Exploring Complexity in Health: An Interdisciplinary Systems Approach
A. Hoerbst et al. (Eds.)

doi:10.3233/978-1-61499-678-1-18

Aspect-Oriented Visualization of the Health Status: An Example in Treatment of Cervical Spine Defect

Yihan DENG[a,1] and Kerstin DENECKE[b]
[a] *Innovation Center Computer Assisted Surgery, University of Leipzig*
[b] *Department of Medical Informatics, Bern University of Applied Sciences*

Abstract. Clinical data is often captured in unstructured texts and scattered in different health information systems. This complicates the aggregation of information in the process of clinical decision making. However, having a quick overview and an efficient representation of relevant aspects of a patient's health status are crucial for this process. While accessing patient data and perusing clinical documents, relevant details need to be discovered quickly. In this paper, we introduce an approach to visualize relevant information from clinical documents by tag clouds. The conventional tag clouds visualize the content of a document using the terms they are containing shown in different sizes with the size calculated based on the term frequency. Important facts and diagnostic results with low occurrence in a text may be ignored by this naïve method. In this paper, we therefore adapt the conventional tag clouds by information extraction and a guidelines-based classification schema, so that the clinical concerns can be visualized more correctly. The aspects are extracted according to a classification schema developed by clinical experts. We evaluate the approach on a set of radiology reports for cervical spine treatment.

Keywords. Visualization, Clinical Guideline, Cervical Spine, Aspects Oriented Retrieval, EHR

1. Introduction

The continuously increasing volume of clinical narratives gathered for patients during the treatment procedure poses many challenges to different stakeholders in the workflow of care providing. These documents provide the basis for physicians that are trying to get a quick access to the patient status and figure out the pre-conditions regarding the patient's medical history. To support in this process, an appropriate visualization of the data could help. There are different approaches for the visualization of text information. The information referring to diagnosis, procedure, and treatment can be highlighted with different effects for the user. Diagrams, treemap, line graphs, pie charts and interactive representation are used to show the connections between entities, semantic groups [1, 2]. The Lifeline [3] and Lifelines2 [4] have illustrated the temporal relation based on alignment, ranking and filtering functions, which facilitate the information seeking in multiple EHRs simultaneously. Some other longitude and trend analysis toolkits such as LifeFlow [5] have enabled the visualization of trends and quality of care by using large

[1] Corresponding Author: Yihan.deng@medizin.uni-leipzig.de.

amount of patient records. The high level patterns of clinical care can be obtained easily through these methods. However, existing visualization methods are mainly based on a structured electronic health record [6] and the visualization is developed to support the quality management and outcome studies in the hospital instead of the diagnostic and therapeutic activities [3, 7, 8]. Therefore, an aspect-oriented visualization based on unstructured patient records for clinical practice is required. Deng et al. [8] have conducted a clinical study on the usefulness and acceptance of tag clouds for individual patient histories. The work approved the usefulness of the cloud representation by providing quick access to unstructured EHR. The additional clinical requirements on a more detailed visualization have been determined. In this paper, we will introduce an approach that extends the simple tag cloud generation process within the same scenario (quick access to the patient status). More specifically, the semantic expressiveness of the word clouds will be improved by aspect extraction combined with the tag cloud visualization.

2. Methods

Our concept of aspect oriented visualization combine information extraction and tag clouds. This combination emphasizes the user-desired aspect and provides the focused details of the aspect in the word cloud, whereas the word cloud is generated for each aspect separately. The important evidence in different aspects can then be illustrated to the user directly. To demonstrate the whole process on a concrete example, we are considering the use case of classifying cervical spine defects. In this decision making process, a physician exploits the information from radiological reports and decides for a defect class. According to the cervical spine defect classification schema developed by Meixensberger and Leimert [9], three aspects are relevant: the defect segment, position and additional pathology. Besides, the diagnosis mentioned in the radiology report is clinically considered to be another important reference for the current defect recognition. Hence, our system will extract information on these aspects from the radiology reports and visualize it in word clouds. The objective is to improve the defect classification process in the clinical practice by this visualization. In this work, we are focusing on texts in German.

2.1. Aspects in the cervical spine defect classification

The three aspects in the classification schema are developed to capture the defect situation of a patient in a rapid way. This defect classification schema is specifically designed for the surgical practice. It covers more defect information and categorizes additional pathological information. Moreover, a much more direct description of the loco regional and anatomical causes facilitates the planning and the feasibility of surgical treatment, whereas focusing on consequences like impression of the dural sac leaves several variables regarding the direct interaction between different kinds of tissues. The segment refers to the affected cervical spine segments. The second point of differentiation is the defect position, i.e. lateral or medial affection of nerval tissue. These two positions may also co-occur at the same time. The third characteristics in the schema pays special attention to additional pathology (functional disorders) that have not been taken into account in the existing classifications: besides compression caused by prolapse of the nucleus pulpous and osseous humps, thickening of ligaments like the yellow

ligaments and the posterior longitudinal ligament as well as structural disorders (listhesis, kyphosis, hyper lordosis or rotation) may also affect viability in terms of surgical intervention. This defect classification schema enables a very detailed description of a defect by using only three axes and it can also be encoded to enable the standardized clinical communication. In addition, the diagnoses mentioned by radiologists can be used as important benchmark for the current judgment. It will therefore be visualized as one additional aspect. In this context, the visualization of the four aspects can largely accelerate the clinical diagnosis and patient triage.

2.2. Generation of aspect clouds

The clouds generation includes three important components. **1) Cervical spine terminology**: For detecting defect aspect features in clinical texts, a list comprising defect terms in German was generated. At a first step, a clinical expert has annotated the terms referring to cervical spine defects in patient records. Then, the terms were categorized into several subcategories such as anatomy, symptom, pathology and position. Meanwhile, the relevant terms in the terminology list were directly linked to existing concepts in Radlex (German), ICD 10 (German) and MeSH 2010 (German) as expansion, so that the coverage of terminology list and matching rate of the concept mapping can be reinforced through the merging of text corpora. Totally, 311 defect specific terms were included into the concept dictionary. **2) Extraction**: The extraction pipeline is established based on an adapted version of UIMA[2] provided by Averbis GmbH[3]. Two types of concept mapping are implemented in the framework: the exact mapping and the segment mapping. They have been configured to our task by including the cervical spine vocabulary. Exact mapping searches only for fully matched terms in the corresponding semantic scope, whereas the segment mapping performs a fuzzy matching considering morphological variations of terms. Further, the contiguous match strategy is employed to obtain the longest match of contiguous tokens within the sentence. Besides, a German negation list is additionally exploited by the concept mapper. Three main types of negations in German were defined, namely post negation ("nicht vorhanden" (non-existent)), pre negation ("frei" (free)) and pseudo negation ("nicht sicher, ob" (not sure whether)). Addressing domain- or task-specific vocabulary, the concept mapper was complemented by additional regular expressions, for example to identify location descriptions expressed as coordination structures. **3) Tag clouds generation**: in general, a tag cloud is a visual representation of the bag of words model, i.e. the tag size reflects the occurrence frequency of the corresponding term. The more often a term appears in one document, the larger the size of the tag in the cloud. We used the OpenCloud[4] tool to generate the tag clouds from clinical documents. In our prototypical system, each aspect (type, position and additional pathology) in our defect schema together with the diagnoses will be presented with one separate sub cloud. All the clouds are generated based on the result of the extraction. The ten most frequent tokens are shown in the tag cloud with size depending on their frequency in the document. Each tag is linked with the corresponding original document. Different clouds are rendered with different colors.

---

[2] https://uima.apache.org/

[3] http://www.averbis.de

[4] http://opencloud.mcavallo.org/

3. Results

By applying the methods as described before, we receive terms relevant for the single aspects. They are visualized in the tag clouds as shown in Figure 1. The four aspects defined based on the classification schema are illustrated in its own sub clouds. Hyperlinks are assigned to the tags in the clouds enabling physicians to see more details in the original text in the patient record. By clicking on a tag, the corresponding text will be opened with highlighted keywords.

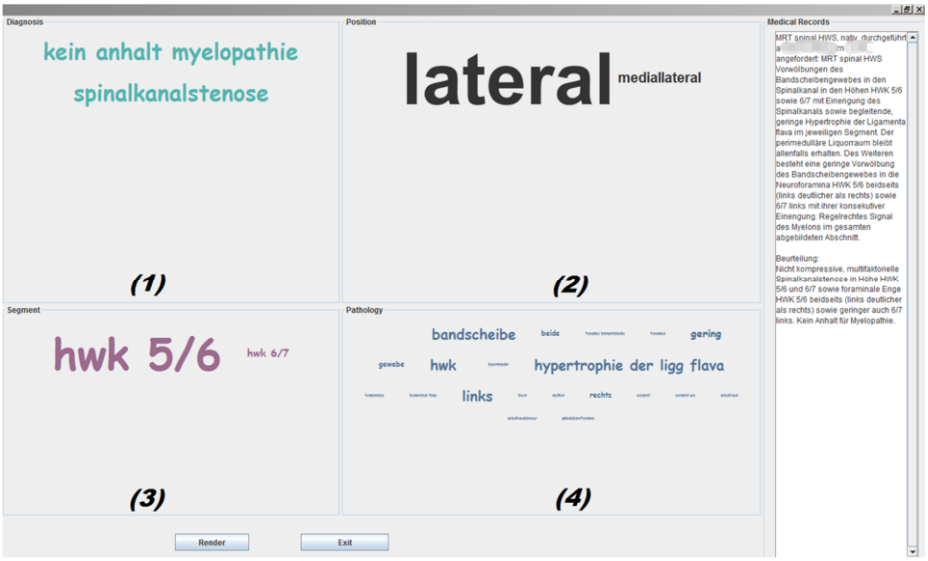

Figure 1. Excerpt of the prototypical visualization system: The four tag clouds reflect the four aspects relevant for the cervical spine defect classification, that are (1) Diagnosis (2) Position (3) Segment (4) Additional pathology

In our evaluation, we assessed the quality of the aspect extraction using an annotated corpus including 100 anonymized radiology reports. Four different aspects were manually annotated, while precision, recall and F1-measure of the extraction were determined. The quality of the extraction is important since it directly influences the usefulness of the cloud visualization. Based on the concept-based mapping and regular expression, we have reached satisfactory result in the aspect extraction. As can be seen in the Table 1, the extraction of position performs best. However, the recall needs still to be improved for the other aspects. The detection of cervical spine segments has the highest miss rate due to the non-uniformed notation and abbreviations of these concepts.

Table 1. Extraction result for the visualization aspects

Extracted Element	Precision	Recall	F1
Segment	100%	85.0%	91.0%
Position	100%	99%	99%
Add. Pathology	100%	97%	98%
ICD-Diagnosis	97%	94%	92.5%

4. Discussion

The errors mainly occurred in the recall (85%) of the extraction of information on the cervical spine segment. After manual inspection, it became clear that varying vocabulary leads to a high miss rate. The regular expressions regarding three variations defined in our extraction pipeline are clearly insufficient to deal with the various notations used by different physicians. As one possible improvement, more possible terms and abbreviations will be collected from the current corpus and will be summarized in regular expressions to increase the recall.

We did not yet performed a structured study of the usefulness of the tag clouds. However, through interviews focusing on the usefulness and relevance of the aspect clouds with clinicians of the neurosurgical department of a university hospital, the effectiveness of the aspect clouds was approved. They confirmed that with this kind of visualization, the detailed patient status which is related to the cervical spine defect can be recognized in very short time. The important evidences can also be found more easily. The document links help the physician to see more details in the original EHR, which provide the physician more flexibility in information seeking. Additionally, this hybrid method enables a more efficient access to the electronic records by patient visits in the outpatient department and enhances the effectiveness of the monitoring patient data in unstructured documents by the clinical collaboration. As a next step, a more comprehensive user study will be organized to figure out the additional user requirements. As technical extensions, the semantic linking between the clouds will be established.

References

[1] Ballstaedt SP. Texte visualisieren (Visualizing texts). In: E.-M. Jakobs, D. Knorr, K-H. Pogner (Eds.): Textproduktion. HyperText, Text, Kontext. Peter Lang Verlag, Frankfurt/Main, 1999. P. 129-40.

[2] Gotz D, Sun J, Cao N, et al. Visual cluster analysis in support of clinical decision intelligence. AMIA Annu Symp Proc. 2011;2011:481–490

[3] G. Kopanitsa, C. Hildebrand, J. Stausberg, K.-H. Englmeier: Visualization of Medical Data Based on EHR Standards. Methods Inf Medicine, 2012,52(1):43-50.

[4] Wang TD, Plaisant C, Quinn AJ, et al. Aligning temporal data by sentinel events: discovering patterns in electronic health records. CHI ‘08 SIGCHI Conference on Human Factors in Computing Systems Proceedings 2008:457–466.

[5] Wongsuphasawat K, Guerra Go´mez JA, Plaisant C, et al. LifeFlow: visualizing an overview of event sequences. SIGCHI Conference on Human Factors in Computing Systems Proceedings 2011: 1747-1756.

[6] A. Rind, T. D. Wang, W. Aigner et al., Interactive Information Visualization to Explore and Query Electronic Health Records. In: Foundations and Trends in Human-Computer Interaction, 5(3), 2013: 207-98.

[7] Klimov D, Shahar Y, Taieb-Maimon M. Intelligent interactive visual exploration of temporal associations among multiple time-oriented patient records. Methods Inf Med. 2009;48: 254–262

[8] Deng Y, Denecke K. Visualizing unstructured patient data for assessing diagnostic and therapeutic history. Stud Health Technol Inform. 2014;205:1158-62.

[9] von Sachsen S, Daenzer S, Freitag S, Meixensberger J, Leimert M, Groll JM, Leimert M, Computer Aided Defect Classification for Model-based Therapy of Cervical Spinal Stenosis. Proceedings of the International Conference on Biomedical Engineering and Systems, 2(1), 2014:126-172