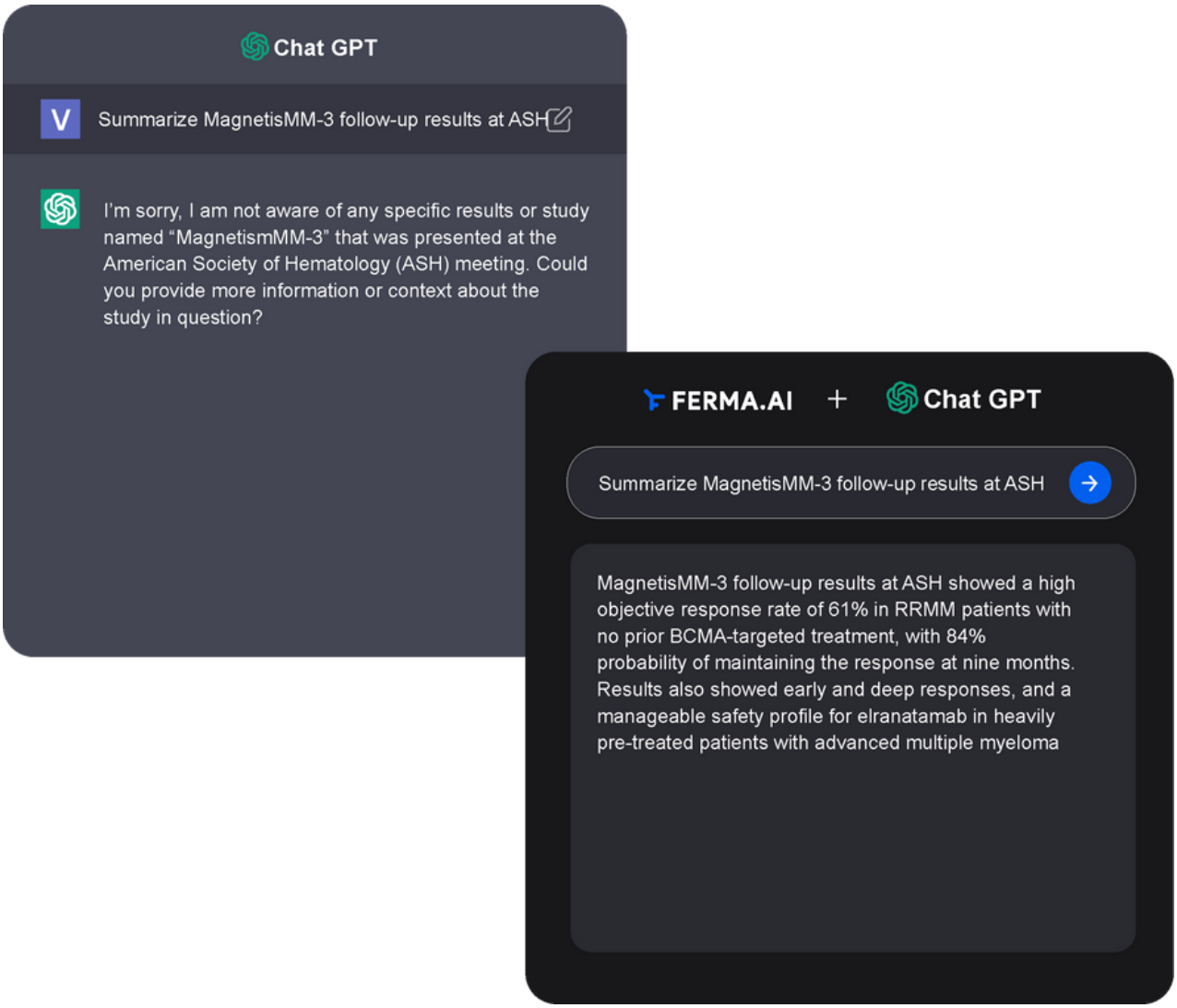
Revolutionizing Pharma: The Power of Large Language Models and ChatGPT

FEBRUARY 2023

FERMA.AI

A product of ZoomRx

CONTENTS



Introduction

In a race to bring innovative therapies to market, the pharmaceutical industry is constantly searching to improve its processes for drug discovery, development, and delivery. Recently, the emergence of impressive large language models (LLMs), such as ChatGPT, has reinvigorated the discussion around artificial intelligence (AI) and its potential to revolutionize existing workflows.

Despite much excitement, it remains to be seen whether or which AI applications will truly disrupt the industry. Due to inherent technical limitations and a history of "trendy" technologies that did not live up to their hype, some remain skeptical about the impact tools like ChatGPT will bring to the pharmaceutical industry.

In this white paper, Ferma.AI showcases the paradigm-shifting potential that LLMs can have across multiple applications and functional areas within pharma. We also demonstrate the transformative potential of combining LLMs with:

- Comprehensive life sciences data sets
- A carefully designed knowledge graph
- Pharma-specific training models and expert oversight

Recent Advances in LLM Technology

The history of LLMs can be traced back to the early days of AI research in the 1950s and 1960s. However, it was not until more recent advances in computational power and data availability that LLMs became practical. Trained on vast amounts of text data, these models produce an understanding of the patterns and structures of human language, allowing LLMs like ChatGPT (and its variants) to engage in natural language conversations with users.

Unlike other chatbots, which are limited to pre-programmed responses, ChatGPT can engage in conversations, answer questions, and generate new ideas on a level that is often indistinguishable from human dialogue. With its user-friendly interface, users can interact with ChatGPT without any prior knowledge of the technology, simply "speaking" to it as any person would to another human.

Barriers to Widespread LLM Use Within Pharma

Despite the recent fanfare, applications for LLMs have been relatively few and narrow to date, and most have occurred outside the pharmaceutical industry. For example, LLMs have been used to improve automated customer service, generate captions for images and videos to help the visually impaired, and even write code.

While the advances in LLM technology demonstrated by ChatGPT are poised to expand the range of applications exponentially, several critical barriers must be overcome to make LLMs commonplace within the pharmaceutical industry.

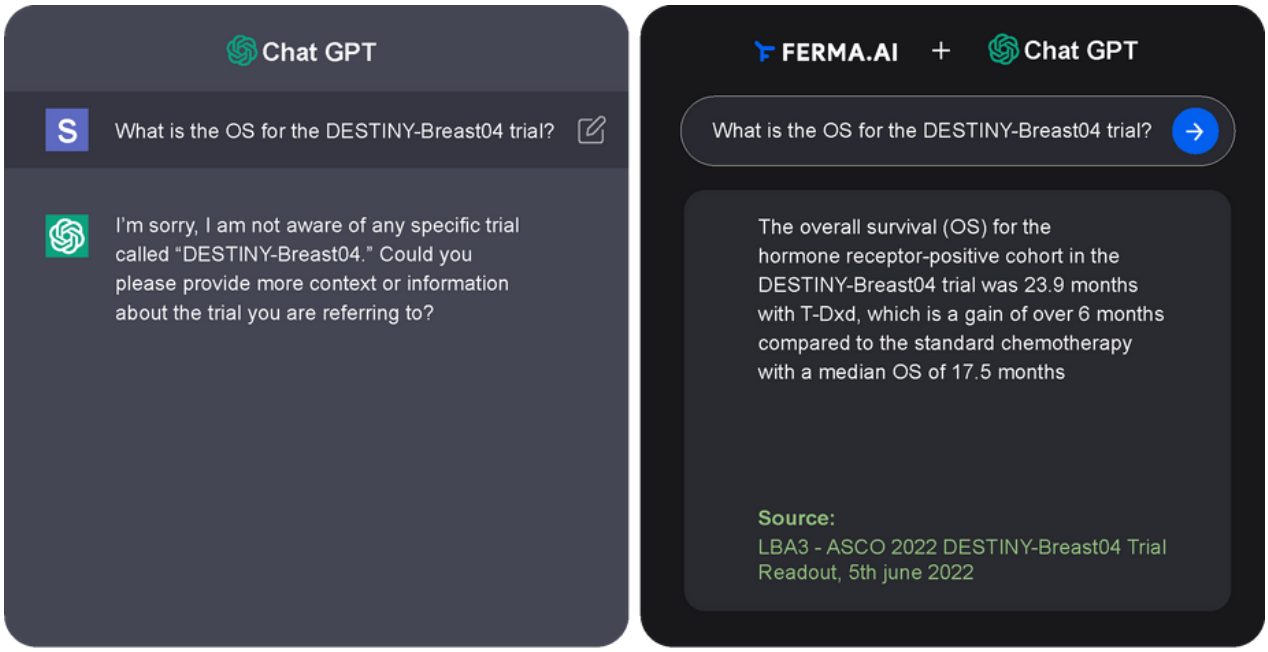
1. Pharma data sets are not easily accessible

The effectiveness of any AI algorithm can be directly tied to the availability of concrete data. Without a robust data set, it is difficult for AI technology to perform as requested. For example, ChatGPT's training data set only runs up until 2020; ask it about any event that occurred in the past three years and it will struggle to report back.

Along those lines, LLMs need millions of different scenarios to discover patterns, learn from experience, and adjust themselves appropriately. However, because pharmaceutical manufacturers work off of proprietary data sources that often lack interoperability, most LLMs have not been trained on data sets specific to the life sciences space.

Ferma.AI solves this problem through a unique data architecture that houses healthcare-specific data, including clinical trials, press releases, news reports,

social media activity, and much more. Further, Ferma.AI can import organizational-specific proprietary data such as market research reports, CRMs, claims, digital marketing activity, and so forth to customize outputs specific to the client. This allows Ferma.AI's approach to incorporate applications specific to the pharma industry that many other LLMs could not.



Ferma.AI's unique data architecture combines healthcare data with proprietary organizational data to provide customized outputs specific to the pharmaceutical industry.

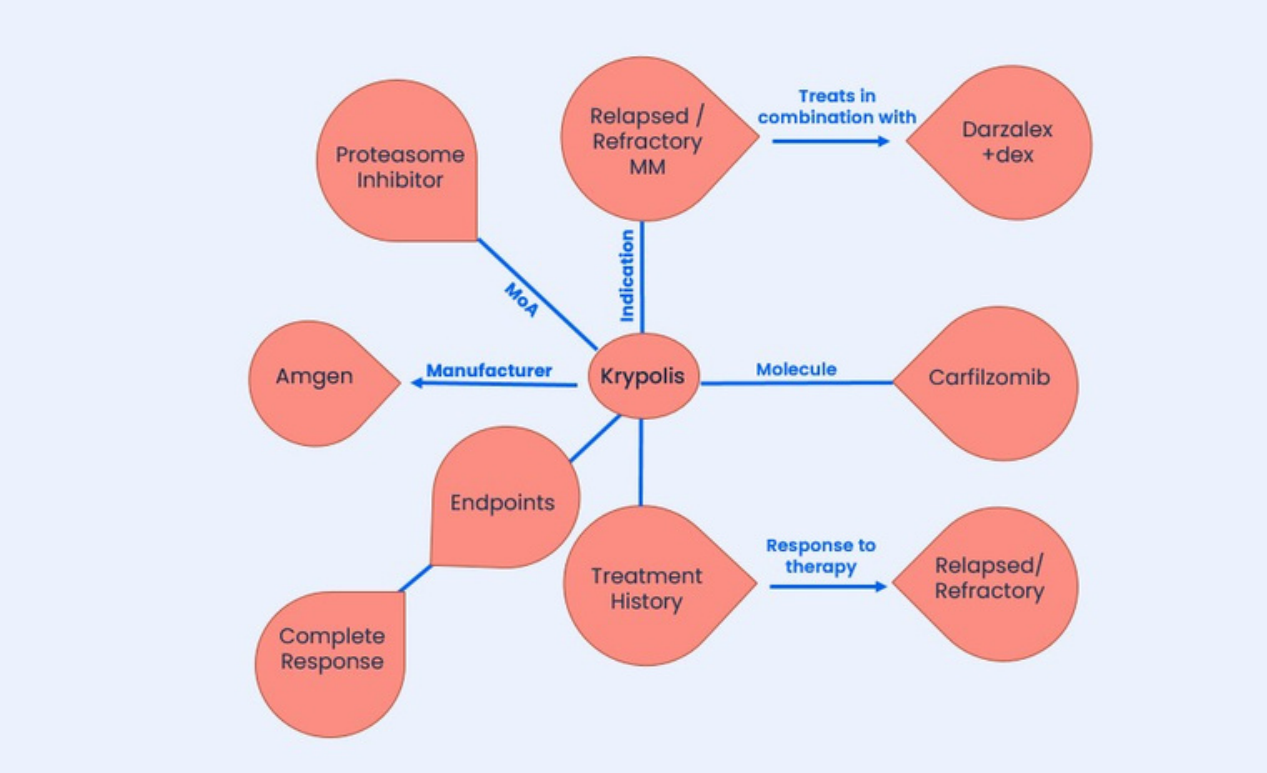
2. Life sciences terminology is inherently complex

Even when a robust data set is available, healthcare data is inherently technical, unsystematic, and chaotic. NCT numbers (an ID used by ClinicalTrials.gov) need to be tied back to clinical studies, generic names to brand names, acronyms to various indications and outcomes, and so forth. As a result, in order to adequately analyze healthcare data, LLMs need human oversight from people who thoroughly understand the life sciences space, its components (e.g., clinical trials, claims, RWE, etc.), and the best approaches to turning these technical data sets into actionable insights.

Over the past 5+ years, Ferma.AI has built a "knowledge graph" of natural language processing (NLP) models that can interpret life sciences data in any

form to drive better, more compelling results. There are three key features to this NLP technology:

- The ability to recognize and extract life sciences data in any format (digital or printed text, PPT slides, video, audio, handwritten notes, etc.)
- The ability to read and understand the language within this data and tag life sciences-relevant keywords
- The ability to create both explicit and implicit connections between those life sciences-relevant keywords to establish structure and enable powerful search



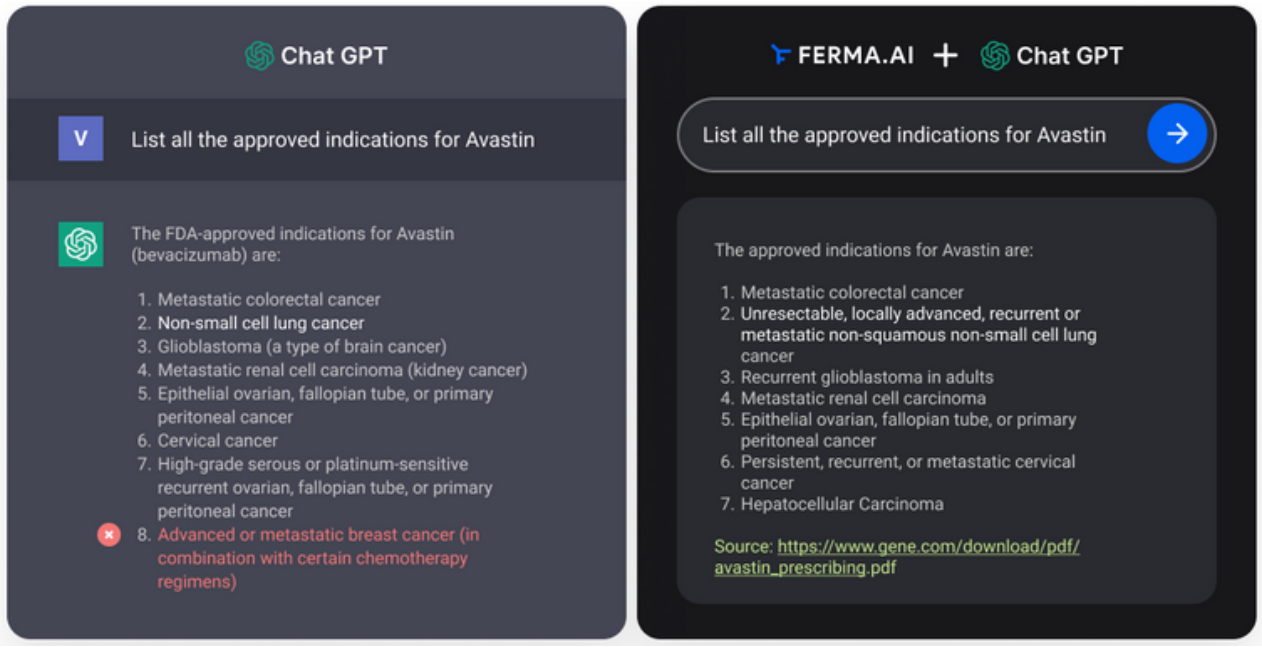
An example Ferma.AI knowledge graph. When Ferma.AI encounters the keyword "Kyprolis," it not only recognizes that term as a brand name but also associates it with its generic alternate name, its manufacturer, its indication, its clinical trial results, and more.

3. Incorrect answers carry significant consequences

The third major hurdle to LLM adoption has been usability and bridging the gap between the capabilities of LLMs and the expectations of users. Not only should the technology be able to understand and answer all types of

questions, but it must also provide accurate, relevant, and rational results in a format that is convenient and easy to understand.

Take, for example, this relatively straightforward and innocuous request to ChatGPT: List all the approved indications for Avastin.



Ferma.AI overcomes ChatGPT's pharma-specific limitations through intensive training, human oversight, and continuous improvement to deliver accurate and accessible outputs.

While ChatGPT's answer may appear accurate at face-value, those who are familiar with Avastin's prescribing information know that there are a couple of errors in its response. In fact, OpenAI, the creators of ChatGPT, openly acknowledges that their model "sometimes writes plausible-sounding but incorrect or nonsensical answers."

These types of suboptimal user experiences can lead to confusion, frustration, and lack of trust. As a result, stakeholders are hesitant to depend on "black box" technology for critical decisions because of the potential consequences attached to misinformation.

Ferma.AI tackles these barriers by providing intensive training to shape the LLM's outputs. With carefully curated training sets, the model generates results

in a format that is easily understood and accessible for end-users. Moreover, life sciences experts provide constant human oversight to ensure the accuracy and relevance of the LLM's responses. This results in continuous improvement of the model's performance, as feedback is incorporated to prevent repeat mistakes and enhance decision-making.

Ferma.AI + ChatGPT and Its Potential Applications

Ferma.AI is a large language model developed by ZoomRx that leverages comprehensive life sciences data sets, a carefully designed knowledge graph, pharma-specific training models, and expert human intelligence and oversight. This approach allows Ferma.AI to be tailored to the specific needs of the pharmaceutical industry, where its unique ability to process and understand complex medical and scientific language sets it apart from traditional models.

Ferma.AI has the potential to transform existing workflows in the pharmaceutical industry by providing quick and accurate information, automating tedious tasks, and making data-driven decisions more accessible. It offers a range of applications, including but not limited to:

Medical Intelligence: Ferma.AI's ability to search and query clinical data and trial results allows pharmaceutical companies to stay up-to-date on the latest advancements in their field. The model can create compact, succinct summaries of key information related to the study design, clinical outcomes, and statistical results, enabling stakeholders to quickly understand the most important aspects of a trial. In turn, this intelligence can be used to inform drug development, identify potential areas of research, and monitor the progress of ongoing clinical trials.

FERMA.AI + Chat GPT

Highlight key differences and similarities between PROpel, MAGNITUDE, and TALAPRO2 trials along with primary endpoints from conferences

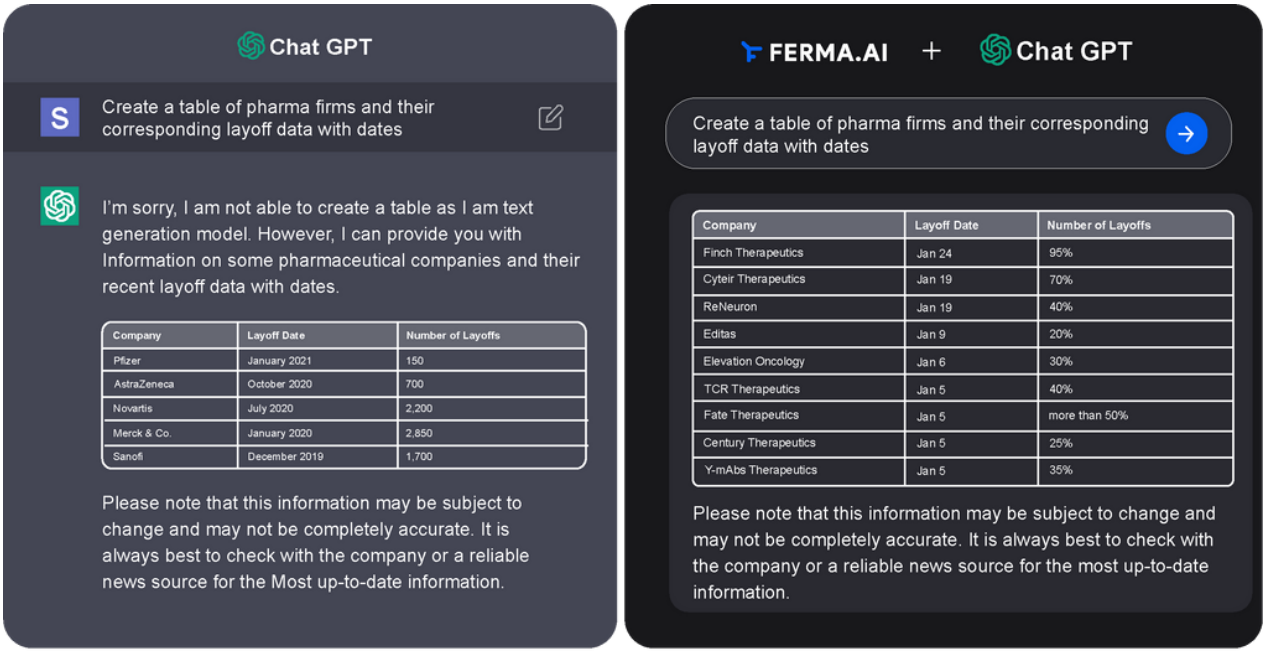
TRIALS	PROpel (Phase3)	MAGNITUDE (Phase3)	TALAPRO2 (Phase3)
NCT ID	NCT03732820	NCT03748641	NCT03395197
Sponsor	AstraZeneca	Janssen Research & Development	Pfizer
Drug	Primary: Olaparib + Abiraterone + prednisone or prednisolone	Primary: Niraparib + Abiraterone Acetate + Prednisone	Primary: Talazoparib + Enzalutamide
Indication	First-line mCRPC	First-line mCRPC	First-line mCRPC
Patient Segments	• Prior taxane at mHSPC (yes/no) • Site of Distant Metastasis	• Prior taxane at mCSPC • Prior ARi for mCSPC / nmCRPC • Prior AAP for L1 mCRPC • HRR+ cohort: BRCA1/2 vs other HRR gene alterations	• DDR Alteration status (positive or negative/unknown) • Prior Abiraterone or taxane-based chemotherapy for CSPC (yes/no)
Primary Endpoints	rPFS	rPFS	• rPFS in unselected patients and in those with DDR deficiencies
Conference Readouts	Median rPFS: 24.8 mon (n=399) vs 16.6 mon (n=397)	Median rPFS in BRCA1/2 subgroup: 16.6 mon (n=113) vs 10.9 mon (n=112) Median rPFS in All HRR+ subgroup: 16.5 mon (n=212) vs 13.7 mon (n=211)	rPFS: Met (not yet disclosed) *results to be presented at ASCO GU23

Sources: Trials (NCT03732820, NCT03748641 & NCT03395197), Conference (ASCO GU22) & Press Releases (Pfizer)

Using natural language processing capabilities, Ferma.AI can generate succinct, compact summaries of the latest clinical data and trial results.

Market Intelligence: Ferma.AI's market intelligence capabilities allow pharmaceutical companies to gain valuable insights into their industry and their competitors. The model's ability to find and synthesize a wide variety of information sources, including internal reports, external intelligence, and social media activity, allows stakeholders to stay informed on industry trends and key insights.

Additionally, Ferma.AI's natural language processing capabilities allow stakeholders to ask specific questions, query a vast data set, and receive relevant answers that can inform strategic business decisions.

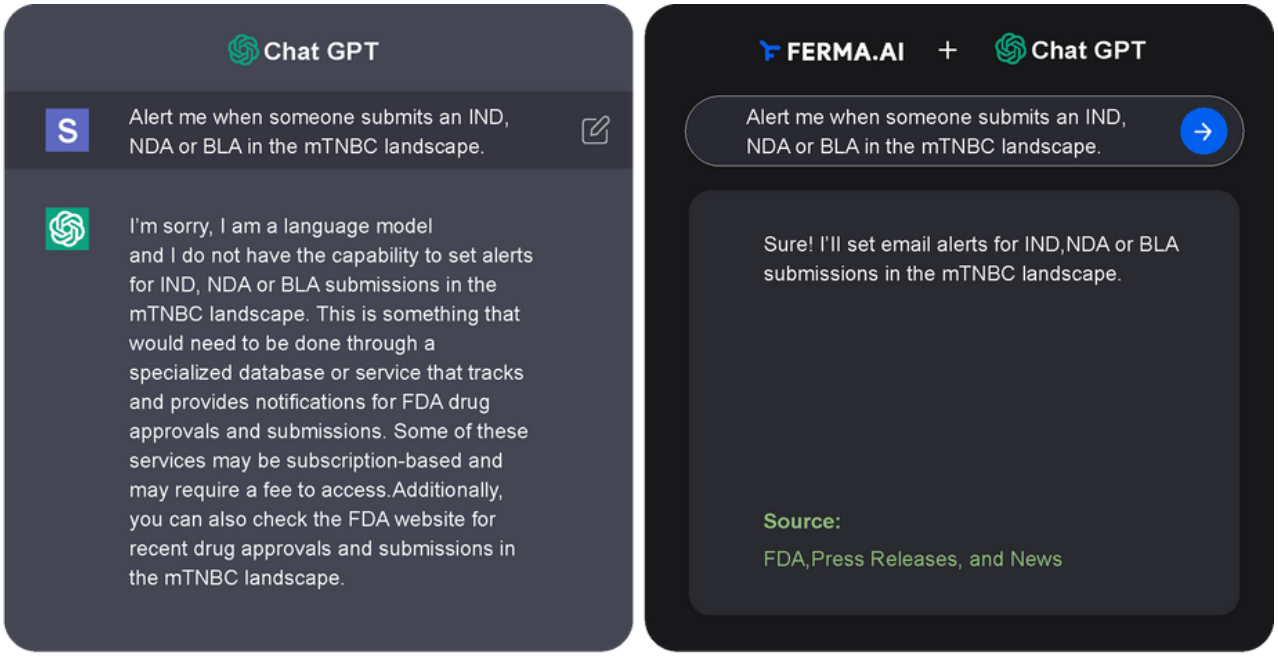


Ferma.AI can query a diverse range of information sources, enabling stakeholders to ask it unstructured questions and yet still receive relevant, convenient answers.

In the example above, Ferma.AI completes a series of tasks normally asked of a human by:

- Combing through news reports to find layoff data
- Identifying the current headcounts of the companies with recent layoffs
- Calculating a percentage of the workforce laid off
- Structuring the output in a useful, convenient table

Real-Time Monitoring: With its advanced natural language processing capabilities, Ferma.AI can be programmed to track relevant news and events specific to the pharmaceutical industry. In particular, Ferma.AI can be programmed to trigger alerts for specific keywords, allowing stakeholders to be informed of key news events as soon as they happen. This capability enables pharmaceutical companies to stay up-to-date on significant market developments, regulatory changes, and new scientific discoveries.



Ferma.AI can be programmed to send real-time alerts for key news and market events.

CONCLUSION

The development and application of LLMs in the pharmaceutical industry is still currently in its nascent phase, with several key hurdles to overcome. However, the potential of this technology across multiple use cases and functional areas is tremendous.

By leveraging comprehensive life sciences data sets, a carefully designed knowledge graph, pharma-specific training models, and expert human intelligence and oversight, Ferma.AI has the potential to revolutionize many pharma-specific workflows. Ferma.AI + ChatGPT's knowledge graph is updated in real-time, ensuring that you have the latest life sciences intelligence. Subscribe to our newsletter and stay ahead of the game.

ZoomRx is actively exploring these areas of potential application, both independently and in collaboration with our clients. We invite you to reach out to us for a live demo and to discuss the potential applications of Ferma.AI for your organization.

SEE A DEMO OF FERMA.AI