Agile Business Process Management in Research Projects of Life Sciences

Silke Holzmüller-Laue[1] and Bernd Göde[2]

[1] Center for Life Science Automation (celisca), University of Rostock, F.-Barnewitz-Str. 8, 18119 Rostock, Germany
[2] Institute of Automation, University of Rostock, R.-Wagner-Str. 31, 18119 Rostock, Germany
{silke.holzmueller-laue,bernd.goede}@uni-rostock.de

Abstract. In life science laboratories the sub-process automation of methods with semi- or full automated, isolated solutions called islands of automation and several IT systems dominate. There are deficits in networking of these sub-processes. The R&D processes in the life sciences research are complex, flexible, unstructured, knowledge-intensive, distributed, and parallel; they use heterogeneous resources, and combine automated, semi-automated, and manual activities in high-variable process chains with a high number of control structures. This characteristic of LSA-processes makes high demands on an integrated process management that contains an interdisciplinary collaborative process control and the documentation of the global process from for example purchasing, sample storage, method development to analytics and interpretation of results as well as the extraction of knowledge.

Using methods, techniques, and tools of business process management (BPM) in the life science automation offers the potential to improve the automation level, the networking and the quality of the life science applications. The authors investigate the suitability of the standard based methods and techniques of BPM for the introduction of a flexible, integrated, and automated process management in the heterogeneous and hybrid systems in life sciences. This article will be focusing on the advances made in distributed workflow automation in highly variable system and application environments of research projects by use of BPM methods and tools.

Keywords: Agile R&D Processes, BPM, Life Sciences, Process Automation, Workflow Management.

1 Introduction

The aim of use of BPM method and tools in the life science automation is an effective automation of the overall processes in fast changing heterogeneous R&D environments and to reduce the developer effort significant. The topic of this paper is to present the idea, the concept, the challenges and the advantages of using business process management methods and tools in the life science automation as well as first results and experiences.

J. Grabis and M. Kirikova (Eds.): BIR 2011, LNBIP 90, pp. 336–344, 2011.

1.1 Challenges in Life Science Automation

Flexible laboratory automation is essential in the most research projects in life sciences. In addition to time constraints and economic reasons, a further goal of automation in pharmaceutical and biotech companies or institutes is the improvement of experimental results and workflow [1]. Research and development in the life science automation have been working with semi- and fully automated islands of automation systems (laboratory robot systems, chemical parallel reactors, biological high throughput screening systems,…) and methods to increase the throughput in the specific sub-processes (high throughput screening / high content screening / high throughput analytics,…) over the last few years [2]-[5]. However complex automation demands on wide-ranging process chains have often met their limits in integration. Isolated solutions for specialized operations such as synthesis and analytics still often dominate in lab automation. There are also complex automation systems, usually robot-assisted, that support more wide-ranging functionalities. Results include complex screening systems for biological testing of potential active agents. The systems available have often only covered a limited part of a continuous and interdependent workflow in lab applications up to now; processes such as synthesis, analysis and biological testing of new potential active agents require different systems that need to be adapted to the application requirements that apply.

Automated systems are usually controlled and monitored via a largely self-contained process control system (PCS) that provides users with specific control panels for process definition and control. Local process control by PCS for heterogeneous components in a robotic system and, increasingly, dynamic local optimization of procedure has become the state of the art [6]-[10]. Process modeling and notation as well as execution languages are only locally valid for certain robotic systems, as they are usually proprietary and manufacturer-dependent.

Several IT-solutions are established for specific parts of data management. These are used parallel for different activities: instrument-specific PCS capture data from analytical measurements (e. g. chromatography, mass spectrometry, optical methods like fluorescence or absorption spectroscopy et al.), bioinformatics systems deal with biological data (genome analysis, sequence analysis, structural bioinformatics), LIMS (Laboratory Information Management Systems) serve the sample, order, research study, and project management, CIMS (Chemical Information Management Systems) register chemical structures and their features, and DMS (Document Management Systems) manage operating procedures (SOP's) [11]. The standardized data exchange between these manifold systems together and with tools for data analysis, statistics, and visualization (like MS Excel, TIBCO Spotfire, MathWorks Matlab etc.) is still insufficiently. The handling of the exponentially growing of data and information flood is often a weakness in the laboratory workflow. Alternative algorithms and calculation rules with a low automation level are used for the evaluation of the data. Data driven decisions are often integrated. The results are available as specific files or reports or in databases. Personal undocumented knowledge reduces the traceability and the reproducibility. The comprehensive obligatory documentation requirement – in sub-processes is taken for granted - cannot be fulfilled [12]. For these reason the process of knowledge extraction in research projects is delayed.

1.2 Previous Works

The authors have been working successfully in several innovation levels of information systems over the past years. One of the first fully web based laboratory information management system (LIMS) with online automation networking was developed and was successfully integrated in a research center of the chemical industry [13] in addition to the use in the center for life science automation (celisca). This LIMS is application-independent by a user-defined parameter library and a method repository. Currently chemical, biological, and preventive medical research groups work with this system. Core functionalities are the administration, the organization and documentation of the laboratory workflows and research projects on user-defined abstraction level and with any structuring. An electronically laboratory notebook (ELN) is integrated for the unstructured process documentation in the development as well as operation phase. The progressive technical automation inclusive the generated data amounts requires an automation of data management, data manipulation and knowledge extraction. Therefore the basic LIMS are expanded from a pure measurement management system to an integration platform. This is a framework solution for a user-defined process description, a configurable process communication for the automated exchange of process data, the process visualization as well as a template based process data manipulation [14],[15]. Additional parts support the compound library management and the document management. In summary it is a data warehouse for all relevant data and information in life sciences applications (raw data, related meta data, derived data, process descriptions, information about chemical structures, storage information,…) expanded by facilities for data manipulation, visualization and communication. Using such system is a basic necessity without that a data interpretation and knowledge extraction is not effective practicable by reason of the enormous data amount in the environment of high throughput screening and high content applications (HTS/HCS).

The limits of this system are in consequence of its interactivity. For example, the detail level of the process description is user-defined. The user has many possibilities to integrate freestyle text, graphics, photos, documents. But it is the decision of the users, in which degree they use this possibilities in a specific process description. Because of this the user has the responsibility for the completeness and the traceability of a process description. However a complete process documentation is very important for the quality assurance and the quality assessment.

A networking of process descriptions is possible but the users apply this across working groups and across applications only rarely. The LIMS is an information and management system for measurements and experiment results. There are two ways of assignment to each other. The manual allocation by import or capture requires a high accuracy of the user to assign the measurements to the correct process instance. The automated way by the process communication uses for example the barcode identification. The configurable process communication allows the data exchange by files, databases and XML. These interfaces offer only the indirect way of device controlling with manual intervention due to a lack of integration facilities on side of devices. Connected to that, a timer-controlled collecting of data from lab instruments and an user-initiated visualization of this data provide only a limited monitoring.

2 BPM-Approach in LSA

Automation is established for the execution of the actual experiment. Preparatory and accompanying sub-processes (storage, maintenance, method development), are mostly time-consuming manual workflows with insufficient process integration. The use of methods, techniques and tools of BPM for agile processes in life sciences moves the overall process into the center of attention. The final goal is to develop an overall system with a comprehensive reproducible process control and process monitoring from the established autonomous automated sub systems (pipetting robot, multi parallel reactors,…) or islands of automation (robot lines that automate complex, high level sub-processes). In addition to these fully-automated sub-processes, manual process steps and activities, which are generating decisions and knowledge, will be integrated. The comprehensive automation of the overall workflow, which covers all involved, any structured automation components across the application fields is the logical further development of the automation of individual working steps. BPMS takes over the role of a systems integrator or rather the role of an integration platform for the process control and support so the interoperability. Means and methods of the library-based process definition and process documentation are needed.

In the interdisciplinary area of life science automation several disciplines cooperate in a common project. It is a major advantage to speak the same language. A documented process description in BPMN provides a uniform basis for the implementation of an automated process and its execution.

The direct controlling of laboratory instruments and a central resource management are realizable based on a service-oriented architecture (SOA). The overall process controlling reduces the down times. Further advantages are accrued for the quality assurance: the kind and the form of the documentation are defined once and are applied for all instances in the same manner. Therefore the reproducibility is guaranteed: there is a detailed process description in form of an executable model and the personal knowledge is integrated in this model. The integration of data analysis tools reduces the need for time-consuming and cost-intensive manual intervention.

2.1 Requirements on BPM in LSA

For an automation of the overall process are required:

- Integration of laboratory instruments by several technical interfaces
- Agile R&D processes with short life cycles require a high flexibility; there are unstructured sub-processes
- The process models are understandable for the process experts and adequate detailed for an process control
- Common repository for process models allows a reusing of sub-process models by adaption and parameterization
- Support of the business process life cycle: process analysis, modeling, execution with resources scheduling (instruments, lab personnel, services, software), monitoring, optimization based on analysis of results

- Integration of the human workflow (user task management) and external partners (e.g. e-mail or SMS to initiate manual intervention in error case or as status information)
- Data-driven decisions
- Rule management (easy to edit business rules, that control process activities, allow a fast adaption of the process und support data-driven decisions)
- Integration of several software products (data acquisition, data analysis, visualization)
- Data manipulation across the instances (common evaluation of experimental series)
- Integration of the document management including document generation, editing and exchange integrated in the message flow
- Support of quality assurance.

2.2 The Concept of BPM Integration

The approach of figure 1 shows an open BPMS, which communicates with any heterogeneous and hybrid components. The displayed SOA-components cover reprehensive important IT-solutions of the hierarchical laboratory automation. A set of adapter and connectors to the several systems is needed to control and exchange data. Sub-processes with integrated connectors, which are stored in the repositories, can easy be combined in new process models by the expert staff. Personal knowledge will be generally available. A BPMS integrates all stakeholder of the process. Thereby the advantages of the BPMS are opened up for the research processes, that are very flexible, unstructured, previously not completely defined, and often data driven. A major challenge is the handling of large data amounts integrated in the message flow of the overall process.

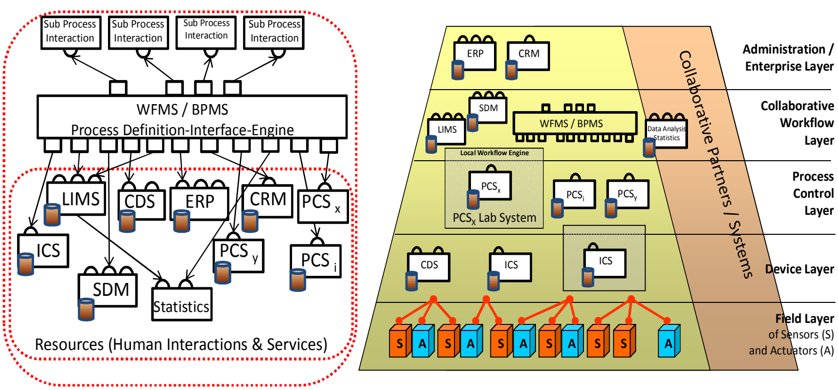

Fig. 1. General view of BPMS as integration platform in life science automation

(Laboratory information management systems - LIMS, Chromatography Data System – CDS, Instrument Control System - ICS, Scientific Data Management – SDM, Process Control Systems – PCS, Enterprise Resource Planning – ERP, Customer Relationship Management - CRM)

2.3 First Results and Future Works - An Example R&D Process

The in the following described models are developed in an evaluation study that investigates the applicability of BPM methods in the life science automation. Workflows of chemical, biological and medical applications are analyzed within this study.

First results are related to a project of the preventive medicine. This following example (Figure 2) describes a telematics platform solution for the fitness and stress monitoring applications based on mobile data acquisition with wireless personal area networks [16]. This system is used in preventive medicine studies, that investigate for example the consequences of the automation for the laboratory staff.

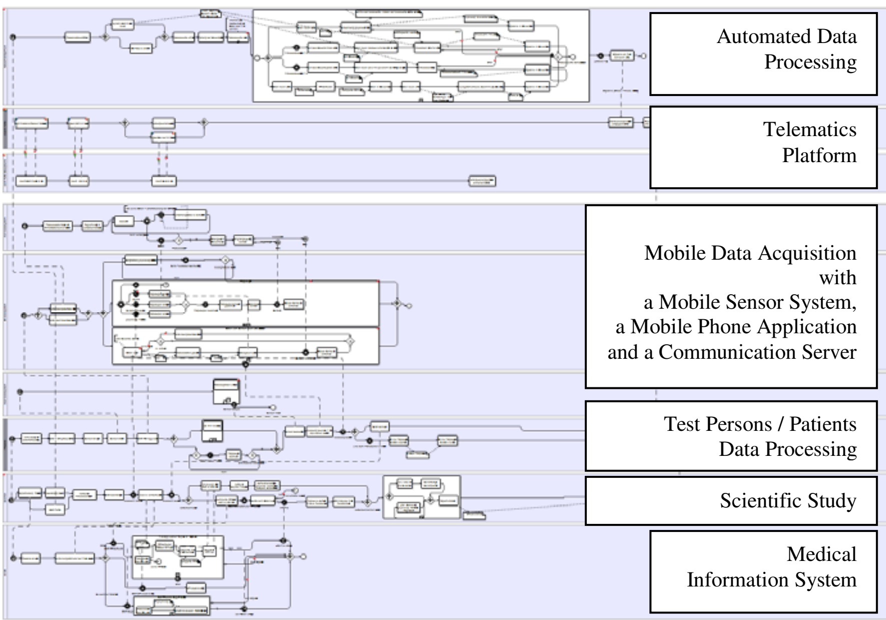

Fig. 2. Model of the overall process of the example telematics application for stress and fitness monitoring

The patient is continuously monitored in the every-day-activities. Patients use their usual mobile phone and a wireless sensor system for the acquisition of several physiological parameters. In addition, test persons are able to add parameters describing subjective assessments during the executed activities, e.g., to document their current physical and psychical state to get an according subjective strain rating. All these data are transferred by wireless communication to a communication server that stores this data in a raw database [17]. Fully-automated, intelligent data processing modules derive secondary information about stress and fitness in in-time calculations (Figure 3). As result for the patient a graphical classification of the calculated value is available on a web-based portal. In a parallel research study the captured and derived data is investigated relating to influences on the physiological parameters, and develop algorithms and models for the calculation of further secondary information.

For the data management within this research study the medical information system (1.2) are used. A first part of this workflow is shown in Figure 4. Manual and user tasks dominate this sub-process now. An increased automation level is preferable.

For the modeling of this processes BPMN are used with the goal to execute. With BPMN 2.0, there is a standard that provides a modeling notation as well as a process execution language.

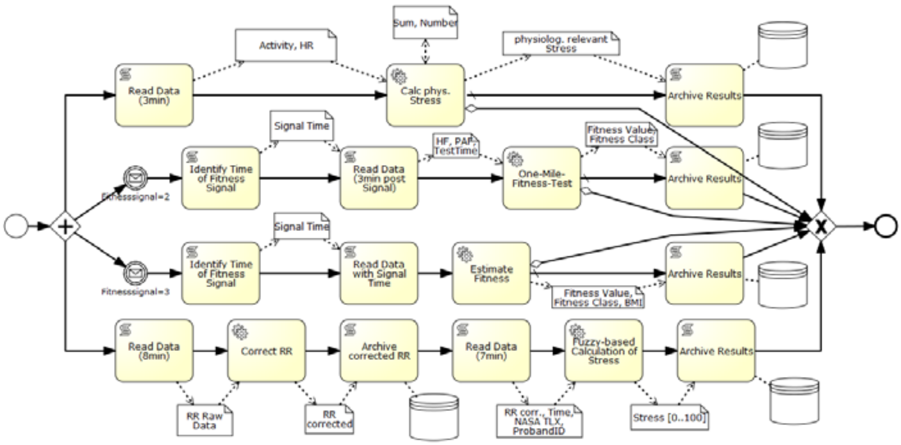

Fig. 3. Model of the sub-process of data processing executed by web services

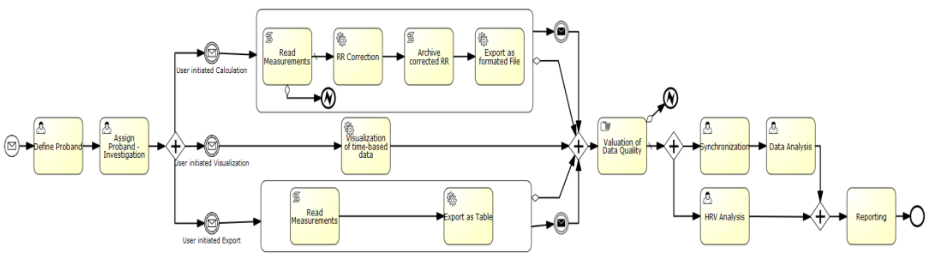

Fig. 4. Model of a part of the scientific study using the information system

In a first step of the study several BPMS are evaluated. Two basically different ideologies are to cite: BPMS based on zero-code-programming and developer-friendly BPMS. For the evaluation study at first the community version of intalio|BPMS [18] and the open source collaboration platform Activiti [19] are used. The modeling with the intalio Designer is intuitive. Simple processes with user task (a dialog editor is integrated), database access, integration of web services and sending of e-mails are implemented comfortable by process experts with few IT skills (SQL and XML is useful). Data structures can be defined and filled with content by mapping. The current version of intalio|BPMS does not support BPMN 2.0 specification. Our investigations will be continued, if intalio|Cloud is ready to use [18]. Activiti is a BPM – platform targeted at business people, developers, and system admins. It is open-source and distributed under the Apache license. Activiti offers many possibilities for use: the engine runs in any Java application, on a server or in the cloud. The tooling around Activiti facilitates an innovative and practical collaboration between business people and developers. Its core is a BPMN 2.0 process engine that has the focus on being light weight and easy to use for Java developers. The components of Activiti support the process modeling, the human task management, the administration, the running, the monitoring, and the collaboration between business people, developers and IT operational people. Thus Activiti is a complete BPMS but also a framework for developer, with that an own application can be realized [19]. In the future works, further life science applications are realized with the focus on scheduling of lab instruments and lab robots, aspects of quality assurance, and knowledge-intensive sub-processes.

Acknowledgments. The German Research Association (DFG) is gratefully acknowledged for funding this project.

References

1. Kalorama Information: The Worldwide Market for Lab automation. New York, USA (2008)
2. Elands, J.: The Evolution of Laboratory Automation. In: Seethala, R., Fernandes, P.B. (eds.) Handbook of Drug Screening, pp. 477–492. Marcel Dekker, New York (2001)
3. Haney, S.A. (ed.): High Content Screening – Science, Techniques and Applications. Wiley-Interscience Series on Mass Spectrometry. Wiley Interscience, Hoboken (2008)
4. Janzen, W.P.: High Throughput Screening - Methods and Protocols. Humana Press Inc., Totowa (2002)
5. Zhang, M., Nelson, B., Felder, R. (eds.): Life Science Automation: Fundamentals and Applications. Artech House Inc., Boston (2007)
6. Moore, K.W., Newman, R., Chan, G.K.Y., Leech, C., Allison, K., Coulson, J., Simpson, P.B.: Implementation of a High Specification Dual-Arm Robotic Platform to Meet Flexible Screening Needs. JALA - Journal of the Association for Laboratory Automation 12(2), 115–123 (2007)
7. Russo, M.F., Sasso, A.: Modeling, analysis, simulation and control of laboratory automation systems using Petri nets: Part 1- Modeling. JALA - Journal of the Association for Laboratory Automation 10(3), 172–181 (2005)
8. Bühl, W., Daniel, J., Höynck, M., Jänicke, W.B., Szarowski, S.: Laboratory resource planning for quality control of pharmaceuticals. Pharmazeutische Industrie 66(11 A), 1430–1434 (2004)
9. Rodziewicz, P., Bell, B.: Overview and architecture of the Java integration framework, hybrid scheduler, and Web-enabled LIMS. JALA - Journal of the Association for Laboratory Automation 9(6), 411–420 (2004)
10. King, J., Gosine, R., Delaney, B., Norvell, T., O'Young, S.: Discrete event control and dynamic scheduling for tele-robotic mining. CIM Bulletin 96(1069), 116–118 (2003)
11. Lämmel, J.: Systemintegration im Bereich der Laborinformatik. LaborPraxis 10 (2004)
12. Dellert-Ritter, M.: LIMS - Aktueller Stand und zukünftige Trends. GIT Laborzeitschrift 54(1), 39–42 (2010)

13. Thurow, K., Göde, B., Dingerdissen, U., Stoll, N.: Laboratory Information Management Systems for Life Science Applications. Organic Process Research & Development 8(6), 970–982 (2004)
14. Göde, B., Holzmüller-Laue, S., Haller, D., Schneider, I., Thurow, K.: Flexible IT-Plattform zur automatisierten HTS-Wirkstoffanalyse. GIT 9, 741–744 (2007)
15. Göde, B., Holzmüller-Laue, S., Rimane, K., Thurow, K.: Integrierte flexible Datenverarbeitung in einem webbasierten LIMS: Idee und Praxis eines Excel-Prozessors in Serverapplikationen. Chemie – Ingenieur – Technik 12, 2043–2049 (2007)
16. Berndt, R.D., Takenga, M.C., Kuehn, S., Preik, P., Stoll, N., Thurow, K., Kumar, M., Weippert, M., Rieger, A., Stoll, R.: A Scalable and Secure Telematics Platform for the Hosting of Telemedical Applications. In: Case Study of a Stress and Fitness Monitoring. IEEE, 13th International Conference on e-Health Networking, Application and Services, Columbia, USA, June 13-15 (accepted, 2011)
17. Neubert, S., Arndt, D., Thurow, K., Stoll, R.: Mobile Real-Time Data Acquisition System for Application in Preventive Medicine. Telemedicine and e-Health 16(4), 504–509 (2010)
18. http://www.intalio.com
19. http://www.activiti.org