Mobile Real-time Analysis of Patient Data for Advanced Decision Support in Personalized Medicine

Matthieu-P. Schapranow, Hasso Plattner
Hasso Plattner Institute
Enterprise Platform and Integration Concepts
August–Bebel–Str. 88
14482 Potsdam, Germany
{schapranow|plattner}@hpi.uni-potsdam.de

Cafer Tosun
SAP AG
Innovation Center Potsdam
August–Bebel–Str. 88
14482 Potsdam, Germany
cafer.tosun@sap.com

Christian Regenbrecht
Charité – Universitätsmedizin Berlin
Institute of Pathology
Charitéplatz 1
10117 Berlin, Germany
christian.regenbrecht@charite.de

***Abstract*—Personalized medicine aims to treat patients specifically with respect to their individual dispositions. For that, researchers and physicians require a holistic view on all relevant patient specifics when making treatment decisions. We present our findings of applying in-memory database technology to enable real-time analysis of individual patient and cohort data. In this contribution, we describes the mobile application "Oncolyzer" that provides a holistic view on individual patient data and enables flexible analysis of cohort data on mobile devices. It opens flexible access to relevant patient data on the hospital campus when time-critical treatment decisions need to be made.**

Keywords*-*Clinical Decision Support; Personalized Medicine; In-Memory Database Technology; Real-Time Analysis; Business Processes

I. Introduction

The amount of data acquired during patient's medical treatments is immense. For example, data of tissue analyses, medical imaging, and haemograms, etc. add up to terabytes of medical data depending on the specific diagnostic approach [1]. In future, we expect the level of detail of therapy data to increase continuously.

Personalized medicine aims at treating patients specifically based on their individual dispositions, e.g. genetic or environmental factors [2]. Thus, we consider the detailed acquisition of medical data as the foundation for personalized therapy decisions. The more fine-grained data are available, the more specific are the gained insights, but the complexity of data processing rises, too.

Modern Hospital Information Systems (HIS) consist of various data sources. Combining data from distributed sources is one of the challenges of Computerized Clinical Decision Support (CCDS) systems. Business Warehouses (BW) or Business Intelligence (BI) systems incorporate specifically prepared business reports in a central system. However, these reports do not allow a free combination of all available attributes.

In the given work, we present our findings of applying in-memory database (IMDB) technology for analysis of clinical data during treatment of patients suffering from

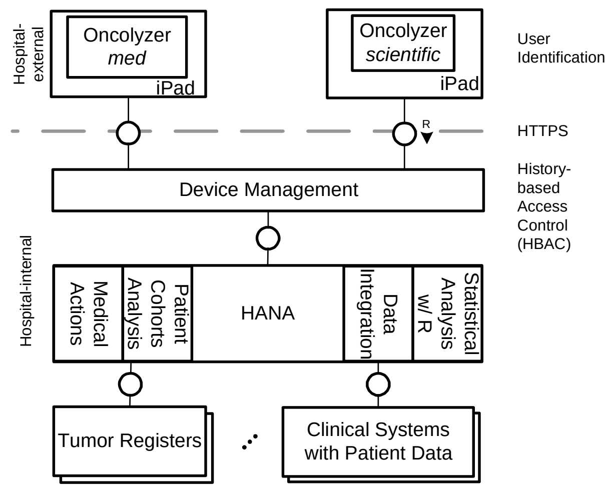

Figure 1. Oncolyzer system architecture: Data from clinical systems is combined and analyzed by the IMDB. The results are accessible by the iPad application.

cancer diseases. Based on the feedback of physicians and researchers, we designed a software system and a specific mobile application as depicted in Figure 1.

The rest of the paper is structured as follows: In Section II our work is set in context of related work. We define requirements for clinical IT systems used in context of personalized therapy in Section III and outline the applicability of IMDB technology in Section IV. Our research prototype Oncolyzer, which is specifically designed for individualized therapy of cancer patients, is presented in Section V. In Section VI, we define innovative business processes that become possible due to the use of IMDB technology for real-time analysis of patient data. We discuss our research results in Section VII and conclude our work with an outlook in Section VIII.

II. Related Work

Requirements engineering for software systems is well defined, e.g. by the International Standards Organization (ISO) [3]. However, requirements for clinical software are only rarely differentiated by existing standards. Coble et al. [4] outline their experiences during designing, prototyping, and testing of clinical software for workstations in the 1990s. We followed the design thinking approach for software engineering of Plattner et al. [5], which builds on viability, feasibility, and desirability of the desired software. In consensus with Coble et al., this approach requires the early and regular involvement of future users of the developed software system, i.e. physicians and medical researchers in our case.

Coble et al. focused on the development for fixed workstations. Our user interviews showed that physicians and researchers do not work at dedicated offices only. In contrast, they have frequently changing working locations, e.g. emergency room, surgery room, intensive care station, etc. As a result, we emphasized the mobility aspect during requirements engineering.

Wright et al. [6] propose the use of Web 2.0 techniques for building CCDS systems, e.g. Wikis for knowledge management or private online forums. They elaborate on the value of online CCDS, but they also discuss various concern for a clinical wide use of Web 2.0 CCDS, such as costs for knowledge management and liability for correct data. We agree that online CCDS can improve clinical knowledge management. However, we also see a need for having the accumulated knowledge at hand at any time. Thus, we focus on converging clinical knowledge from private online sources and the required mobility aspects of physicians and researchers.

Based on our conducted feedback loops, we focused on the selected application perspectives summarized in Table I. A detailed description of design decisions and these perspectives is elaborated in Section V.

III. Requirements

In the following, we define specific requirements for designing clinical software systems. They reflect a selected subset of requirements from the software engineering catalog as defined for product quality in ISO/IEC 9126-1 and specifically revised in context of our work [3].

- **Ease of Use:** Clinical software artifacts must be usable by untrained users, i.e. its user interface (UI) should combine ease of use and functionality.
- **Response Time**: The response time of clinical applications must not exceed an empirical threshold of approx. two second [7]. Our user interviews showed that otherwise the latency outperforms any benefits resulting in the application not being used.
- **Reliability:** Clinical software must be available without unplanned interruptions or malfunctions due to its life-critical purpose.
- **Productivity:** Users of a clinical software solution should be more efficient than performing manual processing steps or using alternative tools.

Table I. Application perspectives and corresponding user groups Physicians (P) and Researchers (R).

Perspective	Group	Sect.
Holistic Patient View	P, R	V-A
Search in Structured and Unstructured Data	P	V-B
Real-time Analysis of Patient Cohorts	R	V-C

- **Scalability:** The system behavior of the designed software must not be affected by the number of concurrent users. Extending existing hardware resources, e.g. number of database servers, should result in a linearly increasing capacity.
- **Data Security:** Clinical data are sensitive and must be accessible by authorized personnel only. Intended or unintended exposure of these data must be addressed during the design of clinical software.

IV. In-memory Technology Building Blocks

We refer to IMDB technology as a toolbox of artifacts to enable processing of enterprise data in real-time in the main memory of server systems. IMDB technology enables CCDS in an interactive way without keeping redundant or pre-aggregated data, which is commonly used in BW or BI systems to improve response times for certain long-running reports [8]. In the following, we outline selected building blocks of the IMDB technology enabling real-time analysis of clinical data.

A. Combined Column and Row Store

Historically, separate database systems for processing of analytical and transactional data evolved. The former store and process data in a row-oriented format, i.e. attributes of one record are stored side by side, while analytical database systems are optimized to scan selected attributes of huge data sets rapidly, e.g. by maintaining pre-aggregated totals. Combining column and row stores improves data access for analytical queries while keeping transactional response times fast.

B. Insert-only

Insert-only defines how newly inserted data are managed. Traditional database systems support four operations for data manipulations, i.e. insert, select, delete, and update of data. The latter two are considered as destructive operations since original data are no longer available after its execution [9, Section 7.1]. In other words, it is neither possible to detect nor to reconstruct the complete history of values for a certain attribute after its execution since only the latest value is permanently stored. This violates legal regulations to permanently store clinical data. Insert-only tables store the complete history of value changes and the latest value for a certain attribute [8]. Insert-only enables tracing

 ISBN: 978-1-61208-252-3

Table II. Comparison of patient database table (top) and its compressed, horizontally three-partitioned pendant with corresponding dictionary (bottom).

Rec.	PatID	Loc.	ICD
1	091487	Colon	C18.4
2	357982	Larynx	C32.0
3	123498	Lip	C00.9
4	998711	Colon	C18.7
5	215678	Rectum	C20.9
6	647912	Rectum	C20.9
7	167898	Mama	C50.9

=

	Rec.	PatID	Loc.	ICD
I	3	123498	2	C00.9
	1	091487	4	C18.4
	4	998711	4	C18.7
II	5	215678	3	C20.9
	6	647912	3	C20.9
	2	357982	1	C32.0
III	7	167898	5	C50.9

+

Val.ID	Value
1	Larynx
2	Lip
3	Rectum
4	Colon
5	Mama

of decisions within the treatment process, e.g. to retrospectively perform analysis when certain treatments were initiated.

C. Lightweight Compression

Lightweight compression refers to a data storage representation, which consumes less space than its original pendant [8]. Due to the fixed data type per attribute, the data domain is within a given interval, e.g. integer values. The given data type defines the max. required data domain for a compressed format. A columnar storage layout supports lightweight compression techniques, such as run-length encoding, dictionary encoding, and differencing [10]. Table II depicts an excerpt of the patient table that is compressed using dictionary encoding.

D. Partitioning

We distinguish vertical and horizontal partitioning [11]. The former refers to the arrangement of individual database columns, which is achieved by splitting columns of one database table in two or more sets of columns. Each of the sets can be distributed on individual database servers [12]. The latter addresses long database tables and its division into smaller chunks of data [8]. Splitting data into equivalently long horizontal partitions supports parallel search operations and improves scalability. For example, a Kaplan Meier analysis requires a full scan of the patient table as depicted in Table II to identify patient records having the identical International Code of Disease (ICD). With a single partition, a single thread needs to access all patient records to check the ICD predicate. The example depicts the major code of the ICD as partition criteria. Thus, the three partitions can be processed in parallel.

E. Multi-core and Parallelization

Parallelization can be applied to various locations within the application stack of software systems. Let us consider the HIS as an enterprise system that needs to serve requests of different users from various departments at the same point in time. Processing multiple queries can be handled by multi-threaded applications, i.e. they do not stall when dealing with more than one query at a time. Operating systems threads are a software abstraction that need to be mapped to physically available hardware resources [13, Chapter 2]. A Central Processing Unit (CPU) core is comparable to a single worker on a construction area. If each query can be mapped to a single core, the system's response time is optimal. If the workload exceeds physical capacities of a single system, the work load can be distributed to multiple servers or blades to achieve optimal processing behavior. From the database point of view, data partitioning supports parallelization since multiple CPU cores on multiple servers can process individual partitions in parallel [14, Chapter 6].

F. Active and Passive Data Store

We distinguish active and passive data stores. The former are accessed frequently and updated regularly, e.g. patient records of patients in the emergency room or on the intensive care unit. Passive data are neither updated nor accessed regularly, i.e. they are mainly used for analytical and statistical purposes or in exceptional circumstances, where specific investigations require them. For example, data stored about patients that have left the hospital and have been invoiced can be considered as passive data. The separation of active and passive data is a special implementation of data partitioning that establishes a memory hierarchy, e.g. fast main memory, solid state disks, hard disks, tapes, etc. [8].

V. Research Prototype

In the following, we share details about our research prototype and its application perspectives as summarized in Table I. The system architecture and its integration within the HIS is modeled in Figure 1. The key component of the architecture is the IMDB HANA that enables real-time statistical analysis of patient cohort data, medical actions, and data from further clinical systems. The mobile devices are connected via a device management layer to the database management system, which combines data from clinical data sources, such as tumor registers or HIS.

 ISBN: 978-1-61208-252-3

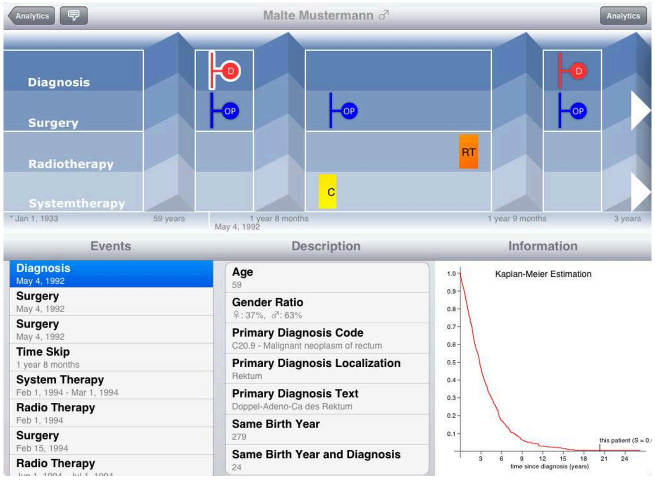

Figure 2. The holistic patient view consists top-down of name and gender details, interactive treatment history combining all relevant treatment information on a single screen as well as a tabular view. The latter contains from left to right details for treatment events, analytical results, and a graphical evaluation of patients with the same primary diagnosis using the Kaplan Meier analysis.

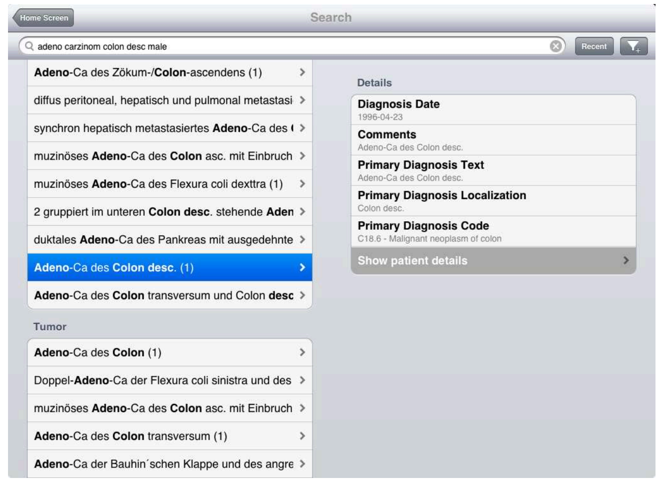

Figure 3. Result set of the search query for "adeno carzinom colon desc male". The search term "carzinoma" is identified by the fuzzy search as equivalent for "carcinoma" and is also substituted by its abbreviation "ca". The result set is also filtered using the value "male" obtained from a search in the structured attribute "gender".

A. Holistic Patient View

The holistic patient view as depicted in Figure 2 combines patient specific data from various clinical databases. For example, the complete treatment history consisting of diagnoses, surgeries as well as radio and system therapies are combined and visualized as a graphical timeline. The IMDB technology performs real-time analysis to identify similarities in the data of the selected patient and data of hundreds or thousands or even hundreds of thousands similar patients. For example, results for patients with the same year of birth, with the same primary diagnosis, or gender ratio are transparently calculated for any selected patient. This environmentally derived information can significantly help to identify sources of diseases, e.g. to identify the connection between TP53 gene mutations and urinary bladder cancer after the nuclear plant catastrophe in Chernobyl in 1986 [15].

The incorporated IMDB is hosted by the central IT department running on multiple high-end servers. In contrast to data warehouse systems, which store pre-aggregated totals to improve long-running queries, IMDB technology performs all calculations on the fly. In other words, once a new patient record has been added to the database, it can be accessed immediately via the Oncolyzer. The new record directly influences calculations and analysis results, e.g. Kaplan Meier analysis.

B. Search in Structured and Unstructured Data

We are convinced that diagnostic reports contain valuable information, although they consist of less standardized free-text documents. As a result, it is hard to access relevant details, e.g. to determine `subscribed drug = Vincristin`. The Oncolyzer is equipped with a combined search for structured and unstructured data, which is supported by the underlying IMDB technology. The application perspective consists of a search box following the Google-like UI. Patient documents from different systems, e.g. diagnosis reports, preliminary diseases, death causes, etc., are searched for the terms entered as depicted in Figure 3.

In addition to an exact match search, fuzzy search also detects similar search terms, e.g. in case of typos [16]. Our research showed that fuzzy search is valuable in clinical contexts as medical terms can be documented in either Latin, English or German or their abbreviations. Fuzzy search accepts a specific degree of fuzziness in the search term, i.e. the result set contains more relevant search results than an exact search for terms.

We further integrated a synonym search, which returns documents that contain synonyms instead of the searched terms. For instance, the abbreviation "ca." is synonymously used for "circa", but it refers to "carcinoma" in clinical documents. A clinical classifications was added to the synonym table, e.g. the ICD and the corresponding tumor location [17]. Thus, a search for the tumor location "rectum" also returns results containing the ICD code "C20".

 ISBN: 978-1-61208-252-3

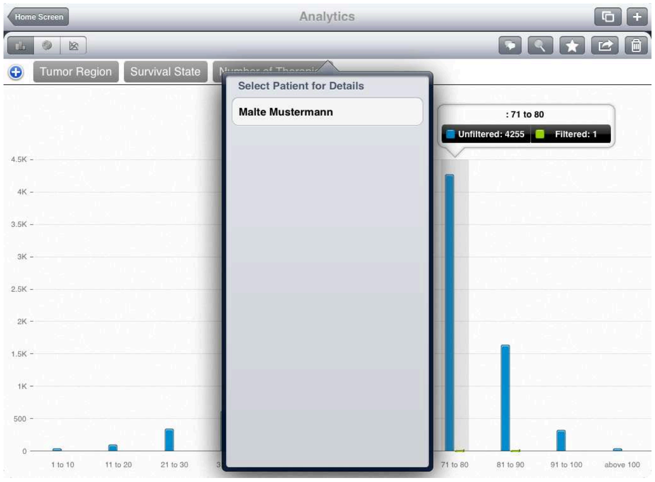

Figure 4. Real-time analysis of patient cohorts: Filters for tumor region, survival state, and number of therapies were applied and a patient in the age group "70 to 80" with the alias "Malte Mustermann" was identified.

Furthermore, stop words for structured search are derived from all values stored as structured attributes. A subsequent search for each stop word is performed in the database table containing the structured attribute. For example, a search for "male adeno ca" is expanded to a combined search for patients having the value "m(ale)" in structured attribute "gender" and associated documents with indications for an adeno-carcinoma.

The combination of search in structured and unstructured data leads to more accurate result set than traditional search tools, e.g. exact match search.

C. Real-time Analysis of Patient Cohorts

Researchers are able to perform individual real-time analysis on patient data on their mobile devices. The analytical view of our Oncolyzer provides the graphical UI for visualization of results of complex analytical queries that are executed by the IMDB system. Thus, researchers are able to identify relevant correlations with external factors, e.g. to prevent recurrences or to form patient cohorts. The quantity of patients that fulfill specific preconditions can be determined by using individual criteria for filtering, such as the patient's age, the gender, kind of tumor, or its localization as depicted in Figure 4. The results are interactively visualized using individual chart types, such as bar, pie or line charts. Clicking on a fragment of the chart shows the list of concrete patients. All patients are linked to their holistic patient view, which connects analysis of cohorts and specific patients.

VI. Innovative Business Processes

In the following, we outline business processes that build on our research prototype. They describe a completely new way of exploring treatment-relevant information. We focus on the involved business entities and their actions. Table III classifies processes, involved entities, and required application perspectives.

Table III. Classification of innovative business processes and affected user groups (P = Physicians, R = Researchers, B = Business managers), category (N = New, E = Extended), application perspective (S = Search view, H = Holistic patient view, A = Analytical view), and corresponding section within this document.

Business Process	Group	Cat.	App.	Section
Evidence-based Treatment	P, R	E	H	VI-A
Building Research Hypotheses	R	N	A, H	VI-B
Pharmaceutical Feedback Loop	R	N	A	VI-C
Federal Bureau of Statistics	P, R	N	A	VI-D
Health Insurance Companies	B	N	A	VI-E
Tumor Board	P, R	E	A, H	VI-F

A. Evidence-based Treatment

Evidence-based treatment decisions are either directly taken by physicians or supported by researchers, such as biologists. It requires the retrieval of all available patient data and the summary of treatment relevant details in the UI [2]. We designed the holistic patient view as described in Section V-A to combine clinical data. In addition to patient specific data, it also contains results of specific analysis results of data of patients with similar indications or primary diagnoses.

B. Building Research Hypotheses

Clinical researchers require flexible ways to test hypotheses and correlations between certain aspects of patient data. BW or BI systems enable the analysis of patient cohorts in a fast way, but involve complex administrative operations for data preparation before accessing results. After patient data is Extracted, Transformed, and Loaded (ETL) into the BI system [18], they are optimized for predefined queries, i.e. expected correlations need to be included in the reports. We introduce a completely new way of analyzing patient data while bypassing ETL processes. We keep all relevant patient data in the main memory in a columnar format without the need to load and prepare data, e.g. from disks, before starting their analysis. As a result, data is accessed at main memory latency time, i.e. 10 ns, instead of about 100 ms required for hard disk access [8]. The analytical view as described in Section V-C is designed for interactive exploration and analysis of patient data. Analyses are processed by the IMDB while the mobile application is connected as remote UI and input device. Researchers can access the holistic patient view for each patient of the result set to retrieve patient-specific details depending on their granted access.

C. Pharmaceutical Feedback Loop

New pharmaceutical products require to pass clinical trials before being approved for regular use [19]. Pharmaceutical manufacturers require patients with very specific indications to participate in clinical trials. Pharmaceutical researchers can analyze patient cohorts in a similar way like clinical researchers using the analytical view. Their access is restricted to anonymized cohort data only, i.e. they are prohibited to access personal details of the holistic view. If a certain number of patients with similar indication is required, but not present, they can use bookmarks to get notifications once relevant patients are present. Instead of performing these analyses multiple times, a push notification automatically indicates new relevant patients. If the patient should be recruited, the pharmaceutical researcher contacts the hospital and the patient is informed by her/his physician. Only if the patient agrees, the contact with the pharmaceutical researcher can be established.

D. Federal Bureau of Statistics

Statistical analyses of cancer diseases are regularly conducted to supervise influencing factors, such as compliance of treatments with clinical guidelines, rate of new or recurrent cases, and an analysis of regional accumulations [20]. The primary step is to obtain relevant data from nationwide hospitals. In Germany, there are clinical tumor registers that contain well documented data about all recent cancer cases. A widely used documentation system is the Giessener Tumor Documentation System (GTDS) that provides interactive tools for documentation, verification, and export of analysis data [21]. In future, we expect the following trends to come up:

- The level of detail and therefore the size of the documented data will increase and
- Consolidated nationwide tumor registers will offer new sources of information, e.g. for research [22].

Combining data from nation-wide tumor registers improves the quantity of available data for evidence-based treatment decisions. We implemented a Kaplan Meier analysis that visualizes fractions of patient cohorts and their survival time after first diagnosis [23, Chapter 9]. It is based on patient cohorts with the same ICD of a single hospital. The outlined business process describes a completely new way to combine data from decentralized tumor registers. In future, the basis for these analyses will include data from national and worldwide tumor registers. Thus, the quality of data used for personalized treatment is improved by combining national and international data sources.

Our contributions prove that IMDB technology can be used to bridge individual data formats without the need of data transformation. Therefore, we make use of database views, i.e. a defined transformation rule that is processed, when the underlying data item is accessed [8]. As a result, views are a transparent way to expose a homogenous data format while original data remains unchanged. On the one hand, views require a portion of designated processing time of the CPU. On the other hand, the use of views reduce the time for integration of new data, i.e. required transformations are eliminated. This is beneficial, if you access only a small portion of data and instant transformations are performed (e.g. 100 relevant patients) while traditional ETL processes need to transform the full data (e.g. of 100k patients).

Exposing analytical access to patient data supports employees of the Federal Bureau of Statistics to create national cancer reports. IMDB technology can reduce the integration effort for combining patient data from clinical tumor registers of different vendors. Furthermore, it enables interactive data exploration to trace any kind of statistical anomalies and to determine their natures. Nowadays, tracing anomalies in the reported data is a time-consuming task since experts of the statistics bureau, of the clinic and the physician who have treated a specific patient, need to be involved. As a result, the release cycles of cancer reports can be reduced, for comparison, the German cancer report 2007/2008 was released in 2012 [20].

E. Health Insurance Companies

Health insurance companies are interested in optimizing the overall time of hospital stays. However, the duration of a specific stay depends on various factors, such as age of the patient, stage of cancer, etc. Periodically, health insurance companies negotiate fixed rates for treatments of specific disease types with hospitals [24]. The rates are aligned with the average national costs for treatment of the specific disease. However, hospitals specialized for the treatment of certain disease types argue that they receive more complex cases compared to small, regional hospitals. As a result, hospitals are also interested in providing in-house analysis of certain disease types to negotiate more adequate rates. Latest analyses of treated patient cohorts suffering from a specific disease type forms the basis for the negotiating process. In addition, it enables a more transparent and holistic view on the treatment process, e.g. to match specific thresholds for the duration of a hospital stay.

F. Tumor Board

In the course of cancer treatments, each individual case is discussed by dedicated experts. They develop a specific treatment plan, discuss alternatives, and regularly evaluate the performance of the chosen therapy, which requires all relevant data of the discussed patient [25]. All decisions of the tumor board need to be documented and communicated, e.g. new treatment decisions to anybody involved in the treatment process.

The holistic patient view of our research prototype provides the tumor board with all relevant patient data. We also added notes for patients and analysis results. Notes can be used to ensure that all treatment relevant decisions of the board are immediately documented and all involved personnel are automatically informed when accessing the patient's data.

VII. Discussion

In the following, we discuss how our work respects the requirements for medical software defined in Section III.

We found that tools for physicians need to match strict response time requirements. Due to the possibility of unexpected events, such as emergencies, long-running surgeries, etc., they need to have access to patient data at any time and location. Physicians frequently change their working locations within the hospital instead of waiting for patients in their office. Thus, there is a need for mobile access to patient data compared in contrast to desktop computers.

We address the response time requirements by applying IMDB technology for data analysis, specific UI design for accessing relevant data, and an optimized data exchange to minimize the amount of exchanged data between servers and mobile device.

Our prototype is designed for use on mobile tablet devices. The mobility requirement forced us to focus on the ease of use, e.g. due to size restrictions of mobile displays. Firstly, we restricted the number of application perspectives to a login, text search, real-time analysis, and the holistic patient screen. During our design evaluation sessions, we constantly received feedback from physician and researcher user groups. We identified that each user group has specific requirements. As a result, we designed the search screen for physicians and the analytical screen for researchers specifically. Both screens consist of a limited number of UI controls. The search screen consists of a search field and a selector for data sources. The analytical screen starts with a selector of primary filter criteria and stored bookmarks.

In comparison, the holistic patient screen is optimized for all user groups. The specifically designed interactive time slider combines patient data from all hospital stays and supports access to huge sets of patient data on a limited screen. The latter enables a detailed drill-down once a certain data entry is selected.

We were able to build an interactive mobile application by exchanging a minimum of patient data to initialize screens. Once the user is drilling-down to a specific data item, detailed data are asynchronously loaded to update the displayed content using asynchronous Ajax calls and JavaScript Object Notation as data exchange format [26], [27]. Thus, the ease of use and the response time behavior of the application are addressed without affecting the quality of available data.

The data quality aspects also depends on the reliability of the application. For that, the Oncolyzer is designed as remote UI only while business logic and complex analysis are processed in the IMDB system. For example, if the battery of the mobile device drains, physicians can switch to any other device, log-on, and access their data. The IMDB system is operated by skilled personnel in the IT department of the hospital. For scaling, new servers are added to the database landscape and data portions are updated without interrupting database operations.

New releases of the Oncolyzer are automatically deployed via device management tools, such as Afaria [28]. The Sybase Unwired Platform (SUP) is used to process user authorizations, perform necessary data transformations, and manage incidents, e.g. when an iPad device gets lost. To prevent unintended data exposure, patient data are never stored on the mobile device. If the application is suspended or closed displayed data are cleared. Unintended access is prevented by built-in locking features of the incorporated devices, e.g. password protection. Access to sensitive patient data is controlled by transparent security extensions of the History-Based Access Control (HBAC) [29]. HBAC logs all user queries, analyzes them for taking access decisions, adapts access rights, and filters the result set accordingly.

The Oncolyzer is designed for very specific use-cases in which the productivity of physicians and researchers was increased, i.e. they access patient data using our application in a unified way. The selected software development process Scrum supported us in releasing new working prototypes at the end of each week, receiving feedback at the beginning of the week, and integrating new features during the remaining time [30]. The tight interaction with future application users helped to validate the results of our development sprints every week.

VIII. Conclusion and Outlook

Treatment of cancer disease is a complex process, which involves different actors contributing their individual piece of information to a complete picture. In context of personalized treatment, we expect the amount of available data per patient to increase further, e.g. by integration of data from worldwide research institutes. We outlined IT challenges in context of personalized treatment and defined requirement for clinical software applications, e.g. strict response time of less than two seconds. To prove the feasibility of our vision, we applied IMDB technology for addressing real-time analysis of clinical data. With regular user feedback, we were able to design a CCDS as mobile application for iPad and Android devices: The Oncolyzer is specifically developed for physicians and medical researchers. We discussed innovative business processes that become possible by integrating the Oncolyzer in the clinical context, e.g. involved industry partners benefit from anonymized analysis results as well as experts in tumor boards.

The given work is an initial move towards taking individualized treatment decisions in the course of systems biology. Our research results depict how to bridge the gap between clinical data in HIS and end users, who require access anywhere at any time.

IX. Acknowledgements

The given work combines the results of the research areas software engineering and translational research on tumor diseases. It is built on the collocation between the Hasso Plattner Institute in Potsdam and the Charité – Universitätsmedizin Berlin. Our results were honored with the 2012 Innovation Award of the German Capital Region. We thank all colleagues, researchers, and partners that were involved in requirements engineering, design, implementation, and evaluation to make this vision come true.

References

[1] J. Manyika *et al.*, "Big Data: The Next Frontier for Innovation, Competition, and Productivity," The McKinsey Global Institute, Tech. Rep., 2011.
[2] K. Jain, *Textbook of Pers. Medicine*. Springer, 2009.
[3] ISO, "9126-1: Software Eng.: Product Quality," 2001.
[4] J. M. Coble, J. Karat, and M. G. Kahn, "Maintaining a focus on user requirements throughout the development of clinical workstation software," in *Proceedings of the Conference on Human Factors in Computing Systems*. New York, NY, USA: ACM, 1997, pp. 170–177.
[5] H. Plattner *et al.*, *Design Thinking Research*, ser. Understanding Innovation. Springer, 2012.
[6] A. Wright, D. W. Bates, B. Middleton, T. Hongsermeier, V. Kashyap, S. M. Thomas, and D. F. Sittig, "Creating and Sharing Clinical Decision Support Content with Web 2.0: Issues and Examples," *Journal of Biomedical Informatics*, vol. 42, no. 2, pp. 334–346, Apr. 2009.
[7] W. Galitz, *The Essential Guide to User Interface Design: An Introduction to GUI Design Principles and Techniques*. New York, NY, USA: John Wiley & Sons, 2002.
[8] H. Plattner and A. Zeier, *In-Memory Data Management: An Inflection Point for Enterprise Applications*. Springer, 2011.
[9] M.-P. Schapranow, "Transaction Processing 2.0," Master's thesis, Hasso Plattner Institute, 2008.
[10] P. Svensson, "The Evolution of Vertical Database Architectures – A Historical Review," in *Proceedings of the 20th Int'l Conf. on Scientific and Statistical Database Management*. Springer-Verlag, 2008, pp. 3–5.
[11] S. Lightstone *et al.*, *Physical Database Design: The Database Professional's Guide to Exploiting Indexes, Views, Storage, and more*. Morgan Kaufmann, 2007.
[12] J. M. Hellerstein, M. Stonebraker, and J. Hamilton, *Architecture of a Database System, Foundation and Trends in Databases*. now Publishers, 2007, vol. 1.
[13] A. S. Tanenbaum, *Modern Operating Systems*, 3rd ed. Upper Saddle River, NJ, USA: Prentice Hall Press, 2008.
[14] J. M. Hellerstein and M. Stonebraker, *Readings in Database Systems*, 4th ed. MIT Press, 2005.
[15] K. Morimura, A. Romanenko, W. Min, E. I. Salim, A. Kinoshita, H. Wanibuchi, A. Vozianov, and S. Fukushima, "Possible distinct molecular carcinogenic pathways for bladder cancer in ukraine, before and after the chernobyl disaster." *Oncology Reports*, vol. 11, no. 4, pp. 881–886, 2004.
[16] F. Färber, S. K. Cha, J. Primsch, C. Bornhövd, S. Sigg, and W. Lehner, "Sap hana database: data management for modern business applications," *SIGMOD Rec.*, vol. 40, no. 4, pp. 45–51, Jan. 2012.
[17] World Health Organization, "International Classification of Diseases," http://www.who.int/classifications/icd/en/[1], May 2012.
[18] J. Singh, *Understanding ETL and Data Warehousing: Issues, Challenges and Importance*. Lambert Acad. Publishing, 2011.
[19] G. M. Ibrahim, C. Chung, and M. Bernstein, "Competing for Patients: An Ethical Framework for Recruiting Patients with Brain Tumors into Clinical Trials," *Journal of Neuro-Oncology*, vol. 104, no. 3, pp. 623–627, 2011.
[20] P. Kaatsch, C. Spix, A. Katalinic, and S. Hentschel, "Krebs in Deutschland 2007/2008," Robert Koch-Institut und die Gesellschaft der epidemiologischen Krebsregister in Deutschland e.V., Tech. Rep., 2012.
[21] *Das Gießener Tumordokumentationssystem*, 2012.
[22] Bundesrat, "Entwurf eines Gesetzes zur Weiterentwicklung der Krebsfrüherkennung und zur Qualitätssicherung durch klinische Krebsregister (Krebsfrüherkennungs- und -registergesetz - KFRG)," http://dipbt.bundestag.de/dip21/brd/2012/0511-12.pdf[1], Aug 2012.
[23] T. A. Lang and M. Secic, *How to Report Statistics in Medicine: Annotated Guidelines for Authors, Editors, and Reviewers*. Med. Writing and Commun., 2006.
[24] A. Frakt, "The Future of Health Care Costs: Hospital-Insurer Balance of Power," *Expert Voices*, 2010.
[25] K. Lehmann *et al.*, "Interdisciplinary Tumour Boards in Switzerland: Quo Vadis?" *Swiss Med Wkly*, vol. 138, no. 9–10, pp. 123–7, 2008.
[26] A. T. Holdener, *AJAX: The Definitive Guide*, 1st ed. O'Reilly, 2008.
[27] D. Crockford, "RFC4627: The application/json Media Type for JavaScript Object Notation (JSON)," http://www.ietf.org/rfc/rfc4627.txt[1], July 2006.
[28] Sybase, Inc., "Afaria: A Technical Overview," http://m.sybase.com/files/White_Papers/Afaria-Technical-WP.pdf[1], 2012.
[29] M.-P. Schapranow, "Real-time Security Extensions for EPCglobal Networks," Ph.D. dissertation, Hasso Plattner Institute, 2012.
[30] R. Pichler, *Agile Product Management With Scrum: Creating Products that Customers Love*, ser. Agile Software Development Series. Addison-Wesley, 2010.

[1] All online references were checked on Dec. 13th, 2012.

 ISBN: 978-1-61208-252-3