Kreuzthaler, M., Bloice, M. D., Simonic, K.-M. & Holzinger, A. On the Need for Open-Source Ground Truths for Medical Information Retrieval Systems. In: Tochtermann, K. & Maurer, H., eds., 10th International Conference on Knowledge

On the Need for Open-Source Ground Truths for Medical Information Retrieval Systems

Markus Kreuzthaler
(Medical University Graz, Austria
markus.kreuzthaler@medunigraz.at)

Marcus Bloice
(Medical University Graz, Austria
marcus.bloice@medunigraz.at)

Klaus-Martin Simonic
(Medical University Graz, Austria
klaus.simonic@medunigraz.at)

Andreas Holzinger
(Medical University Graz, Austria
andreas.holzinger@medunigraz.at)

Abstract:
Smart information retrieval systems are becoming increasingly prevalent due to the rate at which the amount of digitized raw data has increased, and continues to increase. This is especially true in the medical domain, as there is much data stored in unstructured formats which contain "hidden" information within them. By hidden, this means information that cannot ordinarily be found by performing a simple text search. To test the information retrieval systems that handle such data, a ground truth, or gold standard, is normally required in order to gain performance values according to an information need. In this paper we emphasize the lack of freely available, annotated medical data and wish to encourage the community of developers working in this area to make available whatever data they can. Also, the importance of such annotated medical data is raised, especially its importance and potential impact on teaching and training in medicine. As well as this, this paper will point out some of the advantages that access to a freely available pool of annotated medical objects would provide to several areas of medicine and informatics. The paper then discusses some of the considerations that would have to be made for any future systems developed that would provide a service to make the creating, sharing, and annotating of such data easy to perform (by using an online, web-based interface, for example). Finally, the paper discusses in detail the benefits of such a system to teaching and examining medical students.

Keywords: Evaluation, Text Mining, Information Retrieval, Medicine, Health Care
Categories: H.3.1, H.3.3, I.2.7, J.3

1 Introduction and Motivation

Free text consists of natural language, and the automated processing and understanding of natural language – although a research topic for more than 40 years – is still in its infancy [Gell et al., 1976]. Despite the prevalence of advanced

information systems, many organizations still depend on this unstructured plain text as a source of information. This is especially true in the medical domain, where old, hand-written documents are scanned and then read by optical character recognition systems and subsequently fed into information systems, thereby creating very large collections of unstructured text. Medical professionals, therefore, are often confronted with free texts, and these texts play an important role in clinical research, medical accounting, and quality assurance [Holzinger et al., 2007b]. Therefore, making this data more accessible, useful and useable is of high importance [Holzinger et al., 2008]. Consequently, any improvement to information retrieval systems that can assist the end user in handling these free texts is a positive advancement [Kaiser et al., 2007].

In medicine, this is perhaps more difficult to achieve than in other domains; medical texts consist of many abbreviations, synonyms, are written in several languages, and contain significant numbers of linguistic variations. A medical doctor's priority when writing a diagnosis text, for example, is not grammatical correctness and there is no compulsive need for a doctor to pay too much attention to legibility so long as the text is understandable in a medical context. Therefore, the evaluation of such information retrieval systems is a major concern. More often than not, free text retrieval is performed by simply searching the text pool for text patterns specified by the doctor. Purely statistical evaluations of such systems often make use of a ground truth, known as a gold standard, to compare the result returned by a retrieval system with that of a medical expert. Of course, there are a number of gold standard text corpora available to the information retrieval community for researching and testing purposes, but gold standards specifically for the medical domain are rare. What this paper suggests is to design and develop an open source gold standard for use by the information retrieval community working in the medical domain. If such an open source standard were available, it would promote work in this important field, spur development of better information retrieval systems, and generally help to improve medical information systems. Also of importance would be the creation of a pool of tagged medical information “objects” that is freely accessible, and an application that makes creating and publishing these pools easy to perform. Medical objects can be free text, images, electrocardiograms, MRI scans, and other such data.

The paper is organized as follows: Section 2 gives an overview of some of the information retrieval systems currently in use in the medical domain. Following this, common methods and a overview of information retrieval evaluation is given. Next, the development of an open source, community driven, gold standard is discussed, and the advantages that such a system would offer to the information retrieval community are presented. The last section concludes the paper, and presents future research topics.

2 Background: Information Retrieval in the Medical Domain

Information retrieval (IR) systems are being increasingly advocated by medical professionals in order to enhance the quality of patient care and to provide better use of evidence [Hersh and Hickam, 1998]. Much research has been done in the past in

the area of information retrieval system evaluation [Robertson and Hancockbeaulieu, 1992], [Tange et al., 1998], [Brown and Sonksen, 2000] and significant progress has also been made in text mining techniques, in order to cope with the rapidly increasing information overload in the area of medical literature [Sullivan et al., 1999], [Hall and Walton, 2004]. Interestingly, developments in text mining techniques in the area of clinical information systems and medical documentation are rare [Noone et al., 1998], [Holzinger et al., 2007b], [Holzinger et al., 2008]. The application of sophisticated medical information systems amasses large amounts of medical documents, which must be reviewed, observed, and analyzed by human experts [Holzinger et al., 2007a]. All essential patient record documents contain at least a certain portion of data which has been entered in free-text fields and its computational evaluation has very long been the focus of research
[Gell et al., 1976], [Gell, 1983], [Zingmond and Lenert, 1993]. Although text can be created relatively simply by end-users, supporting automatic analysis is extremely difficult [Gregory et al., 1995], [Holzinger et al., 2000], [Lovis et al., 2000]. Often it occurs in practice that relevant relationships remain completely undiscovered, because relevant data are scattered and no investigator has linked them together manually [Smalheiser and Swanson, 1998].

3 Evaluation of Information Retrieval Systems

The standard procedure used to measure information retrieval effectiveness comprises of the following three elements [Harter and Hert, 1997], [Zeng et al., 2002]:

- A document collection.
- A test suite of information requests, expressible as queries.
- A set of judgments for each query-document pair, which defines each pair as either relevant or not relevant.

The common approach to information retrieval system evaluation is based on the exact notion of relevant and non-relevant documents. In the context of information retrieval, relevance describes how well a retrieved set of documents (or a single document) meets the information need of the user. In other words, with respect to a user information need, a document in the test collection is given a binary classification as either relevant or non-relevant [Bemmel and Musen, 1997]. This decision is referred to as a “gold standard” or “ground truth judgment of relevance”. The test collection should be a sample of the kinds of text that will be encountered in the operational setting of interest [Wingert, 1986], and its relevance is assessed according to an information need [Robertson and Hancockbeaulieu, 1992]. Standard textbooks on information retrieval [Baeza-Yates and Ribeiro-Neto, 2006] claim that, as a rule of thumb, fifty information needs is generally considered to be a sufficient minimum. An important aspect of this is that a query for an information retrieval tool is not the information need. Rather an information need can be expressed in terms of a query language for an information retrieval tool. For example, some standard test collections that are often used by information retrieval researchers are the different tracks from the Text Retrieval Conference and GOV2 (a very large web page collection). Once a test collection to be used as a basis to test an information retrieval

system has been chosen, a metric for the system comparison must be decided upon. Basically, this can be separated into two groups of pure statistical performance measures; namely, metrics for unranked retrieval results and ranked retrieval results. Classical information retrieval metrics that are widely used in literature are the Recall, Precision, Fallout, and F-Measure metrics.
A lot of effort has been invested into finding new evaluation measures over the past few years, one of the most famous recently introduced being bpref [Buckley and Ellen, 2004]. Other common information retrieval metrics that are concerned with ranked retrieval results are R-Precision, Precision at k, Mean Average Precision (MAP), and Normalized Discounted Cumulative Gain (NCDG). A good explanation and overview of these and other current information retrieval metrics can be found in standard text books. However, other factors do exist that should also be considered when evaluating an information retrieval system. [Saracevic, 1995] identified six different levels of information retrieval evaluation:

- Engineering level
- Input level
- Processing level
- Output level
- Use and user level
- Social level

The human factor, and human information behaviour, in the context of information retrieval systems are especially important factors to consider when developing such systems. Such systems must be capable of satisfying user needs. Although getting the right answers according to an information need is one of the most important parts of an information retrieval system, human factors should also be considered [Lew et al., 2006].

4 Open Source Ground Truths in the Medical Domain

After giving an introduction to information retrieval and a brief overview about information retrieval evaluation this section will concentrate on the main content of the paper. Therefore, the reasons as to why it is thought that open ground truths might be useful are given. Then some considerations that should be made when creating such a resource are outlined. Finally, a number of ideas are presented regarding the usefulness of gold standards in the context of teaching medicine are provided.

4.1 Motivation and Challenges

The motivation for this paper can be derived from work carried out during a former project where the performance difference between a semantic-based information retrieval tool and a human expert was examined. To do this, information needs in the medical domain were translated into a query language and were compared to the results achieved using the Recall, Precision, Fallout, and F-Measure evaluation metrics.

Because no ground truth was available that contained medical diagnosis free texts in any medical area, it was necessary to create a new one. This resulted in a pool of 3542 annotated diagnosis texts in the field of pathology inflammation that were used as a basis for a web-based evaluation framework. The framework was designed to be as flexible and extensible as possible, offering the ability to add other pools for testing at a later date. However, there are currently no such text pools available in the medical domain that could be used for information retrieval development or testing.

Text-based information retrieval is especially challenging in the medical domain, as the following example text will attempt to illustrate:

MITTELGRADIGE CHRONISCHE GASTRITS (MAGENMUCOSA VOM CORPUSTYP, UEBERGANGSTYP) MIT MITTELGRADIGER AKTIVITAET, KOMPLETTER UND INKOMPLETTER (TYP III) INTESTINALER METAPLASIE, MITTELGRADIGER ATROPHIE DER TIEFEN DRUESEN. ANTEIL EINES TUBULAEREN MAGENSCHLEIMHAUTADENOMS (INTESTINALER TYP; MITTELGRADIGE DYSPLASIE; WHO: GERINGGRADIGE INTRAEPITHELIALE NEOPLASIE). HP NICHT NACHWEISBAR.

The above text epitomizes the types of challenges that are inherent in medical free text analysis:

- The text resembles a memo more than an orthographically correct piece of text. From a doctor's point of view, however, the semantic of the text takes precedence over proper grammar or correct spelling.
- The text contains much domain specific knowledge and words that are would only be properly understood by an expert.
- Medical texts often contain abbreviations. In the text above, WHO stands for World Health Organization (which has a classification scheme for diseases), and the doctor is documenting the classification as GERINGGRADIGE INTRAEPITHELIALE NEOPLASIE. Even more frequent are abbreviations that are only resolvable when the context is known. HP in the context of gastritis stands for helicobacter-pylorii, but is often also used as an abbreviation for haptoglobin.
- Typing errors can further complicate matters when analysing text, as is illustrated by the misspelling of the word gastritis as GASTRITS.

Gold standards can also consist of images as well as text, as medical images can also be tagged with information regarding a particular aspect of a medical image.
It is the opinion of the authors, that a system should be developed that would enable creators of gold standards to make their work available to the community, using an online, web-based application. This application would allow registered users to create gold standard pools, which would contain the medical text, images, or other medical objects (such as electrocardiograms, sonograms, etc.), as well as their annotations. These objects could be used by other research groups, or extended with further annotations to create a more useful gold standard. Each pool could exist in a track,

which would separate the gold standards into their relevant areas of medicine, and would also enable gold standards to be branched if a group decided to move the gold standard into a new direction. The aim of such a system would be that different sets from different medical areas could be made, accessed, and published in a controlled way, while at the same time providing a guaranteed level of quality for the annotations. Furthermore, building gold standards is a time consuming process, and a collaborative system could alleviate some of the aspects of creating gold standards that make it too difficult for a small team to undertake. This proposal is work in progress and will be described in detail in a future paper, where a requirements analysis and framework architecture will be presented.

As mentioned previously, properly annotated medical texts are not freely available, and having access to medical texts form various areas of medicine would encourage the information retrieval community in the following ways:

- It could animate the community into working on these types of text, as they pose difficult challenges.
- Having access to more than one ground truth in a particular domain and in different languages would also make it possible to test and evaluate multi-language information retrieval.
- Because, in the experience of the authors, only semantic-based information retrieval tools tend to have the ability to solve retrieval tasks covering different areas of medicine, it would encourage the community to produce new information retrieval metrics that would take the semantic based nature of these engines into account.

A further challenge is patient privacy. Privacy laws are very strict with regards to patient information, and any gold standard data must be anonymized by the institution before adding it to the gold standard pool. Any institution that uses data that could potentially identify a patient, or data that contains personal health information, would first have to obtain patient authorization. Because applicable laws vary largely from country to country, the topic of patient privacy is large and complex and will be discussed in detail when a formal definition of the collaboration framework is made. Data anonymization is an active field of research, where the focus is on developing scientific methods for making it possible for a data holder to release a version of its private data that guarantees that the individuals who are the subjects of the data cannot be re-identified, while at the same time ensuring that the data remains practically useful [Sweeny, 2002]. A well known concept with respect to data anonymization is k-ANONYMITY [Sweeny, 2002], [Sweeny, 1997], where the level of privacy protection is directly proportional to the size of the anonymized data in which the data set resides. An oft cited study [Sweeny, 2000] showed that 87% (216 million of 248 million) of the population of the United States are identifiable based on only their 5-digit ZIP code, gender, and date of birth. Therefore, the application of best practice methods to protect privacy, while maintaining useful scientific data, is a topic that must be given much thought when developing this framework.

4.2 Ensuring Quality in a Community Driven Gold Standard

If a tool was created that would make available a web-based service to help content creators write gold standards, the single most important factor is to ensure the quality of the annotations made to the medical objects. This could be achieved in a number of ways, such as including a clause that forces any annotations to be reviewed by at least one external medical expert. This is especially true in medicine, as domain expertise knowledge is required to guarantee the quality of an annotated pool. Gold standards in medicine that adhere to a certain level of quality have been discussed in previous work [Geierhofer and Holzinger, 2007], and is one of the most important features that must be considered when wanting to develop a successful annotation framework. A common measure of the agreement between judges is the so-called kappa statistic. As a rule of thumb, a value of 0.8 and above is considered as a good agreement [Manning et al., 2008]. Problems that have to be considered when using this statistics are described in [Geierhofer and Holzinger, 2007].

Another method to help increase quality would be to ensure that all objects are tagged using terms for a single terminology such as the SNOMED Clinical Terms. This would also help when searching is performed across different gold standard pools, and could make it possible to merge similar pools. For example, radiological images that have all been tagged using the same terminology database could be combined to create collections of annotated images.
This is highlighted in the following scenario: An annotator views a medical image and notices inflammation of the linings of the arteries, and tags the document using the SNOMED concept of "Arteritis (disorder)". This is defined in the SNOMED database as having the SNOMED ID of D3-81600, which corresponds to the ICD-9 code of 447.6 – "Arteritis, unspecified". Therefore, all images tagged with the ICD-9 code of 447.6 or the SNOMED ID of D3-81600 would appear in a search for the images that contain this characteristic. By limiting the tagging of objects to a defined terminology, errors in spelling, definition, and terminology are avoided. The same can be applied to any text tagged by the annotator. For example, if the annotator must tag a pathological text with the organ, this organ must be retrieved from SNOMED terminology database to ensure that other texts regarding the same organ appear when searching for this organ. Take, for example, the following text:

Microscopic sections reveal an infiltrating ductal carcinoma with pleomorphic glands infiltrating through the surrounding stroma and fat eliciting a desmoplastic stromal host response. There is good tubular differentiation and the glands are lined by cells showing a mild to moderate degree of atypia characterized by hyperchromatic nuclei with prominent nucleoli. [1]

An annotator might be required to tag the text object with the organ to which the pathology report belongs. In this case, the text refers to a breast tissue pathological examination and the annotator would therefore select "Breast anatomy (body structure)" which corresponds to the SNOMED ID of T-D004F. Therefore, this text could appear in a search for "breast", despite the word not appearing in the text itself.

[1] The Doctor's Doctor: http://www.thedoctorsdoctor.com/pathreports/typical_report.htm

4.3 Teaching Context

Education would stand to benefit a great deal from having more gold standards available to research and teaching institutes. In many different fields of medicine, such as radiology, teaching depends to a large degree on having material available that is annotated and reviewed by experts. Having quick and easy access to annotated medical images, for example, makes it easier and less time consuming to prepare material for teaching medical students how to recognize aspects of radiological images. Annotated radiological images would also be useful for examining students, as teachers can prepare exam material for students using the pool of annotated images, and compare students' answers to the images' peer reviewed annotations.

The advantages for education are not limited to just images, however. Annotated texts would also be useful for teaching students how to interpret medical diagnosis texts, while annotated electrocardiograms, sonograms, and magnetic resonance images could all be used for both teaching and training.

Part of the proposed architecture would be a web-based application that would use the pool of annotated objects to help educators create training material and examinations. While web-based pools of annotated images do exist, such as MyPACS (http://www.mypacs.net), these systems do not allow users to create examinations or teaching material for students, or allow students to use the system to take examinations. Also, by limiting the students' answers to the same terminology used to annotate the gold standard images, the students could be graded in an automatic or semi-automatic way. Therefore the proposed architecture would compliment the teaching and educational process in a more encompassing manner. With 8,913 hospitals around the world using the MyPACS system, it is obvious that there is a real need for such services.

As mentioned previously, more momentum is required to make institutions aware of the advantages of making their gold standards openly available, and to make these institutions more interested in creating gold standards in the first place. As Tim Berners-Lee has recently been advocating, governments, institutions and researchers must make available "raw data now". Making data openly available is beneficial to the entire community of developers and researchers who rely on such data for their scientific work.

5 Conclusion and Future Work

The main conclusion of this paper is that there is a lack of freely available annotated medical texts for use in research, such as in information retrieval research, and it is proposed here that this is the case because no simple and effective framework or application exists that makes this easy to do. For information retrieval, having access to good quality gold standard texts is an essential part of being able to perform analysis in this area. However, the amount of openly available gold standard texts is severely lacking in medicine, despite the fact that many institutions the world over must have created and used gold standards in their research, and that these gold standards do exist.

Therefore, this paper proposed a preliminary web-based framework for the collaborative development of gold standards. A comprehensive outline and proposal

for this framework is the subject of further research, and a paper will be published regarding the design and development of this framework. The framework would help teaching and learning in medicine as properly annotated medical images and texts could be used to teach and examine students, and be used to train doctors. The teaching and training layer of the framework will also be discussed more thoroughly in future work, including a comprehensive review of what aspects of teaching could be aided by such a system, and how a system could facilitate the examining of students in an semi-automated way.

Such a framework, the authors believe, would in fact make the process of creating gold standards quicker, as a community would be involved in its development rather than just an individual institution. Quality assurance, and how to maintain high levels of quality in a community driven gold standard, will also be discussed in detail in a future work. This may even shift the workflow from an institute based method of publishing annotated data, to a community contributed workflow involving institutes from around the world.

As well as this, the paper gave an overview of the theory behind information retrieval and outlined a number of reasons why information retrieval in medicine is a challenging and interesting area to be involved in.

References

[Baeza-Yates and Ribeiro-Neto, 2006] Baeza-Yates, R. and Ribeiro-Neto, B.: "Modern information retrieval", ACM Press, New York, (2006).

[Bemmel and Musen, 1997] Bemmel, J. H. v. and Musen, M. A.: "Handbook of Medical Informatics", Springer, Heidelberg, (1997).

[Brown and Sonksen, 2000] Brown, P. J. B. and Sonksen, P.: "Evaluation of the quality of information retrieval of clinical findings from a computerized patient database using a semantic terminological model", Journal of the American Medical Informatics Association, 7, (2000), 392-403.

[Buckley and Ellen, 2004] Buckley, C. and Ellen, M. V.: "Retrieval evaluation with incomplete information", Proceedings of the 27th annual international ACM SIGIR conference on Research and development in information retrieval, (2004), 25-32.

[Geierhofer and Holzinger, 2007] Geierhofer, R. and Holzinger, A.: "Creating an Annotated set of Medical Reports to Evaluate Information Retrieval Techniques", SEMANTICS, (2007).

[Gell, 1983] Gell, G.: "AURA - routine documentation of medical text", Methods of Information In Medicine, 22, (1983), 63-68.

[Gell et al., 1976] Gell, G., Oser, W. and Schwarz, G.: "Experiences with the AURA Free Text System", Radiology, 119, (1976), 105-109.

[Gregory et al., 1995] Gregory, J., Mattison, J. E. and Linde, C.: "Naming Notes - Transitions from Free-Text to Structured Entry", Methods of Information in Medicine, 34, (1995), 57-67.

[Hall and Walton, 2004] Hall, A. and Walton, G.: "Information overload within the health care system: a literature review", Health Information and Libraries Journal, 21, (2004), 102-108.

[Harter and Hert, 1997] Harter, S. P. and Hert, C. A.: "Evaluation of information retrieval systems: Approaches, issues, and methods", Annual Review of Information Science and Technology, 32, (1997), 3-94.

[Hersh and Hickam, 1998] Hersh, W. R. and Hickam, D. H.: "How well do physicians use electronic information retrieval systems? A framework for investigation and systematic review", Jama-Journal of the American Medical Association, 280, (1998), 1347-1352.

[Holzinger et al., 2007a] Holzinger, A., Geierhofer, R. and Errath, M.: "Semantic Information in Medical Information Systems - from Data and Information to Knowledge: Facing Information Overload", Procedings of I-MEDIA '07 and I-SEMANTICS '07, Graz, 2007a, 323-330.

[Holzinger et al., 2007b] Holzinger, A., Geierhofer, R. and Errath, M.: "Semantische Informationsextraktion in medizinischen Informationssystemen", Informatik Spektrum, 30, (2007b), 69-78.

[Holzinger et al., 2008] Holzinger, A., Geierhofer, R., Modritscher, F. and Tatzl, R.: "Semantic Information in Medical Information Systems: Utilization of Text Mining Techniques to Analyze Medical Diagnoses", Journal of Universal Computer Science, 14, (2008), 3781-3795.

[Holzinger et al., 2000] Holzinger, A., Kainz, A., Gell, G., Brunold, M. and Maurer, H.: "Interactive Computer Assisted Formulation of Retrieval Requests for a Medical Information System using an Intelligent Tutoring System", ED-MEDIA 2000, Montreal, (2000), 431-436.

[Kaiser et al., 2007] Kaiser, K., Akkaya, C. and Miksch, S.: "How can information extraction ease formalizing treatment processes in clinical practice guidelines? A method and its evaluation", Artificial Intelligence in Medicine, 39, (2007), 151-163.

[Lew et al., 2006] Lew, M. S., Sebe, N., Djeraba, C. and Jain, R.: "Content-based multimedia information retrieval: State of the art and challenges", ACM Transactions on Multimedia Computing Communications and Applications, 2, (2006), 1-19.

[Lovis et al., 2000] Lovis, C., Baud, R. H. and Planche, P.: "Power of expression in the electronic patient record: structured data or narrative text?" International Journal of Medical Informatics, 58, (2000), 101-110.

[Manning et al., 2008] Manning, C. D., Raghavan, P. and Schuetze, H.: "Introduction to Information Retrieval", Cambridge University Press New York, NY, USA, (2008).

[Noone et al., 1998] Noone, J., Warren, J. and Brittain, M.: "Information overload: opportunities and challenges for the GP's desktop", Medinfo, 9, (1998), 1287-1291.

[Robertson and Hancockbeaulieu, 1992] Robertson, S. E. and Hancockbeaulieu, M. M.: "On the Evaluation of Ir Systems", Information Processing & Management, 28, (1992), 457-466.

[Saracevic, 1995] Saracevic, T.: "Evaluation of evaluation in information retrieval", Proceedings of the 18th annual international ACM SIGIR conference on Research and development in information retrieval, (1995), 138-146.

[Smalheiser and Swanson, 1998] Smalheiser, N. R. and Swanson, D. R.: "Using ARROWSMITH: a computer-assisted approach to formulating and assessing scientific hypotheses", Computer Methods and Programs in Biomedicine, 57, (1998), 149-153.

[Sullivan et al., 1999] Sullivan, F., Gardner, M. and Van Rijsbergen, K.: "An information retrieval service to support clinical decision-making at the point of care", British Journal of General Practice, 49, (1999), 1003-1007.

[Sweeny, 1997] Sweeny, L.: "Guaranteeing anonymity when sharing medical data, the datafly system", Proceedings of Journal of the American Medical Informatics Association, Washington, DC: Hanley & Belfus, Inc., (1997).

[Sweeny, 2000] Sweeny, L.: "Uniqueness of simple demographics in the US population", Carnegie Mellon University, Laboratory for International Data Privacy, Pittsburg, PA, (2000).

[Sweeny, 2002] Sweeny, L.: "k-anonymity: a model for protecting privacy", International Journal on Uncertainty, Fuzziness and Knowledge-based Systems, 10 (5), (2002), 557-570.

[Sweeny, 2002] Sweeny, L.: "Achieving k-anonymity privacy protection using generalization and suppression", International Journal on Uncertainty, Fuzziness and Knowledge-based Systems, 10 (5), (2002), 571-588.

[Tange et al., 1998] Tange, H. J., Schouten, H. C., Kester, A. D. M. and Hasman, A.: "The granularity of medical narratives and its effect on the speed and completeness of information retrieval", Journal of the American Medical Informatics Association, 5, (1998), 571-582.

[Wingert, 1986] Wingert, F.: "An Indexing System for Snomed", Methods of Information in Medicine, 25, (1986), 22-30.

[Zeng et al., 2002] Zeng, Q., Cimino, J. J. and Zou, K. H.: "Providing concept-oriented views for clinical data using a knowledge-based system: An evaluation", Journal of the American Medical Informatics Association, 9, (2002), 294-305.

[Zingmond and Lenert, 1993] Zingmond, D. and Lenert, L. A.: "Monitoring Free-Text data using medical language processing", Computers and Biomedical Research, 26, (1993), 467-481.