zensar

Accelerated Drug Discovery Enabled by **AI-Powered Molecular Intelligence**

Case Study

An RPG Company

Overview

Shortening the product development timeline

A reputed pharmaceutical company, operating in more than 50 countries, was facing significant challenges in its product development process, particularly related to impurities and overall process efficiency.

Zensar's brief: Reduce the existing product development lifecycle by over 30 percent and enable optimizations and efficiencies, using these key moves:

- Accelerate the literature search and data extraction process by leveraging AI.
- Enhance process efficiency by integrating various internal software systems.

Beyond the brief: Working closely with the client, we put together a phased AI-implementation strategy to demonstrate quick wins and build momentum.

Challenges

Need to cut effort, time, and risk

The IT department was exploring ways to help the business address these major concerns:

- **Information abundance:** With over 10,000 papers returned by a single keyword search, critical findings were often lost and efforts were duplicated across therapeutic teams.
- **Turnaround time:** There was increasing pressure to reduce the overall product development lifecycle by 30 percent.
- **Slow go/no-go decisions:** Manual reviews were creating delays in hit-to-lead and lead-optimization cycles.
- **Compliance and risk:** Missing patent, safety-signal, or regulatory updates presented the risk of derailing programs.
- **Ineffective process:** The manual search process was not just error prone and time consuming, it was ineffective. It resulted in an accumulation of similar contextual information from multiple sites, creating complex data redundancies.

Solution

Leveraging AI to optimize pharmaceutical product development

Having worked with multiple clients in the pharmaceutical space, we were keenly aware of the challenges involved in managing a hybrid model of in-house and outsourced clinical trials, as well as tackling the issue of counterfeit drugs, particularly in oncology product development. We were also mindful of the fact that pharmaceutical researchers often lose over a quarter of their time sifting through thousands of publications for every new molecule.

Building on our understanding of this industry, we worked closely with the client's team to deploy a comprehensive solution with a three-pronged approach:

[A] Automating content ingestion

We built and implemented an E2E molecular drug intelligence platform that automates the process of continuously harvesting content from multiple sources, extracting fine-grained insights, and answering domain-specific questions through a chat interface with citations.

Utilizing a combination of API and web-scraping pipelines, we enabled the system to carry out these key tasks:

- Gather articles, patent documents, clinical dossiers, conference abstracts, and other relevant documents.
- Clean and store the content in a secure vector repository for faster retrieval.
- Enhance performance on scientific documents, patents, and dossiers, employing a fine-tuned SciBERT model.

[B] Enabling deep document understanding

We designed the system to identify and extract key elements from scientific literature, including pharmaceutical ingredients, bioequivalence data, regulatory compliance details, formulation guidelines, safety profiles, patent information, and manufacturing processes.

To enable effective data extraction and structuring, we deployed a multimodal approach:

- Involved IT and domain experts to discuss integration points and software tools.
- Deployed domain-tuned NLP models to tag entities and extract bioequivalence metrics, safety endpoints, and much more.
- Enriched the embeddings with the use of, but not limited to, MeSH, ChEMBL, and patent ontologies.

[C] Deploying an enterprise grade multi-agent RAG engine

We deployed a RAG engine that combines dense semantic search with a large language model (LLM) to draft answers and provide inline citations, ensuring traceability for regulatory audits.

These are the key features of the engine:

- Searches the internet for relevant documents and information, if information is unavailable in the knowledge base.
- Ranks the retrieved information based on feature importance.
- Evaluates the information with the LLM performing the role of a judge, achieving consistency, depth, scalability, and cost-effectiveness.
- Allows for detailed information retrieval within specific areas such as API, bioequivalence (BE), formulation, patent, pharmacokinetics, safety covariates, and more.

Impact

Greater business competitiveness

- **Efficient molecular reviews:** Up to 50 percent time savings on data extraction and summarization for analysts.
- **Faster candidate prioritization:** Early visibility into safety, IP, and manufacturability helped shave months off discovery timelines.
- **Regulatory confidence:** Reduction in the risk of overlooking compliance or patent constraints, enabled by citation-backed answers.

Business outcomes: The solution delivered the potential to save up to $1.5 million annually and accelerate discovery timelines by 30 percent.

zensar

An RPG Company

At Zensar, we're 'experience-led everything.' We are committed to conceptualizing, designing, engineering, marketing, and managing digital solutions and experiences for over 145 leading enterprises. Using our 3Es of experience, engineering, and engagement, we harness the power of technology, creativity, and insight to deliver impact.

Part of the $4.8 billion RPG Group, we are headquartered in Pune, India. Our 10,000+ employees work across 30+ locations worldwide, including Milpitas, Seattle, Princeton, Cape Town, London, Zurich, Singapore, and Mexico City.

For more information, please contact: **info@zensar.com | www.zensar.com**