ChemAxon 2014 User Meeting Report – extracted: Migration from ISIS

A series of extracts from Wendy Warr's UGM meeting report focusing on specific issues of interest to the industry. The complete report is available at http://tinyurl.com/EUGM14

...from the conclusion

> *In my September 2013 conclusion I said that ChemAxon was probably the market leader in mainstream chemical structure handling; we have now nearly reached the time when the word "probably" can be removed from that statement. Although there are companies that can compete with parts of the ChemAxon portfolio, and some of them perhaps perform better in a certain niche, there is no single company that offers the same breadth of solutions or equal value for money....*

Migration from ISIS

The first talk, from Edith Richter of **Boehringer Ingelheim, had the theme of managing a relationship to get best results.** Two partners, ChemAxon and Boehringer Ingelheim, have a thousand differences, but one goal: to achieve migration of the ISIS platform to ChemAxon tools. ChemAxon's mission is to enable scientists to manage their chemical and related data *via* intuitive, powerful and cost effective informatics tools, developed together with customers and partners. Boehringer Ingelheim's philosophy is to serve patients, attract and retain talent, and act with integrity, honesty, transparency, fairness and full regulatory compliance.

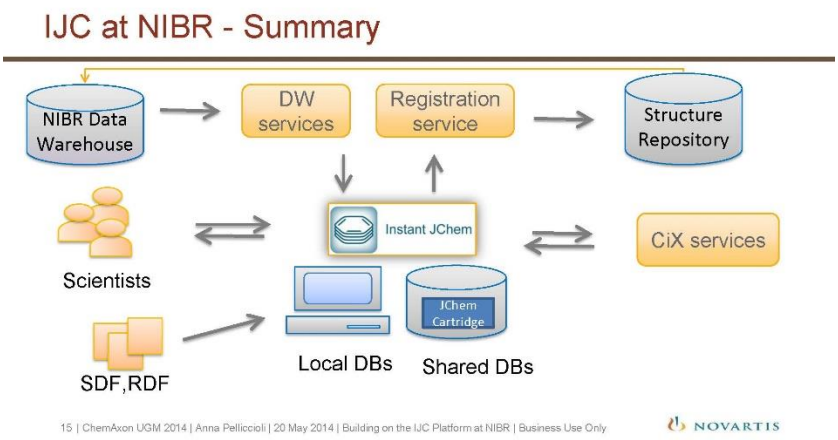
ChemAxon is 16 years' old and has 140 employees. Boehringer Ingelheim has 47,000 employees at multiple sites, and it has 1,400 IT staff globally.

The company made the decision in 2012 to migrate from ISIS to the ChemAxon tool set as the chemistry infrastructure platform. The first phase of the migration, in 2013, focused on cartridge technology and basics; the second, in 2014, focuses on desktop applications. The migration necessitates big changes. For a small company like ChemAxon making a change is akin to turning round an agile motor boat but for Boehringer Ingelheim, with its large number of users and their global distribution, implementing software changes is inevitably more complex. For example, the company has standards and policies such as the limited number of versions of Oracle and Linux that must be used for all systems. New things have to fit into the existing structure. ChemAxon does try to respond when told about these challenges.

Another challenge is horizontal and vertical communication. A big company has project lead, enterprise architect, business consultant, software developer, and database administrator functions, and many others. In a small company one person can do database administration, networking, licensing etc. A ChemAxon employee is more focused on a product (e.g., JChem Base or Calculators) than on a function. Support is now handled differently: the two companies are working together on the PC to get a solution. Roadmaps are relevant: planning what functionality appears in each release. Boehringer Ingelheim may not accept all releases. Dependencies must also be considered. Boehringer Ingelheim uses not just ChemAxon tools, but also IDBS software and Spotfire. If one of these ChemAxon partners uses JChem version 6.1, Boehringer Ingelheim cannot implement JChem version 6.3. Instant JChem, JChem for Excel and Marvin are horizontal tools across all databases: there are dependencies here too. The best approach is to talk and explain, listen and be aware, and stay enthusiastic and dedicated.

Novartis Institutes for Biomedical Research (NIBR) introduced IJC to replace ISIS in 2011. Anna Pelliccioli said that they had hundreds of little ISIS databases and a few big ones. The data had to be exported as SDfiles or RDfiles *via* IJC into local and shared databases under JChem Cartridge. It took two years to complete the migration. IJC was deployed *via* Java Web Start to make software updates and bug fixes easy. *Ad hoc* customisation for single, shared projects was done using Groovy scripting, but the code incurs a maintenance burden. NIBR-specific features for the whole user community were implemented with Java plug-in technology. Here support is easier as the code is stored in one place.

Preparative laboratory databases, analytical data and related structures, the Radioactive Compound Inventory, Building Block Archive, Chiral Separation Database and NIBR IP Priority Database (including ChemAxon's Markush Search Plugin) are now shared, searchable databases. Scientists collaborating in teams use IJC to enter data into the databases or retrieve data from them. Groovy scripts were used to create a custom application for the Radioactive Compound Inventory.



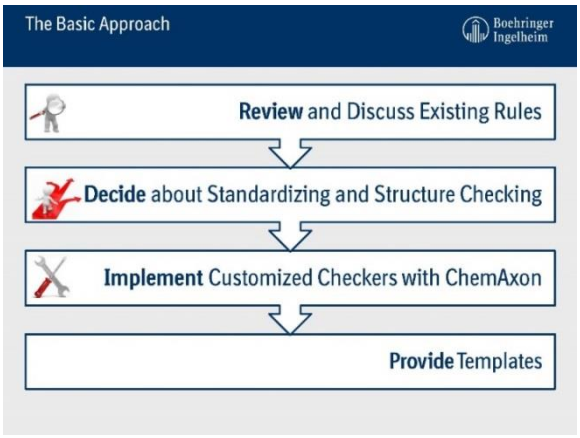
Some large structure collections that do not fit into the data warehouse are nevertheless useful to a wide community. These include the Vendor Sample Database (more than 10 million structures), legacy combinatorial chemistry library collections, and Accelrys' Available Chemicals Directory (ACD). In these cases a data administrator maintains the JChem databases, and scientists access them using IJC. The SciQuest enterprise resource management (ERM) and IJC share the same ACD data in Basel, and ChemAxon's Test to Production metadata migrator transferred the IJC project from Basel to other ERM instances in Shanghai, Singapore, Cambridge (US) and Emeryville. Work is in progress on converting the Accelrys Metabolite database.

Assay data, batch registration and calculations required in-house, Java plug-in development. Assay data is retrieved from the NIBR data warehouse using Web services and the Data Analysis and Reporting Tool (DART); forms are created and the data are annotated. IJC sits on local databases and shared databases, and links the scientists with DART and the Web Services. Scientists also access the batch registration service through IJC. They validate data against the registration service (which accesses the structure repository), highlight errors, fix the data in grid view, register from within IJC and get back identifiers. Calculating values for QSAR modelling, and other physicochemical properties, is done using the same framework as that currently used for Spotfire and KNIME access. The scientist enters a SMILES or compound number in IJC and the calculation is done *via* the Cheminformatics Framework (CiX), a platform where models are accessible to the whole NIBR community *via* Web Services.

The session ended with another talk from Edith Richter, this one more technical than strategic. As part of the migration from ISIS to ChemAxon, **Boehringer Ingelheim replaced Cheshire with ChemAxon Standardizer and Structure Checker** facilities on client and database levels. In an analogy

with word processing, standardising is equivalent to translation, Structure Checking is equivalent to spell-checking, and Structure Checking with Fixing is equivalent to auto-correction.

In four steps, Boehringer Ingelheim reviewed and discussed its existing rules, decided about Standardizing and Structure Checking, implemented customised checkers with ChemAxon, and provided templates. Some of the features that Boehringer Ingelheim required were not initially available, so ChemAxon carried out some customisation. The final templates can be called by both JChem Cartridge and Marvin tools, enabling an identical structure checker, compliant with internal guidelines, across different tools.



Packages provide functions for single- and multi-step checks and fixings, plus functions for automated checks. Database tables store the default checker and standardising configuration; customised results messages for checks; and dedicated structures for tests. Basic checkers are implemented as abbreviated groups and pseudo atoms etc. Substructure checkers work with substructures defined as SMARTS. These can be set individually in a very flexible approach. Customised checkers were provided from consultancy projects with ChemAxon; Boehringer Ingelheim was happy with the service and found the checkers easy to implement both for the database cartridge and Marvin tools.

Archive of all presentations (slides and video): chemaxon.com/ugm-archive/2014-eu/

Related presentations from previous User Meetings:

- **Migrating away from ISIS with ChemAxon's help.** Anna Pelliccioli, NIBR (2012 EUGM)
- **Integrating Standardizer and Structure Checker into corporate environment.** István Rábel, et al, ChemAxon. (2013 EUGM)
- **Great Migrations! – Approaches to moving your chemistry**. Michael Dippolito, DeltaSoft (2013 EUGM)
- **Delivering Instant JChem to the Masses - A User Perspective**. Stephen Swanson, GlaxoSmithKline (EUGM 2012)
- **Deploying Instant JChem on an Enterprise Scale.** Brett Hiemenz, GlaxoSmithKline (2011 USUGM)
- **Instant JChem as a Cloud Hosted LIMS**. Adam Idone, Chromocell (2013 US UGM)