

Enabling Semantic Search for Pharmaceutical Research and Development



The Challenge

A major pharmaceutical research and development company faced difficulty creating regulatory reports and files based on years of drug experimentation data. Their regulatory intelligence teams and drug development chemists spent dozens of hours searching through hundreds of thousands of documents to find past experiments and their results in order to fill out regulatory compliance documentation. The company's internal search platform enabled users to look for documents, but required exact matches on specific keywords to surface relevant results, and lacked useful search filters. Additionally, due to the nature of chemistry and drug development, many documents were difficult to understand at a glance and required scientists to read through them in order to determine if they were relevant or not.



The Solution

EK collaborated with the company to improve their internal search platform by enhancing Electronic Lab Notebook (ELN) metadata, thereby increasing the searchability and findability of critical research documents, and created a strategy for leveraging ELNs in AI-powered services such as chatbots and LLM-generated document summaries. EK worked with the business stakeholders to evaluate the most important information within ELNs and understand the document structure, and developed semantic models in their taxonomy management system with more than 960 relevant concepts designed to capture the way their expert chemists understand the experimental activities and molecules referenced in the ELNs. With the help of the client's technical infrastructure team, EK developed a new corpus analysis and ELN autotagging pipeline that leveraged the taxonomy management system's built-in document analyzer and integrated the results with their data warehouse and search schema. Through three rounds of testing, EK iteratively improved the extraction of metadata from ELNs using the concepts in the semantic model to provide additional metadata on over 30,000 ELNs to be leveraged within the search platform. EK wireframed 6 new User Interface (UI) features and enhancements for the search platform designed to leverage the additional metadata provided by the autotagging pipeline, including search-as-you-type functionality and improved search filters, and socialized them with the client's UI/ User Experience (UX) team. Finally, EK supported the client with strategic guidance for leveraging their internal LLM service to create accurate regulatory reports and AI summaries of ELNs within the search platform.



The EK Difference

EK leveraged its understanding of the capabilities and features of enterprise search platforms, and taxonomy management systems' functionality, to advise the organization on industry standards and best practices for managing its taxonomy and optimizing search with semantics.

Enabling Semantic Search for Pharmaceutical Research and Development

Furthermore, EK's experience working with other pharmaceutical institutions and large organizations in the development of semantic models benefited the client by ensuring their semantic models were comprehensively and specifically tailored to meet their needs for the development of their semantic search platform and generative AI use cases. Throughout the engagement, EK incorporated an Agile project approach that focused on iterative development and regular insight gathering from client stakeholders, to quickly prototype enhancements to the autotagging pipeline, semantic models and the search platform that the client could present to internal stakeholders to gain buy-in for future expansion.



The Results

EK's expertise in knowledge extraction, semantic modeling and implementation, along with a user-focused strategy that ensured that improvements to the search platform were grounded in stakeholder needs, enabled EK to effectively provide the client with a major update to their search experience. As a result of the engagement, the client's newly established autotagging pipeline is enhancing tens of thousands of critical research documents with much-needed additional metadata, enabling dynamic context-aware searches and providing users of the search platform with insight at a glance into what information an ELN contains. The semantic models powering the upgraded search experience allow users to look for information using natural, familiar language by capturing synonyms and alternative spellings of common search terms, ensuring that users can find what they are looking for without having to do multiple searches. The planned enhancements to the search platform will save scientists at the company hours every week from searching for information and judging if specific ELNs are useful for their purposes or not, reducing reliance on individual employee knowledge and the need for the regulatory intelligence team to rediscover institutional knowledge. Furthermore, the company is equipped to move forward towards leveraging the combined power of semantic models and AI to improve the speed and efficiency of document understanding and use. By utilizing improved document metadata provided by the auto-tagging pipeline in conjunction with their internal LLM service, they will be able to generate factual document summaries in the search platform and automate the creation of regulatory reports in a secure, verifiable, and hallucination-free manner.

Enterprise Knowledge (EK) is a services firm that integrates Knowledge Management, Information Management, Information Technology, and Agile Approaches to deliver comprehensive solutions. Our mission is to form true partnerships with our clients, listening and collaborating to create tailored, practical, and results-oriented solutions that enable them to thrive and adapt to changing needs.

Our core services include strategy, design, and development of Knowledge and Information Management systems, with proven approaches for Data and Information Management, Knowledge Graph Implementation in support of NLP, ML, and AI initiatives, Taxonomy Design, Project Strategy and Road Mapping, Brand and Content Strategy, Change Management and Communication, and Agile Transformation and Facilitation. At the heart of these services, we always focus on working alongside our clients to understand their needs, ensuring we can provide practical and achievable solutions on an iterative, ongoing basis.