MAPS – A service-oriented, customizable, multi-purpose LIMS

Dieter Kopecky[1,*], Otto Weichselbaum[2] and Silvia Fluch[1]

[1]AIT Austrian Institute of Technology GmbH, Health&Environment Department, Bioresources, Konrad Lorenz Strasse 24, 3430 Tulln, Austria

[2]SEW Software Engineering Weichselbaum, Heindlstrasse 19, 4040 Linz, Austria

ABSTRACT

Summary: *MAPS* is a LIMS that was developed to be applicable in multiple fields of life sciences (like genomics, genetics, proteomics) and for various applications. It uses an innovative, flexible data model and comes with numerous technologies for adaptation, integration, and extension. In this way, *MAPS* especially meets the frequently changing requirements of R&D laboratories.

Availability and Implementation: *MAPS* is freely available at http://www.picme.at/maps. It can work with any Hibernate-supported DBMS at the data layer and has been used in routine applications with PostgreSQL 8.3. JBoss 4.2 and SEAM form the application layer, while JSF and ICEfaces are used at the presentation layer. All major current browsers are supported by the system.

Contact: dieter.kopecky@ait.ac.at

Supplementary information: Further details about *MAPS* can be found at http://www.picme.at/maps.

1 INTRODUCTION

Sample management in medical and biological laboratories is increasingly challenging as the number of samples grows along with the advances in analytical technologies. Especially within small-to-medium-size Research & Development (R&D) laboratories, finding and establishing a Laboratory Information Management System (LIMS) capable of adapting to the various, frequently changing requirements is critical. The need for permanent adaptation of lab protocols and analysis workflows contradicts the usually non-recurring tailoring of an off-the-shelf LIMS to lab requirements during product launch. Furthermore, the broad spectrum of commercial to open-source software applied in modern laboratories requires concepts for seamless integration.

2 METHODS AND RESULTS

MAPS (Material Administration and Preparation System) is a new concept of LIMS, which tries to meet these requirements with an innovative data model as well as a modern service-oriented architecture on top of state-of-the-art web technologies.

2.1 Data Representation

MAPS uses a hybrid data representation strategy which tries to benefit from the advantages of structured data storage and document management. Every material and every container – these are the two central storage items – can be described by an arbitrary number of documents, which themselves consist of an unrestricted number of attributes. While in traditional document-based applications documents consist of unstructured content, in *MAPS* the attributes of a document are required to be specified by name, data type, and range of appropriate values. Apart from all basic data types, the system allows the storage of any kind of files and media formats particularly, georeferences, URLs, and literature references. Users can add data types by building on the existing ones, including the creation of enumeration types and composite types (for complex values like annotation data). This hybrid data organization comes with the flexibility of document management to store any type of information, and the strictness of structured data storage to preserve data consistency and integrity. The possibility to search for any type of information stored in the database is another valuable effect resulting from this data model.

2.2 System Customization

MAPS is customized via a graphical administration user interface, which allows the specification of container types and geometries (racks, boxes, plates, vials, tubes, etc.), of material types (DNA, RNA, blood, tissue, etc.), and of consumable supplies, amongst others.

Laboratory processes and standard operating procedures are defined with a set of graphical wizards supporting this step. The so-called task definitions are described by necessary preconditions for their execution, modifications to containers, materials, and any document attribute during execution, as well as the supplies being consumed. Both, manual and robotics-driven task definitions are supported, including the possibility to specify lab-specific mapping protocols and command file generation.

MAPS supports multiple laboratories or projects with an arbitrary number of users in terms of virtual databases within a single installation. All definitions (including documents and task definitions) are private to a virtual database, but can be shared with other virtual databases steered by a powerful permission system.

2.3 Technologies for Integration

MAPS comes with three different technologies to integrate legacy systems, application software, and external devices.

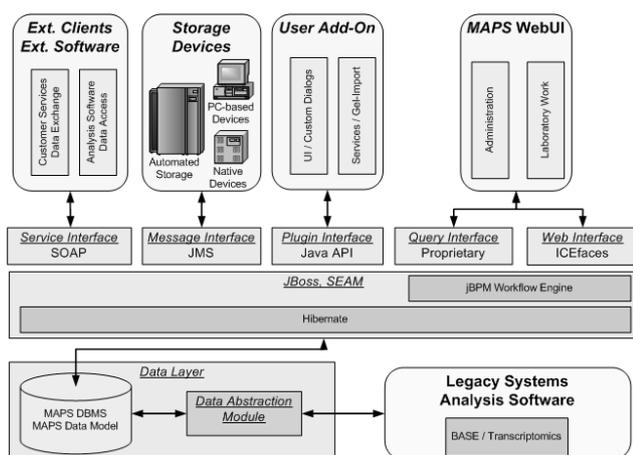
[*]To whom correspondence should be addressed.

Fig. 1. System Architecture of *MAPS*

First, there is a data abstraction module, which simulates the expected data model to a legacy system or analysis software, whose data repository partly overlaps with that of *MAPS* and whose complete data storage requirements can be fulfilled by the underlying database. The data abstraction module uses DBMS-specific features to realize this functionality, resulting in a rather tight integration and therefore a certain amount of dependency. The Microarray Analysis Workbench BASE2 (Vallon-Christerrson *et al.*, 2009) was integrated with *MAPS* this way.

Second, *MAPS* is able to use JMS (Java Messaging Service) to communicate with external software or devices. JMS is an open messaging standard and therefore supported by many vendors to facilitate message exchange between integrated systems. We have used JMS to set up the communication between *MAPS* and an automated Brooks (Nexus, Inc.) -20/-80° sample storage device for 1.5 million samples. In this way, the end user can access the organization and reporting capabilities of the storage device seamlessly, including container storage and retrieval, without the need to work with two different applications. JMS is intended to be the data exchange interface with PC-based devices.

Third, *MAPS* offers a SOAP (Simple Object Access Protocol) interface which offers services to query information on samples stored within the LIMS as well as to enter new samples into the system. These services can also be made available across organizational borders, so that, for example, customers are able to query information about their samples via Internet. Appropriate security mechanisms are used here to prevent unauthorized access.

2.4   Extending the Basic System

Apart from the organizing and documenting processes, queries against the LIMS database, reports having to be generated from a LIMS, as well as methods to analyze data stored in the LIMS change over time. *MAPS* provides two mechanisms to extend the capabilities of the basic system to meet these needs.

MAPS comes with a proprietary query interface allowing to query the database for any kind of stored information as well as to create reports from these data. The user can add new queries for search and report generation with the administration component described above. In this way, users can immediately react to upcoming or changing reporting needs.

The most powerful interface is the Plugin mechanism providing access to the core functionality of *MAPS*. This Plugin Interface allows a Java developer to hook into the core *MAPS* services and to add new functionality to the system. This can either be exten-

sions to the graphical user interface (like customized data import or export dialogs) or services being scheduled for iterative execution in the background of the core system. An example for such a background service is an automatic agarosis gel evaluation component (Bajla *et. al*, 2005) which regularly checks for available image data from the gel analysis system, creates an interpretation for the available images, and stores the results into the *MAPS* database.

2.5   Implementation Details

MAPS is a modern J2EE application making extensive use of the SEAM framework and the Hibernate API. Therefore, MAPS can work with any Hibernate-supported DBMS at the data layer and has been developed and used in routine with PostgreSQL 8.3 und JBoss 4.2. The ICEfaces framework was used to implement a modern AJAX-based user interface giving the user the impression of a usual desktop application. By applying these technologies, *MAPS* can run on many server operating systems including Microsoft Windows, Linux, and Mac OS X. Furthermore, *MAPS* supports all major web browsers, including Mozilla Firefox, Microsoft Internet Explorer, Google Chrome, and Apple Safari.

The graphical end user interface is entirely rendered dynamically from the system and document specification which was stored into the *MAPS* database during system customization. Therefore, no adaptations to the user interface are necessary in case of changing laboratory processes or data structures.

3   DISCUSSION AND CONCLUSION

A great deal of LIMS are currently available on the market. A compilation at http://www.scientificcomputing.org is listing as many as 200 commercially and freely available implementations. Still, many of the open systems – including recently published ones – are tailored at a specific purpose, e.g. (Van Rossum *et al.,* 2010) or (Voegele *et al.,* 2007). The numerous commercial implementations are quite flexible but often go beyond the financial scope of many R&D laboratories. Available toolkits for automated generation of a LIMS (Huang *et al.*, 2011) provide a cheap and quite new approach, but the resulting systems usually do not fulfill the complex requirements of the users. *MAPS* is a new LIMS which is able to deal with complex data requirements (like data sets from Next Generation Sequencing) due to its flexible data model and to adapt to upcoming and changing requirements due to its numerous technologies for adaptation, integration, and extension. MAPS has been applied in some large-scale projects with varying demands, including cattle sample genotyping and a large plant gene repository centre (http://www.picme.at).

REFERENCES

Bajla, I. *et al.* (2005) An Alternative method for electrophoretic gel image analysis in the GelMaster software. Comput Methods Programs Biomed, doi:10.1016/j.cmpb.2004.09.007.

Huang, Y.W. *et al.* (2011) WIST: Toolkit for rapid, customized LIMS development. *Bioinformatics*, doi: 10.1093/bioinformatics/btq676.

Vallon-Christersson, J. *et al.* (2009) BASE - 2nd generation software for microarray data management and analysis. BMC Bioinformatics, doi:10.1186/1471-2105-10-330.

Van Rossum, T. *et al.* (2010) SLIMS—a user-friendly sample operations and inventory management system for genotyping labs. *Bioinformatics*, doi:10.1093/bioinformatics/btq271.

Voegele, C. *et al.* (2007) A Laboratory Information Management System (LIMS) for a high throughput genetic platform aimed at candidate gene mutation screening. *Bioinformatics*, doi:10.1093/bioinformatics/btm365.