doi:10.3233/978-1-61499-808-2-180

From a Content Delivery Portal to a Knowledge Management System for Standardized Cancer Documentation

Danijela SCHLUE [a,1], Sebastian MATE [b], Jörg HAIER [c], Dennis KADIOGLU [d], Hans-Ulrich PROKOSCH [b] and Bernhard BREIL [a]
[a] *Niederrhein University of Applied Sciences, Krefeld, Germany*
[b] *Medical Informatics, Univ. of Erlangen-Nürnberg, Erlangen, Germany*
[c] *The Nordakademie, Elmshorn, Germany*
[d] *University Medical Center Mainz, Germany*

Abstract. Heterogeneous tumor documentation and its challenges of interpretation of medical terms lead to problems in analyses of data from clinical and epidemiological cancer registries. The objective of this project was to design, implement and improve a national content delivery portal for oncological terms. Data elements of existing handbooks and documentation sources were analyzed, combined and summarized by medical experts of different comprehensive cancer centers. Informatics experts created a generic data model based on an existing metadata repository. In order to establish a national knowledge management system for standardized cancer documentation, a prototypical tumor wiki was designed and implemented. Requirements engineering techniques were applied to optimize this platform. It is targeted to user groups such as documentation officers, physicians and patients. The linkage to other information sources like PubMed and MeSH was realized.

Keywords. Oncological documentation, knowledge management, meta-data-repository (MDR), Kano model, wiki.

1. Introduction

Tumor documentation in Germany is very heterogeneous [1], which results in different documentation practices depending on the clinic, region or information system used. This is also underlined by the German Hospital Federation in their research concerning quality in the national cancer plan for Germany [2, 3]. Possible uncertainties in the interpretation of terms (e.g. primary tumor, diagnosis date) lead to problems in analyses created by clinical and epidemiological cancer registries. These differences are a result of the respective specialization of cancer registries, varying quality assurance requirements and diversely used documentation systems.

Existing resources like handbooks for oncological documentation are mostly limited to technical parameters, such as field definitions and value domains [4]. But evaluation and also re-use of existing routine data for analyses require conceptual agreements on terms [5]. Up until now, knowledge systems, which provide structured,

[1] Corresponding Author.

free and independent information access to oncological terms, hardly exist. An international approach, which addresses this topic, is the terminology system of the National Cancer Institute (NCI) with their Enterprise Vocabulary Services [6]. Due to the increasing need for valid information, it is necessary to have one single repository as a national reference for routine cancer documentation, quality assurance documentation and scientific purposes. This repository should contain terms, definitions, comments about usage and interpretation - but no individual patient data. We designed and implemented a content delivery portal with commenting functionality and semantic links to international term definitions to increase transparency in medical documentation and to contribute to a common understanding for physicians, documentation officers and for laypersons. This system behaves similar to a wiki, although the creation of new content is limited to experts in the field of oncological documentation to assure a high quality of information.

Our objectives were to collect and analyze user requirements to improve the existing system from the user's point of view, to recognize and classify user requirements with regards to their effect on user satisfaction and to increase transparency of the oncological documentation for physicians, documentation officers and for laypersons. Requirements engineering provides many techniques for defining and prioritizing requirements [7]. The evaluation and classification technique suggested by Kano is well suited for prioritizing requirements artefacts.

2. Methods

To create a user-focused concept for the system, we analyzed the existing documentation handbooks (GEKID, ADT), coding guidelines for oncological documentation (OnkoZert) and information repositories. Together with partners from four German Comprehensive Cancer Centers (Münster, Hamburg, Dresden, Frankfurt), we defined an approval and validation process for the submission of new terms. First, terms from the existing sources were collected and analyzed concerning representation, comprehensibleness, evidence, confirmability and availability of source data. Second, a common data field definition with value domain, plausibility checks and documentation notes was created.

The Samply.MDR [8] was used as a technical base. This metadata repository already contained a raw structure of the ADT/GEKID-terms. The system is based on common web technology and its design consists of a layered architecture with data access, service and presentation layer. Within the latter one, a flexible design for different views (depending on the target group) was realized. The graphical user interface was created with JavaServer Faces 2.2 [9] and has been extended by portal-specific functionality, such as a navigation tree and a navigation bar to recognize the classification order for a term. We implemented new Java classes for the commenting functionality and extended the generic slot functionality with specific multiline text fields to ease the collection and management of term characteristics containing continuous text.

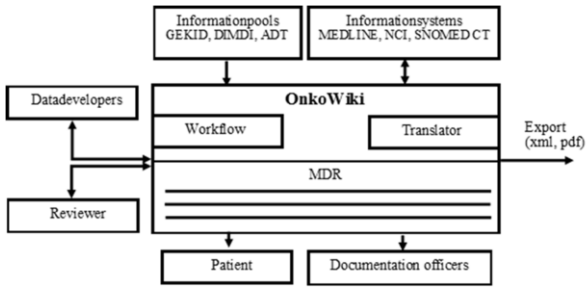

Figure 1. Architecture of the content delivery portal. Three layers allow independent access from different user groups to data from oncological information pools.

Figure 1 shows a simplified, high-level representation of the system architecture and the interaction with users and other applications, which calls services implemented within the application's service layer.

One focus was a user-oriented design, suitable especially for the main target group of documentation officers. There are physicians and scientists who have own requirements. To involve users and to implement their requirements, different methods were conducted: a pre-test in July and August of 2016, the Kano method for the prioritization process, requirements collection in an internal content management system and a main test in February 2017.

The questionnaire, which was specifically developed by the medical experts for the pretest, was based on the reference model for process description PAS[2] 1032-1, published by the German Institute for Standardization (DIN). We received complete questionnaire replies from six users.

Following a user-oriented design with focus on the documentation officers in the university hospitals with clinical cancer registries, we got insight regarding technical requirements definitions and prioritization of user requirements with Kano-classification. These requirements definitions are mainly based on contact with project workers and analysis of scientific literature. The requirements data collection comprised of an e-mail survey among medical experts and documentation officers in addition to telephone interviews. In the beginning, a set of 46 system features and requirements was identified. Informatics experts drafted a questionnaire based on the results of structured surveys with six categories: functions for compilation of terms, term usage, data, ergonomics, interfaces and documentation. A questionnaire related to the Kano model contains a set of question pairs for each requirement. The question pair includes a functional question, which captures the user's response if a product has a certain attribute, and a dysfunctional question, which captures the user's response if the product does not have that attribute. Questionnaires are presented to the future users, project partners from four German Comprehensive Cancer Centers (Munster, Hamburg, Dresden and Frankfurt) and a group of medical documentarists from Essen. The corresponding requirement class was identified by using statistical analysis of the survey results for each requirement.

After prioritization and application of changes to the user interface, we analyzed the common design and requirements again. To avoid unstructured proceeding during realization of requirements, we collected the newly arrived requirements and suggestions

[2] Publicly Available Specification

to alter the design provided by the users and summarized their feedback. In an internal content management system we analyzed more than 50 requirements. For each requirement, users filled in the category (preference, must be, bug) and prioritization of that item (high, medium, low). The comments of web application developers and the time required for the realization was captured. Agreement by all four centers was required to implement the specified offer.

Having the revised version of the web application available after the pre-test, medical experts started the main evaluation in February 2017. The user feedback was considered during the adaptation of the questionnaire, which was reduced to seven categories. The maximum number of questions was set to five per category.

3. Results

The developed data model is generic and enables the detailed specification of terms. The repository currently contains 126 oncological terms in different grades of complexity, definitions and comments about usage stored in 16 additional attributes. The terms are divided in 18 groups based on the ADT/GEKID dataset. Each term is described by its data type, domain, category and code. It is also possible to add synonyms and links to other data sources. The first prototype contains the management of terms (term creation, term representation) and user management. Currently it is possible to manage and view existing terms and review specific information about data type, plausibility checks and documentation hints with focus on medical experts. A flexible model with different views as a part of the web application's environment was developed and is available at: www.onko-wiki.de.

Usability evaluation in the pre-test phase resulted in a heterogeneous picture. The responses varied considerably and showed the importance of a structured approach that involves all key actors in the requirements analysis.

The Kano survey was planned to be distributed within the oncological network in order to receive enough answers and to prioritize the requirements from a technical point of view. Four out of ten interviewed persons (one physician, two documentation officers and one person in the field of education) returned completed questionnaires. We identified the corresponding requirement classes for each requirement. The most relevant category is the "Must-be"-category, currently containing 23 items. According to the Kano method an absence of these requirements will lead to user dissatisfaction. Seven items belong to the A-category, correlates linearly with user satisfaction. A low degree of fulfilment of these attributes reduces user satisfaction.

During the development phase, the users remained informed. The 45 proposals of their communicated feedback or change requests collected in the internal content management system were recorded within three months**.** All 13 requirements with high prioritization were implemented. Jointly agreed changes led to an adjustment of the system and the structure of the portal. The evaluation based on the answers of 23 persons obtained to the present show that the user acceptance score increased significantly.

4. Discussion

We implemented a prototype of the tumor wiki, which allows information retrieval of existing terms as well as adding new elements by a selected number of medical experts.

A pre-test with a limited number of medical experts assessed the software and gave valuable input for further optimization to follow a user-oriented design. A disadvantage of the Kano classification is the length of the questionnaire. Each requirement is represented so that a complex application will result in a longer list of questions. There might also be a misunderstanding while reading the requirements. Therefore, it is important to discuss the results in the user group while offering them the chance to give feedback. This study will help to establish a wiki featuring moderation and an approval concept to provide oncological terms in a quality-assuring way. The integration of the MDR helps to establish the nation-wide character of the platform as (1) further possibly relevant data elements can be retrieved from the MDR and included into the approval process and (2) portal-specific information (e.g. documentation hints) can be helpful for any other third party, which is using the MDR.

With regards to the content, the ADT-datasets build a very good source for our portal, but have to be extended to fulfil all subordinated purposes. We aim to connect it to other important knowledge bases, such a link to MEDLINE, which display other studies using the same term definitions. All terms are mapped to concepts defined by the NCI [6]. In addition, all terms should contain references to international terminology systems (especially to SNOMED CT or LOINC) to allow international comparisons. A term submission process should be extended. Medical societies can suggest new terms to be included into the wiki. After formal evaluation of a coordinator, the term will be provided to a team of medical oncology experts to rate and comment on the new entry. After a two-stage commentary phase and a final approval of experts and coordinator, the term will be added to the wiki and can be accessed by the different target groups.

References

[1] Blatt K, et al. Aufwand-Nutzen-Abschätzung zum Ausbau und Betrieb bundesweit flächendeckender klinischer Krebsregister. *Gutachten Prognos AG* (2010).

[2] German Hospital Federation, Quality assurance, http://www.dkgev.de/dkg.php/aid/7691/cat/42 (last access 13.03.2017).

[3] Beckmann M, Sell C, Aydogdu M, Brucker SY, Fehm T, Janni W, Kreienberg R, Kümmel S, Neumann M, Scharl A, Schleicher B, Wallwiener D, Wöckel A, Fasching PA, Lux MP, Documentation Time and Effort and Associated Resources for Patients with Primary Breast Cancer from Diagnosis to End of Follow-Up - Results of a Multicentre Validation. *Gesundheitswesen* (2015).

[4] Altmann U, Handbuch für die Anwendung der Basisdokumentation basierend auf den neuen ADT-Bögen, https://www.uni-giessen.de/fbz/fb11/institute/imi/ag/akkk/datAkkk/BAsisdoku (last access 23.03.2017).

[5] Ries M, Prokosch H U, Beckmann M, Bürkle T, Single Source Tumor Documentation. Reusing Oncology Data for Different Purposes, *Onkologi*e **36(3):136-41** (2013)

[6] NCI Enterprise Vocabulary Services (EVS) http://www.cancer.gov/research/resources/terminology (last access 01.02.2017).

[7] Pohl K, Requirements Engineering. Fundamentals, Principles and Technics, *Springer-Verlag Berlin Heidelberg* (2010), 632-643.

[8] Kadioglu D, Weingardt P, Ückert F, et al., Samply.MDR – Ein Open-Source-Metadaten-Repository, *German Medical Science GMS Publishing House* 2016, doi:10.3205/16gmds149.

[9] JavaServer Faces http://www.javaserverfaces.org/specification/ (last access 08.03.2017).