

MDL's chemistry registration domain service

The next generation in chemistry registration

“MDL's middle-tier technology is a solid environment for building a variety of services used by applications.”

Cindy James, MDL Project and Process Management

Chemical registration is the first essential informatics step in any discovery process involving small molecules.

A sophisticated set of business rules controls how structures and data are registered, how they are searched, and how information is managed at higher levels. The n-tier MDL® Isentris™ discovery informatics platform provides the foundation for a new generation of registration systems in which business rules are managed in a middle tier capable of hosting a variety of new domain objects and services. Leveraging its unmatched industry experience, which includes the pioneering of advanced corporate chemistry registration systems for over 20 years, MDL has built and is successfully implementing a new registration service that is producing strong customer interest.

MDL Isentris—a modern informatics technology that is extensible to support current and future industry requirements—is also the platform on which MDL has constructed a new chemistry registration domain service that provides:

- A standardized registration capability that can be reused across multiple applications (e.g., traditional single compound registration, electronic laboratory notebooks, parallel and combinatorial synthesis systems, externally purchased libraries, etc.)
- Extensibility to accommodate virtually any registration application and database with many user-configurable options
- The ability to easily handle compounds that cannot be described as a single “structure” (e.g., mixtures, formulations, and structure alternatives¹)

Chemistry registration applications and services

In a typical registration process, a scientist enters structures and data into a registration application, and the software automatically applies the appropriate business rules, performs a uniqueness evaluation, assigns structure/batch ID based on uniqueness, inserts structure/batch data, associates IDs with non-structural data, inserts non-structural data, and returns the results of the processing.

The registration application software that implements this process follows a user-defined workflow to enter chemical structures and associated data into multiple tables in accordance with corporate business rules. The registration application includes functionality for drawing structures, querying data, and browsing results, using existing MDL products such as MDL® Draw, MDL® Direct (data cartridges), MDL® Core Interface, and additional focused services such as the new chemistry registration domain service, which extends the middle-tier framework of MDL Core Interface.

“MDL's middle-tier technology is a solid environment for building a variety of services used by applications,” says Cindy James of Project and Process Management who is managing the project through its first iterations. “For example, MDL's new chemistry registration domain service consists of server-side software that evaluates a finite data set—a structure, a compound, or batch sample—based on established business rules governing structure normalization and uniqueness. If the data are accepted, they go into a predefined database structure. Programmers using this new domain service appreciate its open API, as well as its configurability and extensibility to fit in virtually any registration application and database.”

A flexible, multi-functional registration service

The chemistry registration domain service works together with MDL Core Interface to provide sophisticated business logic and a common application programming interface for inserting structure-based discovery data into a predefined database. The service manages compound and batch objects independently in accordance with established business processes, automating the generation of compound and batch IDs. Structure normalization and novelty check tools apply the business rules for determining structure uniqueness.

The functionality of the chemistry registration domain service is constantly evolving to provide the flexibility required by MDL customers. In the first iterations, the service includes the following customer-requested functionality:

- Accepts single or multiple batches via API or file
- Supports both synchronous and asynchronous registration
- Interfaces with MDL® Cheshire scripts or Cheminformatics Business Rules Manager (CBRM)
- Executes MDL Cheshire scripts for structure normalization
- Performs duplicate/uniqueness check on normalized structures
- Returns a corporate ID for each compound and a batch ID for each batch
- Inserts both structures and data into the chemical registration database
- Returns messages (status, error, warnings, Cheshire, etc.)

(continued on page 10)

¹ An example of a structure alternative would be two reactions, each yielding a separation of a mixture (of diastereomers or enantiomers). The result is two samples characterized by structure A or B, which are difficult to distinguish without additional analysis or experiments. Both need to be registered and both should get different compound IDs. The database structure of the chemistry registration domain service is capable of distinguishing between compounds characterized by the same “A or B” structure.

(continued from page 7)

Expanding middle-tier chemistry services

MDL has observed a strong market demand for additional middle-tier services similar to the chemistry registration domain service. All the attributes of a modern informatics technology like MDL Isentris—openness, flexibility, scalability, and extensibility—are driving the development of chemical warehouse, calculation, combiChem, experiment management, procurement, and literature services, to name but a few.

Aleksandar Ruzicic of Global Biopharma Solutions says, “MDL will standardize other middle-tier services to provide essential business logic to MDL customers. In the context of a registration solution, MDL has already built an enumeration domain service based on MDL Cheshire for chemists performing parallel or combinatorial chemistry experiments. Another area MDL is currently exploring is an experiment management service that optionally includes the ability to register reactions.”

Charles Buse, Ph.D., of Solution Design and Implementation and the scientific and technical leader of the project, observes, “The new chemistry registration domain service, maintained and supported by MDL, provides a robust, server-side foundation for any corporate registration application. In addition, it will effectively support any related workflow application that requires compound/batch registration into a corporate database—including an electronic laboratory notebook application.”

Towards a viable chemistry electronic laboratory notebook

A corporate chemistry registration application, usually centered around compound and batch information, is the cornerstone of most chemical

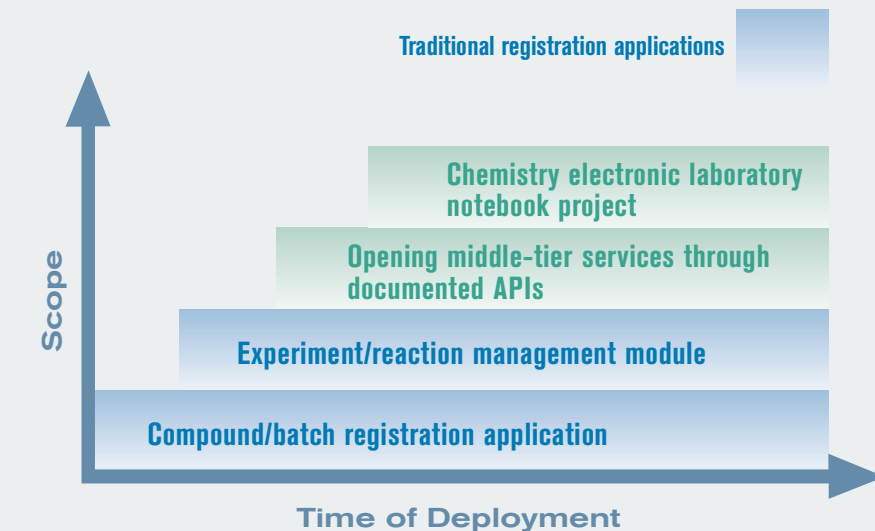


Figure 1. Compound/batch registration and experiment/reaction management capabilities are the cornerstone of chemical R&D. By opening middle-tier services through documented APIs, organizations can extend these capabilities to encompass compound sourcing, logistics, electronic notebook, and other applications.

R&D activities. Organizations may subsequently add components to extend the capabilities of these applications to manage experiments and reactions. When considering the introduction of chemistry electronic laboratory notebooks, organizations should investigate how they can best tap into the functionality of these existing application components.

Mario Dimayuga, Ph.D., of Global Corporate Solutions stresses that the most effective way to do this is to redesign registration applications following the n-tier architecture paradigm. With reference to Figure 1, Dimayuga asserts that R&D organizations need to “expose the specific business logic for reuse by opening its underlying functionality as middle-tier services

and making these services available through documented and accessible APIs.” Dimayuga continues: “This serves as the distinct mechanism through which all registrations are processed. This applies not only to electronic notebooks, but also to compound sourcing and logistics applications. Without this single mechanism, scientists would be able to enter compound/batch information into the corporate database through various unregulated mechanisms, making it impossible to maintain a reliable and coherent corporate chemistry database.”

A White Paper on the chemistry registration domain service is available from your Account Manager, who can also put you in touch with an MDL expert for further discussions. ■