Development and Practical Use of a Medical Vocabulary-Thesaurus-Dictionary for Patient Empowerment

Marco Alfano
Anghelos Centro Studi sulla Comunicazione
Palermo, Italy
Dipartimento di Matematica e Informatica
Università di Palermo, Italy
marco.alfano@anghelos.org

Biagio Lenzitti
Dipartimento di Matematica e Informatica
Università di Palermo, Italy
biagio.lenzitti@unipa.it

Giosuè Lo Bosco
Dipartimento di Matematica e Informatica
Università di Palermo, Italy
Dipartimento SIT, Istituto Euro-Mediterraneo di Scienza e Tecnologia, Palermo, Italy
giosue.lobosco@unipa.it

Davide Taibi
Istituto per le Tecnologie Didattiche
Consiglio Nazionale delle Ricerche
Palermo, Italy
davide.taibi@i td.cnr.it

ABSTRACT

Health empowerment can be obtained through an informative and educational intervention to increase one's ability to think critically and act autonomously. Medical texts are usually written by professionals and can be difficulty understood by non experts who do not have the same skills and vocabularies. Thus, it would be desirable to have an online medical vocabulary-thesaurus-dictionary that can help a non expert to easily find the consumer equivalent of medical (technical) terms and additional consumer information. To this end, we have developed an online multilingual medical vocabulary-thesaurus-dictionary by interconnecting different online sources, i.e., medical vocabularies to create a list of technical terms, consumer health vocabularies (CHVs) for translating technical terms into their consumer equivalents and consumer dictionaries for finding explanations of the terms. In addition, we have built an online editor that allows to add new medical terms (with the related consumer information) and modify existing consumer terms and definitions. Furthermore, we have built some practical applications, on top of the medical vocabulary-thesaurus-dictionary, in order to facilitate the empowerment of patients, or non-experts in general. The applications are located at the data, information and knowledge levels of the 'knowledge pyramid' that, in our case, contains the empowerment at the top level.

CCS CONCEPTS

• Information systems → Data management systems • Information systems → Information systems applications → Collaborative and social computing systems and tools • Applied computing → Life and medical sciences → Health informatics • Applied computing → Education → E-learning

KEYWORDS

E-Health, Patient Empowerment, Plain Language, Medical Vocabulary, Medical Dictionary, Consumer Health Vocabulary.

1 Introduction

An empowered patient is a person who is aware of his/her health conditions and collaborates with health professionals for making the best health decisions concerning his/her health [1]. Health empowerment can be obtained, among others, through an informative and educational intervention to increase one's ability to think critically and act autonomously.

Medical information is usually conveyed in the form of medical reports, medicine leaflets, scientific papers and so on. This information is usually written by professionals (physicians, medical researchers, etc.) who use their own language and communication style but it is often read by patients, or non-experts in general, who do not have the same skills and vocabularies of the experts and can have difficulties in text comprehension [2], [3]. A language that can be understood by almost anyone, made of common everyday terms, should be possibly used in writing medical content that is also read by non-experts. Of course, it is not always easy for the medical experts to write in plain language because of the risk of losing precision and the time, often required, to simplify concepts.

For a better understanding of medical texts, a user may need some external help (online or not) such as vocabularies, dictionaries and thesauri [4]. The user will usually look at those resources in an unestablished order to find the technical terms, their synonyms and additional information so to be able to understand the whole text. However, this approach assumes that the additional resources are readily available and this is not always the case. Moreover, it is a disorganized process, time consuming, and can lead to dispersion and, ultimately, to an information overload that may create more confusion.

To help a generic user in understanding a medical text, it would be desirable to have a single resource, such as an online medical vocabulary-thesaurus-dictionary, where the user can easily find the consumer equivalent of any medical (technical) term and additional information in a simple language. To this end, we have developed an online multilingual medical vocabulary-thesaurus-dictionary by interconnecting different online sources, i.e., medical vocabularies to create a list of medical (technical) terms, consumer health vocabularies (CHVs) for translating technical terms into their

consumer equivalents and consumer dictionaries for finding explanations of the terms. It contains both English and Italian resources but other languages can easily be added if the related vocabularies, thesauri and dictionaries are available. In addition, to control and modify the information gathered automatically and contained in the medical vocabulary-thesaurus-dictionary, we have built an online editor that allows to add new medical terms (with the related consumer information) and modify both consumer terms and definitions of existing terms.

As a final step, we have built some practical applications on top of the medical vocabulary-thesaurus-dictionary in order to facilitate the empowerment of patients and non-experts, in general, through data, information and knowledge by loosely following the different levels of the 'knowledge pyramid' [5] that, in our case, contains the empowerment at the top level as a representation of wisdom. In particular, for the data level, we have built an online tool that, for each medical term, provides the consumer equivalent and explanation and gives the possibility to suggest modifications. It can also be used to collaboratively improving the information provided by the medical vocabulary-thesaurus-dictionary. For the information level, we have developed a tool that accepts, as input, any medical text and highlights its medical (technical) terms complementing them with the consumer terms and further explanations. Finally, for the knowledge level, we have developed a medical quiz environment for evaluating and improving knowledge of medical terms.

There are other systems in the literature that try to simplify medical texts for non experts [6], [7], [8], [9], [10]. Our system, to the best of our knowledge, is unique since it connects the automatic aspects of creating the medical vocabulary-thesaurus-dictionary and processing the medical texts with the human intervention for checking and improving the consumer terms and explanations. Moreover, it presents great flexibility to develop different tools aimed to facilitate patient empowerment, as described above.

The paper is organized as follows. The second section describes the details of the multilingual medical vocabulary-thesaurus-dictionary we have developed. The third section presents the practical uses of the vocabulary-thesaurus-dictionary that follow the knowledge pyramid. The final section presents some conclusions and future work.

2 A Multilingual Medical Vocabulary-Theasurus-Dictionary

As said in the Introduction, when a generic user reads a medical text written by a professional he/she has often difficulties in understanding it because of the used technical language. He/she may then need some external help to understand the technical terms, find familiar synonyms and get additional information. This external help comes in the form of different resources (online or not) such as vocabularies, dictionaries and thesauri:

1. A medical vocabulary is a selective list of words and phrases used in the medical field. It is usually created for professionals and, for this reason, it contains all the technical terms used in the medical field. Thus, it can be used to find the technical

terms in a medical text. One of the most comprehensive medical vocabulary is the 'Unified Medical Language System (UMLS)', a large collection of multilingual vocabularies created and maintained by the US National Library of Medicine [11]. It mainly uses a 'Concept Unique Identifier – CUI' (a unique identifier for each concept) to create a mapping among these vocabularies and thus allows translation among the various terminology systems. It may also be viewed as a comprehensive thesaurus and ontology of health and biomedical concepts.

2. A medical-consumer thesaurus contains synonyms and antonyms of medical terms. It can be used to find consumer synonyms of the technical terms. Consumer health vocabularies (CHVs) have been created for translating medical terms and concepts in their equivalent for consumers and they can be very useful for translating medical (technical) terms into consumer ones [12]. One of the best known examples of CHV is the 'Open Access Collaboratory Consumer Health Vocabulary (OAC-CHV)', created and maintained by the Consumer Health Vocabulary Initiative [13]. It is a relationship file that links commonly used terms to associated medical terminology represented by the UMLS. The OAC-CHV contains different fields among which:
    - 'Term': The term as found in the text;
    - 'Concept Unique Identifier' (CUI): The unique identifier of a concept as found in the UMLS;
    - 'CHV Preferred Name': The preferred consumer term as defined in the Consumer Health Vocabulary.

3. A health-consumer dictionary gives information about the meaning of the terms in a consumer language. It can be used to find additional information on the technical terms. There are many sites specifically created for health consumers that use a language that can easily be understood by them. Examples of online dictionaries in English are WebMD [14] and MedlinePlus [15] and online dictionaries in Italian are Ok Salute [16] and Dizionario della Salute [17].

Of course, there are numerous resources of each type and it is worth to mention that there is not a clear distinction among these categories of resources. In fact, more information can be concentrated in a single resource, such as technical terms, synonyms and definitions that can be found in a single vocabulary.

We have developed an online multilingual medical vocabulary-thesaurus-dictionary by integrating different online resources belonging to the categories described above. To this end, we have built a database by using medical vocabularies to create a list of medical (technical) terms, consumer health vocabularies (CHVs) for translating the technical terms into their consumer equivalents and medical consumer dictionaries for finding supplementary information on the terms. The database contains both English and Italian resources so that a translation of terms between the two languages can also be automatically performed. The connection among technical terms, consumer equivalents and consumer definitions, as implemented in our database, is shown in Fig. 1. Presently, the database contains around 125.000 technical terms and 60.000 consumer terms and/or definitions.
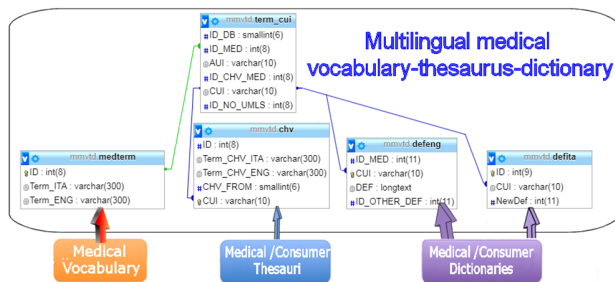
Figure 1: Data connections in the medical vocabulary-thesaurus-dictionary

The initial use of the developed vocabulary-thesaurus-dictionary provided us with satisfying results but a human intervention was required in order to improve the obtained results for consumer terms and explanations. This is a consequence of the limited availability of consumer terms and definitions (compared to the medical terms) and of the non-correct consumer equivalent of a medical term (mainly for the Italian language, since it is a translation of the English OAC CHV term). To this end, we have built an online editing tool that allows to add new terms (and related information) and to modify existing consumer terms and explanations. It allows searching for terms by indicating the medical name, its consumer equivalent or the related CUI. If the term is found, the consumer equivalents and definitions (both in English and in Italian) are shown and can be edited. If the term is missing, the tool allows to insert it together with the related consumer terms and definitions. Note that the use of this tool is only allowed to the management team of the medical vocabulary-thesaurus-dictionary through an authentication phase. This is important in order to guarantee the integrity of the information stored in the database.

3 Practical Uses of the Medical Vocabulary-Theasurus-Dictionary

The knowledge pyramid represents the relationship among data, information, knowledge, and wisdom [5]. The applications we developed, built on top of the medical vocabulary-thesaurus-dictionary, loosely reflects the three levels of data, information and knowledge of the pyramid and are designed to guide users towards the highest level of the pyramid that, in our vision, is his/her empowerment in the medical domain (Fig. 2).
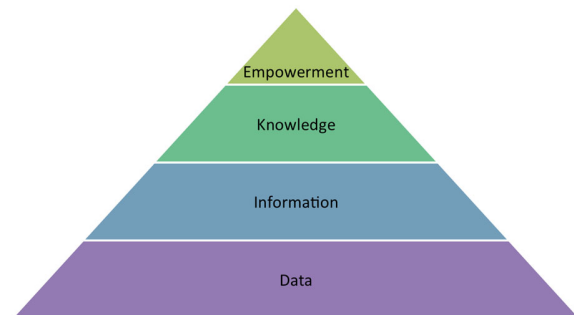
Figure 2: Knowledge pyramid with empowerment at the top level

With regards to the data level, we have developed an application that provides users with consumer information about single medical terms and receives suggestions from users to update the stored data. For the information level, we present an application that supports users in understanding a medical text by providing consumer terms and explanations of all technical terms of the text. Finally, for the knowledge level, we present an application to deliver questionnaires specifically designed to improve users knowledge of medical information. All these applications contribute to foster user awareness on the medical domain thus reinforcing his/her empowerment. The three applications are described in details in the following subsections.

3.1 Collaborative M-Search

We have built a web application, accessible at http://www.math.unipa.it/simplehealth/med-vtd/, that allows any user to specify a medical term (both in Italian and in English) and provides the information connected to that term, i.e., the CUI, the consumer term and its explanation, both in Italian and English (Fig. 3). The application can also be used for translating technical or consumer terms between Italian and English or vice-versa. A non-authenticated user can propose changes to the medical vocabulary-thesaurus-dictionary database by suggesting new consumer terms or explanations through the proper forms (Fig. 3). The suggested modifications will be notified to the management team, which will decide whether to include the suggestions into the medical vocabulary-thesaurus-dictionary.

Figure 3: Collaborative M-Search used with the encephalon term.

In practice, as depicted in Fig. 4, the suggestions of non-authenticated users will be stored in a temporary database and the management team will be informed about the new submission.


Figure 4: Collaborative scheme for updating the medical vocabulary/thesaurus/dictionary.

If validated, the new data will be automatically moved to the permanent database by means of the editing tool. Note that such a connection realizes a collaborative improvement of the medical vocabulary-thesaurus-dictionary where anyone can propose new terms or changes to the existing ones.

3.2 Simple

We have implemented a system, available at http://www.math.unipa.it/simplehealth/simple/, that analyses medical texts both in English and Italian, highlights the technical terms and provides consumer translations of the terms together with supplementary consumer information [18]. Fig. 5 presents a medical report where the technical terms are highlighted after being processed through the system. Next to each technical term, there is one infobutton that, when pressed, shows a tooltip with the consumer translation of the term and some additional information. Notice that our system does not create any change in the original text by replacing the word or inserting an explanation in the text because, in our opinion, this could disorient the user. It only provides a translation and additional info, in a tooltip, on request, leaving the user fully in charge of his/her navigation path through the original text.
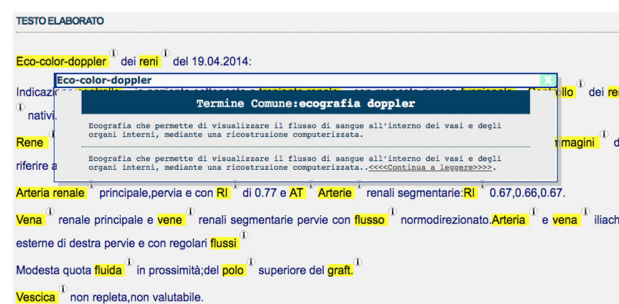
Figure 5: Medical report with highlighted technical terms and consumer information

3.3 S MedQuiz.

The structured information related to medical terms included into the medical vocabulary-thesaurus-dictionary has also been used for the development of an application, MedQuiz, that supports questionnaire construction in the medical domain. The usage of medical vocabulary-thesaurus-dictionary is particularly suitable to create multiple choice and matching questions. For the first type, respondents have to select the correct common term out of a multiple-choice list. For the second type, users have to associate a set of consumer terms to a set of medical terms. MedQuiz supports the construction of a questionnaire by selecting the type and number of questions together with the number of possible answers. The application randomly selects the information required to build the questionnaire from the medical vocabulary-thesaurus-dictionary. The list of created questions is shown to the instructor who can exclude one or more of them and substitute with new generated questions. Once the instructor confirms the list of questions, MedQuiz provides the question set in a format that can be imported by the most common quiz platforms. Fig. 6 shows a snapshot of the quizzes imported, as xls files, into the Savsoft quiz platform [19]. It can be accessed at http://www.consalute.it/savsoft/index.php/quiz/open_quiz/.
The questionnaire can be delivered in medical contexts (e.g., waiting rooms in hospitals in family-doctor clinics) to encourage

patient empowerment, or can be used in informational websites where content related to medical issues can be delivered along with questionnaires aimed at reinforcing the learned information. Moreover, MedQuiz can be used not only to support patient empowerment but also in an educational context, for example in language courses, where the mix of technical and familiar language is topical.

Figure 6 Match-the-column quiz

4 Conclusions and Future Work

In this paper we have presented an online multilingual medical vocabulary-thesaurus-dictionary that facilitates empowerment of patients, or non-experts in general, through some practical applications built on top of it, loosely following the data, information and knowledge levels of the 'knowledge pyramid'.

The preliminary use of the medical vocabulary-thesaurus-dictionary and the related applications have provided us with satisfying results but, as explained above, some improvement were needed mainly in terms of consumer terms and consumer definitions. To this end we have realized a collaborative environment with the editing tool and we plan to advertise the use of the medical vocabulary-thesaurus-dictionary in the social media and ask for public collaboration. Moreover, we are in the process of performing a thorough test session on the Simple application with a community made up of physicians and patients so to have it evaluated by both experts and non-experts. Finally, with respect to MedQuiz, we plan to directly deliver the questionnaire to the respondents without using an external quiz platform so that data related to answers can be collected inside the application for subsequent analysis.

As a future step, the medical vocabulary-thesaurus-dictionary will be enhanced by adding new medical vocabularies, thesauri and dictionaries. Other languages will also be added in order to create an European multilingual vocabulary-thesaurus-dictionary similar to the one presented in [20] but with a number of terms one hundredfold and the addition of the terms definitions. Moreover, we plan to add more information on the technical terms, beside the definitions, by linking the most suitable web pages classified on the basis of their level of health information and used language [21]. Finally, we are in the process of developing a graphical framework to build health consumer-oriented advanced services and allow users to easily deal with health information [22].

ACKNOWLEDGMENTS

REFERENCES

[1] All Party Parliamentary Groups on Global Health. "Patient empowerment: for better quality,more sustainable health services globally". http://www.appg-globalhealth.org.uk/reports/4556656050

[2] Keselman, A. and Slaughter, L., "Towards consumer-friendly PHRs: patients' experience with reviewing their health records", AMIA Annual Symposium Procs., pp.399–403, 2007.

[3] Seedorff, M. & Peterson, K., Incorporating Expert Terminology and Disease Risk Factors into Consumer Health Vocabularies. Pacific Symposium in Biocomputing, pp.421–432, 2013.

[4] Alfano, M., Lenzitti, B., Lo Bosco, G., and Perticone, V., "Facilitating text understanding for e-learning users", Proc. of International Conference on e-Learning (e-Learning'14), 11-12 Sept. 2014, Tenerife, pp. 83-88, 2014.

[5] DIKW pyramid. https://en.wikipedia.org/wiki/DIKW_pyramid.

[6] Kandula S, Curtis D, Zeng-Treitler Q. A semantic and syntactic text simplification tool for health content, in AMIA Annual Symposium Proceedings. American Medical Informatics Association, 2010, pp: 366–370.

[7] Keskissarkka R. Automatic text simplification via synonym replacement, Ph.D. dissertation, Linkoping, 2012. Available at: http://liu.diva-portal.org/smash/record.jsf?pid=diva2%3A560901&dswid=-9669. Accessed: 29/09/2017.

[8] Jonnalagadda S, Gonzalez G. Sentence simplification aids protein- protein interaction extraction, in The 3rd International Symposium on Languages in Biology and Medicine, Hyatt Regency, Seogwiposi, Jeju Island, South Korea November 8-10, 2009, pp. 8–10.

[9] Zeng-Treitler Q, Goryachev S, Kim H, et al. Making texts in electronic health records comprehensible to consumers: a prototype translator. AMIA Annu Symp Proc. 2007; 11: 846-50.

[10] Abrahamsson E, Forni T, Skeppstedt M. Medical text simplification using synonym replacement: Adapting assessment of word difficulty to a compounding language. Proceedings of the 3rd Workshop on Predicting and Improving Text Readability for Target Reader Populations (PITR) @ EACL 2014, pp: 57–65, Gothenburg, Sweden, April 26-30, 2014.

[11] Unified Medical Language System, http://www.nlm.nih.gov/research/umls/.

[12] Zielstorff, R.D., "Controlled vocabularies for consumer health", Journal of Biomedical Informatics, 36(4-5), pp.326–333, 2003.

[13] Consumer Health Vocabulary Initiative, http://consumerhealthvocab.org/.

[14] WebMD, http://www.webmd.com/a-to-z-guides/common-topics/default.htm.

[15] Merriam-Webster – Medline Plus' medical dictionary, http://www.merriam-webster.com/medlineplus/.

[16] Ok Salute, http://www.ok-salute.it/dizionario-medico.

[17] Dizionario della salute, http://www.corriere.it/salute/dizionario/.

[18] Alfano, M., Lenzitti, B., Lo Bosco, G., and Perticone, V., "An Automatic System for Helping Health Consumers to Understand Medical Texts", Proc. of HEALTHINF 2015, 12-15 Jan. 2015, Lisbon, pp. 622-627, 2015.

[19] SavsoftQuiz. http://savsoftquiz.org/

[20] Multilingual Glossary of Technical and Popular Medical Terms in Nine European Languages, http://users.ugent.be/~rvdstich/eugloss/welcome.html.

[21] Alfano, M., Lenzitti, B., and Lo Bosco, G., "A web search methodology for health consumers", Proc. of ACM International Conference on Computer Systems and Technologies (CompSysTech'14), 27-28 June 2014, Ruse, pp. 150-157, 2014.

[22] Alfano, M., Lenzitti, B. Lo Bosco, G., Taibi, D. A Framework for Opening Data and Creating Advanced Services in the Health and Social Fields. Proc. of ACM International Conference on Computer Systems and Technologies (CompSysTech'16), 23-24 June 2016, pp. 57-64.