Patient Empowerment by Increasing the Understanding of Medical Language for Lay Users*

V. Topac, V. Stoicu-Tivadar
Department of Automation and Applied Informatics, "Politehnica" University of Timișoara, Timișoara, Romania

Keywords
Terminology, language understanding, patient empowerment, NLP

Summary
Background: Patient empowerment is important in order to increase the quality of medical care and the life quality of the patients. An important obstacle for empowering patients is the language barrier the lay patient encounter when accessing medical information.

Objectives: To design and develop a service that will help increase the understanding of medical language for lay persons.

Methods: The service identifies and explains medical terminology from a given text by annotating the terms in the original text with the purpose of adapting medical text assigned by medical personnel for the assisted patients. b) Into a dedicated web site that can adapt the medical language from raw text or from existing web pages.

Result: The output of the service was evaluated by a group of persons, and the results indicate that such a system can increase the understanding of medical texts. Several design decisions were driven from the evaluation, and are being considered for future development. Other tests measuring accuracy and time performance for the fuzzy terminology recognition have been performed. Test results revealed good performance for accuracy and excellent results regarding time performance.

Conclusion: The current version of the service increases the accessibility of medical language by explaining terminology with a good accuracy, while allowing the user to easily identify errors, in order to reduce the risk of incorrect terminology recognition.

1. Introduction

Patient empowerment is defined as helping people to discover and use their own innate ability to gain mastery over their disease or status [1] – by providing education for informed decision-making, assisting patients to weigh costs and benefits of various treatment options, setting self-selected behavioral goals, and providing information about the importance of their role in self-management.

The assessment of the improvement resulted by applying the principles of patient empowerment seems to be difficult, due to the lack of standards. The models identified are not universally accepted [2]. However, the obvious way to assess the impact of the patient empowerment is to use questionnaires [3] for a statistic analysis of the level of satisfaction of the patients themselves, and/or of the medical personnel involved. The results of this research also used questionnaires to evaluate the impact of the tool we developed.

The research and related work presented in this paper considers the difficulty of understanding medical language and information a key limitation of patient empowerment. It is commonly known that medical language is very often hard to understand for lay people. Given this the communication between doctors and patients can suffer especially when dealing with remote communication that can appear in systems like tele-care systems or web page based communication. A research project, using a specialized classifier, tried to evaluate how easy it is for regular people to access data expressed in medical language reached the following conclusion: “The classifier was then applied to existing consumer health Web pages. We found that only 4% of pages were classified at a layperson level, regardless of the Flesch reading ease scores, while the remaining pages were at the level of medical professionals. This indicates that consumer health Web pages are not using appropriate language for their target audience” [4]. This can affect in a grate manner the accessibility of the patients to their health information. Having a bad understanding of their health status may have a bad influence on their heath evolution. Empowering the patients with more understanding of the medical information related to them will strongly reduce this risk.

The classic solution in this area is language interpretation done by human interpreters. The presence of the interpreter

* Supplemental material published on our website www.methods-online.com
makes it possible for the patient and provider to achieve the goals of their encounter as if they were communicating directly with each other. There are several international institutions like IMIA (International Medical Interpreters Association) [5] that are providing standards and frameworks for medical interpreters.

We propose a Natural Language Processing (NLP) based tool to annotate medical terminology identified in raw texts or web sites.

1.1 Similar work

Text mining is a technique known to be used on medical content for purposes like automatic classification or information retrieval [6].

Classic machine translation tools can be used for translating medical content too, but these tools are very dependent on the training data. The text resulting from this process was evaluated in a research to be mostly incomprehensible [7]. A step forward solving this kind of accessibility issue is given by research and tools analyzing the level of accessibility of specialized language. One research [8] proposed a framework to inform the design of an “interpretive layer” to “mediate” between lay (illness model) and professional (disease model) perspectives.

Probably the most closely related research projects are:

a) A NLP solution, simplifying medical text by replacing difficult terms with synonyms and/or reducing sentence size, has been recently developed, having good results in terms of readability increase [9].

b) A tool identifying and explaining terminology from reports of electronic health records [10].

The tool developed has some major differences compared to the existing work. Compared to [9] and [10], our tool does not replace the terminology in the original text, but it annotates the text with terminology explanations. Also, the use of fuzzy matching techniques was chosen because the tool is designed to work on natural language, and because we wanted to keep this project as much as possible language independent (the only dependencies are the dictionary for a specific language and optionally the incorrect matching repository, that can be used with this tool). The drawbacks of using a strictly controlled terminology system on natural language was also acknowledged in some other research [11], stating: “Any controlled terminology will necessarily lack the richness of detail available from the vocabulary of a natural language; this loss of this detail is one of the trade-offs for having data in a computable form.” These terminology tools mainly falls into the vocabulary or glossary type, as categorized in [12].

2. Methods

2.1 Tool for Explaining Medical Terminology

Before designing this tool, the authors studied the standards [13] given by the IMIA [5] association. These standards have been used as a set of guidelines for this research considering issues like: interpretation, cultural interface, ethical behavior.

To perform interpretation of specialized language and terminology labeling a NLP tool was designed. The tool is composed of three modules:

- Text Parser (processing raw text or HTML from a given web address)
- Intelligent Terminology Dictionary (glossary, fuzzy matching and incorrect matching)
- Terminology Labeling (explanation presentation)

The architecture of the application, illustrating the interaction between modules is illustrated in Figure 1. The modules are being presented in more details below.

A) **Text Parser.** Accepts as input raw text or HTML (retrieves HTML content from a given web address). It normalizes the input by handling lower and upper case, punctuation and special characters. Then it iterates through all the words and verifies each word against the dictionary. If recognized as term, the word in the original text is annotated with an explanation. The architecture of the application is illustrated in Figure 1.
B) **Intelligent Terminology Dictionary**
(Vocabulary). The problem with linguistic data, especially natural language processing, is that it is dealing with uncertain information. A method of dealing with this kind of data is using error-tolerant methods like fuzzy string matching. In many cases when dealing with text, the back-end data storage solutions are databases. Fuzzy usage has been used in database information manipulation, FSQL (FuzzySQL) or SQLf getting closer to standardization [14]. Not the same thing is true when coming to in memory fuzzy data structures. This project uses a fuzzy data structure that was designed especially for this terminology interpretation. A detailed presentation of this novel data structure named FuzzyHashMap (FHM) can be found in article [15], and the sources of the data structure project are available as open source at [16].

**Fuzzy Medical Dictionary.** We want to identify, in plain text, medical specific terms. The terms have to be identified even though they are not in the canonical form. For this a FHM was used to build a medical terminology dictionary.

In order to explain how this will work, we consider we are parsing the following phrase: 

"... in diabetic diet recommendations ..."  

### 2.1.1 The Initial Implementation

Each word is checked against the dictionary. When arriving to "diabetic" term, as presented in Figure 2, the dictionary will search by firstly pre-hashing the term. The hash code for the resulted string "diab" is computed, and it points to the "diabetes" entry. The Levenshtein distance [17] (which is the default approximate matching algorithm in FHM) between "diabetes" and "diabetic" is 2, which is default threshold value in FHM.

The default FHM settings were set empirically after several tests, based on the best results; this is the complete set of settings:

- Pre-hashing function: Substring(0,4) – (first for letters of the word). This function was chosen for speed reasons. However, in order to reduce the false positive results, we have revised this, and used all 4-sized subwords of the original word (detailed bellow).
- Fuzzy matching algorithm: Levenstein Distance
- Distance threshold value: 2

### 2.1.2 The Revised Implementation (Reducing False Negatives)

Pre-hashing keys only by first n letters (first 4 in our case) lead to false negative results when the first letters of the inflected term (the term as found in the text) were different than the first ones of the canonical term.

In order to avoid this and to reduce the false negative rates, the pre-hashing was changed to consider all n-long (empirically we chose n = 5 as default) subwords of the original term.

So considering the above example, the word "diabetes", the subwords used in pre-hashing for populating and searching in the map were:

"diabe", "iabet", "abete", "betes"

Although this had a side effect over the performance of the map, the impact was too small to be perceptible over the general speed and usability.

Coming back to the example, the word **diabetic** has been associated to the term **diabetes** from the dictionary. In conclusion, the FHM enables the identification of terms that are not in their canonical form, in a very efficient way.

### 2.2 Reducing False Positives
(Incorrect Matches)

Using fuzzy string matching, and improving the pre-hashing function had a big impact on reducing false negative rate, but it has also increased the rate of false positives (mapping terms on words that are not related to the specified term). For example term "relapsed" was incorrectly mapped on word "related".

In order to decrease the false positive rates, an incorrect matching repository was created. The repository was created by training the system (we developed a dedicated module for training) with data that was not medical related; Identified matches were saved together with their frequency of appearance.

For example for Romanian language we have trained the system with texts like the Romanian constitution, some Romanian novels, geographical and historical texts, sport, politics and technical articles. By this, we wanted to identify those words that are outside of medical language, but when using fuzzy matching, they are mapped on medical terms.

After training the system with texts counting all together more than 200,000 words, we have gathered 248 incorrect matches, with word frequencies ranging from 37 to 1.

We have revised the identified matches, and found that 45 of them were actually correct matches. So in the end 203 out of 248 matches (82%) were saved as incorrect matches.

Here is one example of top ranked incorrect matching identified:

```
word = "internet" term = "internist"
freq = "26"
```

When used for explaining medical terminology, any approximate matching identified by the tool is checked against this incorrect matching repository, and if found
here, the match is considered invalid (is not considered a medical term). This helped reduce the rate of false positive, more details about this are presented in the Tests and Results chapter.

2.3 Terminology Labeling – Explanation Presentation

After medical terms were identified and annotated (tagged), the explanation of the term has to be added to the original text. For this, we enable multiple explanation presentation types, depending on user preferences: it can be inserted inline in the text, tooltip over the text or footnote in the text.

Because the main usage of this tool is to explain terminology from existing web sites, for presenting explanations we have considered the guidelines from Web Content Accessibility Guidelines 2.0 (WCAG), Sufficient Techniques 3.1.3 – Unusual Words [18].

Also the tool can be used by to tag the terms in the text in order to prepare it for other NLP tools, like translation tools. The tagging is done based on specification from Internationalization Tag Set (ITS) Version 1.0, section 6.4 – Terminology.

"The Terminology data category is used to mark terms and optionally associate them with information, such as definitions. This helps to increase consistency across different parts of the documentation. It is also helpful for translation." [19] This is an example of this type of markup:

...in <quote its:term="yes">diabetic</quote> diet...

2.3.1 Languages and Dictionaries Used

Currently the terminology tool is able to explain Romanian and English medical terminology. For the Romanian vocabulary, we used data from a small and lay user friendly dictionary available online [20], while for the English vocabulary we used data from a free dictionary plus data from [21]. Adding vocabularies for new languages and increasing and improving the existing vocabularies and incorrect matching repositories is an ongoing task, having direct impact on the quality of the results.

2.4 Using the Terminology Tool in TELEASIS Project

The TELEASIS project [22] has developed a pilot tele-assistance network with homecare electronic integrated services, allowing tele-assistance of the elderly, at their residence, based on the most recent IT&C technologies, with a medical and as well, a social target. The service-integrating tele-assistance system grants elders the opportunity to benefit from healthcare at home, to enjoy an improved personal lifestyle.

The TELEASIS system suggests several ways to ensure this, like access to a central medical information database, access to additional communication channels and a service explaining medical language. TELEASIS system is offering patients access to their health data, reports and additional medical information. All this data is stored in an information and content database. Enrolled medical stuff or other power user can add documents to this database, and can set the access rights for patients or groups of patients. In this way each patient can access different documents. While allowing the patients to access medical information proves to be useful, as reminded in the introduction, the patients may encounter big difficulties in understanding that information. For this, we integrated the presented service into the TELEASIS system, allowing the patient to get the medical information explained.

Since this service has been proven efficient in other use cases, we considered that TELEASIS users will also benefit from it. However we have not evaluated the efficiency of the service with real subjects (tele-assisted elderly persons) in this particular use case.

2.5 Using the Terminology Tool in a Dedicated Web Page & Web Service

In order to make this terminology tool available for any user, a web site and service has been developed. The tool can be accessed at the address specified at [23]. It allows users to see explained terminology from raw text, or from an existing website.

2.5.1 Using the Service for a Custom Text

The case when the user introduces a custom text, and obtains as output the text having terminology explained is illustrated (using a screen capture) in the Supplemental Figure 1 (see online material).

In the settings area, for advanced usage, the term recognition approximation level (fuzzy matching level) can be set by the user.
2.5.2 Using the Service to Adapt an Existing Web Page

If the user wants to see terminology explained on an existing web site, he can enter the URL of the website in the input area. In this case the terminology service is acting as a mediator between the original web site, and the page displayed in the browser. As a result, the web service adds the explanation of recognized terminology as tooltip over the term.

The case when the end user is browsing a web page via the terminology service [23] is presented in Figure 3.

The adapted web page is displayed in the browser having terminology explained. A portion of the adapted web page, having medical terminology identified and explained as tooltip is illustrated in Figure 4.

In order to add more flexibility and to enhance usability, the user can choose how the definition of the recognized term will be displayed; for this there are two options: a) to mark the term as underlined, and add the definition of the term as tooltip (this is the default display option); b) to append the definition in parentheses in the original text.

Behind the scenes, the system is powered by a Java Servlet on the server side that is transcoding the original web site, redirect links, applying special styles, processing and adapting language/terminology. All the terminology recognition and annotation is done by a dedicated module developed in Java. On the client side HTML, a JSP page and JavaScript is used in order to create and manage the user interface. The project is running on AppEngine, the cloud infrastructure from Google.

3. Evaluation Methods

The main aspect considered for evaluation was message understanding increase:

For this we decided to perform tests on content from a web site offering medical information for end users and interviews with doctors [24]. The text chosen was about the symptoms and possible treatment of thyroid cancer [25]. We selected it because we believed that this kind of web sites (educating end users about medical problems), and this kind of content, explaining symptoms and treatment can benefit more (in terms of patient empowerment) when the terminology is labeled. Also this content was rich in medical terms not so popular and known by lay persons.

We used a questioner (formally called Q1) to evaluate the message understanding increase. The first items from the questioner were asking few details about the participant: age, education level and knowledge in medical field. The last item was used for avoiding biased results (filtering out participants that had strong medical knowledge, since they are not in the target audience). Age and education level were asked in order to analyze the results classified on groups of age and education.

The questioner was distributed online using mainly a popular social network (Facebook) and some other means (direct e-mails to some participants). On Facebook social network, friends of the authors have shared the questioners, so that among participants there are mostly persons unrelated to the authors, but there could also be friends of the authors (we considered this less relevant for the test, and used other means to avoid biased results).

The questioner contained two texts (generically called text A and B) about thyroid cancer symptoms and treatment. One of the texts was in the original form, and the other had terminology annotated. The questioner was designed so that it will randomly display text A annotated and text B as original, or text B annotated and text A as original. This was done in order to avoid influences over the results due to an eventual difference of difficulty between text A and B. The participants were asked to rephrase both texts, using their own words. Then an expert in the area evaluated the level of understanding in both rephrased texts, and decided whether or not the annotated text leads to a better understanding.

In the last question, participants were asked about the impact of terminology annotation.

Another questioner (formally called Q2) was designed to investigate user preferences over explanation presentation type, and the impact over reading performance.

The evaluation of terminology recognition accuracy, usability and latency was also tested and it is detailed in the “Tests and Results” section.

4. Tests and Results

Several tests were performed, focusing on message understanding increase, terminology recognition accuracy, latency and usability.

4.1 Message Understanding Increase

Questioner Q1 described above was done in Romanian language, and had 41 participants. From this, only answers from 37 participants were taken into consideration, since the other 4 responded affirmative when asked if they had strong medical knowledge.

Most of the participants were having higher education: 30 with higher education and 7 with secondary school or high school. One could argue that participants with higher education can lead to biased results, but the results of the test proved that this category can benefit more from terminology adaptation. Also, a big part of the consumers of online health information is composed of persons with higher education. Age distribution was: between 20 and 25 years 14 participants, between 25 and 30 years 16 participants, and over 50 years 7 participants.

On the answer to the question about the impact of the explained terminology over...
the message understanding the participants responded as indicated in Table 1.

4.1.1 Rephrasing Analysis

After analyzing the rephrased texts, we found that in 11 cases (29%) the annotated text was better understood than the text without annotations.

Here we present few other trends we have noticed based on the results, age and education:

- The terminology annotation had a smaller impact over understanding increase for participants with lower education compared to those with higher education. This result was unexpected, but since there were only 7 participants with lower education, this finding should be validated with more participants.
- Participants over 50 years tend to reuse medical terminology in the rephrased text, although they were asked to use their own words.

4.1.2 Reading Performance

In the other questioner (Q2), when asked about the impact of the explained terminology over the reading ease, the participants responded as indicated in Table 2.

4.1.3 Terminology Recognition Accuracy

Terminology recognition accuracy represents the precision of terminology recognition in medical texts (in both, raw text or web pages). To evaluate this we used the default settings of the system (fuzzy matching level = 2). In this test we were interested in finding the rates of correct terms matching, false positives and false negatives. The evaluation was done by one of the authors on a web page [26] containing 2616 words.

We have tested both the initial implementation (using first 4 letters for pre-hashing), and the revised implementation (using a 5-long subwords for pre-hashing, and validation with incorrect matching repository). The results for the test done using the initial and revised implementation are listed in Table 3.

By comparing results from Table 3, one can see that changing pre-hashing lead to a small improvement for false negative rate, while the same change, coupled with the validation with Incorrect Matching Dictionary had a big impact over false positive rate.

The relative big number of false negatives is mainly related to the size of the medical vocabulary used. Increasing the number of terms in the vocabulary will decrease the number of false negatives.

4.1.4 Latency

Since this project can be used to mediate web pages browsing, speed is important, in order to maintain a decent browsing experience. After testing several web pages, an average of 3.5 seconds were needed (additionally to original load time of the web page) in order to load the adapted web page (the pages tested had approximately 2000 words per page). We concluded that this is a reasonable time, preserving a decent browsing experience.

4.1.5 Usability

The usability tests were performed by the authors. The main focus was on the web page adapter, seeking the best way to present the explanation of the terms and other feedback. Currently the system can present the explained terminology inline (inserted into the text) or as tooltip over the term. In the questioner Q2, we asked the participants what would be their preferred method to present the explained terms. The results are listed in Table 4.

5. Discussion

One of the most important questions while designing this tool was whether we should
only identify and explain medical terms (labeling) or we should make a complete translation (rephrasing), removing medical terminology and adding lay alternatives. We chose to do the first one mostly based on an interesting research [27] that concluded that a translation to lay language may decrease the confidence level of the message for patients. So by only labeling terminology, we try to maintain a fair balance between preserving message authority and improving message understanding.

The evaluation outcome shown that by recognizing the medical terms and adding the definition of the terms the understanding of the message can increase. However, the evaluation was done on a small medical text. It would be interesting to see if/how the results can change when performing the test on a long text. We expect the outcome would be even better, since it is hard to make a resume of a short text (we could see a big difference in the number of people who answered that the tool helped them understand the message, and the number of understanding increase reflected in the summaries of the two versions).

We were surprised to identify in the results a trend indicating that participants with lower education were less helped by terminology annotation. Maybe for this category it is better to replace the terminology with synonyms. However this hypothesis needs to be further tested, since in this test there were only 7 participants from this category.

Related to usability, from the evaluation done with Q2 we deduced that we should enable our web page adapter to offer the terms explanation by request (click or right click on term).

As a general rule, we have decided that in all cases the term in the canonical form will be shown before the definition. This is done in order to reduce the risks of misleading the reader in case of false positive terms recognition (due to fuzzy matching). This way the reader can easily identify and ignore errors.

6. Conclusions

Patient empowerment is important in order to increase the quality of life of the patients. This research and the suggested solutions contribute to this by supplying to the people in need, but not only, a more understandable and accessible information, even if that information contain terms difficult to understand, as medical ones.

The tool developed in this research is available for everyone at [23] and can be used in several ways, being customizable.

The current results are encouraging, showing that this tool can increase the understanding of medical language. This also shows that fuzzy matching can be used for terminology recognition, in order to increase the recognition rate.

This work will continue focused on improving the medical text understanding, improving existing vocabularies and adding vocabularies for more languages.

Also, we started working on a mechanism for obtaining, validating and using feedback from users related to fuzzy matching accuracy.

Acknowlegdment

Part of this work was supported by a Romanian National research grant called “TELEASIS – Complex system, on NGN support for home care and teleassistance dedicated to elderly individuals”.

References