

HEALX CASE STUDY

Structured quality data for machine learning predictions.

ABOUT HEALX

Healx is focused on accelerating the development of treatments for rare diseases. By integrating AI with deep pharmacology and patient insight, Healx can repurpose existing drugs, translating therapies into the clinic within 24 months.

healx.io

INDUSTRY

Drug Repurposing

USE CASE

Machine Learning & AI

High-quality data is the first step for training Machine-Learning (ML) and Artificial Intelligence (AI) algorithms, but obtaining this information is difficult as most knowledge about drugs exists within scientific publications in an unstructured text format.

Drugbank’s data solutions solve this problem by providing detailed, machine-readable information that lets customers explore AI strategies for drug discovery, development, and repurposing and incorporate extensive scientific knowledge to make their customer’s models more powerful.

DrugBank interviewed Richard Smith, Senior Software Developer at Healx, an AI-powered biotech company from Cambridge (UK), to learn how a Drug Repurposing company utilizes DrugBank data in their products.

TELL US A BIT ABOUT HEALX

Healx is focused on accelerating the development of treatments for rare diseases. We integrate AI with deep pharmacology and patient insight to repurpose existing drugs, translating therapies into the clinic within 24 months. Our process dramatically reduces the time and cost compared with conventional drug discovery methods.

To achieve this goal, we have developed Healnet, the most comprehensive AI-based drug discovery platform for rare diseases, with the objective of translating 100 rare disease treatments towards the clinic by 2025.



HOW DOES HEALX INTEGRATE ARTIFICIAL INTELLIGENCE INTO ITS PRODUCTS?

Healnet integrates a vast amount of biological, clinical and therapeutic information from structured data sources and natural language processing (NLP) of scientific literature.

From this unique knowledge graph of rare disease data, Healnet then maps more than a billion data points connecting diseases, patients and drugs. This approach allows for new connections to be made between related diseases (in terms of biology, the symptoms and drugs that may be effective on them) and enables Healnet to predict potential treatments.

HOW DOES DRUGBANK GET USED AT HEALX, AND WHY DID YOU CHOOSE TO USE DRUGBANK DATA?

DrugBank is one of many data sources we integrate into our internal databases and knowledge graph. We use a wide range of data to train our drug repurposing algorithms, including drug properties, targets, indications, clinical trials and drug-drug interactions from DrugBank.

For our purposes the quality and breadth of data is important but also the organisation and format of downloaded files. DrugBank's clear separation and structure of fields means data can be easily imported into a database or graph structure. Use of standard identifiers (such as UniProt protein identifiers or ICD 10 disease classification) and controlled vocabularies enables much simpler integration with data from other sources.

//

DrugBank's clear separation and structure of fields means data can be easily imported into a database or graph structure.

RICHARD SMITH
Senior Software Developer at Healx

We also benefit from the level of detail in curation. For example drug indications have a flag for off-label use and controlled vocabulary terms defining if the drug is used for prevention, management, symptomatic relief and so on for a particular disease.

What would you say to other companies looking to integrate structured pharma data into their products?

For AI and machine learning applications the choice and quality of input data is as important as the algorithms themselves. It's essential to make considered decisions about the data that goes into an algorithm so you can understand the value of predictions that result.