A Digital Platform That Transforms Tox Safety Knowledge Into Insights

Juan Guerrero, Eamonn O'Donnell, Vernon Brunton, Fahim Alam, Murhaf Sousli, Matthew Bogdanffy, Birgit Fogal, Charles Wood, Warren Ku, Zhichao Liu

Abstract

Boehringer Ingelheim introduces NDS Studio, a platform designed to enhance drug safety evaluations and predictive analytics. NDS Studio features an extensive safety data repository, including over 1,400 curated internal toxicology reports and datasets for more than 680 compounds. It integrates resources like SendExplorer, TG-GATES, PubChem, PubMed, and ToxCast, and offers advanced search features and adaptive data visualization choices. The platform also introduces a new technology, ToxGPT, for efficient interaction with toxicology data. NDS Studio represents a promising tool for nonclinical drug safety evaluation in pharmaceutical development.

Introduction

Drug safety is a critical concern in the pharmaceutical industry, with a significant proportion of drugs failing during preclinical and clinical phases due to safety issues. These failures not only increase attrition rates but also delay the progression of potential therapeutic candidates, leading to escalated development costs. Over the years, it has become evident that streamlining the challenging task of procuring and reviewing data to identify potential toxicology issues could greatly improve drug safety evaluations and introduce considerable savings in overall drug development costs.

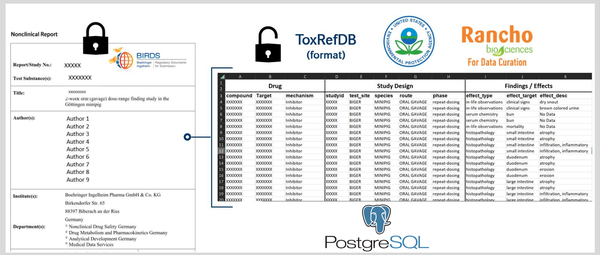
Addressing this challenge, Boehringer Ingelheim's Digital Innovation Unit introduces NDS Studio, a cutting-edge platform designed to connect users to all safety-related data in one place. The primary objective of NDS Studio is to enhance drug safety evaluations and support safety predictions, facilitated by the AcuteTox predictor, a Deep Learning model for Small molecules. The Acute Oral Toxicity (AOT) model, trained on around 30,000 labeled molecules, is deployed as an ensemble of algorithms predicting the GHS category, achieving a Matthews Correlation Coefficient of 0.55.

NDS Studio, currently in the development phase, is a lighthouse use case for the Digital Innovation Unit. It boasts an extensive safety data repository, including over 1,400 curated internal toxicology reports and datasets for over 680 compounds. The platform also introduces ToxGPT, a GenAI-powered tool that enhances user interaction with toxicology data through a fine-tuned LLM model, the ToxChatBot.

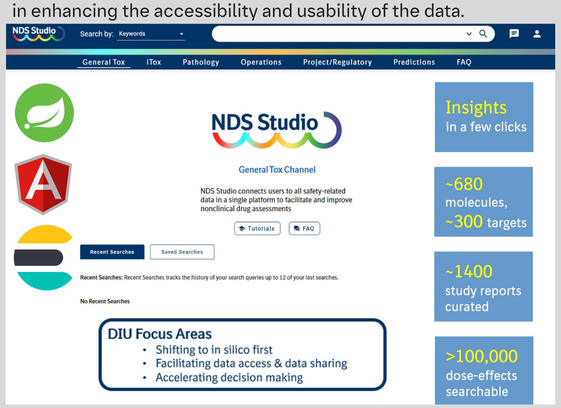
In summary, NDS Studio represents a promising tool for nonclinical drug safety evaluation in pharmaceutical development, aiming to tackle the significant challenges faced by the industry.

Methodology

The project's methodology started with curating data from BIRDS Toxicology reports, which were initially semi-structured. A rigorous process was undertaken to convert this data into a structured format, involving data extraction, cleaning, and organization. Following this, a relational data model was created using PostgreSQL technologies to provide a comprehensive structure for data analysis. Rancho Biosciences, a global leader in data curation services, assisted throughout the process, ensuring the successful transformation of the data and the creation of the relational data model. Their expertise was key in enabling effective use of the data for further research and analysis.



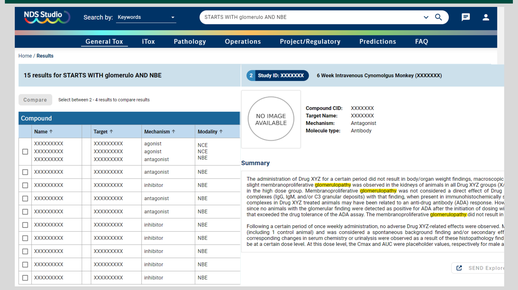
In addition to the data transformation and modeling, the project also utilized various software components to enhance its functionality. Java Springboot served as the robust backend, ensuring efficient data management and processing. On the frontend, JS Angular was employed to create an intuitive and responsive user interface. Furthermore, Elastic Search was integrated into the system to provide powerful keyword and boolean search capabilities, enabling users to easily navigate and explore the dataset. These technologies were instrumental in enhancing the accessibility and usability of the data.



Results

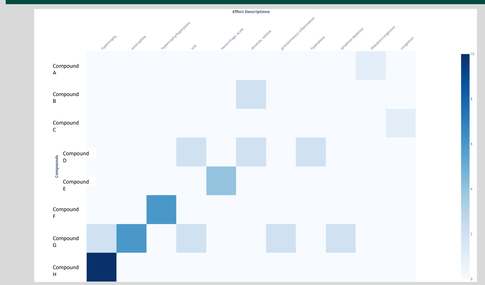
"In cyno tox studies with my NBE compound, glomerulonephritis in kidney was seen that appears treatment-related."

"Have we seen glomerulonephritis in cynos for other NBE studies?....."



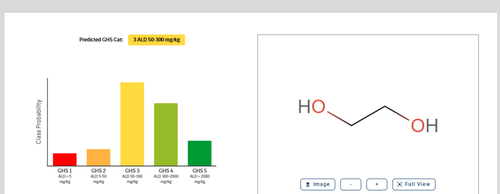
"In cyno tox studies with my compound, adrenal gland changes were seen that appears treatment-related."

"What types of adrenal gland effects have we seen in cyno tox studies?



"I would like to predict the Approximate Lethal Dose (ALD) range for the new compound that we are developing..."

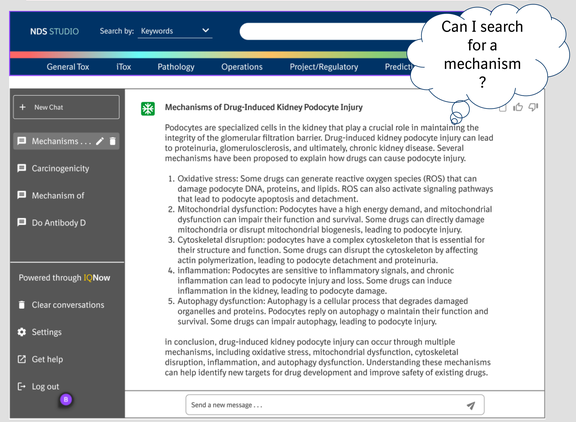
"If I enter the following SMILES string, what GHS category will be predicted?"



Results Cont.

"I would like to leverage GenAI to find more information about a particular mechanism..."

"If navigate to NDS Studio's ToxGPT can I get the information that I'm looking for ?"



Conclusion & Future Work

Boehringer Ingelheim's NDS Studio represents a significant advancement in drug safety evaluations and predictive analytics in the pharmaceutical industry. By harnessing an extensive safety data repository and advanced technologies, NDS Studio provides a comprehensive platform for nonclinical drug safety evaluation.

Looking ahead, the integration of WoE.AI into NDS Studio is anticipated. WoE.AI is a tool that gathers and summarizes information from the public domain, providing scientists with a concise overview for evaluating New Molecular Entities (NMEs). Furthermore, the development of AutoTSA is on the horizon. AutoTSA aims to automate the generation of Target Safety Assessment (TSA) reports, utilizing an AI agent to produce high-quality draft TSAs efficiently. This tool also taps into the public domain to gather all the necessary information, streamlining the process and enhancing the capabilities of NDS Studio.

As NDS Studio continues to evolve, it promises to revolutionize the way drug safety evaluations are conducted, potentially becoming an indispensable tool in the pharmaceutical development process. With the future integration of tools like WoE.AI and AutoTSA, NDS Studio is poised to become even more powerful and efficient.