The System Architecture of the Open European Nephrology Science Centre

G. Lindemann, D. Schmidt, T. Schrader, M. Beil, T. Schaaf, H.-D. Burkhard

Abstract—The amount and heterogeneity of data in biomedical research, notably in interdisciplinary research, requires new methods for the collection, presentation and analysis of information. Important data from laboratory experiments as well as patient trials are available but come out of distributed resources. The Charite Medical School in Berlin has established together with the German Research Foundation (DFG) a new information service center for kidney diseases and transplantation (Open European Nephrology Science Centre - OpEN.SC). The system is based on a service-oriented architecture (SOA) with main and auxiliary modules arranged in four layers. To improve the reuse and efficient arrangement of the services the functionalities are described as business processes using the standardised Business Process Execution Language (BPEL).

Keywords—Software development management, Business data processing, Knowledge based systems in medicine

I. INTRODUCTION

THE domain of Nephrology covers the diagnostic and treatment of kidney diseases including the organ replacement (dialysis, transplantation). Illnesses of the kidney cause a considerable cost factor in the German health service and worldwide. Considering Germany, so annually more than 2.5 billion Euros are necessary for the financing of dialysis treatments, whereby the incidence of the dialysis-requiring kidney insufficiency rises annually around 5-8%. The transplantation of a kidney as alternative to dialysis is reduced by the small availability of donor organs. Nevertheless the mentioned domain is well circumscribed and in relation to other medical topics small. The typical problems of research in biomedical disciplines are apparent: a limitation of the amount of cases of rare diseases, heterogeneous data types with varying quality of automatically and "by hand" generated data, and decentralized data retention inhibit the scientific work.

G. Lindemann, D. Schmidt, H.-D. Burkhard: Humboldt University of Berlin, Institute of Computer Science, visitors: Rudower Chausse 25, Unter den Linden 6, 10099 Berlin, Germany.
Contact: [lindeman,hdb]@informatik.hu-berlin.de, danilo.schmidt@charite.de
http://www2.informatik.hu-berlin.de/~lindeman (German site)

T. Schrader: Department of Pathology, Charite Universitätsmedizin Berlin, Chariteplatz 1, 10117 Berlin, Germany.
Contact: thomas.schrader@charite.de

Although both standards for the processing of clinical information (HL7, DICOM) and Ontologies (UMLS) exist, their integration is insufficient into a scientifically oriented information system. Therefore the establishment of a centre for research information is for the development of new treatment concepts in medicine of crucial importance.

The main task of OpEN.SC is the implementation of a metadata repository for clinical data, data of studies, literature and virtual slides. The data will be stored and proceed by an intelligent data management tool - the Intelligent Catalogue. The structured data will be specific presented for different purpose types such as looking for a rare disease or a specific diagnosis or a distinct pattern in the virtual slides of a case. The system works as a data centre to improve the availability of standardised raw data including digitalised glass slides (Virtual Microscopy) and works as an interface for scientific publications. Moreover, in future this meta-data respository can serve e.g. as an e-learning system for young physicians in the mentioned field [8].

The Open European Nephrology Science Centre – OpEN.SC - is a co-operation project of the Charité, University Hospital Berlin, the Humboldt-University and Free University of Berlin and includes in the core the four Departments for Nephrology (Charité: Campus Mitte, Campus Berlin Buch, Campus Virchow Klinikum and Campus Benjamin Franklin), the Institute of Pathology, Charité Campus Mitte, the Institute of Medical Informatics, and the Lab of Artificial Intelligence of the Institute for Computer Science of Humboldt-University.

Every one of the four mentioned nephrology departments treats 3000 – 4000 medical cases per year. The availability of all medical cases of Berlin's university nephrology departments, the sum of about 12.000 cases, would open completely new possibilities of detection of rare disease patterns and quality management. Concerning the data of the project partners, their properties and their preferences, the service can promote the creation and organisation of a science network with national and international partners. Many relations and basic collaboration exist on national and international level e.g. Barcelona (Spain), Linz (Austria), Lodz (Poland) and Paris (France) yet.

In a first working phase we developed a structured system architecture based on a Service Oriented Architecture (SOA) proposed by Krafzig [2]. In this paper we describe in short the main modules of the whole system. In a second part we introduce the underlying business processes and conclude with remarks to current project processes.

II. SYSTEM COMPONENTS AS MODULES

The design of the system architecture of OpEN.SC concerns the various changing requirements over the time and domain specific needs. The project structure is based on distributed teams with very specific views, different needs and requirements as well as from technological point of view on various information systems at the backend. The processes of patient documentation and collecting data for medical research differ highly in every clinical department. Therefore the system architecture implements the model of a Service Oriented Architecture (SOA) using various web services (Fig. 1).
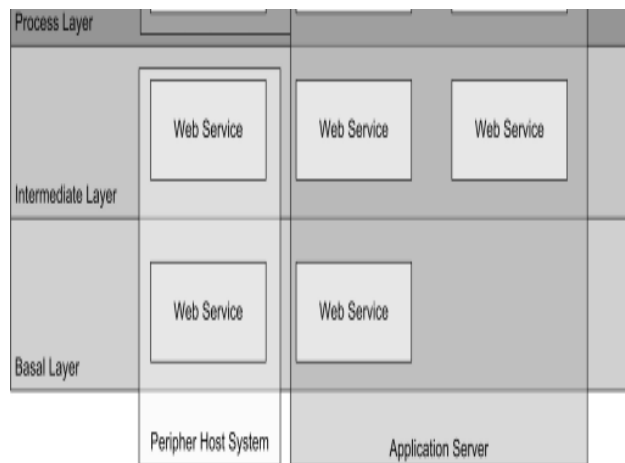

Fig. 1: General SOA Architecture (modified from Krafzig [2])

The OpEN.SC Meta Data Repository consists of some main web service modules in the intermediate and process layer and some auxiliary modules in the basal layer. The main modules implement the core functionalities: collection, store and intelligent presentation of data and support the collaboration of scientists. Let us mention the main modules:

- OpEN.SC-Portal-Module – This module is located in the enterprise layer and is responsible for the user interaction. A graphical user interface presents all structured information about the cases, the images and the members of the network.

- Security Role Module – A specialized Web Service controls and coordinates the access of information and services due to a fine granulated user right management. The rules of use will be designed at the Medical Advisory Board and implemented in this module at the intermediate layer.

- Case Description Module – Different sources of data will be collected and analysed within OpEN.SC. A main module is the Case Description Module. Every event in the clinical database will be handled as an event called case, independent from the patient identity.
  The different row data from for instance **TBase©** [4],[12] letter of the physicians and Virtual Microscopy will be analysed using techniques for data mining. The new cases (events of the clinical databases) will modified to a flexible, network like structure which is the basis for the case based retrieval.

- CRN-Service Module [6]– This module realises the Case Retrieval Net (CRN). A Case Retrieval Net is a network which represents the cases as graphs with weighted edges. The Network consists of nodes to identify the cases (case descriptors) and another node type to represent information units to characterise the case.
  Every case descriptor is linked via edges of relevance to the nodes of information units. Between the information units are edge weights as expression of the result of a similarity function between two information units calculated during the retrieval process. As retrieval process the Spreading-Activation Algorithm will be used. The structure of a CRN is most appropriate and flexible to add new cases and information units.

- Virtual Microscopy Module – A cornerstone of the OpEN.SC is the storage of complete digitalised histological slides [9]-[12].

Some further modules are essential for the whole system and reflect the basic technology for instance:

- BPEL (Business Process Execution Language) Compensating Transaction Module – a module responsible for the orchestration of the implemented web services

- Web Service Addressing Module – this module handles the communication layer of the web services.

III. BUSINESS PROCESSES IN OPEN.SC

One software engineering goal is the implementation of a generic software architecture for reuse in other scientific fields with the same technological problems. A lot of scientific processes are stunted due to the distributed resources of data and the difficult problems of the integration of various data bases and data types [1],[3],[5]. In establishing a meta data repository the difficulties to hold a high level of data quality should be concerned.

Based on the described software architecture a set of atomic web services can be effectively arranged and reused not only for various purposes but also for the same purpose but different resources and environments under two conditions:

1. The functionality and used cases are described and understood as business processes. This means the utilisation of an additional abstraction layer in system analysis and design.

2. A standardized description language is used to enhance the system independence of a web services software model. Such language is the Business Process Execution Language (BPEL) [7].

Beside the aspect of modeling the BPEL file can be processed by an engine to arrange and coordinate the web services (orchestration of web services). The BPEL-engine is a web service itself.
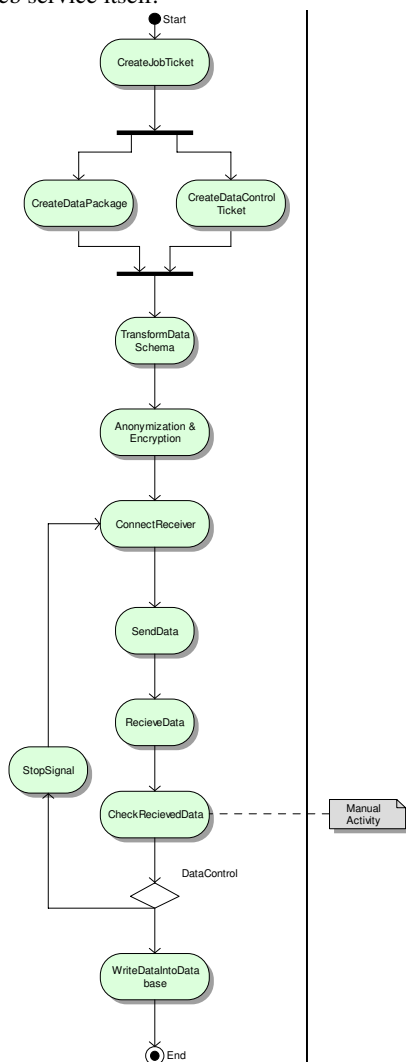
Fig. 2: BPEL activity steps

For the functionality of the data entry a complete automatically data upload is not appropriate for the different resources and data types. For some steps the manually controlled correction and confirmation are necessary. The different steps can be modeled as a process consisting of automatic steps and manual interaction.

To visualise the data entry process the different activity steps are modeled with UML shown in figure 2. In the abstraction layer of a process different web services are embedded: security rule module and OpEN.SC-Portal module. Some auxillary atomic modules are appropriate for this functionality: encryption/decryption, anonymisation via monikers, and validation.

These modules can be reuses for other data resources and types. The process is independent of the operating system, the type of resources and data structure.
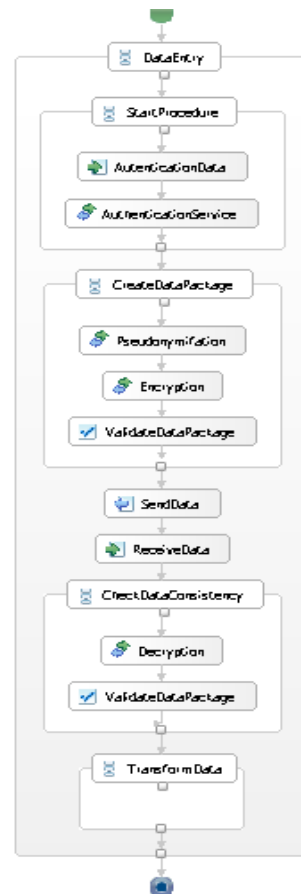
Fig. 3: BPEL-Model for the data entry process

The result of the process modeling is a description file with the embedded services like a chain. The BPEL engine executes this file, invokes the web services in the right sequence, controls the data flow, and monitors the complete process.

Beside the complex data entry procedure at least two further focuses will be modeled by BPEL:

1. the knowledge distribution,
2. the collaboration.

Knowledge distribution as well as collaboration consists of various activity steps using different web services. Knowledge distribution comprises a process chain with the security rule module, the CRN module for data analysis and structuring, and the virtual microscopy module to manage large scaled image data files (called virtual slides). The use case is to find new patterns and interesting cases to create a new scientific question, study or evaluation.

The process of collaboration is highly flexible and depends on various factors: the specific personal background and preferences, the aims and motivation to use the meta data repository.
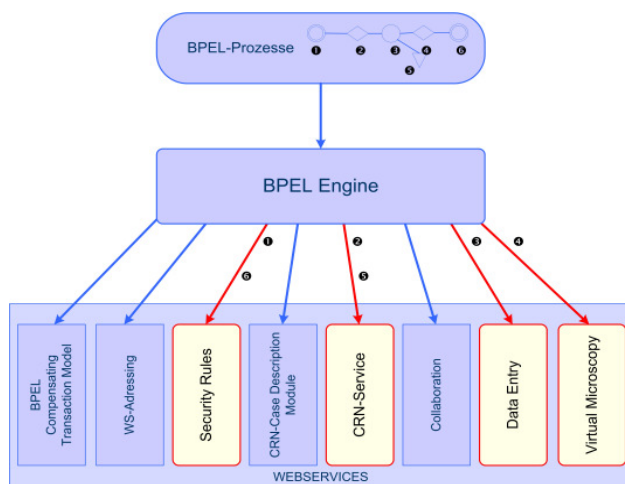

Fig. 2: Orchestration of the web services by BPEL engine

Many decision points, manually interaction and dependencies of other services and actions of user reflects the omplexity.

IV. CONCLUSION

In information systems for purposes a highly flexible software architecture is essential for user acceptance, and usability. Beside the complexity of data heterogeneity and distributed resources the target user group has various and varying conceivabilities and requirements. The system architecture should cover these specific situation by enhancement of flexibility and scalability. The software oriented architecture and the additional abstract layers of business processes form the basis for these requirements. The functionalities are modeled as business processes using the standardised language: Business Process Execution Language.

BPEL allows to arrange and orchestrate atomic web services for specific needs independent from the system environment. Software components as web services can be reused for the various purposes. The development of the Open European Nephrology Science Center includes the data entry, storage and processing. The close relation between clinical data, pathology report and virtual slides inaugurates completely new possibilities in data analysis and scientific work. The access to a large data pool will promote also the collaboration between the scientific participants.

REFERENCES

[1] A. Colman , LD. Pham, J. Han, J-G.. Schneider. Adaptive application-specific middleware. Proceedings of the 1st workshop on Middleware for Service Oriented Computing (MW4SOC 2006) MW4SOC '06. 2006

[2] D. Krafzig, K. Banke, D Slama „Enterprise SOA. Service Oriented Arcietecture Best Practices" Prentice Hall PTR, 2004

[3] Y.S. Kuo, L. Tseng, H.C. Hu, N.C. Shih. Document management: An XML interaction service for workflow applications. Proceedings of the 2006 ACM symposium on Document engineering DocEng '06. 2006

[4] G. Lindemann, L. Fritsche. Web-Based Patient Records - The Design of TBase2. In Bruch, Köckerling, Bouchard, Schug-Paß (Eds.) New Aspects of High Technology in Medicine; Seiten 409-414; 2000

[5] Q. Li, D.K.W. Chiu, P.C.K. Hung, S.C. Cheung. Flows and views for scalable scientific process integration. ACM International Conference Proceeding Series; Vol. 152 archive Proceedings of the 1st international conference on Scalable information systems.2006

[6] M. Minor, C. Biermann. Case-Acquisition and Semantic Cross-Linking for Case-Based Experience Management Systems. In Du Zhang, Taghi M. Khoshgoftaar, and Mei-Ling Shyu, editors, Proceedings of the 2005 IEEE International Conference on Information Reuse and Integration (IRI-2005), pages 433 -- 438, IEEE Systems, Man and Cycbernetics Society (SMC), 2005

[7] OASIS. (2006-11-30) WSBPEL. Available: http://www.oasis-open.org/committees/document.php?document_id=18714&wg_abbrev=wsbpel

[8] A.-B. M. Salem. "Intelligent Technologies for Medical e-Learning." In: Proceedings of the Third International Conference on Intelligent Computing and Information Systems. Cairo, Egypt, 2007

[9] T. Schrader, P. Tietz, C. Tennstedt, M. Schwabe, T.-N. Nguyen-Dobinsky, P. Hufnagl. Attributing images of interdisciplinary medical examinations in databases. E J Pathol 032 032-006. 2003

[10] T. Schrader, S. Niepage, T. Leuthold, K. Saeger, S. Hellmig, P. Hufnagl „Implementation of a pathology report compiler with integrated diagnostic path functionality in the Diagnostic Virtual Microscopy" Pathol Res Pract 200[4], p. 356, 2004

[11] T. Schrader, S. Niepage, T. Leuthold, K. Saeger, K. Schluns, P. Hufnagl, K. Kayser, M. Dietel. The diagnostic path, a useful visualisation tool in virtual microscopy. Diagn Pathol. Nov 8;1(1):40. 2006

[12] K. Schröter, G. Lindemann, L. Fritsche: TBase2 – A web-based Electronic Patient Record. Fundamenta Informaticae 43, 343-353, IOS Press, Amsterdam, 2000

Gabriela Lindemann has studied mathematics at Humboldt University of Berlin and received her diploma and PhD degree there in the field of applying discrete optimization methods in VLSI circuit design. With the foundation of the Institute of Computer Science of Humboldt University she entered a research position at the Lab of Artificial Intelligence there. Since more than 15 years she has been working in several applications of AI in medical informatics.
During the last years Dr. Lindemann has published in the fields of Multi-Agent Systems, Case-Based Reasoning and medical computer science. Moreover she was co-organizer and has been a member of programming committees of a lot of international conferences. She is a member of the academic senat of Humboldt University of Berlin.

Danilo Schmidt studied computer science at Humboldt University of Berlin. In 2006 he got his Diploma degree. With the beginning of the year 2007 Dipl.-Inf. Danilo Schmidt starts his business career within the project OpEN.SC. Within this project he is especially responsible for the part of the specification and implementation of the "Intelligent Catalogue", a tool which allows a fast access to the meta-data repository and what is based on principles of Case-Based Reasoning coming from Artificial Intelligence.

Thomas Schrader is a board certified pathologist and a computer scientist in the Department of Pathology at the Charite Medical School in Berlin, Germany. He is the program director of the Open European Nephrology Science Center and of various other projects in Telemedicine including in Virtual Microcopy.

Michael Beil received a PhD in Biomathematics at the University Paris 7, France. He teaches computer engineering at the Lausitz University of Applied Sciences, Germany.

Thorsten Schaaf received his diploma degree in computer science from the Technical University of Berlin (TUB) and a Ph.D. degree from the medical faculty of Humboldt University and the Free University of Berlin (FUB). He is currently leader of a bilateral research project with the Karolinska Institutet in Stockholm, Sweden, dealing in business measurement tools for the improvement of quality in health care.

Hans-Dieter Burkhard is head of the Artificial Intelligence group in the Institute of Informatics at Humboldt University of Berlin. He has studied Mathematics in Jena and Berlin, and he has worked on Automata Theory, Petri Nets, Distributed Systems, VLSI Diagnosis and Knowledge Based Systems. Current interests include Cognitive Robotics, Distributed AI, Agent Oriented Techniques, Case Based Reasoning, Socionics, AI applications in Medicine.
Prof. Dr. Hans-Dieter Burkhard is Vice President of the RoboCup Federation and fellow of the ECCAI.