Interactive Query Builder (IQB): A Lightweight Ad-hoc Assay and Chemistry Data Retrieval Tool

Jared Milbank[$], Dimitri Bondarev[&], Mary Bradley[$]

[$]FORMA Therapeutics, Watertown, MA; [&]Saber Informatics, Burlington, MA

Abstract

Interactive Query Builder (IQB) is a fast **ad-hoc query** tool complementary to our project datamart. It has an intentionally minimalistic yet functionally powerful user interface for querying enterprise assay and chemistry data across multiple databases.

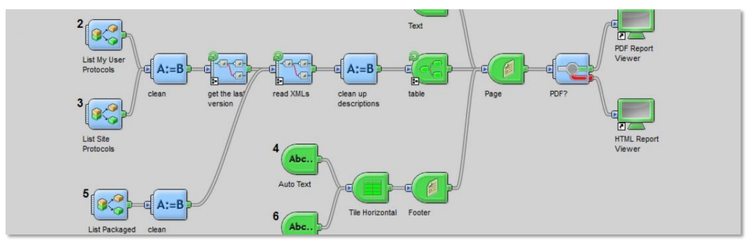
IQB was designed and implemented collaboratively with chemists and screening biologists. Functionality features were prioritized using real-time feedback on usage patterns. The iterative development approach resulted in a fast, lean and elegant application that precisely addresses the needs of our researchers.

Project Goals

- Chemists and screening biologists needed a straightforward, flexible **ad-hoc query** tool that would be complementary to our project datamart. A key requirement was that this tool have an *uncomplicated* user interface.
- Must be able to perform validation and verification of data prior to use in summarized form in project marts.
- Must be able to profile specific compounds for cross-project screening.
- Must be able to generate detailed cross-project datasets such as ADME and PK properties.
- Should have a powerful, wizard-like UI for federated (multi-source) queries against enterprise assay and chemistry data, *without* exposing the complexity of the underlying data sources needlessly to users.
- The software should require no installation, and only minimal or no user training.
- Must link out to additional reporting and visualization tools (such as Vortex).
- Must store user queries for visual, iterative refinement of query parameters.
- Must store query result sets for later retrieval.
- Must be able to run long queries asynchronously and notify users by email on completion.

Approach

IQB was built using **Pipeline Pilot** and the free **Web Design** component collection.

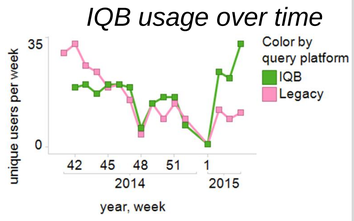


It is an interactive Web 2.0 application with a responsive screen layout that supports different screen sizes.

Collaborative, iterative development with chemistry and screening biology scientists.

We acted on **usage patterns** based on analysis of system logs and **User Surveys**. This identified key users, slow-running queries and UI improvements to overcome inefficient query definitions.

IQB usage over time



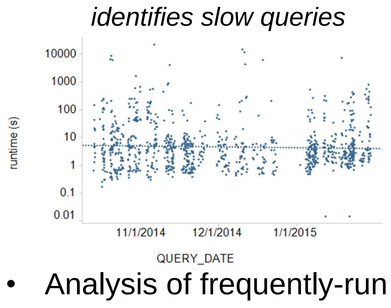
Innovations

IQB was designed and implemented collaboratively with chemists and screening biologists.

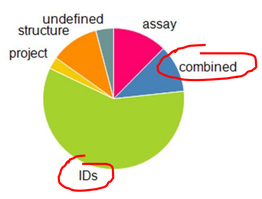
New functionality prioritized with user surveys and real-time, quantitative feedback on usage patterns.

- Online survey of existing and additional features helped drive prioritization of development efforts. For example, advanced query management features were prioritized differently by users and the development team. Analysis of the survey allowed the most user-relevant features to be brought forward in priority and others that were deemed less valuable to be moved into phase II.
- Analysis of slow-running queries uncovered cases where the UI was used to define queries in inefficient ways, and we updated the UI to offer more targeted options.

Query runtime vs. date identifies slow queries

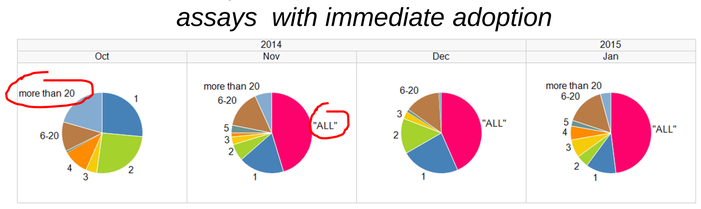


Query type analysis highlights significant and unusual cases

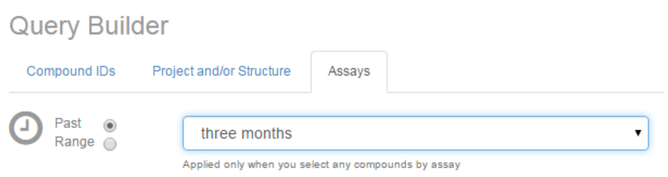


- Analysis of frequently-run queries allowed the UI to be streamlined to make those more straightforward.

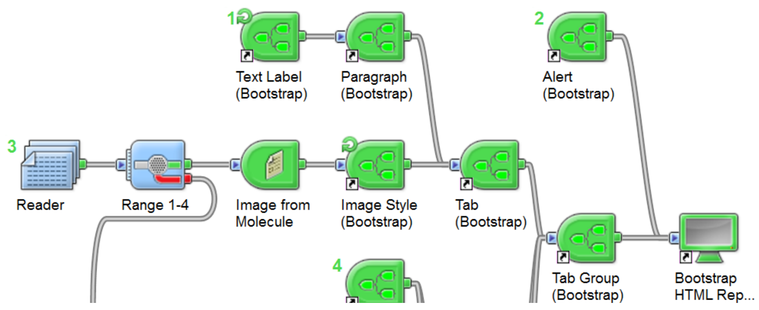
Number of assays in queries identifies use case for "all" assays with immediate adoption



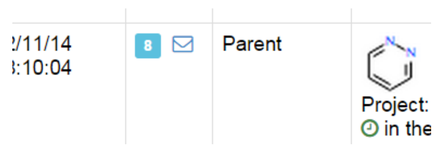
Past timeframe became key for saved queries



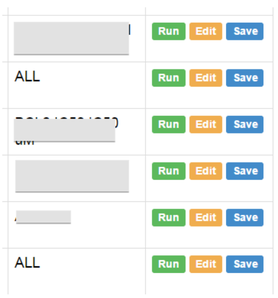
Using Pipeline Pilot and the Web Design components allowed us to focus on scientists' needs rather than implementation details and for that reason we were able to **deliver a production application to our users in weeks.**



IQB sends application alerts to its users: result sets as email attachments, query completion messages, etc. So the UI is no longer limited to a web browser: those email notifications contain **live links** that when clicked, **run as live queries on real-time assay data.**

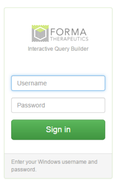


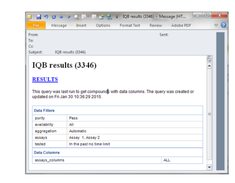
IQB stores all query versions, so users can modify their queries interactively. Intermediate results and the query versions used to generate the results are also stored. Any stored query **can be re-run on real-time data** with a simple click.



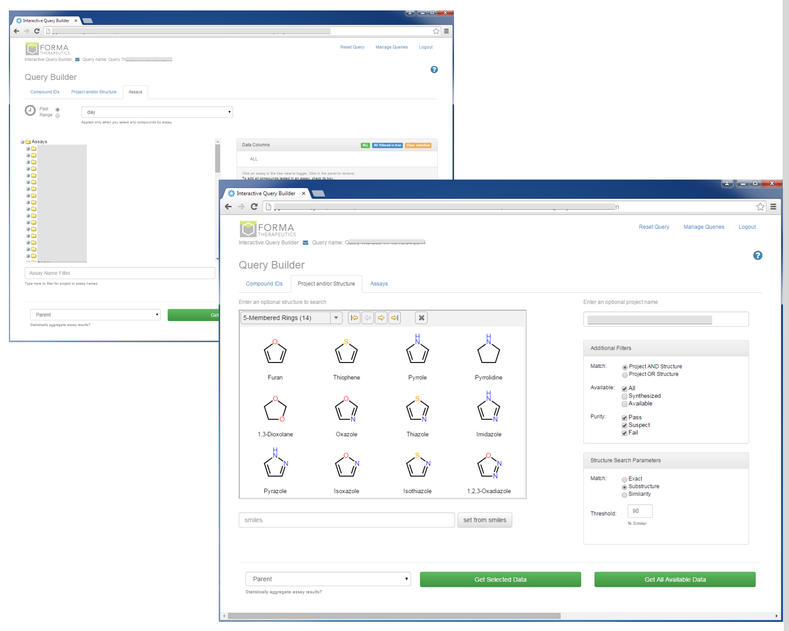
User experience is key to productivity

Consistent application across web and mail

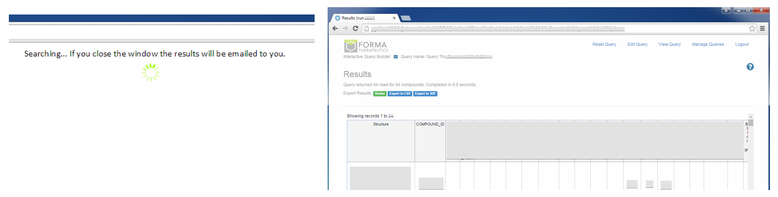




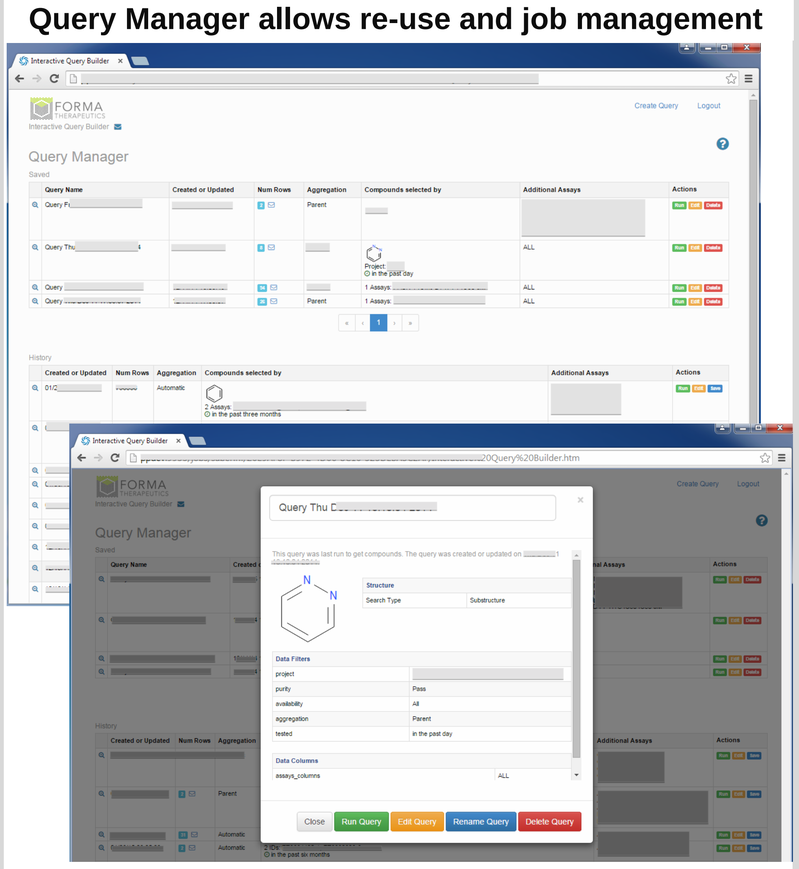
Query Builder to define by structure, assay or ID



Results sent to analytics platform of choice



Query Manager allows re-use and job management



Conclusions

- Designed and built a straightforward, flexible ad-hoc query tool.
- Close collaboration with users (scientists) is key to successful software design.
- The iterative development approach resulted in a fast, lean and elegant application that precisely addresses the needs of our researchers.
- Chemists and biologists run ad-hoc queries against enterprise assay and chemistry data from their browsers (no software installation).

For information on the **free** Web Design components for Pipeline Pilot please contact Saber Informatics LLC, a scientific software solutions company. http://saberinformatics.com