Integrated Repository of Nonclinical Studies: A Case Study for a Centralized, Searchable Registry of all Nonclinical Studies for Drug Safety and Toxicology Research

Semantic Integration, Safety Data Repository, Xbiom Nonclinical, Standards Governance, SEND Review, SEND Data Visualization, Study Data Harmonization

In this case study we describe how a major Biopharma ranked in the top three, uses Xbiom Nonclinical as a centralized, shared searchable repository of all ongoing, current and legacy studies across disparate study designs serving toxicologists, drug safety and early development researchers.

Background

A top ranked Biopharma needed a central hub where all safety and toxicology analysis can be done on all their current and legacy studies across disparate drug candidates that were evaluated or developed into marketed products. The company wanted the option to search for studies, groups of subjects or specific subjects across studies, as well as species and molecules that displayed specific dose or pathology dependent patterns or signals.

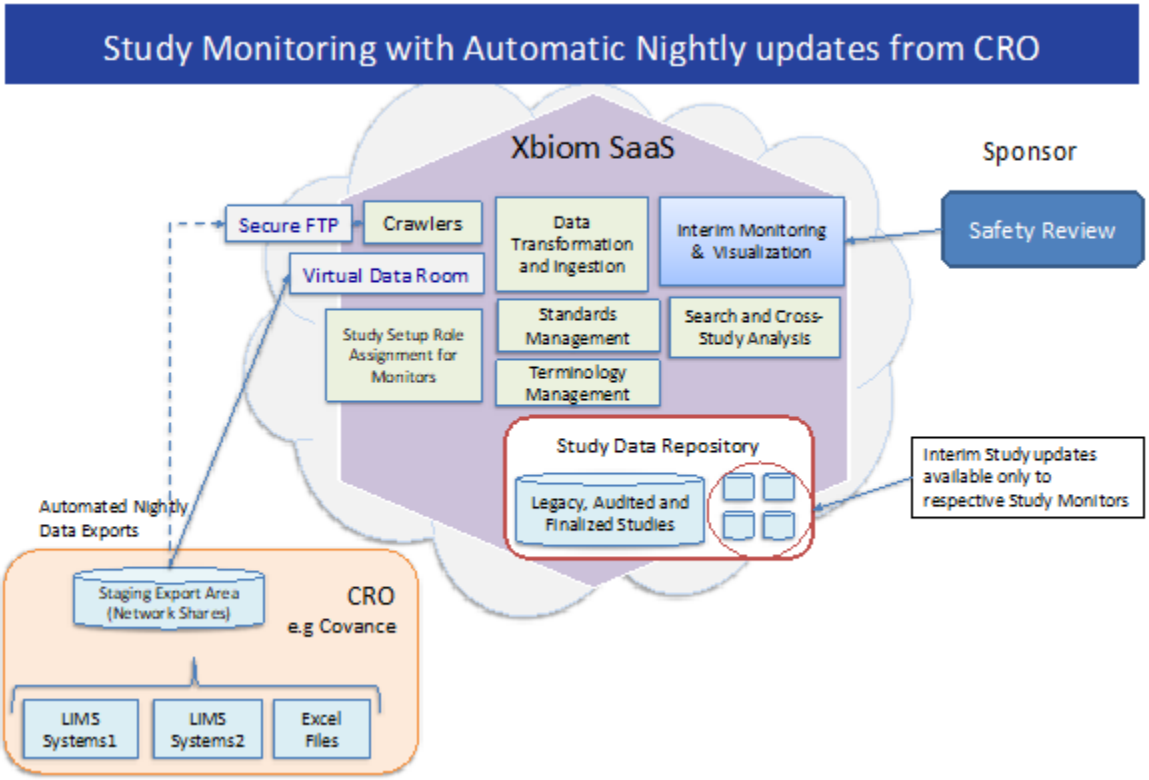
Challenge

Nonclinical study data is available in disparate forms including archived digital sources, study reports with printed tabulations, and standardized data such as in various SEND IG versions coded in various control terminology releases. Given the disparate terminologies and data models in each study, it can be a major challenge to search across these studies to identify dose related, or exposure related signals that are comparable across these studies. Therefore, transforming the studies into a universal data model that is normalized and harmonized for terminology is essential before the data is searchable. Adding on-going studies so they can be monitored in the context of other completed studies is also important for this company.

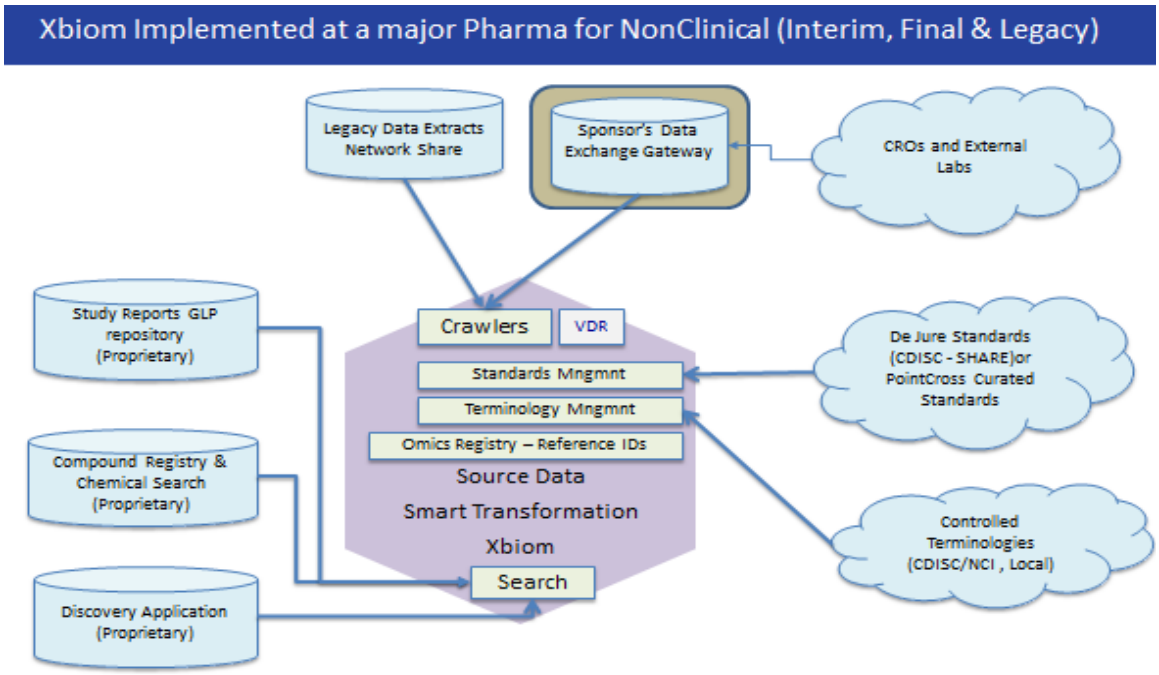
Solution

PointCross Xbiom was chosen for this searchable repository currently holding about 6000 studies including a variety of study designs, molecule families, and species. The Smart Transformation of Xbiom is used to transform and ingest source data from multiple sources including raw outputs from LIMS from on-going studies, SEND data, and custom formatted archived data from legacy studies. Ingestion workflows and data flows are automated with dashboards to support data

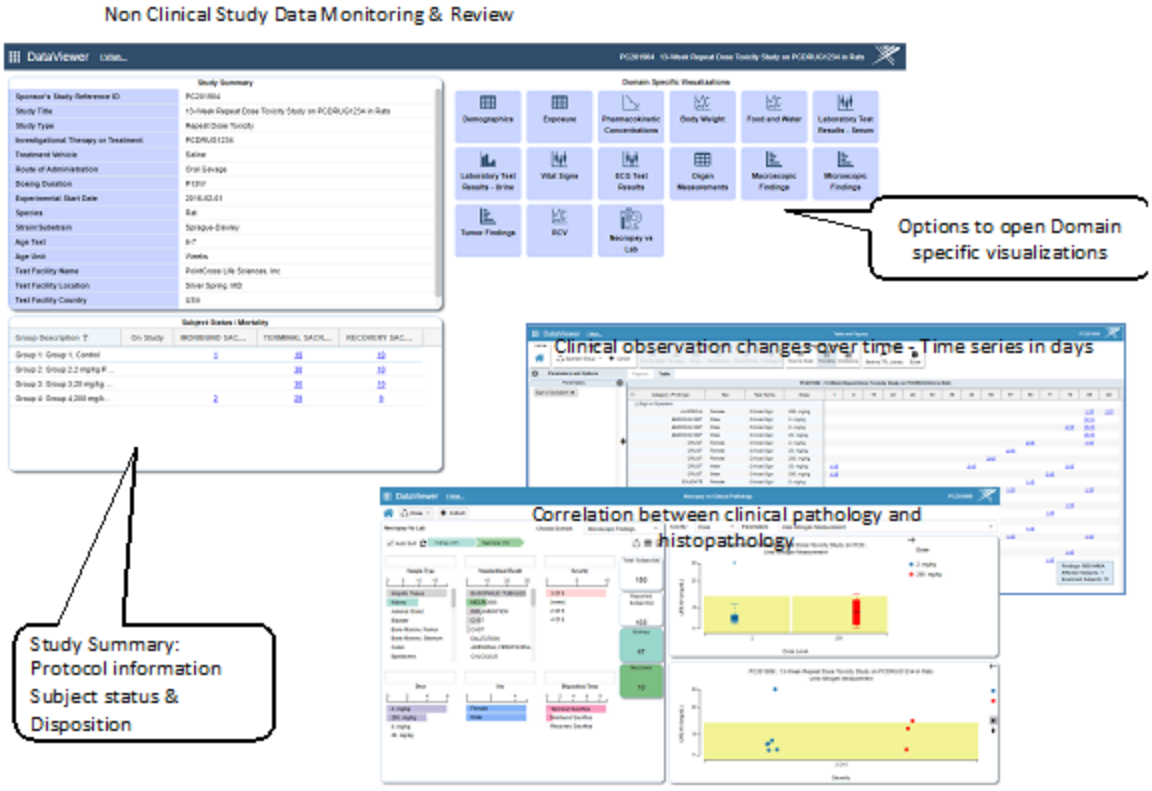
managers and QC managers. Once ingested, the studies are available for individual review and annotations of Table, Listings and Figures for collaboration and decisions.



Once the completed studies have been loaded, the studies are then added to the normalized and harmonized repository where they are indexed for search or made available through a linked object file for graph queries. Since early discovery and exploratory assays are an important part of the early development, these studies are integrated with discovery application and studies are available for search.



The addition of open O-Data APIs for researchers to run remote queries and analyze search results with Python, R and other programming tools they prefer has helped extend the power, versatility and usage of the system. The onboard TFL collaboration environment along with Jupyter notebooks add an additional dimension of flexibility to the research tools allowing, as an example, stratified cohorts and results for comparison against other pre-defined groupings.



Results

The extensive integrated repository of pre-clinical studies is now a central hub for all drug safety and toxicology research that is able to relate compound families, dose response and pathology translationally. Over 6000 studies are currently part of this internal registry.