painter, Jeffery L., et al. "Enhancing drug safety documentation search capabilities with large language models: a user-centric approach." *2023 International Conference on Computational Science and Computational Intelligence (CSCI)*. IEEE, 2023.
- LLMs to enable automatic report generation:
  - Painter, J. L., Chalamalasetti, V. R., Kassekert, R., & Bate, A. (2025). Automating pharmacovigilance evidence generation: using large language models to produce context-aware structured query language. *JAMIA open*, *8*(1), ooaf003.

JAMIA Open, 2025, **8(1)**, ooaf003
https://doi.org/10.1093/jamiaopen/ooaf003
Research and Applications

AMIA OXFORD

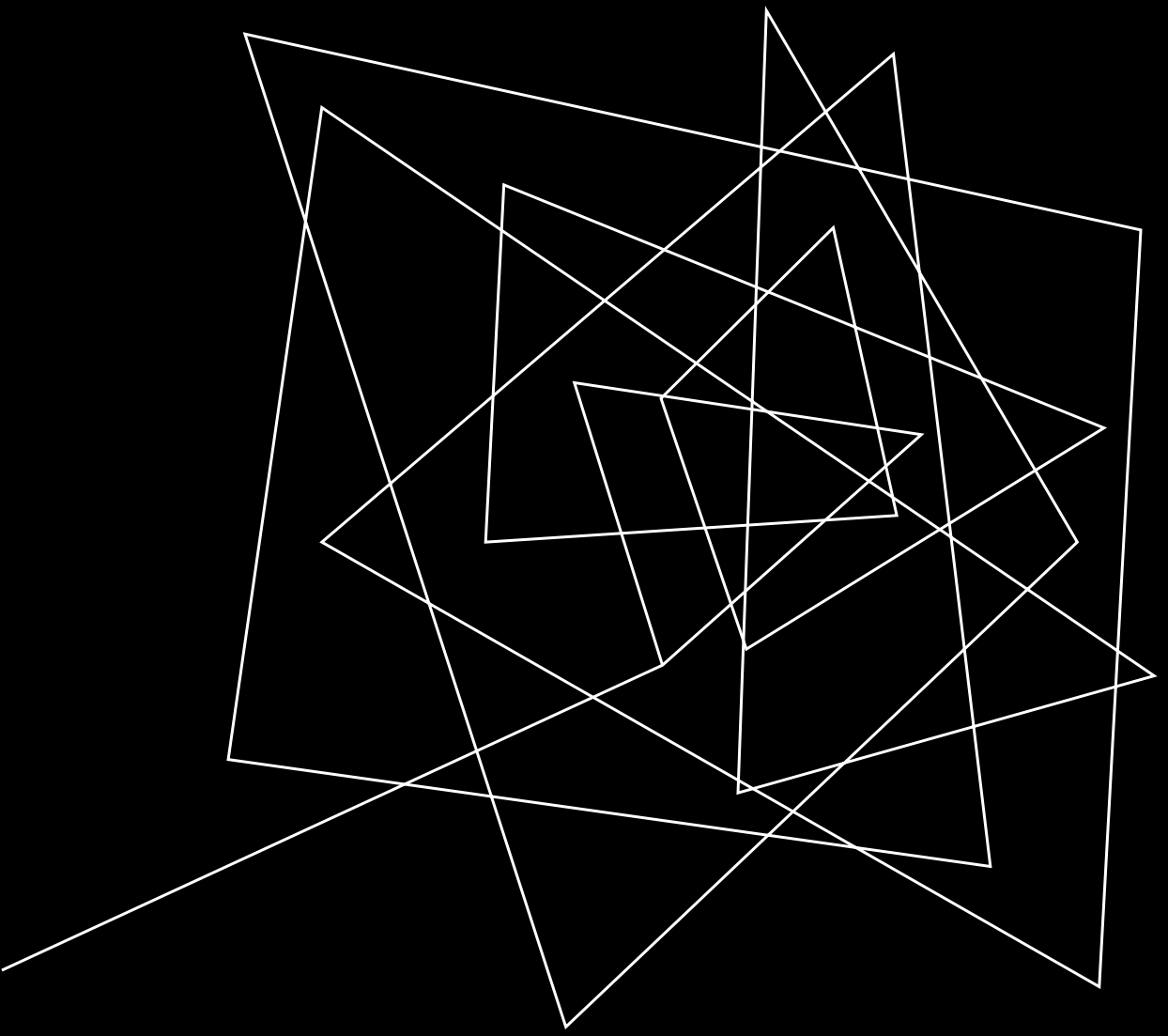
Research and Applications

Automating pharmacovigilance evidence generation: using large language models to produce context-aware structured query language

Jeffery L. Painter, MS, JD[1,*], Venkateswara Rao Chalamalasetti, MCA[1,2], Raymond Kassekert, MBA[3], Andrew Bate, PhD[4,5]

[1]GlaxoSmithKline, Durham, NC 27701, United States, [2]Tech Mahindra, Plano, TX 75024, United States, [3]GlaxoSmithKline, Philadelphia, PA 19104, United States, [4]GlaxoSmithKline, London WC1A 1DG, United Kingdom, [5]London School of Hygiene and Tropical Medicine, London WC1E 7HT, United Kingdom

*Corresponding author: Jeffery L. Painter, MS, JD, GlaxoSmithKline, Durham, NC 27701, United States (jeffery.l.painter@gsk.com)



MULTI AI AGENTS FOR DRUG SAFETY

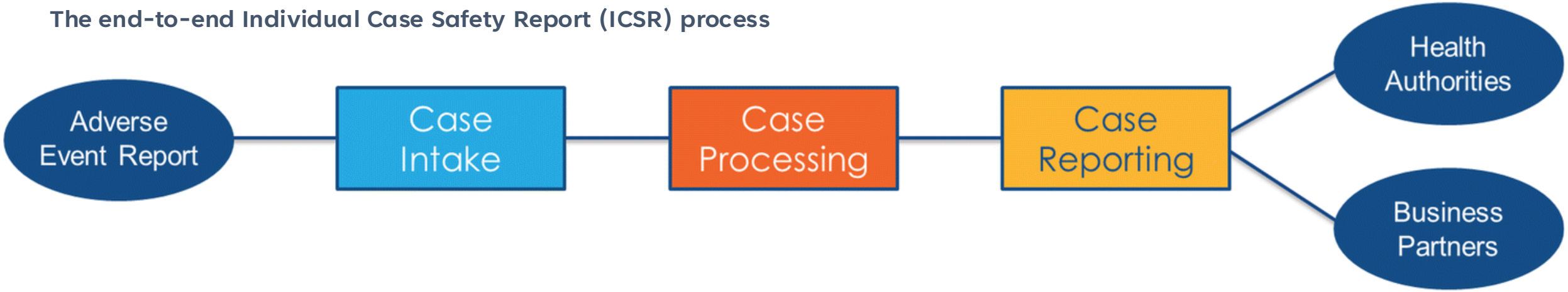
Venkateswarao Chalamalasetti

WHY MULTI AI AGENTS IN PHARMACOVIGILANCE (PV)

Vision for Multi-AI Agents in Pharmacovigilance

The vision is to revolutionize pharmacovigilance (PV) by deploying a network of intelligent, autonomous AI agents that seamlessly manage the entire drug safety lifecycle from case intake to regulatory reporting—enhancing efficiency, accuracy, and patient safety on a global scale.

The end-to-end Individual Case Safety Report (ICSR) process



Data Formats:

- Unstructured: emails etc.
- Semi Structured: CIOMS, Clinical forms etc.
- Structured: ICH E2B XML

User Roles: Intake Specialist, Intake Manager, Data Entry Specialist , Medical Review, Report Submissions etc.

Business Tasks: ICSR Triage, translation, duplicate checks, follow ups, narrative generation, drug coding, event coding , medical assessment, QCing, submission, archive, reconciliation etc.

Challenges:

- Growing Volumes
- Inefficient Data Handling
- Increasing cost
- Compliance Risks

Ghosh, R., Kempf, D., Pufko, A. et al. Automation Opportunities in Pharmacovigilance: An Industry Survey. Pharm Med 34, 7–18 (2020). https://doi.org/10.1007/s40290-019-00320-0

LARGE LANGUAGE MODELS (LLM) CAPABILITIES AND LIMITATIONS FOR PV

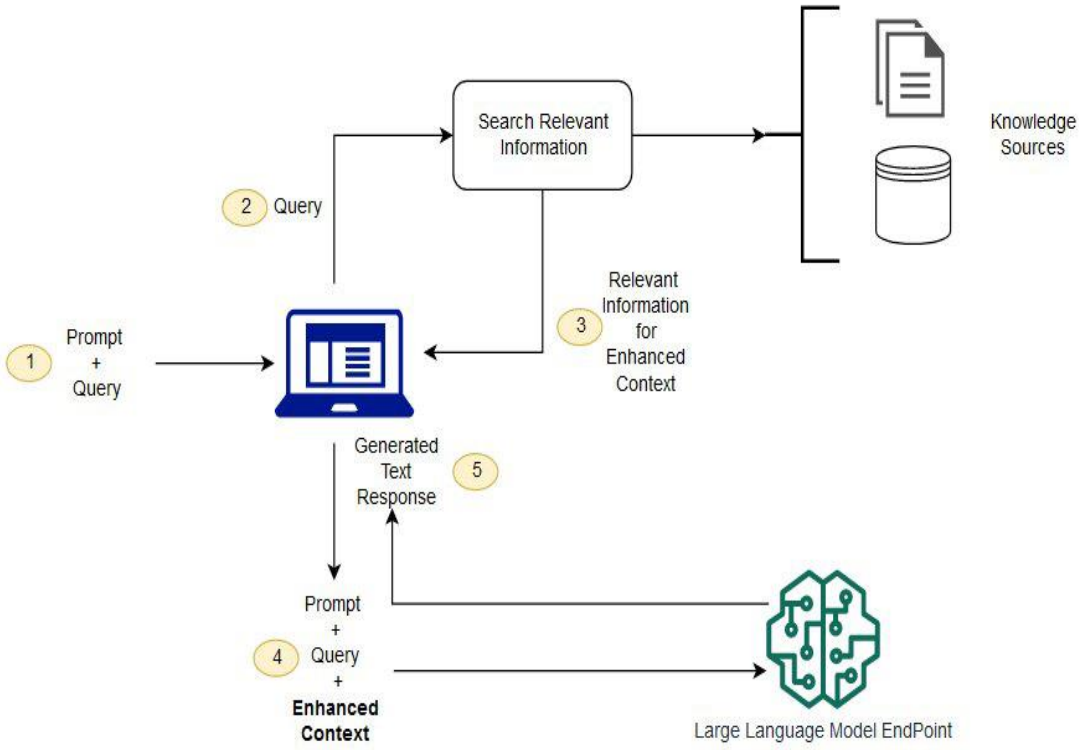
Prompt engineering is the process of designing and refining instructions to guide AI models like ChatGPT to produce specific, desired outputs. A **prompt** is natural language text describing the task that an AI should perform. A clear, precise, and detailed instructions to maximize the AI's effectiveness in solving tasks.

RAG with LLMs - *Could AI Make Business Decisions*

Input Prompt:

Can you extract patient, event and country from below text "John Doe, a US resident, filled a prescription for Hypertensinex in Canada but experienced adverse effects after returning to the US. The symptoms started in the US, and the case was reported by a doctor in New York. Based on these details," also determine the most likely Country of Occurrence for this adverse event?

(Synthetic Data Example)



Business Guidance, Rules and SOPs (Standard Operating Procedures)

(Context Window)

Model Response:

From the text provided:
- Patient: John Doe
- Event: Experienced adverse effects after taking Hypertensinex
- Country of Reporter: United States (doctor in New York)

The most likely **Country of Occurrence for this adverse event is the United States**, as the symptoms started after the patient returned to the US.

LLM Capabilities

Understanding ,Text Generation, Summarization, Text Classification, Data Extraction Question Answering, Translation, Code Generation etc.

LLM Limitations

Finite context length, don't have state/memory, Inability to Perform Actions, Hallucinations, Lack of Real-time Data Access. Lack of Explainability , probabilistic etc.

*Figure Referenced from https://doi.org/10.48550/arXiv.2402.06196

AUTONOMOUS AI AGENT, CHARACTERISTICS AND OVERVIEW

AI Agent: An artificial entity capable of perceiving its environment, make decisions and taking action can be considered an agent.

AI Agent characteristics:

- **Autonomy:** Agents independently perceive their environment, make decisions, and take actions without relying on external instructions.
- **Perception:** Agents are equipped with sensory capabilities that allow them to gather information about their environment through the use of sensors.
- **Decision-making: Agents make decisions** based on perceived information, selecting appropriate actions to achieve their goals.
- **Action: Agents perform actions** that alter the state of their environment.

How LLM agents extend the capabilities of base LLMs through several key components:

Planning Module: Breaking down complex tasks into manageable subgoals

Memory Module: Short-term memory and Long-term memory

Tool Module: Refers to any instrument an agent can utilize, such as API, software interfaces or services.

Environment: The context or surroundings in which the agent operates and makes decisions.

Objective: The primary goal, denoted as the Objective, represents the terminal state or condition that the agent must achieve.

* Figure Referenced from https://doi.org/10.48550/arXiv.2401.03428

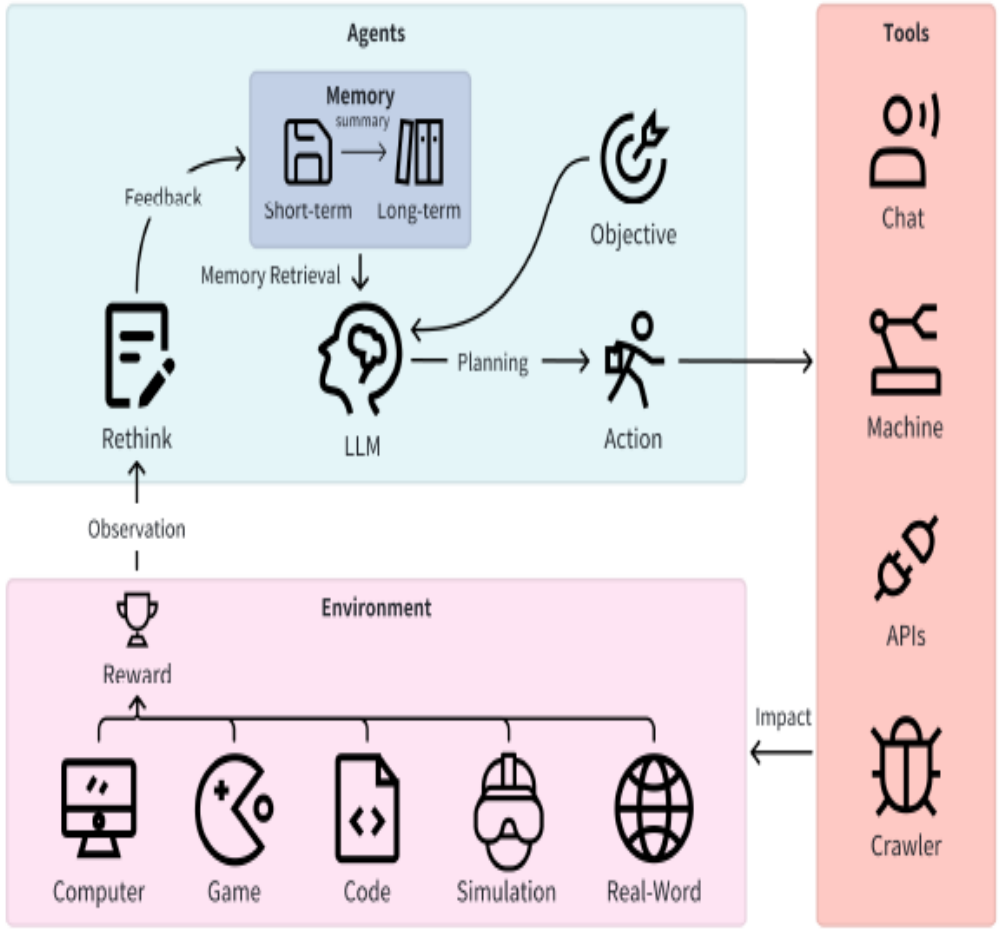


Figure 2: Overview of LLM-based agents

THE RELATIONSHIP BETWEEN LLM-BASED AGENTS AND ENGINEERING ASPECTS

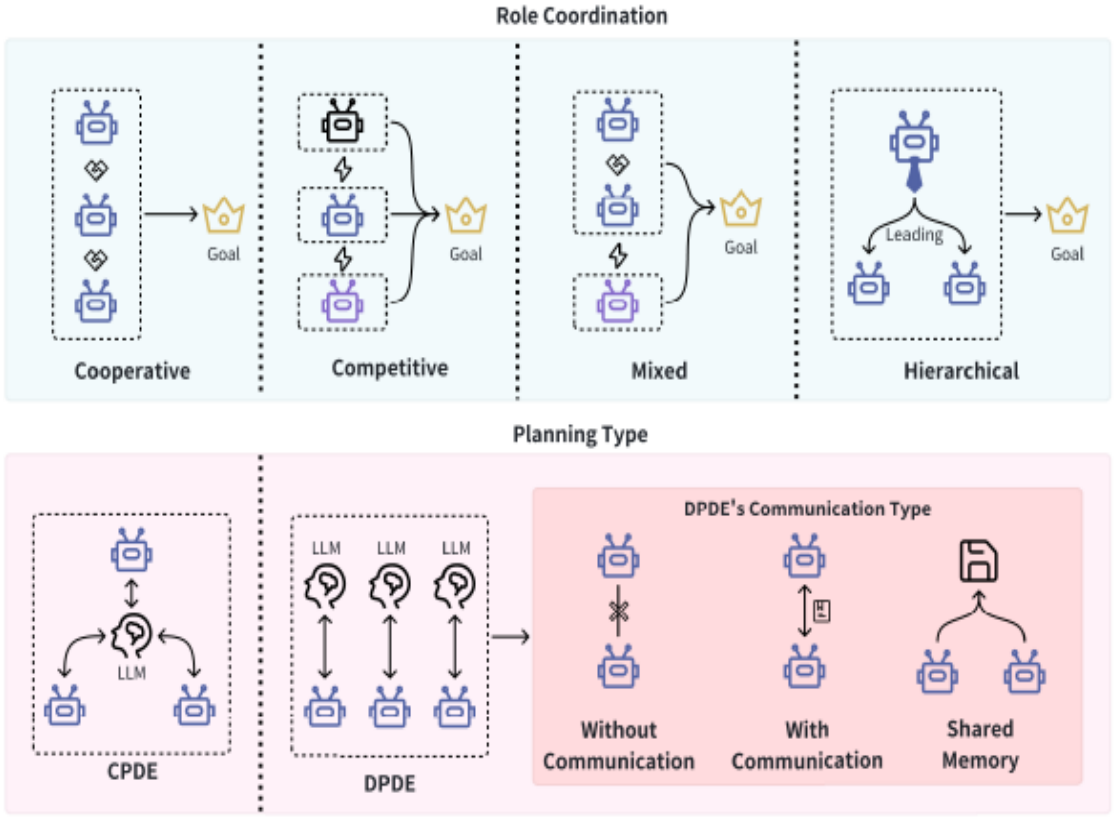


Figure 3: The Relationship between LLM-based agents

Agent Configuration: roles and capabilities of each agent in the workflow. Set up prompts, context, and constraints for each agent to guide their behavior.

Agents Type: Heterogeneous or Homogeneous.

Task Mode: Cooperative-like or Competitive-like.

Knowledge Sharing: Agent Level, Scenario Level, or Task Level.

Coordination: Cooperative, Competitive, Mixed, and Hierarchical;

Interaction Patterns: sequential, parallel, or hierarchical interaction patterns

Orchestration: centralized orchestrator or a decentralized system where agents communicate directly

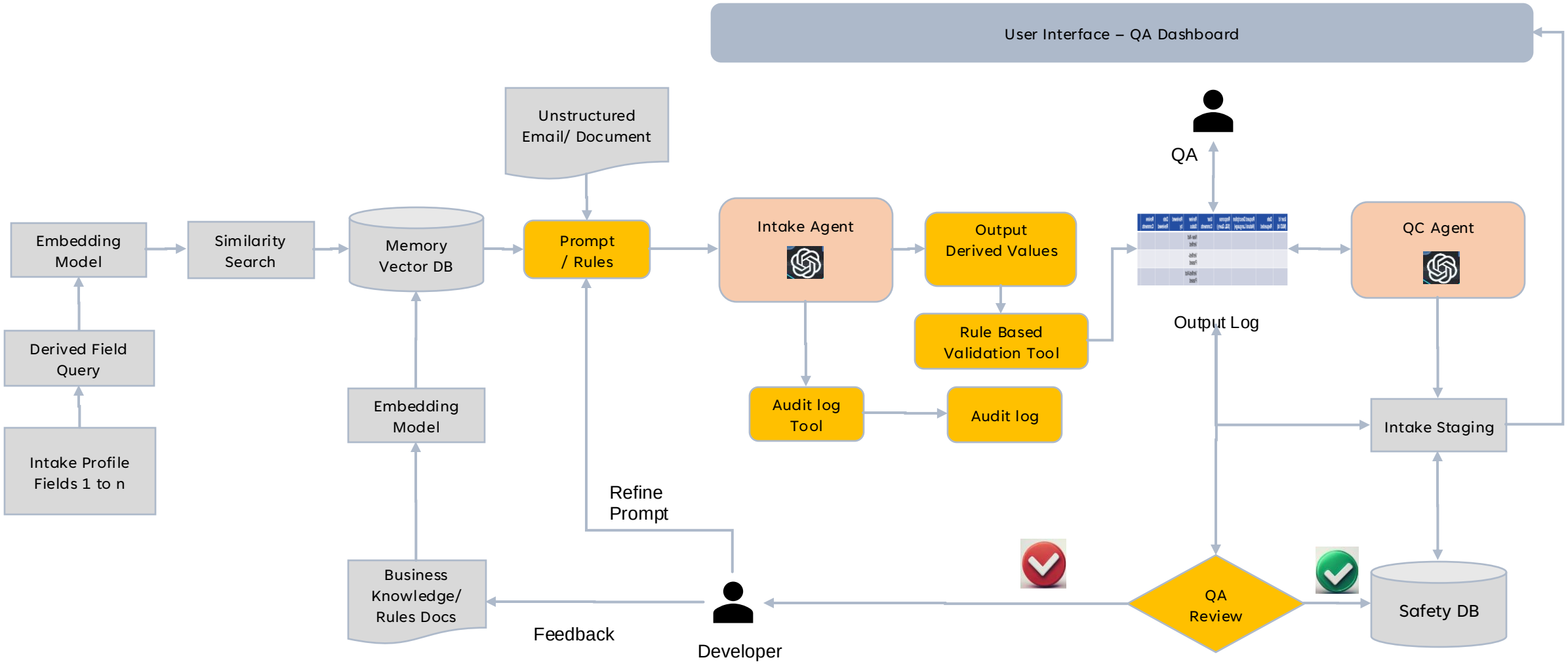
Planning Type: Centralized Planning Decentralized Execution (CPDE) and Decentralized Planning Decentralized Execution (DPDE).

Integration with External Systems: connections to databases, APIs, and other external tools

Feedback Loops: learning from past interactions or human feedback.

*Figure Referenced from https://doi.org/10.48550/arXiv.2401.03428

A PROPOSED QC FRAMEWORK FOR MULTI AI AGENTS WITH HUMAN-IN-THE-LOOP REVIEW



Jeffery L Painter, Venkateswara Rao Chalamalasetti, Raymond Kassekert, Andrew Bate, Automating pharmacovigilance evidence generation: using large language models to produce context-aware structured query language, *JAMIA Open*, Volume 8, Issue 1, February 2025, ooaf003, https://doi.org/10.1093/jamiaopen/ooaf003

THE FUTURE OF PHARMACOVIGILANCE (PV) AUTOMATION USING MULTI-AI AGENTS

- **End-to-End Autonomous Case Processing:** A team of specialized AI agents handles the entire adverse event (AE) reporting pipeline from intake to submission with minimal human intervention.

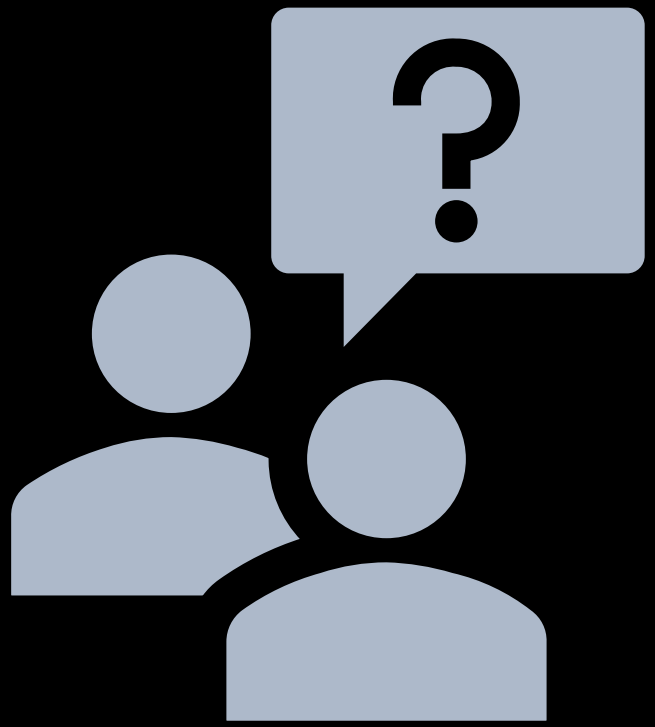
Potential AI Agents for PV:

- **PV Master Agent:** Orchestrates the workflow, develops case processing plans, and delegates tasks to child agents.
- **Intake Agent:** Receives and categorizes incoming adverse event reports.
- **Data Entry Agent:** enters data into a safety database.
- **Duplicate Check Agent:** Identifies and merges duplicate cases.
- **Translation Agent:** Translates documents received in different languages.
- **Coding Agent:** Applies standardized medical codes (e.g., MedDRA) to the case.
- **Narrative Generation Agent:** Creates narratives for cases based on extracted data.
- **Case QC Agent:** Performs in-stream quality checks on ICSR case data to ensure data accuracy and completeness.
- **Report Submission Agent:** submits regulatory reports in required formats to health authorities and partners.

Recommendations:

- **Data quality improvement:** Invest in data cleaning, standardization, and enrichment processes to ensure reliable data for AI training and implementation.
- **Process simplification:** Break down complex business processes into smaller, more manageable steps to facilitate automation.
- **Pilot projects:** Start with smaller, less critical processes to test and refine AI capabilities before scaling up.

- **Vision for 2030 and Beyond:** By 2030, multi-AI agents could make PV fully proactive predicting AEs before they occur, automating 60 to 70% of case processing, and autonomous AI agents that seamlessly manage the entire drug safety lifecycle from case intake to regulatory reporting—enhancing efficiency, accuracy, and patient safety on a global scale. While LLM agents demonstrate potential for automation, human expertise remains essential for oversight and ensuring compliance with regulations.



Q&A

THANK YOU

Jeffery Painter
Senior Director, Safety & Quantitative Innovation

Venkateswara Rao Chalamalasetti
Sr. Manager and SME in Drug Safety #Data & Systems | Exploring #AI/#ML in #Pharmacovi...

Slides available at https://github.com/jlpainter/ATOI_2025/

GSK

Home | Content Hub | Challenges | Working with us | Who's with us

GSK.ai

GSK uses AI to discover transformational medicines

AI is the key to interpret genetic datasets so we can understand the 'language' of the cell and develop medicines with a higher probability of success