Enhancing Patient Safety through Human-Computer Information Retrieval on the Example of German-speaking Surgical Reports

Christof Stocker*, Leopold-Michael Marzi†, Christian Matula‡,
Johannes Schantl*, Gottfried Prohaska§, Alberto Brabenetz§ and Andreas Holzinger*
* Research Unit Human–Computer Interaction
Institute for Medical Informatics, Statistics & Documentation,
Medical University Graz, Auenbruggerplatz 2, 8036 Graz, Austria
Email: c.stocker,j.schantl,a.holzinger@hci4all.at
† Legal Department, University Hospital of Vienna,
Währinger Gürtel 18-20, 1090 Vienna, Austria
‡ Department of Neurosurgery, University Hospital of Vienna,
Währinger Gürtel 18-20, 1090 Vienna, Austria
§ IBM Austria,
Obere Donaustraße 95, 1020 Vienna, Austria

***Abstract*—In view of the high number of deaths and complication rates of major surgical procedures worldwide, surgical safety is described as a substantial global public-health concern. Naturally, patient safety has become an international priority. The increasing amount of electronically available clinical documents holds great potential for the computational analysis of large repositories. However, most of this data is in textual form and the clinical domain is a challenging field for the appliance of natural language processing. This is particularly the case if you deal with a language other than English, due to the little attention from the international research community. In this project, we are concerned with the utilization of a German-speaking operative report repository for the purpose of risk management and patient safety research. In this particular paper we focus on the description of our information retrieval approach. We investigated the thought process of a domain expert in order to derive his information of interest and describe a facet-based way to navigate this kind of information in the form of extracted phrases. Initial results and feedback has been very promising, but a formal evaluation is still missing.**

I. Introduction

Electronic patient files contain large portions of data which has been entered in a non-standardized format, often referred to as "free text" [1]. An important subset of patient-related (and mostly textual) data are operative reports, which are documentations of surgical procedures. Surgery, as an essential branch of medicine, is concerned with the treatment of injuries as well as disorders of the human body by incision or manipulation [2]. It's outcome is influenced by various factors, including the surgeons' skills, techniques used, incision length, and supplies used.

Following each surgical procedure, operative reports are created for the purpose of documentation and billing [2]. They contain information on the indications of the procedure, pre- and post-procedure diagnoses, complications, findings during the procedure, as well as the detailed description of the procedure [2]. It is very reasonable to assume that the information content of operative reports offer a lot of potential for secondary uses such as surgical (safety) research, decision support, and quality improvement. However, the manual review of such a data repository, even within just one single hospital, is an unpractical and unscalable solution [2]. The worldwide volume of surgery is large and in view of the high number of death and complication rates of major surgical procedures, surgical safety is now described as a substantial global public-health concern [3].

In our ongoing project we are concerned with the utilization of an operative report repository in the German language for the purpose of risk management to improve patient safety and thus limit malpractice exposure. The first goal is to implement a scalable and expandable information retrieval and analytics tool that can be used to identify, assess, and review potential risk factors, hazards, and problematic cases. Most importantly we want this system to achieve these goals through user inclusion, which means that the functionality should ideally be a byproduct of the user being able to interact with the data. In this paper, we are going to focus on the description of our approach for the Information Retrieval (IR) functionality.

A promising step towards the identification of adverse events has been the use of triggers or indicators of harm within the medical record system [4]. However, while this works reasonably well for known harms, it does less well for spotting rare or unusual events [5], which is exactly the kind of information we are looking for.

We follow the assumption that there is a significant amount of unreported cases in surgery, in which mistakes or other hazardous factors did not lead to obvious harm or their reporting [6]. This could have happened for various reasons. Let us state a couple examples: (1) A mistake made in surgery was caught and corrected early enough to prevent any notable harm to the patient. Of course this would be transcribed within

1529-4188/14 $31.00 © 2014 IEEE
DOI 10.1109/DEXA.2014.53

the operative report, but other than that this information would be lost. (2) In high risk areas, such as neurosurgery, it is not uncommon that adverse events occur that cause harm to the patient. This need not be the anyone's fault in particular. Sometimes, such an event results in a legal case, which would lead to the information of the case potentially being relevant for patient safety research being documented, but most of the time - so our hypothesis - it doesn't. However, the operative report should contain this information either way.

Naturally, in the first step we are interested in a way to *identify* such cases without much human effort. We approached this problem by investigating how a trained medical professional would filter out non-relevant reports, to then design an Information Extraction (IE) system that mimics this process. However, the extraction of information out of text remains a challenging task (for a recent overview on text mining methods see [7]). Although text can easily be *created* by medical professionals, the support of (semi-) automatic analyses is extremely difficult and has challenged researchers for many years [8] [9]. The language constraint (German mixed with Latin vocabularies) raises additional challenges such as the lack of qualitative, publicly available domain- and semantic resources.

To work around current natural language processing (NLP) limitations, we build on the faceted navigation idea of Human-Computer Information Retrieval (HCIR) [10] to follow the principle behind HCI-KDD [11]. What this essentially means is that we designed the system to utilize the user's domain knowledge in order to overcome (or work around) computationally difficult problems by including him or her into the IR process. However, instead of using the facet tree to categorize documents, it is used to organize and navigate extracted information. For this purpose we decided to build upon an existing expandable content analytics tool that provides IE support as well as an applicable user interface for faceted navigation [12].

II. Background

Modern NLP systems utilize approaches from sub areas such as information theory, linear algebra and probabilistic modelling [13]. The major tasks of a NLP system include sentence boundary detection, tokenization, part-of-speech (POS) tagging, IR, IE, named entity recognition, and morphological decomposition [14].

IE is a part of NLP and describes processes to find information automatically from data sources. Therefore it returns information and not information sources like information retrieval systems do. Grishman defines IE as the process of automatically identifying and classifying instances of entities, relations and events in a text, based on some semantic criterion [15]. Typical task are name-, entity-, relation- and event extraction.

In contrast to IE, IR is the process of finding relevant information sources (in this case documents) to a specific search request [16], and is a part of Human-Computer Interaction (HCI). The provided user-interface of an IR system is crucial for the usability and therefore the interface design has to be thought-out [17]. This gave birth to HCIR, which has a strong focus on user inclusion in the IR process [10].

A grand challenge is to answer the question "*What is interesting?*" and Beale (2007) describes this "interest" as a perspective on relationships between data, which is influenced by tasks and prior knowledge [18]. Thus "interest" is essentially a human construct. Consequently, a fruitful approach is to combine HCI and knowledge discovery (KDD) in order to enhance human intelligence by computational intelligence [19]. The main contribution of HCI-KDD is to *enable* end-users to *find and recognize* previously unknown and potentially useful, usable, and interesting information. This approach is based on the assumption that the domain expert possesses explicit domain knowledge and by enabling him or her to interactively look at his data sets, he may be able to identify, extract and understand useful information, as to gain new - previously unknown - knowledge.

III. Related Work

The clinical domain has presented itself as a challenging field for the appliance of NLP technology. It's barriers include the lack of access to shared data, the lack of annotated datasets for training and benchmarking, insufficient conventions and standards for annotations, and the lack of user-centered development and scalability [20].

Quite recently a generalizable approach to a clinical IR system, referred to as automated retrieval console (ARC), has been proposed [21]. It's design is based on the hypothesis that supervised machine learning algorithms using robust enough feature sets are capable of delivering acceptable performance across a variety of clinical IR applications [21]. The authors identified the reliance on custom software and rules or heuristic development as one of the most substantial barriers for clinical IR adoption. To overcome this problem, they implemented the system to be easily configured for a specific task by providing a user interface for corpus annotation, model selection, and performance calculation. ARC is an UIMA-based system. UIMA stands for Unstructured Information Management Architecture and is an OASIS standard used by systems that analyse unstructured information such as text [22]. ARC utilizes the output of other UIMA-based NLP systems, such as the Clinical Text Analysis and Knowledge Extraction System (cTAKES), to derive the available feature set.

It has been discussed, that we need new approaches for rapid adaptation of NLP systems to new applications, with less dependence on deeply annotated data [20]. One could argue that for some applications, the answer to "what is interesting?" might not be well defined, or even still a matter of research [18]. It is not too far fetched to argue the need for the end user to interact with the features - or more generally speaking with the information. This leads to a different question, namely "What information is representative for the document content?"; a question that might have different answers, depending on the research domain and targeted audience.

Risk management, for example, can be defined as the structured process of identifying, assessing, controlling, and reacting to risks arising from various factors [23]. It is thus concerned with the organisational tactics to fix problems [24]. In contrast to risk management, patient safety is more about changing the work culture of the organisation than about the problem itself [24]. In other words, it is proactive rather than reactive and has the ability to support risk management efforts through new ways of understanding how things go wrong. Early reports underscored the importance of patient safety as a key dimension of quality and identified information technology as crucial for achieving this goal [6]. Since then, patient safety has become an international priority [5]. An important subarea of safety research is concerned with surgical safety, which is described as a substantial global public-health concern [3].

Since the early 1990s researchers have been exploring the potential of administrative data in quality assessment and patient safety research [4]. In 2002, the Agency for Healthcare Research and Quality (AHRQ) developed and released the so-called patient safety indicators (PSIs) [25], a tool designed for screening administrative data for patient safety events and medical errors [4].

A major limitation of such a system has been described as the reliance on structured and coded data [26], which is also usually collected for other purposes such as billing [5]. One problem concerning structured data in the electronic health record is consistency. The ability to code structured data can vary widely between clinicians, leading to wide variations in the accuracy and completeness of this data [26]. Furthermore, concerns exists about the validity of administrative codes [27].

Quite a few authors were exploring the idea that free-text searches of electronic medical records could yield additional information for the identification of patient safety. One interesting study by Murff et al was carried out in 2003 [28]. The authors used keyword queries to detect adverse events in electronic screening discharge summaries. However, the results lead to the conclusion that simple keyword queries are not specific enough to be practical. On the other hand, the incorporation of NLP techniques was hypothesised to be a promising next step.

Another study by Murff et al in 2011 was focused on the evaluation of an NLP approach using an electronic medical record for the identification of postoperative surgical complications [29]. The results showed some cases, such as post-operative myocardial infarction, where the PSI algorithm had excellent test characteristics that were not improved through the NLP approach. On the other hand, the NLP approach in general showed significantly greater sensitivities with only a small reduction in specificities. Another advantage that was mentioned was the flexibility of the NLP approach to meet the individual institutional needs at a relatively low programming effort. The author's results lead them to the conclusion that using NLP with an electronic medical record can greatly improve post-operative complication identification compared with the PSIs [29].

Other related investigations concerning surveillance of adverse events using NLP, carried out by Melton & Hripcsak [30] and Penz et al [31] respectively, reached similar conclusion. The first study used the NLP system MedLEE to construct an adverse event detection system for discharge summaries. They found that NLP outperformed traditional adverse event detection methods and thus concluded that it is an effective technique for detecting a broad range of adverse events [30]. The second study utilized two methods, a NLP program and a phrase-matching algorithm, for the identification of adverse events related to the placement of central venous catheters. To put the results into perspective the authors compared them to a sample of manually reviewed records. They found that methods, such as phrase matching, can be more sensitive than administrative data in identifying patients with devices.

Concerning potential NLP limitations, Wang et al [2] investigated the sub-language used to describe actions within the operative report. In their comprehensive study, the authors demonstrated the need to construct surgical domain-specific semantic resources for IE from operative notes.

IV. Methods

Our ongoing mission is the reduction of risk factors (be it known or yet unknown) to improve patient safety and limit malpractice exposure. Experience has indicated that a lot of risk factors could be of local origin. We hold the hypothesis that there is a significant amount of unreported cases in the surgical domain, in which mistakes or other hazardous factors did not lead to obvious harm or their reporting. To be able to verify such type of assumptions, data analysis has to be performed. Given the large amount of data a typical hospital generates, performing such an investigation by hand would be an unscalable solution [2]. Furthermore, it is essential to continuously monitor the data for change, even after potential solutions are identified and implemented.

This lead us to the conclusion that the first step towards an integrated solution needs to be the implementation of an exploratory analytics tool that is capable of dealing with large amounts of structured and unstructured data in the surgical domain. The structured data contains the meta-data associated with each operation, such as their dates and durations. The unstructured data is the textual content of the operative reports. It includes the indications of the procedure and the detailed description of the procedure.

The system is being build around an existing UIMA-based tool and framework called IBM Content Analytics with Enterprise Search (ICAwES) [12]. It offers out-of-the-box NLP functionalities such as part-of-speech tagging, custom dictionaries, a scalable system architecture, a graphical end-user interface, and an Eclipse-based NLP development suite with deployment functionality. The end-user interface supports various ways to visualize and interact with the data, with the most practical being the facet view.

We operate on the assertion, that a domain expert would be able to identify documents that are interesting for the identification of risk factors or adverse events, if he would process them manually. Thus, the first work package was to identify and formulate the thought process of such an expert. In other words, we focused on identifying what type of information in the text is interesting or relevant for the domain expert. In this process, a couple different information of interest were identified. Those include unanticipated events in the description section, strong deviation of typical operation duration or report length, used equipment, and heavy emphases.

In order to identify phrases of unanticipated events and heavy emphases, further called *triggered-phrases*, several rules were defined in IBM Content Analytics. These rules utilize a dictionary annotator, to find words identifying the beginning of a phrase, and a POS tagger to identify the end of the phrase and to define combinations of POS tags allowed in between the phrase.

The beginning of a *triggered-phrase* is identified by using two dictionaries which are manually filled with common German trigger words and their flexions. To date, they contain a total of 30 words that are similar to the English counterparts *surprisingly*, *suddenly*, *extreme*, and so forth. To be more precise, the first dictionary contains those trigger words that indicate surprise or unexpected events. Examples: *plötzlich*, *überraschend*, *seltsam*, *erstaunlich*, *unvorhergesehen*, *unerwartet*. The second dictionary contains the trigger words indicating strong emphasis, such as: *enorm*, *extrem*, *mächtig*, *massiv*, *speziell*, *schwer*, *schwierig*, *völlig*, *heftig*.

Keyword matching alone, however, did not deliver enough information. To overcome this, we utilized the triggers in combination with the POS tags to extract associated phrases. One benefit of the German language is that nouns generally start with a capital letter. We used this property to identify nouns not present in a typical dictionary in order to complement POS.

At all we defined eight rules to match *triggered-phrases*. While all rules are triggered by a *trigger-word*, and take a noun (in one case also a verb) as final word of the phrase, they differ mainly in the combination of words allowed in between trigger and final word. For instance, a rule is defined to match a case where the trigger word is followed by a noun without words in between. A more complex rule for instance can match also cases like "*massive und komplexe weite Blutung*", where the *trigger-word* is followed by a connective word and two adjectives and the phrase ends with a noun. In a similar way the six other rules are defined.

With this approach, only the trigger words have to be known beforehand. Of course, the extracted information of the indication section and the procedure-description section had to be displayed separately, since the context of the phrases is very different. Using the build-in lemmatization, the extracted phrases were normalized.

Examples for extracted normalized phrases from the indication sections are "*insbesonders ... etwaig alternativ ... Option*", "*insbesonders ... potentiell ... Option*", "*insbesonders ... potentiell ... Problematik*". All these phrases show cases where the patient was especially advised to consider alternative options. Of course this need not mean anything in context with risk management. It does, however, provide good clues for the user on the surgery's difficulty.

Examples for positive phrases extracted from the procedure-description section are "*völlig ... problemlos ... Blutstillung*", and "*völlig ... zufriedenstellend ... Verhältnis*". Both phrases indicate positive findings, which make them less interesting for safety research. This can be useful for filtering, since the user can also select phrases and exclude the associated documents from the search query. In contrast, two examples for negative phrases are "*plötzlich ... Blutung*" and "*enorm ... geschwollen ... Hirn*". The first phrase indicates that it came to a sudden bleeding, while the second phrase shows that the brain was enormously swollen; both indications for potentially interesting cases.

V. Discussion

For a system designed for exploratory content analytics, evaluation can be challenging. Since the first step of our project focuses on the use of generic indicators, rather than medical terms and phrases, recall and precision tell only half the story. For this scenario, we think there are two main ways to get meaningful results aside from the usual relevance measures. First, we need to conduct a user study to assess the tool's interface and usefulness in the eyes of the targeted end-users. Secondly, we need to evaluate if our exploratory analysis approach really helps the end-user identify documents that lead to new insight.

In contrast to other studies, we are concerned with German-speaking operative reports. The biggest restriction in the project design was the lack of publicly available ontological resources. We aimed to find practical ways to work around this restriction by investigating the thought process of domain experts when manually analyzing such documents. We realize that this is an iterative process that needs clear evaluation going forward.

The triggered-phrases approach seems very promising as it opens the option for it to deal with unseen data. This might be because the triggers don't change much, at least based on the data we have analyzed so far. Of course, the phrases are only very weak indications and are in this constellation only useful in an interactive approach. Their purpose is to offer the end user insights and filter methods to navigate the data collection. Also, in this first step, manual analysis of the individual cases can not be completely eliminated, but hopefully more focussed.

VI. Conclusion

In our ongoing project we are concerned with the utilization of an operative report repository in the German language for the purpose of risk management to improve patient safety and thus limit malpractice exposure. Experience has indicated that a lot of risk factors could be of local origin. We suspect that there might be a significant amount of unreported cases in the surgical domain, in which mistakes or other factors did not lead to obvious harm or their reporting. However, the manual analysis of such large data repositories is an unscalable solution. For this purpose we are developing a analytics system that is capable to support the domain expert in the data analysis process.

Battles and Lilford categorize patient safety research initiatives into three stages: (1) the identification of the risks and hazards; (2) the design, implementation and evaluation of the patient safety practices; and (3) maintaining vigilance to ensure the practices continue and the patient safety culture remains in place [5]. Considering this categorization scheme, this project is currently in stage one. In this paper we focused on the description of our approach to the IR functionality.

The biggest restriction was the lack of publicly available ontological resources in the German-speaking medical field. We approached this problem by aiming to somehow work around this fact through close communication with the domain expert in the design phase. The main contribution of this paper is the description of a lightweight, trigger-based HCIR approach for the purpose of patient safety research, that does not depend on annotated data or ontological resources, but simple trigger dictionaries.

The feedback thus far has been very enthusiastic. However, to really achieve hard evidence we need more data and time for testing and evaluation. We move forward towards more secondary data inclusion, such as the information about which surgery resulted in a legal case.

ACKNOWLEDGMENT

The authors would like to thank the Software Group of IBM Austria for their support; in particular Karin Maurer and Michael Grosinger.

REFERENCES

[1] A. Holzinger, R. Geierhofer, and M. Errath, "Semantische informationsextraktion in medizinischen informationssystemen," *Informatik Spektrum*, vol. 30, no. 2, pp. 69–78, 2007.

[2] Y. Wang, S. V. S. Pakhomov, N. E. Burkart, J. O. Ryan, and G. B. Melton, "A Study of Actions in Operative Notes," in *AMIA Annual Symposium Proceedings*, vol. 2012. American Medical Informatics Association, Jan. 2012, pp. 1431–1440.

[3] T. G. Weiser, S. E. Regenbogen, K. D. Thompson, A. B. Haynes, S. R. Lipsitz, W. R. Berry, and A. A. Gawande, "An estimation of the global volume of surgery: a modelling strategy based on available data." *Lancet*, vol. 372, no. 9633, pp. 139–44, Jul. 2008.

[4] C. Zhan and M. R. Miller, "Administrative data based patient safety research: a critical review," *Quality and Safety in Health Care*, vol. 12, no. Suppl 2, pp. ii58–ii63, 2003.

[5] J. B. Battles and R. J. Lilford, "Organizing patient safety research to identify risks and hazards," *Quality and Safety in Health Care*, vol. 12, no. Suppl 2, pp. ii2–ii7, Dec. 2003.

[6] D. W. Bates, "Detecting Adverse Events Using Information Technology," *Journal of the American Medical Informatics Association*, vol. 10, no. 2, pp. 115–128, Nov. 2002.

[7] A. Holzinger, J. Schantl, M. Schroettner, C. Seifert, and K. Verspoor, "Biomedical text mining: Open problems and future challenges," in *Lecture Notes in Computer Science LNCS 8401*. Heidelberg, Berlin: Springer, 2014, pp. 271–300.

[8] J. Gregory, J. E. Mattison, and C. Linde, "Naming notes - transitions from free-text to structured entry," *Methods of Information in Medicine*, vol. 34, no. 1-2, pp. 57–67, 1995.

[9] C. Lovis, R. H. Baud, and P. Planche, "Power of expression in the electronic patient record: structured data or narrative text?" *International Journal of Medical Informatics*, vol. 58, pp. 101–110, 2000.

[10] G. Marchionini, "Toward human-computer information retrieval," *Bulletin of the American Society for Information Science and Technology*, vol. 32, pp. 20–22, 2007.

[11] A. Holzinger, *HumanComputer Interaction & Knowledge Discovery (HCI-KDD): What is the benefit of bringing those two fields to work together?* Heidelberg, Berlin, New York: Springer, 2013, pp. 319–328.

[12] A. Holzinger, C. Stocker, B. Ofner, G. Prohaska, A. Brabenetz, and R. Hofmann-Wellenhof, "Combining HCI, Natural Language Processing, and Knowledge Discovery - Potential of IBM Content Analytics as an Assistive Technology in the Biomedical Field," in *Lecture Notes in Computer Science, LNCS 7947*. Springer, 2013, pp. 13–24.

[13] C. D. Manning and H. Schütze, *Foundations of statistical natural language processing*. Cambridge, MA, USA: MIT Press, 1999.

[14] P. M. Nadkarni, L. Ohno-Machado, and W. W. Chapman, "Natural language processing: an introduction." *Journal of the American Medical Informatics Association : JAMIA*, vol. 18, pp. 544–551, 2011.

[15] R. Mitkov, *The Oxford Handbook of Computational Linguistics (Oxford Handbooks in Linguistics S.)*. Oxford University Press, 2003.

[16] S. M. Meystre, G. K. Savova, K. C. Kipper-Schuler, and J. F. Hurdle, "Extracting information from textual documents in the electronic health record: a review of recent research." *Yearbook of Medical Informatics*, vol. 47, no. Suppl 1, pp. 128–144, Jan. 2008.

[17] R. Baeza-Yates and B. Ribeiro-Neto, *Modern Information Retrieval*, 463rd ed. New York: ACM press, 1999, vol. 463.

[18] R. Beale, "Supporting serendipity: Using ambient intelligence to augment user exploration for data mining and web browsing," *Int. J. Hum.-Comput. Stud.*, vol. 65, no. 5, pp. 421–433, May 2007.

[19] A. Holzinger, "On knowledge discovery and interactive intelligent visualization of biomedical data - challenges in humancomputer interaction & biomedical informatics," in *DATA 2012*, 2012, pp. IS9–IS20.

[20] W. W. Chapman, P. M. Nadkarni, L. Hirschman, L. W. D'Avolio, G. K. Savova, and O. Uzuner, "Overcoming barriers to NLP for clinical text: the role of shared tasks and the need for additional creative solutions." *Journal of the American Medical Informatics Association : JAMIA*, vol. 18, no. 5, pp. 540–543, 2011.

[21] L. W. D'Avolio, T. M. Nguyen, W. R. Farwell, Y. Chen, F. Fitzmeyer, O. M. Harris, and L. D. Fiore, "Evaluation of a generalizable approach to clinical information retrieval using the automated retrieval console (ARC)." *Journal of the American Medical Informatics Association : JAMIA*, vol. 17, no. 4, pp. 375–82, Jan. 2010.

[22] D. Ferrucci, A. Lally, K. Verspoor, and E. Nyberg, "Unstructured Information Management Architecture (UIMA) Version 1.0, OASIS Standard," OASIS, Tech. Rep., Mar. 2009.

[23] J. Makki, A.-M. Alquier, and V. Prince, "Ontology Population via NLP Techniques in Risk Management," *International Journal of Humanities and Social Sciences*, vol. 3, no. 3, pp. 212–217, 2009.

[24] B. J. Youngberg, Ed., *Principles of Risk Management and Patient Safety*, 1st ed. Jones & Bartlett Publishers, 2010.

[25] *AHRQ Quality Indicators Guide to Patient Safety Indicators*. Rockville, MD: Agency for Healthcare Research and Quality, 2003.

[26] K. Cresswell, A. Majeed, D. W. Bates, and A. Sheikh, "Computerised decision support systems for healthcare professionals: an interpretative review." *Informatics in Primary Care*, vol. 20, no. 2, pp. 115–128, Jan. 2012.

[27] P. S. Romano, H. J. Mull, P. E. Rivard, S. Zhao, W. G. Henderson, S. Loveland, D. Tsilimingras, C. L. Christiansen, and A. K. Rosen, "Validity of selected AHRQ patient safety indicators based on VA National Surgical Quality Improvement Program data." *Health services research*, vol. 44, pp. 182–204, 2009.

[28] H. J. Murff, A. J. Forster, J. F. Peterson, J. M. Fiskio, H. L. Heiman, and D. W. Bates, "Electronically screening discharge summaries for adverse medical events," *Journal of the American Medical Informatics Association*, vol. 10, pp. 339–350, 2003.

[29] H. J. Murff, F. FitzHenry, M. E. Matheny, N. Gentry, K. L. Kotter, K. Crimin, R. S. Dittus, A. K. Rosen, P. L. Elkin, S. H. Brown, and T. Speroff, "Automated identification of postoperative complications within an electronic medical record using natural language processing." *Journal of the American Medical Informatics Association*, vol. 306, no. 8, pp. 848–855, Aug. 2011.

[30] G. B. Melton and G. Hripcsak, "Automated detection of adverse events using natural language processing of discharge summaries." *Journal of the American Medical Informatics Association*, vol. 12, no. 4, pp. 448–457, Jan. 2005.

[31] J. F. E. Penz, A. B. Wilcox, and J. F. Hurdle, "Automated identification of adverse events related to central venous catheters." *Journal of Biomedical Informatics*, vol. 40, pp. 174–182, 2007.