Simplifying the flow of drug discovery data

Dr Jonathan M.R. Davies, Product Development Manager, IDBS; Dennis Curran, Director of Business Consulting, IDBS;
IDBS Inc, 750 US Highway 202, Suite 200, Bridgewater, NJ 08807, USA.
+1 908-429-2900; dcurran@idbs.com

Abstract

Regardless of research disciplines, scientists need to easily reach the information pertinent to their research. Ideally this data access is easy. Researchers also need the ability to 'move the data around' to gain a better view or different perspective. This data manipulation needs to be straightforward. Incorporating the varying views and information required by different scientific disciplines is a considerable challenge.

This poster discusses a solution that enables extensive cross discipline analysis within a single application. Based within an intuitive, interactive and configurable platform, IDBS's SARview reporting tool enables extensive data access, complex data pivoting and multiple presentation formats in a way that allows organizations to simplify the flow of data.

Introduction

The scope of today's laboratory research has broadened in both scientific and organizational directions. The volume and complexity of the data relevant when making research decisions has significantly increased. Organizations have extended themselves through collaborations, partnerships, and utilization of contract research organizations (CROs). The result is that drug discovery is no longer a linear process. It is a set of multi-threaded pathways of parallel activities which are intended to lead to a commercial drug.

How do you simplify something that has dramatically increased in difficulty? How do you create and distribute the knowledge gained from your research programs?

IDBS believes in building tools that make scientific data easily accessible for any scientist, manager, or informatician. This is necessary even when the information is not in a form, format, or location convenient for the use of the 'data consumer'.

Workflow

Several aspects of the issues surrounding data access and analysis need to be considered, captured in the generalized workflow shown in Figure 1.

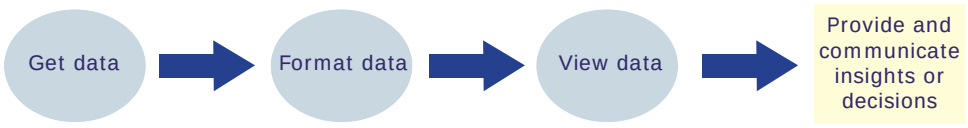

Figure 1. Typical high level view of the workflow of data analysis

Get Data - Unfortunately the data isn't always in one place and can come in different forms and from different sources – compound properties, biological data and context, graphical results and structures. A researcher needs to access different types of data to research different issues.

Format Data - Biological data can be complex with many relationships between pieces of data, which need to be maintained and presented to the user in alternative formats.

View Data - Looking at data from a fresh angle can provide new insights to researchers, who need to be able to fine-tune a view to correspond with their mental model of the data.

Communicate insights or decisions - Decision support may require computer-assisted methods or could be a scientist simply "eyeballing" the data to reach a hypothesis.

I want it to look like that

Data presentation is not simply putting a table of numbers into a table. Individual researchers need to examine the data from multiple perspectives in order to get the complete picture. Biological and chemical data together add to this complexity.

Researchers don't want to send their datasets to an expert in order to reformat the information into a new layout or with a new pivot. The ability to navigate, bin, and format the data without calling complex algorithms stored in some computer in the centralized intelligence bureau simplifies the workflow.

The example below shows screening inhibition data recorded at a number of targets. By dragging the target field, the data can be pivoted differently. The concentration context does not vary and so can be removed by simply dragging the field to the 'trash can'.

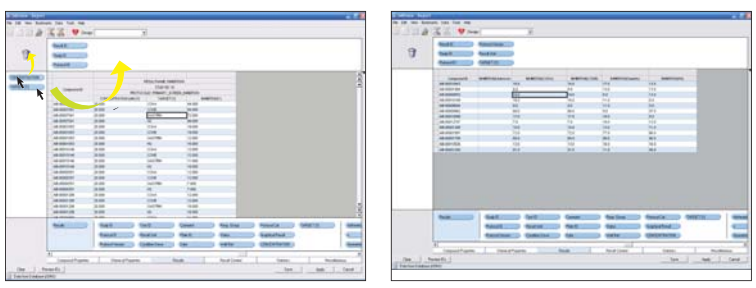

Figure 2. The interactive pivoting capabilities of SARview 6

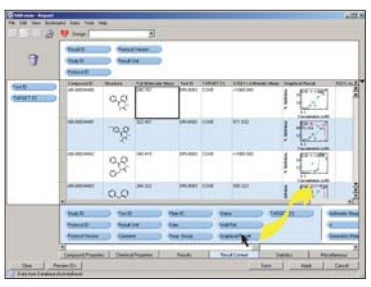

Figure 3. What you want is what you get

Generating averaged results and adding structures and curves is easy in SARview. Simply drag and drop.

Data access

Every organization has slightly different needs when accessing drug discovery data and retrieving the relevant data can be a considerable effort for the researcher and the informatics support team. The nature of databases and the requirement for the same database to capture a variety of laboratory test data often makes data seem inaccessible.

SARview solves this problem with a configurable data access layer, shown in Figure 4. Using a simple drag and drop paradigm, SARview manages various result types, graphics, and compound properties. All data access happens behind the scenes. Because SARview supports XML metadata-driven access to data, informaticians can easily direct and organize the availability of data to user groups, leaving data analysis to the researcher.

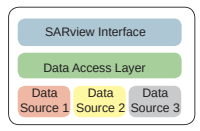

Figure 4. SARview's one interface supports multiple data sources

Different views give different news

Scientists need to see the same data in a variety of views to draw their own insights and conclusions. An application that provides a simple and intuitive approach to switching views and navigating the data set can provide significant benefit to the decision and reporting process. SARview includes built-in designers which display the data interactively as the user makes design decisions, offering three different viewing modes with sorting, filtering, and bookmarking functions.

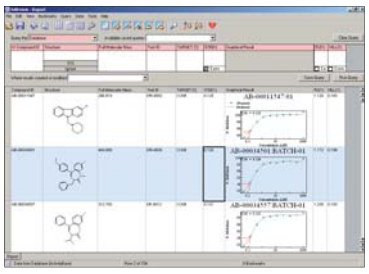

Figure 5. Tableview – a traditional SAR view that can be filtered and sorted.

- **Tableview** gives a traditional SAR-type tabular view of data, with compounds displayed in rows and associated chemical and biological properties in columns, enabling interactive sorting and iterative filtering of the data
- **Tileview** provides a view that maximizes the number of structures and curves displayed. This type of view is excellent for project review meetings where compounds are assessed for progression
- **Profileview** drills down into the data for any single compound. This view may be more beneficial for a medicinal chemist and can automatically retrieve additional data to what may be available in the report

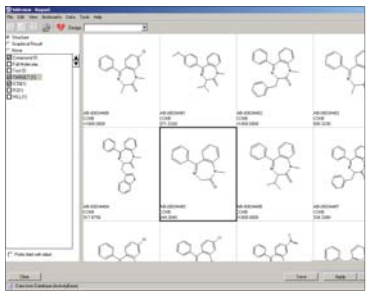

Figure 6. Tileview's interactive designer shows data as you design

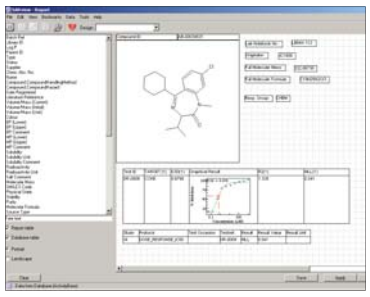

Figure 7. Profileview's interactive designer shows data as you design

A picture says a thousand words

Many effects upon data patterns and relationships are investigated visually. SARview includes an advanced graphics capability that allows scientists to create and filter pictures of the data that can incorporate up to seven dimensions. The interactive graphical visualizations include scatter plots, bar charts, histograms, and line charts – all of which can be viewed simultaneously and filtered dynamically, including bookmarking of key data across all views in synchrony.

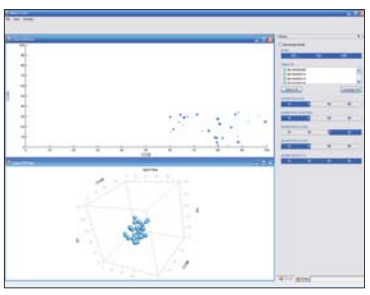

Figure 8. Interactive visualization of data in SARview.

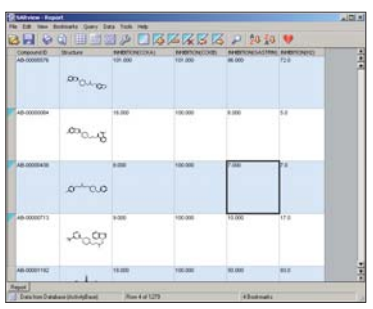

Figure 9. Bookmarking in one view is synchronized with other views.

It's good to share

IDBS brings the most common elements of working with research data together into a single reporting application. The application creates a 'portable report' of the information that can be accessed offline from the data system. Offline working also means that data can be easily shared between different sites that do not all have access to the common data repository, for example, when research is outsourced to a CRO.

A scientist makes observations and decisions while they work with their data. SARview allows the researcher to make comments and annotations about the information and have that persist in the report. By supporting a portable exchange format that allows templates of views to be shared via email or a file share system, SARview makes the sharing of an informative data view easy.

Conclusion

IDBS's SARview reporting tool simplifies the management of data for an organization's biologists and chemists, managers and administrators, and informatics professionals, in a single interface. By providing both easy access to data and flexible visualization, SARview allows researchers to access and navigate a large variety of data, structures, and graphs to extract the meaning they need from discovery data. Allowing control over the way data is viewed, SARview delivers efficient presentation that maximizes effective insight into drug discovery data.

One integrated framework for all your discovery data

www.idbs.com