

# nature

[ **inside**view ]



Profile Feature as seen in *Nature* 19 November 2020

# CRACKING THE CODE FOR THE IDEAL DRUG DISCOVERY DATA VAULT

A conversation with **BARRY BUNIN**, chief executive officer of Collaborative Drug Discovery



CDD Vault's web-based software platform helps scientists securely manage drug discovery data, while sharing select information for collaborations. It democratizes the drug discovery process, leveling the informatics playing field for scientists, technicians and managers. The platform contains four suites of tools: data registration, visualization, inventory, and electronic lab notebook. Barry Bunin explains how this ability to selectively share data can spur partnerships.

## What do the tools do?

Data registration, the first tool, handles the core uploading of a scientist's drug discovery data. The visualization tool allows you to do calculations flexibly and in real time, including dose response curves, scatter plots and histograms, to help users see which assay appears more selective to a specific target. Our inventory tool helps track the amounts of compounds, reagents and screening plates, making re-ordering easy. Our fourth tool, the electronic lab notebook, saves your data automatically several times a second and can be scaled to include as many users as you want.

## What makes these tools easy to use?

We've made the design intuitive and versatile for all types of drug discovery workflows. Using CDD Vault is more like using online banking and LinkedIn rather than something that requires an IT professional to manage. CDD Vault has given scientists the ability to quickly and securely capture drug discovery screening data.

## What makes searching so quick?

You can use natural-language queries of data, as well as the metadata around a particular procedure. You can search any objects. You can use controlled vocabulary we call 'pick lists' if you want to give extra structure to the data. You can handle raw and highly processed data, simple and complex Boolean

queries, as well as chemical structure and sequence searches to refine data sets. If you are just running an assay, you can pull down a list of terms and it will help automatically write the text for you. Of course, if you have text, you can do more sophisticated searches on the protocols, all with a single simple interface. All four modules for handling both data and protocols are in the fully integrated, secure, collaborative platform.

## Do you need to know any programming languages to use CDD Vault?

No. All CDD Vault's technology has simple, easy-to-use web-based interfaces that any scientist can start using immediately. Of course, the better you understand your science, the smarter you'll use it. But you don't need to be a bioinformatics or cheminformatics expert to use these tools. You just log in and go.

## How do CDD Vault's tools make it easier for people within an organization to work together?

Technicians can upload data from instruments or other databases. Managers can see what their group has accomplished. Scientists can see what compounds have been made or screened. And informaticians can control searches at the command line.

## How do the tools help collaboration?

We summarize our approach with three words: archive, mine

and collaborate. Archiving and mining complex data simply is the prerequisite for collaborating. We help scientists lower the bar in terms of getting their data, which makes it super easy to get quality answers out of their data.

## How can you promote both openness and security in the same platform?

We've put a lot of emphasis on promoting secure collaborations. Once you perform a search, you can draw a 'lasso' around the results and choose how, when, and with whom you want to share it. For example, you can share either the logic of the search or the results of the search.

## How do such collaborations play out?

Let's say you have a research organization on the other side of the world that's involved with one-third of your project. They would only see the third you grant access to. You can scale up your collaborations without scaling up your IT headaches and overhead.

## What about combining public and private data?

CDD Vault supports the full continuum of sharing, including private only, private-to-private and public data sharing.

## How does the company promote collaboration?

Part of our mission has been to build a large community of highly controlled data sharing, with an equal commitment

to both commercial efforts as well as often neglected drug discovery efforts. We make it easy to control when to share data publicly, say after patents or publications. For example, if you make your data available to the public, we will host the data you've chosen to share for free.

## What are areas that are especially robust for such free data sharing?

We've had collaborations on public health issues, like drug discovery for African trypanosomiasis, schistosomiasis, tuberculosis and malaria. CDD Vault has been used in collaborations such as More Medicines for Tuberculosis (MM4TB) with the European Union, the TB Drug Accelerator Programme with the Bill and Melinda Gates Foundation, and the Neuroscience Blueprint in collaboration with the NIH.

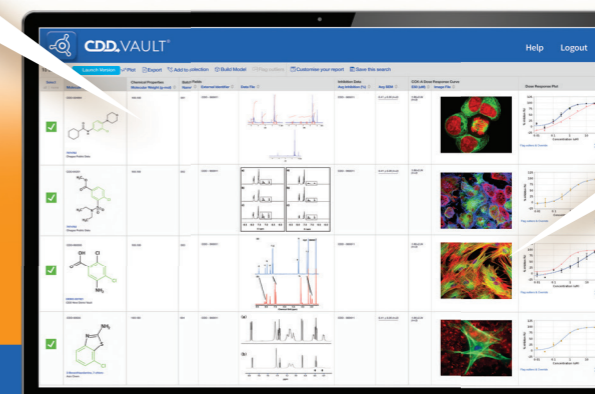
## Why is data sharing so important to your company?

We can do good for the marketplace as well as for inspiring humanitarian projects. Our goal was to create tools that allow scientists to advance their own research, while simultaneously advancing health and discovery. This is often accomplished with CDD Vault.



## Discover more with CDD Vault

CDD Vault is a complete platform for drug discovery informatics, hosted through an intuitive web interface.



Helps your project team manage, analyze, and present chemical structures and biological assay data.

## Smart Software Saves Time®

### Electronic Lab Notebook

Capture experiments digitally

Archive and search all of your experiments with ease and collaborate securely.

### Inventory

Keep track of compounds

Track and locate materials in your inventory whether in-house or shared externally.

### Activity & Registration

Store & organize your data

Store, organize and analyze your data with ease. Share and collaborate securely across project teams.

### Visualization

Plot data sets & mine them

Analyze large data assets to find interesting patterns, activity hotspots and outliers.

### BioHarmony™ Drug Data Store

Structured data on *any* drug

Real-time semantic data streams on your drugs of interest. Standardized, up to date, and ready to use.

