

Towards Improving Search Results for Medical Experts and Laypersons

Karin Friberg Heppin and Anni Järvelin

University of Gothenburg, Språkbanken, Department of Swedish, Box 200, 405 30 Gothenburg, Sweden

karin.friberg.heppin@svenska.gu.se,
anni.jarvelin@svenska.gu.se

Abstract. In a domain such as medicine, it is important that individuals' information needs are met with information on a suitable level of difficulty and expertise. This paper focuses on facilitating medical information access through reformulating queries and re-ranking result lists utilizing features typical for the language written for professionals or for laypersons. The aim is to produce result lists where the ranking is better suited for the expertise level of the user. We will explore the possibility of using features such as trigger phrases for query reformulation and document length, average word length or compound ratio for re-ranking.

The Swedish medical IR test collection, MedEval, from Språkbanken, University of Gothenburg, will be used to find features specific for professional language and lay language and to study the effectiveness of these features in reformulating queries and re-ranking search results based on the target group. The test collection contains 42,250 documents from the medical corpus MedLex, collected from all types of written medical information found in electronic format, except patient records. The collection contains 62 topics. In total, 7,044 documents have been assessed both for relevance to these topics and for target group.

Our experiments will be based on earlier explorative studies on medical expert and lay language where some features were identified. It was found that documents written for professionals tended to have more tokens per document, longer words, and more compounds than lay documents

The assessed documents were run through a perl program which counted the frequencies of occurring multiword units (MWUs). The frequencies of individual MWUs were much higher in the expert documents. For the doctors, medical phrases dominate the MWUs while the patients' documents mostly contained general language units. The most frequent patient MWUs from the medical domain, were more frequent in the doctor documents and could therefore not be said to be typical for patient documents. Many frequently recurring MWUs are not specific for any topic. However, they may be seen as trigger phrases indicating target group. Such trigger phrases will be used in the reformulation of the queries.

We believe that the phrases typical for a user group can be used to reformulate queries and that the likewise typical features can be useful for calculating

target group scores for medical documents and further for improving the ranking of the documents to better match the expertise of the user.

Keywords: Information Retrieval; Medical User Groups; Query Reformulation; Reranking

1 Introduction

In a domain such as medicine, it is important that individuals' information needs are met with information on a suitable level of difficulty and expertise. Laypersons and experts might need information, about the same topics, but not in the same form. Search engines do a fairly good job of finding documents to satisfy the range of information needs people may have for a query, but do less well in discerning individuals' search goals.¹

This paper focuses on the need to facilitate medical information access through reformulating queries and re-ranking result lists utilizing features typical for the language of the documents for professionals or for laypersons. The aim is to produce result lists where the ranking is better suited for the expertise level of the user. We will explore the possibility of using features such as trigger phrases for query reformulation and document length, average word length or compound ratio for re-ranking.

2 Materials and Methods

The Swedish medical IR test collection, MedEval, from Språkbanken, University of Gothenburg will be used to find features specific for professional language and lay language and to study the effectiveness of these features in reformulating queries and re-ranking search results based on the target group.²

The test collection contains 42,250 documents from the medical corpus MedLex,³ collected from scientific journals, health care communicators, newspapers, patient FAQs etc., in short, all types of medical information found in electronic format, except patient records. The collection contains 62 topics. In total, 7,044 documents have been assessed both for relevance to these topics and for target group: medical professionals or laypersons.⁴ 3,272 documents were assessed to have medical professionals as target group; 4,334 documents to have laypersons as target group; and 562 of these were assigned to both categories. For a document to be assigned to both categories it must have appeared in the pool of more than one topic and further have been assigned different target groups for the different topics.

In addition to the assessed documents, the target group of the documents can in many cases be reliably identified from the document source: scientific journals are written for medical professionals, while health care counselling web sites and discussion forums are typically for laypersons. Thus more documents can be added to the sets of documents used for extracting the target group features.

To test the reliability and effectiveness of target group features in re-ranking results for the correct target group, a set of 20 topics from the MedEval test collection will be

used. The effect of the re-ranking based on the target group features will be tested for both target groups using both medical and common vocabulary in the queries.

3 Preliminary Results and Discussion

Our experiments will be based on the explorative studies on medical expert and lay language described in Friberg Heppin⁴ where some promising features were identified. It was found that documents written for professionals tended to have more tokens per document, longer words, and more compounds than lay documents (see Table 1).

The assessed documents were run separately through MWT (multi-word term), a perl program which counts the frequencies of occurring multiword units using both lexical and statistical calculation.⁵ The version used here was later enhanced and provided to us by Jussi Karlgren (personal communication).

Table 1. Features found in the MedEval test collection and in the documents assessed to have target groups doctors or patients

	Entire collection	Assessed documents	Doctors	Patients
Tokens per document	307	715	988	561
Average word length	5.75	6.04	6.29	5.73
Ratio of compounds	0.098	0.114	0.128	0.098

Table 2. Potential doctor trigger phrases

Frequency	Swedish MWU	English equivalent
844	hos patienter med	in patients with
500	ökad risk för	increased risk of
459	av patienter med	of patients with
440	för behandling av	for treatment of
410	kan leda till	can lead to

Table 3. Potential patient trigger phrases

Frequency	Swedish MWU	English equivalent
316	läs mer om	read more about
218	ställa din fråga	put your question
94	om du har	if you have
70	när behandlingen är	when the treatment is

There were differences in both types and frequencies of multiword units (MWUs) found in the documents for the two target groups. The frequencies of individual MWUs were much higher in the doctor documents. For the experts, medical phrases dominate the MWUs while the patients' documents contain general language units. The most frequent patient MWU which could be said to be medical, *att drabbas av* 'to be afflicted by', has no more than 95 occurrences in the patient documents while it has 255 occurrences in the doctor documents. Another of the most frequent patient phrases *när det gäller* 'when it comes to', is also more common in the doctor documents, 574 for doctors and 102 for patients. Thus these cannot be said to be MWUs typical for the patient documents.

Many frequently recurring MWUs are not specific for any topic. However, some of them may be seen as trigger phrases indicating target group. Potential trigger phrases for doctors and patients can be seen in Table 2 and 3. Such trigger phrases will be used in the reformulation of the queries, for example by adding or reducing ranking scores of documents where the phrases are present.

4 Conclusion

Our preliminary studies have identified a few target group specific features and trigger phrases in professional and lay medical documents. These features and phrases were extracted from a relatively small number of documents and could be improved by using a larger number of documents. We believe that these phrases can be used to reformulate queries and that these features can be useful for calculating target group scores for medical documents and further for improving the ranking of the documents to better match the expertise level of the user.

Acknowledgements

We would like to thank the Centre of Language Technology at the University of Gothenburg for providing funding for this research.

References

1. Jaime Teevan, Susan T. Dumais, and Eric Horvitz. 2010. *Potential for personalization*. ACM Trans. Comput.-Hum. Interact., 17(1):4:1–4:31, April.
2. Karin Friberg Heppin. 2011. MedEval – A Swedish Medical Test Collection with Doctors and Patients User Groups. *Journal of Biomedical Semantics*, 2(Suppl 3):S4.
3. Dimitrios Kokkinakis. 2004. *MEDLEX: Technical report*. Technical report, Department of Swedish, University of Gothenburg.
4. Karin Friberg Heppin. 2010. *Resolving Power of Search Keys in MedEval a Swedish Medical Test Collection with User Groups: Doctors and Patients*. Ph.D. thesis, University of Gothenburg
5. John S. Justeson and Slava M. Katz. 1995. Technical terminology: Some linguistic properties and an algorithm for identification in text. *Natural Language Engineering*, 1(1):9–27.