

Revolutionizing Drug Discovery: 90% Accuracy Achieved with AI-Driven Biomedical Data Curation



Challenge

Pharmaceutical companies face the challenge of rapidly annotating vast volumes of biomedical data required to train AI models. The complexity of mapping biomedical terminology to standard ontology IDs further complicates the process, necessitating domain expertise and precise entity identification.

Solution

We deployed a workflow tool equipped with validation rules and an annotation framework, supported by subject matter experts in life sciences. The solution facilitated accurate biomedical entity identification, ontology mapping, and data normalization, ensuring high-quality datasets for AI model training.

Impact

- Achieved over **80%** automation in data curation processes.
- Attained **90%+** accuracy in entity annotation and data curation.
- Accelerated pre-clinical drug discovery by processing over **5,000** abstracts with high domain complexity.

About Straive

Straive helps operationalize the data → insights → knowledge → AI journey with its deep domain expertise, process knowledge, and tech and analytics capabilities. Serving a diverse range of industries—including science and research publishing, information services, EdTech, life sciences, and banking and financial services—Straive boasts a global client base spanning over 30 countries. Our strategically positioned resource pool operates across seven countries, including the Philippines, India, the United States, Nicaragua, Vietnam, the United Kingdom, and Singapore, where the company is headquartered.

