How deep is the sea: Web Intelligence for Patient Education

Mauro Mazzocut¹, Marialuisa Antonini², Ivana Truccolo¹, Paolo Omero³, Emanuela Ferrarin⁴, Raffella Gandelli⁴, Nicolas Gruarin¹, Carlo Tasso³

¹ CRO Aviano IRCCS National Cancer Institute - Scientific & Patients Library, Aviano, Italy
² infoFACTORY srl, Udine, Italy
³ University of Udine- Artificial Intelligence Laboratory, Udine, Italy
⁴ CIFAV onlus- Drug Information Service, Aviano, Italy

Corresponding author: Mauro Mazzocut, mmazzocut@cro.it

Abstract

Introduction: Web users have turned into information producers who share with others contents and opinions. The literature reports that patients use the Web for searching health information, and also for sharing their feelings and experience with others experiencing the same condition. Peer-to-peer communication has relevant effects on patients’ decisions about their own clinical path. That’s why we think that the social Web is full of implicit knowledge hidden in the semantic relationships among users, in their opinions expressed by tags and comments. Web Intelligence (WI) is a new direction for scientific research that combines Artificial Intelligence (AI) and advanced Information Technology to develop a next generation of Web empowered products, system, services and activities.

Aims: We aim to present the results of a preliminary WI survey made in cooperation with the AI-laboratory University of Udine (WI-Lab). We focused on Complementary and Alternative Medicine (CAM) conversations available on the Web.

Methods: We can get an insight of the actual patients perspectives on relevant cancer topics by analyzing the Italian language online conversations of health consumers. We focus on CAM and we exploit WI in order to map/identify the main topics, online sources, and relationships among users.

Discussion: The innovation is the application of a WI approach to the oncologic field in order to highlight the evidence available online about untold perspectives shared by cancer health consumers.

Key words: patient education; web intelligence; social media; internet; health literacy.
Introduction
Web users have turned into information producers who share with peers contents and opinions. The literature reports that patients use the Web for searching health information, and also for sharing their feelings and experience with others going through the same condition (1).

Peer-to-peer communication has relevant effects on patients’ decisions about their own clinical path (2,3). That’s why we think that the social Web is full of implicit knowledge hidden in the semantic relationships among users, in their opinions expressed by tags and comments (4–8).

We focus our survey on Complementary and Alternative Medicine (CAM) as we hypothesize that this topic is largely discussed among patients themselves, or among their relatives and friends (9), but it is not considered much when interacting with doctors (10,11).

The statistics related to CRO Patients Library & Drug Information Service users show that there is an increasing request for information about CAM from 2008 to 2013 (12).

Peer-to-peer Web conversations may give us some insight of what patients think about CAM and cancer rumours. There is evidence that poor information affects health outcomes (5,13) and misinformation affects the decision makers policy as well as public debates (14,15).

Web Intelligence (WI) is a new direction for scientific research that combines Artificial Intelligence (AI) and advanced Information Technology (IT) to develop a next generation of Web empowered products, system, services and activities (16–19).

Aims
We aim to present the results of a preliminary WI survey made in cooperation with the University of Udine WI-Lab. We focused on Complementary and Alternative Medicine online conversations.

Methods
Various specific WI techniques have been exploited to extract meaningful knowledge from the information stream of peer-to-peer Web conversations about CAM.

CAM keywords set
We identified a set of keywords describing the known CAM therapies and their context. We used scientific and medical sources (20–22); specific booklets for patients published by Volunteer Oncologic Associations (23,24), and public materials easily accessible on the Web in Italian language (25–27).

We found 255 therapies. For every therapy we found related synonyms and, if present, reference to a specific philosophy or religion; names of people, organizations or places; book titles; anatomy or physiology terms; substances or drugs; principles or processes; methods; instruments or tools; effects or side effects; diseases.

Eligibility criteria
We included Websites, forums, blogs, communities, social networks of any kind of topic containing discussions about CAM and addressed to the general public.

In particular we decided to include Facebook pages and groups and YouTube channels, when allowed by the privacy settings.

We decided to exclude Twitter, because the length of the messages is too short to express a complex opinion about the topic taken into account.

The search was limited to the sources published from January 1st 2013 to May 31st 2014. The search was also limited only to sources using Italian language. These online sources have been analyzed by using WI tools.

Materials
More specifically we used ifMONITOR, a Web based system that allows the automation of Web sources harvesting, indexing, filtering and classification. ifMONITOR has been developed by the AI-Lab of the University of Udine.
**Web search**

We submitted to Google and Bing some generic search query in Italian language, terms as “cure”, “therapies”, “treatments”, “remedy”, “cancer”, “tumor”, and “leukemia”. This produced a amount of results.

**Filtering and Analysis of the Web sources**

The large submitted Web query identified many non pertinent sources caused by the synonymy or polisemy phenomena (i.e. “Cancer” is also a Zodiac sign). The collected sources have been then analyzed by means of the specific ifMONITOR module devoted to automatic text extraction, in order to separate the pertinent from the non pertinent discussions.

ifMONITOR identified in the text the occurrence of keyword and sequences of two-four words present in the CAM keywords set initially provided. The selected sources have been indexed by Website name, date, post or comment author, and CAM items extracted from the CAM keywords set. This first quantitative analysis allowed us to identify the most discussed CAM topics and most important Websites including texts about CAM. A sample of 1500 documents from relevant Web sources related to CAM therapies have been automatically tagged by ifMONITOR and then revised and qualitatively evaluated by a team of WI experts from the WI-Lab.

The team manually associated semantic tags to the content of each conversation by exploiting terms extracted from the CAM keyword set. Moreover the human analysis allowed:

- to measure the popularity of each CAM therapy or drug, by considering the effectiveness/ ineffectiveness perceived by people through a numerical score (ranging from 1 to 5)
- to estimate the sentiment (positive, neutral, or negative) of the most discussed CAM therapies mentioned in the conversations. Furthermore, the team explored in more details posts, opinions, and comments related to specific pros and cons within topics of interests.

The quantitative and the qualitative analysis performed on the online conversations allowed to:

- Identify the CAM most discussed therapies;
- Identify the most relevant Web sources about CAM conversations;
- Identify the most discussed topics about every CAM therapy identified during the analysis;
- Measure the popularity and sentiment about the identified CAM therapies;
- Identify the pros & cons issues in the discussions on the CAM therapies.

These results will be compared to the statistics related to CRO Patients Library & Drug Information Service users, in order to detect possible consistency or inconsistencies.

**Discussion**

The innovation is the application of WI approach to the oncology field, in order to increase the evidence of untold perspectives shared by cancer health consumers’ (4).

A Web conversation analysis, based on WI approach, can: enhance new strategies for patient education activities, tailored on health consumers needs both in digital and traditional environments; improve doctor-patient relationship; positively affect the patients course of treatment and healing (7,28,29). The final results will be presented and discussed at the 14th EAHIL Conference.
REFERENCES


<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>22.</td>
<td>Consiglio Nazionale FNOMCeO. Linee guida per la formazione nelle medicine e pratiche non convenzionali riservate ai medici chirurghi e odontoiatri della FNOMCeO [Internet]. FNOMCeO; 2009. Recuperato da: <a href="http://www.omceomb.it/public/upload/Normativa/Linee_guida_FNOMCeO_formazione_nelle_medicine_e_pratiche_non_convenzionali.pdf">http://www.omceomb.it/public/upload/Normativa/Linee_guida_FNOMCeO_formazione_nelle_medicine_e_pratiche_non_convenzionali.pdf</a></td>
</tr>
</tbody>
</table>