LTIMindtree

Point of View

Life Sciences
Enhancing Trust and Safety of Generative AI

Background

Generative AI has rapidly gained popularity due to its remarkable ability to produce human-like content and its adaptability across various use cases. Large Language Models (LLMs) are trained on huge corpus of textual data, enabling them to perform multiple tasks through in-context learning. Users can leverage these pre-trained LLMs to perform novel tasks by providing just a few examples of problem-solution pairs. However, one of the major challenges in utilizing generative AI is that the reasoning behind the generated responses is often inaccessible to humans. The inherent opacity of these models make it challenging to interpret their outputs and draw meaningful conclusions for effective decision-making.

In the life sciences industry, where safety and trust are vital, this opacity becomes a critical issue. Given that Gen AI models are probabilistic in nature, understanding and addressing the challenges of explainability is crucial for making LLMs more trustworthy and safer. The necessity to ensure that AI-generated content adheres to strict safety standards and ethical guidelines is particularly pressing in this field, where the consequences of errors can be significant.

Leveraging Gen AI in Life Sciences

The life sciences industry is rapidly advancing, and Gen AI offers a significant opportunity to boost productivity and efficiency across the value chain—from drug development to commercialization. Companies are actively exploring the potential of Gen AI in the life sciences industry such as biomedical text generation, development and management of standard operating procedures (SOPs), semantic search, and knowledge mining.

These innovations promise to accelerate drug development, improve patient outcomes, and streamline various processes within the industry.

As life sciences organizations embrace these technologies, they are also confronted with the growing importance of data privacy and information security. Regulatory pressures and consumer demands for robust data protection are driving organizations to prioritize these concerns. Ensuring the security of sensitive information, including product research data and personally identifiable information (PII) of patients, is crucial when integrating Gen AI into their existing systems.

A significant risk exists when sensitive data is used to train open source LLMs, as users with the right prompts can potentially retrieve this information, leading to data breaches and vulnerabilities. Therefore, it is critical for companies to incorporate stringent data safeguards during the design of these systems to create robust and secure applications.

Moreover, LLMs predict the next word in a sequence without truly comprehending the meaning behind the words. This can result in highly fluent and convincing responses that may deviate significantly from the ground truth. The black-box nature of LLMs further complicates the process of discerning how they arrive at specific responses. In a field like life sciences, where compliance and patient safety are paramount, quality decision-making based on factual data is essential. The potential inaccuracies in LLM-generated responses highlight the importance of critical evaluation, as even small errors can have major implications on drug quality and patient safety.

LTIMindtree's Approach

With the increasing adoption of Generative AI, LTIMindtree suggests that life sciences organizations must pivot to prioritize both data privacy and response trust to ensure quality. While the integration of Gen AI brings substantial benefits, it also introduces risks related to data privacy and transparency. Addressing these risks is crucial to preserving the trustworthiness and reliability of AI systems.

LTIMindtree can assist companies in mitigating these risks through various frameworks and scores designed to enhance data privacy and transparency. By implementing a data privacy and trust-integrated framework, alongside multiple guardrails, LTIMindtree can help companies ensure that their LLMs perform optimally across various dimensions, including data privacy, scientific accuracy, retrieval quality, and output integrity.

Output Monitoring

Quality Gates

Given that generative AI models are inherently generic and often trained on open-source data lacking in scientific and medical domain knowledge, there is a need for stringent output monitoring. LTIMindtree recommends employing conventional AI algorithms and natural language processing (NLP) techniques to verify the outputs generated by these models, ensuring that they meet the highest standards. These standards are set by LTIMindtree's team of life sciences domain experts who possess deep knowledge of the field.

Traditional AI, with its deterministic nature, can be used to incorporate medical knowledge through simple rules that validate the output. For instance, if a model classifies certain protein names as 'biomarkers' when they should be classified as 'treatments,' domain experts can develop rules to enhance and validate the accuracy of the output. This approach ensures that the outputs generated by Gen AI models are not only fluent but also factually correct and contextually relevant.

Claim Verification

Effective output monitoring also involves implementing quality gates and robust claim verification processes. LTIMindtree proposes several strategies to improve the quality and trustworthiness of Gen AI outputs, thereby facilitating quality decision-making in life sciences

Knowledge graphs: Knowledge graphs are a powerful tool that enhances LLMs' understanding by providing a rich context stored in network of nodes and edges. Incorporating domain-specific knowledge graphs increases LLMs' awareness of medical concepts and clinical terms, thereby improving the accuracy of their responses. Knowledge graphs also increase the reasoning capabilities of LLMs, enabling them to interpret facts, answer complex questions, and trace back to the source document, thus enhancing the reliability of the outputs.

Built-in prompts: Utilizing built-in prompts is another effective strategy to guide the generation of responses and ensure they align with the desired context and scientific validity.

Advanced RAG pipelines: Retrieval-Augmented Generation (RAG) pipelines are designed to ensure that responses generated by LLMs are grounded in actual data, making them more accurate and reliable. These pipelines retrieve specific documents or pages that are directly relevant to the query, thereby increasing the trustworthiness of the outputs.

Human oversight: Incorporating human human-in-the-loop is crucial for reviewing and validating AI-generated outputs. This approach ensures that the outputs meet the required standards of accuracy and reliability, particularly in a domain where the consequences of errors can be significant.

Data Privacy Framework

LTIMindtree's Canvas.ai Platform plays a critical role in safeguarding sensitive data by restricting its transmission to public LLMs. The platform moderates over 50 data formats to assess whether organizational data is at risk of compromise. Techniques such as data obfuscation and data redaction are employed to protect confidential product and research-related data. These methods ensure that the data remains functional for legitimate users while being obscured from unauthorized access. By implementing these data privacy measures, life sciences organizations can prevent inadvertent exposure of sensitive information, thereby reducing the risk of data breaches.

Robust Metrics

In the safety-critical domain of life sciences, it is crucial to detect risks in LLM-generated responses and measure them against key metrics defined by life sciences domain experts. Generative AI has the potential to produce variable responses for the same input, which necessitates a thorough annotation and assessment process. Responses must be evaluated on multiple factors, such as scientific factuality, context retrieval, accuracy, and coherence. Built-in prompts can be used to score these responses and set threshold values, ensuring that they align with the provided context and scientific validity.

For instance, in knowledge base question-answering, providing relevant context to LLMs is vital for generating accurate answers. Evaluating whether the right context is retrieved involves measuring various score indices such as:

- **Q-index:** Measures the quality of the retrieved document
- **R-index:** Tracks the number of times a document has been opened or reused and the number of users engaging with it
- **Chunk length:** Balances the signal-to-noise ratio by providing sufficient information to LLMs without redundancy

Conditioning LLMs for Life Sciences

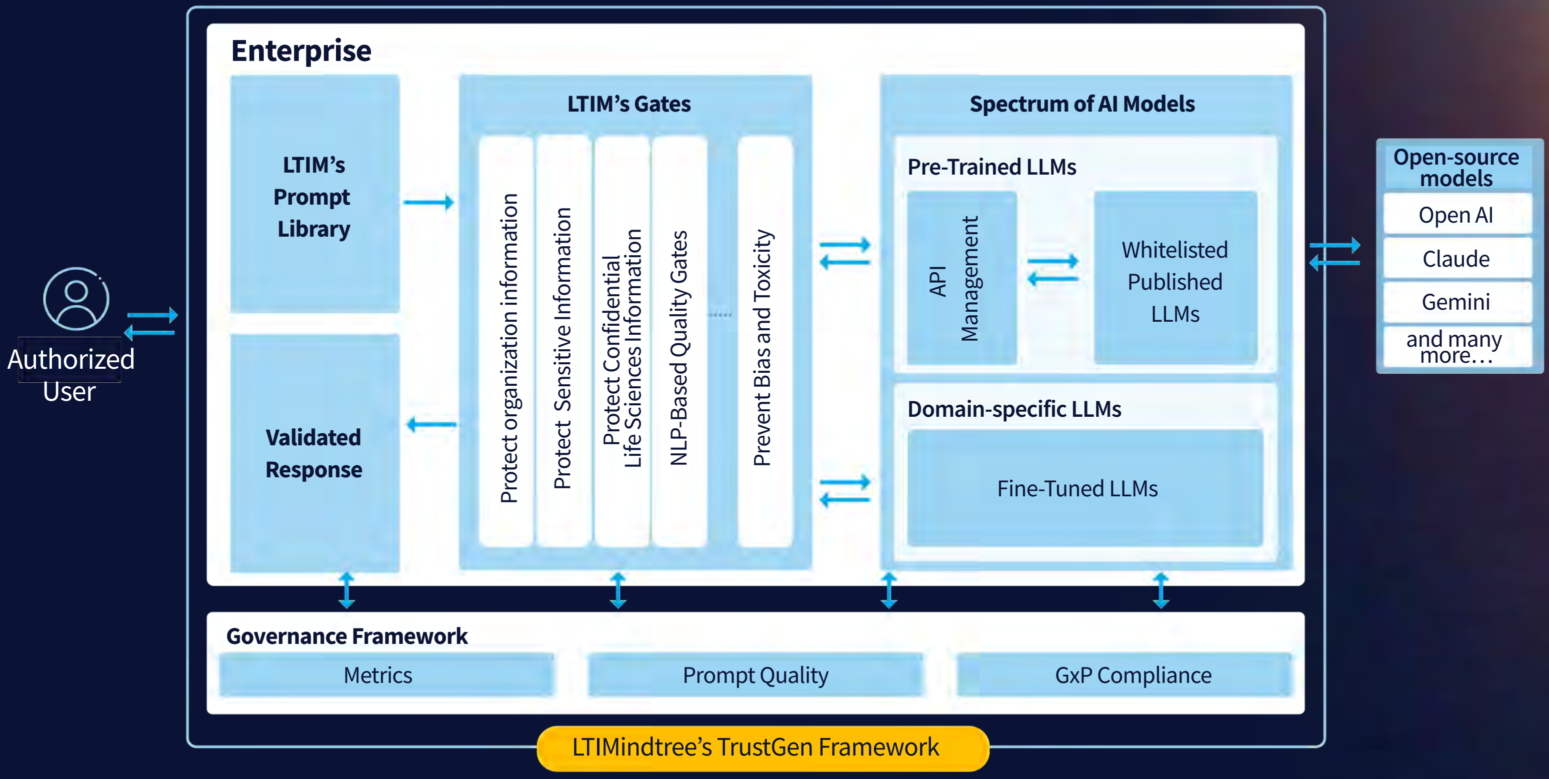
Conditioning LLMs to perform domain-specific tasks is essential for achieving high accuracy and reliability in the life sciences domain. This process involves advanced prompting techniques and fine-tuning processes tailored to the unique requirements of the field.

a. Prompt engineering

LTIMindtree can assist life sciences companies by providing ready-to-use prompt templates that have been validated and refined by domain experts. These templates are designed to serve various purposes, such as literature review summarization and knowledge base question answering.
By utilizing advanced prompting techniques, such as few-shot prompting and chain-of-thought, we strive to enhance the performance of LLMs for domain-specific tasks. These templates are systematically tested and adjusted to refine outputs and modify LLM behavior to meet intended use.

b. Fine-tuning

Vanilla LLMs, which are trained on generic corpus text, may struggle with complex biological texts. Fine-tuning these models with targeted datasets reduces bias and adapts the model to domain-specific tasks. By updating pre-trained weights with clinical knowledge, LLMs can generate text that reflects the learnings from these datasets, thereby enhancing performance across various clinical tasks, such as biomedical text generation, text mining, and clinical summarization.



Success Stories

A global pharmaceuticals and medical devices conglomerate needed to identify duplicate entries among healthcare professionals' data from different sources. Traditional master data management (MDM) algorithms weren't effective in this case. LTIMindtree developed a customizable solution using language understanding and embeddings to identify duplicate entries, significantly reducing the manual effort required for data review. The solution identified and eliminated 70% of duplicate entries that were previously marked as unique by existing software. By implementing best-in-class privacy and security guardrails, our solution provides heightened trust and robust protection of sensitive information while improving overall accuracy and data quality.

Conclusion

Integrating Gen AI in healthcare and life sciences presents unique opportunities for innovation and efficiency. However, it also introduces challenges that must be addressed to fully harness its potential. LTIMindtree offers a comprehensive approach to mitigate these challenges, ensuring that organizations can maximize the benefits of Gen AI. By focusing on data privacy, transparency, and trust, we empower life sciences organizations to drive innovation and improve patient care.

About the Authors

Siddharth Joshi

Principal Director—Life Sciences

Siddharth is a seasoned professional with over 25 years of experience in IT, supply chain, and manufacturing. His successful stewardship of major accounts, coupled with collaborative engagements with CIOs and VPs of Fortune companies, has resulted in the delivery of numerous transformative programs. With a strong background in both domain expertise and IT, Siddharth has excelled in digital consulting, solution architecture, program management, and application development and maintenance.

Freny Rambhia

Subject Matter Expert—Life Sciences

Freny collaborates closely with clients, offering valuable insights, guidance, and support to deliver innovative solutions. With a proven track record of addressing diverse customer needs, she actively contributes to projects that drive business outcomes and she consistently exceeds expectations. Her passion lies in exploring and studying new technologies.