Characteristics and Web Genres of Health-Related Websites

Csongor Alexandra

Department of Languages for Specific Purposes, Faculty of Medicine, University of Pécs

Introduction: This paper explores some general features of health-related websites in Hungary and also aims to investigate the typical web genres represented by examining a corpus of 50 web pages.

Material and method: The most prominent categories of health related and medical websites were summarized and compared to a survey in Romania in 2010. To answer questions about the web genres associated with the purpose of health, a corpus was constructed and analyzed based on the lists of a Hungarian web directory. The web pages in the sample were classified into specific genres.

Results: The most common categories in the health section are alternative medicine, lifestyle issues, services and dental health. The results suggest that commercial and educational sites dominate this field. Besides organizational and business homepages, the use of topical homepages is characteristic and other genres appear embedded in the texts, like Frequently Asked Questions (FAQ's) reports, newsletters, discussion forms and articles.

Conclusions: Genre analysis can be applied to the specific context of digital documents. The study of genres is essential in describing the language use of the Internet and the results suggest that some web genres are more represented for providing health related information than others.

Keywords: health-related websites, web directory, web genres, homepage

Introduction

According to a survey in 2001, scholarly research, shopping, cultural art activities, reading the news and learning about health issues were the most common purposes for which users searched the web [1]. This fact justifies that there is a need for studying medical and health-related websites. The aim of this study is to investigate some general characteristics of Hungarian medical and health-related websites as well as to identify the typical web genres that are associated with this subject area.

The study of genre is extended from verbal and written communication towards digital communication. In most definitions a genre is characterized by having similar content, form and intended communicative purpose [2]. Researchers of digital documents build on the definition of Orlikowski and Yates, who defined genres as "social type of communicative actions, characterized by a socially recognized communicative purpose and common aspect of form" [3]. The first characterization of web genres was performed by Crowston and Williams, who identified reproduced genres from the traditional media and emergent or novel genres [2].

A wide spectrum of genres appear on the web, which include existing genres that moved to the electronic medium without changes and also new genres that emerged as a result of digital communication like homepages, hotlists, blogs etc. [4]. There are numerous surveys about digital genres, however, a well-established genre concept and categorization of Internet texts has just started to develop. The present study is intended to be a contribution to this process.

Material and method

With the help of 5 web directories and one collecting portal, links related to the topic of health were collected. The use of search engines and simple search terms were excluded as entering 2 search terms (health and treatment) into search engines resulted in an unmanageable number of links (more than 822,000). The number of links to websites listed under the health section of the 5 web directories and in the collecting portal (startlap.hu) varied from 118 to 1036. The most common categories under the section of health were summarized and compared to a survey in Romania in 2010. Finally, a corpus was constructed with the help of the web directory, which provided the highest number of links. 50 websites were selected with purposive sampling chosen from each category of the web directory under the health section and were examined in March 2011.

Sites that were not in Hungarian or unusable pages that were not available for technical or other reasons were excluded from the sample. The web pages selected were analyzed in order to categorize the web genres represented in the corpus. The aim is to explore which genres are typically associated with the purpose of providing health information.

Results

Based on the links collected the most common topics of health-related websites were analyzed. The sites were classified according to their content in the 26 most frequent categories: alternative medicine, medical science, diseases and symptoms, psychiatry, sensory organs, men, fitness, dental

Table I.   The composition of the corpus

Genres identified	Number of websites in the corpus	Proportion of websites in the corpus (%)
Organizational/business homepages	15	30
Topical homepages	12	24
Product information	9	18
Tutorial pages	3	6
Link lists	2	4
Self-help pages	2	4
Blog sites	2	4
Search result lists	2	4
Forum	1	2
Article	1	2
Personal homepage	1	2

health, weight loss, children, pharmaceutics, information and cyclopedia, addictions, hospitals, media and publications, women assistant services, professions, self-improvement, animals, cosmetology, organizations, sexology, services, nutrition and products. The remaining categories and medical specialties found in the web directory with the most links were subcategories of the originally listed items. To sum up, the categories include links about a wide spectrum of health issues including diseases, symptoms, medical science, health care, institutions, pharmaceutical products, alternative medicine and lifestyle. Several categories cover a wide area of health related topics. The first three categories with the highest number of sites were alternative medicine, lifestyle or services, and dental health.

The second major question of the study is to assess the web genres that are the most characteristic when searching for health related information. In order to categorize the genres represented, a special corpus consisting of 50 websites was constructed. The classification of websites based on their genres in the present corpus is shown in Table I, which represents the number and proportion of the genres identified. Based on the findings, the most typical genres used are organizational, business and topical homepages, and product information appears to be the third most common in the corpus. These three genres comprise 72% of the corpus. The proportion of the other genres is low as compared to the first three most represented genres. Tutorial pages comprise 6% of the whole corpus. Link lists, self-help pages, blogs and search results lists represent 4 % of the websites, respectively. There was only one example for each of the following genres: forum, news article and personal homepage. Other genres can be identified as part of the homepages like Frequently Asked Questions (FAQ's), newsletters, discussion forms, articles and reports establishing a genre system or genre hierarchy.

Discussions

The results concerning the thematic orientation of health related websites correspond to a survey conducted in Romania in 2010 based on a sample of 317 medical websites [5]. The most prominent categories regarding the number of sites were complementary and alternative medicine sites,

multispecialty sites and dental health is ranked in the third place [5]. The first and third categories are similar when comparing the results of the two surveys, and the second category covers a wide range of topics in both cases.

The results suggest that the majority of websites are commercial or at least contain some commercial information, and their communicative purpose is to publish information about certain organizations or products. Topical homepages are very common, which can be the result of the query method using a thematic link catalogue. On the other hand, topical homepages are useful in navigating and searching for topics and are extensively used on the world-wide web. The websites analyzed are mostly for non-professionals, and in most cases they provide educational or commercial information. There is only one article in the corpus that can be an example of a scientific category, however there may be scientific genres embedded in the other homepages as well.

With the help of genre analysis researchers aim to understand how virtual communities use communication. It is suggested by linguists that certain genres are more appropriate and relevant for certain purposes and tasks than others [1]. The genre classification of the present corpus is based on the study of Roussinov et al., who suggested a genre repertoire of web pages associated with specific purposes [1]. The aim of the present study is to propose a genre palette of Hungarian health-related websites. The use of the Internet provides a particularly interesting setting for conceptualizing the notion of genre. The readers encounter a number of different genres, which are linked together forming a genre hierarchy or a more complex genre system [2].

The study has some important limitations. Firstly, the representativeness of the research is to be verified, however as the most popular web directories were used, the results are likely to reflect common trends. Secondly, Internet texts are in a constant change and this also means that web genres, like all kinds of genres are dynamic in nature.

Conclusions

The world-wide web provides an enormous amount of information and the Internet is an important tool for selecting and buying products as well as for collecting information on a particular topic. The Internet has an impact on people's decisions upon choosing a medical treatment or taking some kind of medicine. To sum up, health is one of the top purposes for which Internet users search the web, so it is essential to survey this field of language use. The most common categories in the health section of the Hungarian web directories and the most commonly used collecting portals are alternative medicine, lifestyle issues, services and dental health. According to the results, promotional and educational sites for non-professional readers are the most represented in the corpus of 50 web pages. It would be useful to extend this study in order to investigate scientific medical information on the Internet.

Further research can help find the relevant information effectively and facilitate the use of the Internet for patients and health-professionals.

References

1. Roussinov D, Crowston K, Nilan M, et al. – Genre based navigation on the Web. In: Proceedings of the Thirty-fourth Annual Hawaii International Conference on System Sciences. Hawaii,2001,vol. II: 97-109
2. Crowston K, Williams M – Reproduced and emergent genres of communication on the World-Wide Web.The Information Society 2000, Vol. 16, No.3,: 201-216
3. Orlikowsky WJ, Yates J – "Genre Repertoire" The Structuring of Communicative Practices in Organizations. Administrative Science Quarterly 1994, 39,4: 541-574
4. Shepherd M, Watters C – "The Evolution of Cybergenres". In Proceedings of the Thirty-First Annual Hawaii International Conference on System Sciences-Volume 2, 1998, 97-109
5. Nadasan V, Ancuceanu R, Tarcea M, Grosar CM, Ureche R – General characteristics of the Romanian medical webscape. Acta Medica Marisiensis, 2011, Vol 57, No 2 (In Press)