Design and Prototypical Development of a Web Based Decision Support System for Early Detection of Sepsis in Hematology

Andreas Wicht
Heidelberg University Hospital,
Institute of Medical Biometry
and Informatics
Im Neuenheimer Feld 305
Heidelberg, 69120 Germany
Andreas.Wicht@med.
uni-heidelberg.de

Gerrit Meixner
German Research Center for
Artificial Intelligence (DFKI)
Trippstadter Strasse 122
Kaiserslautern, 67663 Germany
Gerrit.Meixner@dfki.de

Ulrike Klein
Heidelberg University Hospital,
Department of Hematology
and Oncology
Im Neuenheimer Feld 410
Heidelberg, 69120 Germany
Ulrike.Klein2@med.
uni-heidelberg.de

ABSTRACT
Physicians do not always make optimal decisions. Computer based clinical support systems are intended to provide clinicians with decision aids, but their practical impact remains low. We introduce a software architecture which might overcome key barriers and present the prototypical implementation of a web based knowledge module for early detection a life-threatening medical condition, sepsis.

Keywords
Clinical decision support, Knowledge-based systems in medicine, hematology, sepsis, fever, web-based application, knowledge maintenance, ESGOAB.

INTRODUCTION

The information overload physicians are confronted every day with makes it impossible for them to keep up with all the information and knowledge that would be potentially useful in making optimal clinical judgments. Empirical studies have shown that physicians do not always make optimal decisions [17] [6]. Clinical decisions are often made under time pressure, without having all information and knowledge needed in the right place at the right time. Computer-assisted *clinical decision support systems* (CDSS) are intended to provide *"clinicians, patients or individuals with knowledge and person-specific or population information, intelligently filtered or presented at appropriate times, to foster better health processes, better individual patient care, and better population health. CDS interventions include alerts, reminders, and order sets [...]"* [11].

Although research has been done in the field of CDSS for decades, the practical impact remains low for several reasons, e.g. [3]:

- Systems failed to cover an entire medical domain
- Poor practicability and integration into the clinical workflow
- Poor availability of digital patient data
- Poor acceptance

Within the ESGOAB[1] project, which will be described in more detail in the next section, we are trying to overcome the weaknesses mentioned above. In this paper we will briefly describe the ESGOAB software architecture, which provides an electronic health record (EHR) and is designed to provide the base to interact with knowledge modules. We focused on two specific clinical challenges: supporting the physicians' order entry process (CPOE) and supporting early detection of sepsis (a serious and life-threatening medical condition) on patients with hematological underlying diseases. In this paper we will describe the second challenge. We introduce our conceptual design of the sepsis knowledge module and present the currently implemented web-based prototype.

Project Background

The ESGOAB project is a 2-year public funded joined research project between two scientific partners (Heidelberg University Hospital, DFKI) and two industrial partners (COPRA System GmbH, Dosing GmbH).

A survey at the Department of Hematology and Oncology of the Heidelberg University Hospital has shown a range of problems concerning various applications of Information and Communications technologies (ICT) within the hospital (response rate = 70.5 %, 36 of 51 of the medical personnel)

[1] ESGOAB = „**E**ntwicklung einer **S**oftwareumgebung zur **G**enerierung von **o**rganisationsspezifischen **A**nwendungen zum **B**ehandlungs-prozessmanagement"; *english*: Development of a Software Environment for Generation of Organization-Specific Applications for Treatment Process Management

[20]. The study revealed problems concerning e.g., time consuming searches, redundant data entries, use of various software applications to perform the various tasks, different user interfaces, and only marginal decision support. However we found a promising openness towards CDSS [20]. In general the staff was open-minded towards new Information Technology (IT) systems (88 % indicated to be "rather open-minded" or "open-minded"), concerning CDSS, the potential benefit was assessed by the majority (72 %) as "rather high" or "high", as well as prospects of success (53 %). At least 47 % rated *reliability* as well as *acceptance* as "rather high" or "high".

The ESGOAB project aims at encapsulating various data sources like e.g., hospital information system (HIS), laboratory data or drug information systems (see Figure 1) into one integrated software system which provides one consistent user interface. The second main aim is the development and integration of knowledge bases into the ESGOAB system.

The ESGOAB system is based on a 3-layer software architecture concept. A data collector (layer 1) is responsible to capture and gather data from various existing sub systems which are illustrated below the three layers. Adapters for each of the sub systems are transforming data into defined structures. The knowledge carrier (layer 2) analyzes incoming data from the data collector layer or handles requests triggered by user interactions. The knowledge carrier consists of various knowledge bases (e.g., about drug information) which are connected following a modular concept. If there are e.g., new blood values these will be evaluated by the knowledge carrier and according hints or warnings will be immediately displayed by the visualizer (layer 3) following a defined alerting concept, taking into account the importance, severity etc. of the alert [4].

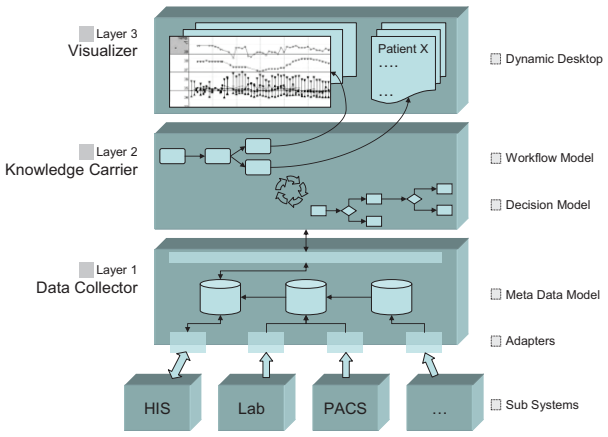

Figure 1: 3-layer architecture of the ESGOAB system

At the end of the ESGOAB project we expect an increase of efficiency of tasks and processes at the Department of Hematology and Oncology at the Heidelberg University Clinic (due to time-savings), a reduction of performance stress, and optimization of clinical decisions, which altogether may improve quality of care.

Related work

Research in the field of medical knowledge-based systems has been done since the early 1960s. Many different research approaches have been explored, but yet the degree of impact of clinical decision support systems still remains low [5]. Only few knowledge-based systems are widely used day-to-day, such as automated electrocardiogram (ECG) interpretation [18]. However, systems utilizing a broad variety of individual patient data had to fail due to poor availability of digital data. Providing an EHR eliminates this obstacle [3].

Research in the field of monitoring and analyzing vital signs in Intensive Care Units (ICU) for early warning of patient deterioration or sepsis were done by [15] and [12]. However, including microbiological findings as well as the accurate handling of the specifics of the hematological patients remained unconsidered.

Clinical Background

The treatment of patients with hematological diseases has advanced enormously in the past years. Nevertheless such infections pose a serious life-threatening risk for these immunocompromised patients. Beside various other clinical and laboratory parameters, fever is an essential factor, which indicates a manifest or beginning infection. Therefore a refined assessment of the body temperature is needed. The responsible physician has to distinguish between innocent fever as immunologic reaction, fever of unknown origin, fever caused by bacteremia or the onset of a severe sepsis. Hereby assists the combination of lab-values, microbiological findings and vital signs. The emphasis and valuation of the combination of single-values and the experience of the doctor partly determine the treatment course and the outcome of the patient. Clinical studies demonstrated that the survival probability of patients with sepsis depends most essentially on the period of time between diagnosis and start of effective antibiotic treatment [7]. Sepsis is not only a problem of hematological patients. It's rather a challenge for the population. Severe sepsis is considered to be the most common cause of death in non-coronary critical care units. Approximately 150.000 persons die annually in Europe and more than 200.000 in the United States [1]. The problematic nature of a timely recognition is not that data is missing, but it is detached from one another, generated at various places and different times. The responsible physician has to link the separated information for the plurality of patients. Ward rounds, printouts with lab values, calls from the microbiologist and signs from monitors serve as instruments for this connection. This work has to be done by the medical personnel even in the future, but the model we present in this article offers the convincing advantage of automated joining of relevant data and usable presentation, resulting in efficient and faster decisions.

APPROACH

Conceptual Requirements

When deploying and operating knowledge-based systems, a weak point is often poor practicability, in particular in terms of maintenance. Either the knowledge model is implemented statically, there is no way experts of the certain domain (in the present case hematologists) can modify the model on their own. This leads to the fact that each adaptation has to be done by a software engineer. Or the user interfaces do not provide intuitive means to modify the knowledge base; thus users have to be instructed and the system becomes error-prone. Between designing and using knowledge-based systems, a long-lasting cyclic process of modeling, testing, adapting, and retesting of the core engine has to be passed through. While operating knowledge-based systems the focus shifts towards maintenance issues. Maintaining the knowledge within the system is critical to successful delivery of decision support [3]. In this context practicability plays an important role. Being aware of this issue, easy knowledge maintenance was an important goal. The clinical expert should be able to modify the underlying knowledge model without extensive training. It has to be simple and intuitive to use. Furthermore it should be possible to test the constructed or adapted model right away. Beyond that, the knowledge model should be generic, so it might be useful for other diagnostic problems.

Knowledge Engineering Process

Knowledge Engineering is the systematic approach for the development of knowledge based systems. The process may be divided into two main phases (Figure 2): *knowledge acquisition* and *knowledge operationalization* [14]. It should be noted, that typically the process of acquisition and operationalization is not a linear process but rather a cyclic process characterized by continuous, iterative refinement.

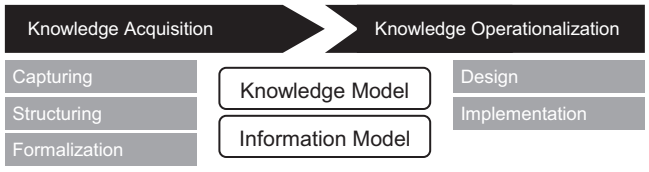

Figure 2: The Knowledge Engineering Process

Knowledge acquisition is the process of capturing, structuring and formalizing knowledge. The result of the acquisition phase is a *knowledge model* and an *information model* which both serve as a base for the system design of an implementation [5]. Sources of knowledge are typically domain experts, medical literature and patient data repositories. Our approach was based on expert opinion and literature research.

Information Model

A specification of the kinds of information that were required was created, including the data format and the taxonomy. The resulting information model – which will be implemented as an object-oriented data model – provides us the flexibility to use the same implementation (objects) in two kinds of settings:

1. Interactive data retrieval with the user (module execution)
2. Running in background through a web service retrieving data from the EHR

Knowledge Model

Sepsis accompanies with several symptoms, such as fever, increased heart rate, low blood pressure etc. Further important parameters are signs of infections such as specific blood values and microbiological findings. The sequence of appearance and the severity of these manifestations differ from patient to patient. We have to deal with fuzzy and uncertain information. However, some signs are more important than others and certain value ranges are supporting sepsis more than other diagnoses. So the idea was to design a decision model, which balances between a set of differential diagnoses and specifies the one which can be explained best by the observed findings. The *set covering model* is a potentially useful approach, which was introduced by [13], as well as the more abstract view on multiple diagnose problems by [9]. Our approach is based on the set covering model, extended by the possibility to define parameters which contradict certain diagnoses since we experienced a further need for accuracy.

Two sub modules based on a rule engine were required to handle two specific problems:

1. Interpretation of microbiological findings: Presence of an infection or suspected contamination?
2. Interpretation of white blood cell count (leukocytes): May we take this parameter into account?

Implementation

Initially, the basic idea was demonstrated using a prototype realized in Microsoft Excel (Figure 3). Taking advantage of quick implementation possibilities, this prototype helped us to refine our knowledge model.

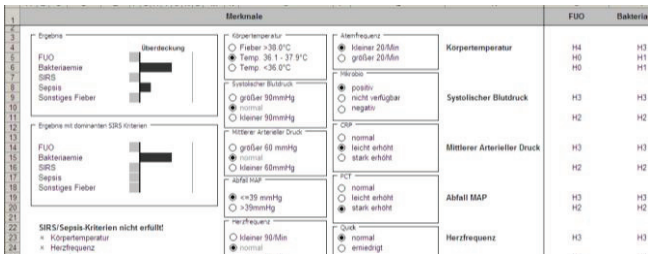

Figure 3: Prototype developed with Microsoft Excel (result presentation in the upper left corner, value selection in the mid-area)

The decision to proceed with the development of a *web-based* application was made due to the following main reasons:

- Easy Accessibility from all clients via the clinic's intranet; no need to install software on clients
- Quick deployment of new versions

We used a *XAMPP* [2] installation on Windows including an *Apache* 2.2.14 web server and a *MySQL* database system. We used the scripting language *PHP* and the Ajax toolkit *xajax* [19].

The application is based on the Model-View-Controller (MVC) design pattern; the PHP code architecture follows an object-oriented approach. The application is implemented using the open source relational database management system MySQL, with use of the *InnoDB* storage engine. An initial database model was designed using the database-modeling tool *MySQL Workbench* [10]. The model was refined iteratively during the implementation of the application.

WEB-BASED PROTOTYPE

Application Structure

The application is composed of three modules: *Sepmod*, *Leukomod* and *Mibimod* (Figure 4). Sepmod represents the generic core model, which implements the weight model as specified before and interacts with the sub modules Leukomod and Mibimod. Leukomod is based on a rule engine, which is responsible for the white blood cells' assessment. Mibimod is also based on a rule engine and performs the assessment of the microbiologic findings.

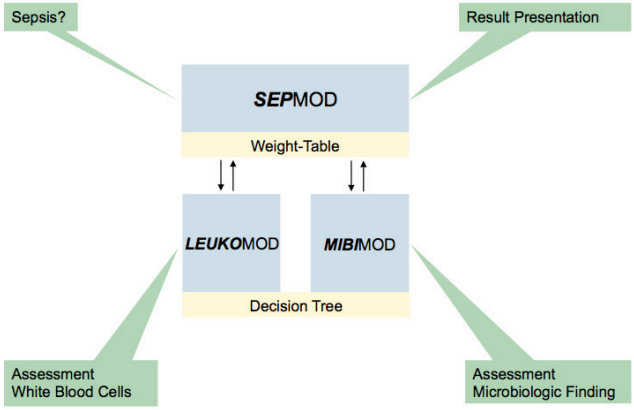

Figure 4: Structure of the application

The web-front-end provides the following two main sections:

- *Maintenance*: Provides editors for each module for construction and modification of the knowledge base.
- *Module-Execution*: Interactive tool, which requests for data or selection of options by the user and presents the results.

The interactive tool is basically designed for testing purposes. The impact of adaptions of the knowledge base can be explored right away.

Knowledge Maintenance Section

Sepmod

In this section the user can create a new and edit existing knowledge models. Created models can be loaded, saved, and deleted. The core element is the weight table (Figure 5).

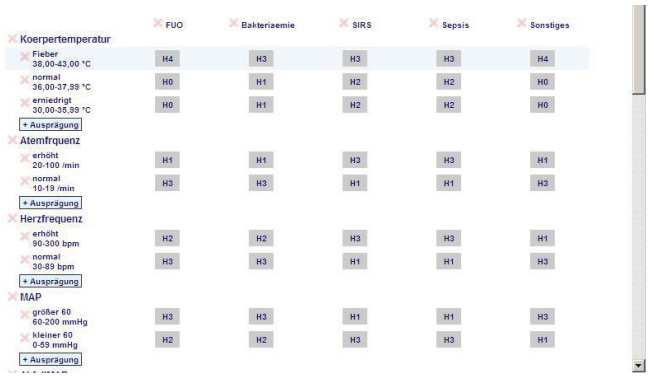

Figure 5: Screenshot of a weight table (columns: diagnoses, rows: parameters and values)

The process of creating a new model facilitates the following steps:

1. Add *diagnoses*
2. Add *parameters*
3. Add *value* ranges for each parameter
4. Create *weight relations* between value and diagnosis
5. Define *Equivalencesets*
6. Define *Minimalsets*

These steps are described in more detail below.

Step 1: Firstly we need to add diagnoses, by providing its *name* and optionally a *description*. For each diagnose added, the table will be expanded by one column.

Example definition:

Name: `SIRS`
Range: `Systemic Inflammatory Response Syndrome`

Step 2: In the second step, symptoms or "parameter" can be defined, specifying its *name*, *type, unit*, *value range*, *validity*, *importance*. The *type* tells us, which kind of data we deal with, e.g. integer, float or special medical classifications like the *anatomic therapeutic classification* (ATC) code. The *value range* defines the valid value range for the parameter and is used to perform plausibility checks. Regarding the process of diagnosis, an important and relevant issue is always the time context. *How long can I rely on a measured value?* The answer depends on each parameter. In the present model we therefore define for each parameter a time frame (*validity*), entering a numerical value for the absolute time (minutes, hours or days) within this parameter remains valid or in other words we can assume that the measured value may be used for the diagnostic assessment. If a parameter exceeds the defined time frame, it will be ignored and treated, as it would be not available. Alternatively we can define a relative time frame such like "Valid until next measurement". Some parameters may have a more significant importance than other. Thus for each parameter the importance can be defined, choosing from given weights "very important", "fairly important", "important".

Example definition:

Name:	*bodytemperature*
Type:	*float*
Unit:	*°C*
Range:	*30.0 to 43.0*
Validity:	*6 hours*
Importance:	*very important*

Step 3: In the next step, for each parameter, several *values* have to be defined. The values are specified by its name and its value range.

Example definition:

Name:	*fever*
Range:	*38.0 to 43.0*

Step 4: Adding diagnoses, parameter and its values results in a *count(diagnoses) x count(values)* matrix. For each pair of a value and a diagnosis we can now define a weight relation between a specific value and a diagnosis by clicking on the corresponding button. A weight relation is the symbol for the strength a specific value supports a diagnosis. The currently implemented model supports five different weights, ranging from H0 (value does not support the diagnosis or even contradicts) to H4 (value strongly supports the diagnosis or is even essential) depending on the currently used weight model which can be defined in a separate section. For each weight symbol the value can be selected through a slide control.

Step 5: In the section called *equivalencesets* we define sets of previously defined parameters, which are clinically equivalent (Figure 6). In other words all of these parameters support a specific context diagnosis, such as *low blood pressure*.

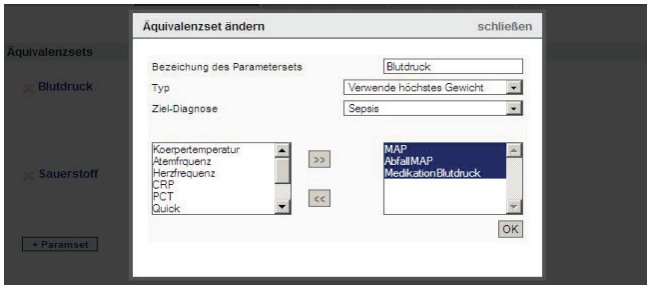

Figure 6: Screenshot of the equivalenceset dialog

In this case we would define that low blood pressure exists if at least one of these blood pressure parameters takes a specific value (and a corresponding high weight) or in this special case we also assume low blood pressure if vasopressors (substances which result in an increase in blood pressure) are given what we can specify through a list of ATC codes.

Step 6: A decision model might be sophisticated but its accuracy highly depends on the available parameters. So it only makes sense to perform a diagnostic assessment if at least a certain set of parameters is available. In the section minimalset the clinical experts can define, which parameters they consider as being essential for performing a profound assessment. It is possible to define more then one set.

Leukomod

Leukomod is based on a rule engine. The rules were defined by clinical experts. In the current version they are implemented statically and can be activated or deactivated through the web-front-end in the section leukomod/rules. Further we can adapt various parameters (like thresholds) of the rules. For the assessment, if the white blood cells may be included, the underlying disease is of high importance. A dynamic list of diagnosis codes (*ICD*, *International Statistical Classification of Diseases and Related Health Problems*) can be maintained in the section diagnoses.

Mibimod

Mibimod is based on a rule engine. A dynamic list of germs which support the suspicion of a contamination can be maintained in this section. The rules of this module are currently implemented statically.

Module Execution Section

For testing purposes the web-front-end provides a section to run the modules. The current version has one section to run each of the modules separately and one, which encapsulates all three of them. These execution modules are implemented as interactive tools, requesting each parameter step-by-step.

For each input field a check for plausibility is performed while entering data, considering two main issues

- Valid characters (depending on the data type)
- Valid value range (as defined in the model).

To keep track of entered values, we show breadcrumbs horizontally across the top of the input form (Figure 7).

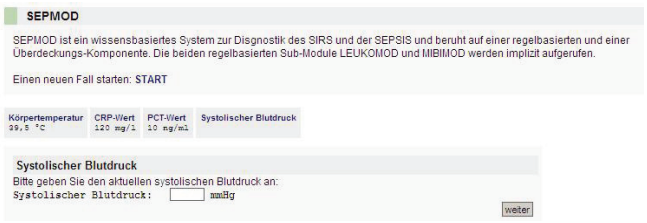

Figure 7: Entering values using the interactive mode

The user can easily go back to previously entered values and change them if necessary. Each breadcrumb item shows the name of the parameters as well as the entered value and unit. Once all values are entered, the system performs the assessment and presents the results, giving explanations and a visualization of the results (Figure 8). Presenting results in an appropriate way means finding a balance between (1) simple, clear and aggregated representations, supporting e.g. quick task completion and (2) comprehensive and detailed representations, which are needed for making comprehensible decisions.

Figure 8: Result presentation (diagram and textual explanation)

DISCUSSION & FUTURE WORK

We identified two important and promising factors that will help overcome key barriers limiting more widespread use of CDSS: Firstly we did not experience the "physician resistance" using decision support systems; they are rather convinced of their usefulness as described in the introduction. Secondly we can take advantage of the software infrastructure we presented which has the potential to improve clinical processes and to integrate and interact with sharable knowledge modules conceived to support the physicians in their decisions. We also introduced a web based prototype of a knowledge module for early detection of sepsis. We believe to contribute important elements to lift computer-aided decision support into widespread practice.

However, our approach still needs to be refined and evaluated. It has to be shown that our knowledge modules are able to provide accurate and traceable support, characterized by high sensitivity. A test concept as well as test scenarios will be defined. Further alerting concepts will be tested in practice since empirical studies have shown that too many generated alerts could lead to "alert fatigue", whereas "non-interruptive" alerts only have low impact and are not effective [8].

ACKNOWLEDGMENTS

This work was funded by the German Federal Ministry of Education and Research under grant number 01 IS 09027 C. The responsibility for the content of this publication lies with the authors.

REFERENCES

1. Angus, D.C., Linde-Zwirble, W.T., Lidicker, J., Clermont, G., Carcillo, J., and Pinsky, M.R. Epidemiology of severe sepsis in the United States: analysis of incidence, outcome, and associated costs of care. *Critical care medicine 29*, 7 (2001), 1303-1310.
2. Apache Friends. XAMPP distribution. Available at http://www.apachefriends.org/xampp.html.
3. Bates, D.W., Kuperman, G.J., Wang, S., et al. Ten commandments for effective clinical decision support: making the practice of evidence-based medicine a reality. *Journal of the American Medical Informatics Association 10*, 6 (2003), 523–530.
4. Durieux, P. Electronic Medical Alerts – So Simple, So Complex. *New England Journal of Medicine 352*, 10 (2005), 1034-1036.
5. Greenes, R.A. (Ed.). *Clinical decision support*. Elsevier Academic Press, Amsterdam; Boston, 2007.
6. Kohn, L.T. (Ed.). *To err is human*. National Academy Press, Washington, DC, 2000.
7. Kumar, A., Roberts, D., Wood, K.E., et al. Duration of hypotension before initiation of effective antimicrobial therapy is the critical determinant of survival in human septic shock. *Critical care medicine 34*, 6 (2006), 1589-1596.
8. Lo, H.G., Matheny, M.E., Seger, D.L., Bates, D.W., and Gandhi, T.K. Impact of non-interruptive medication laboratory monitoring alerts in ambulatory care. *Journal of the American Medical Informatics Association 16*, 1 (2009), 66–71.
9. Miller, R.A. A Heuristic Approach to the Multiple Diagnoses Problem. *Proceedings of the 6th Conference on Artificial Intelligence in Medicine in Europe*, Springer-Verlag (1997), 187–198.
10. MySQL Workbench visual tool. Available at http://www.mysql.com/products/workbench/.
11. Osheroff, J.A., Teich, J.M., Middleton, B., Steen, E.B., Wright, A., and Detmer, D.E. A roadmap for national action on clinical decision support. *Journal of the American medical informatics association 14*, 2 (2007), 141-145.
12. Pilz, U., and Engelmann, L. Früherkennung von septischen Krankheitsverläufen durch Einsatz eines Expertensystems. *Intensivmedizin und Notfallmedizin 43*, 7 (2006), 575-579.
13. Reggia, J.A., Nau, D.S., and Wang, P.Y. Diagnostic Expert Systems Based on a Set Covering Model. *International Journal of Man-Machine Studies*, (1983), 437-460.
14. Spreckelsen, C., and Spitzer, K. *Wissensbasen und Expertensysteme in der Medizin*. Vieweg + Teubner, Wiesbaden, 2008.
15. Tarassenko, L., Hann, A., and Young, D. Integrated monitoring and analysis for early warning of patient deterioration. *British Journal of Anaesthesia 97*, 1 (2006), 64-68.
17. Warner, H.R., Sorenson, D.K., and Bouhaddou, O. *Knowledge engineering in health informatics*. Springer, New York; Berlin; Heidelberg, 1997.
18. Willems, J.L., Abreu-Lima, C., Arnaud, P., et al. Evaluation of ECG interpretation results obtained by computer and cardiologists. *Methods of Information in Medicine 29*, 4 (1990), 308-316.
19. xajax Project. xajax framework. Available at http://www.xajax-project.org/.
20. Zuehlke, D., Meixner, G., and Klein, U. Smart Medical Software Systems For Dummies? The Case For A User-Centered Systems Design. *Proceedings of the 3rd International Conference on Health Informatics 2010 (Valencia ESP, 2010)*, 350-354.