Framework Based on Ontologies for Palliative Care of Patients with Breast Cancer

Denis CEDEÑO[1,*], Miguel VARGAS-LOMBARDO[2]

[1] Technological University of Panama, Computer Programming Department, Panama City, Panama
[2] Technological University of Panama, Research Group Electronics Health and Supercomputing, Panama City, Panama
E-mails: denis.cedeno@utp.ac.pa; miguel.vargas@utp.ac.pa

* Author to whom correspondence should be addressed; Tel.: +507 244-0377

Abstract

The use of information technology solutions has become an innovative way to treat patients in hospitals. The creation of a framework based on the development of ontology to structure information on patients with breast cancer is developed. The main aim is to find patterns of information that can be drawn after analyzing the data using text mining techniques. The obtained knowledge can be used for decision-making by health specialists or can make the subject of research for improvement of the health care system.

Keywords: Framework; Ontology; Text mining; Knowledge

Introduction

In the Republic of Panama, medical information (e.g. medical charts) [1] is not stored electronically, and this behavior affects the health system of the country. Patient information is stored on different media making any medical administrative task very difficult [2]. This problem affects the country's health system, since there are no data are available for decision-making.

Application of Information and Communication Technology (ICT) in issues affecting healthcare is increasing. The main advantage of ICT is that it allows efficient alternative diagnoses and proficient monitoring of patient healthcare. For specialists this translates into an improvement in access to relevant information, leading to high-quality healthcare services, lower costs and better patient participation in modern healthcare systems [3].

Health Information Technology (HIT) tools and technologies have the potential to improve the health care systems and could lead to high-quality healthcare services in modern health systems [4]. Knowledge was and still is a powerful strategic tool for a health organization, analyzers, and healthcare policy makers in regards of extracting and managing information for innovation. Human knowledge [5,6] is thus a powerful strategic tool for organizing and managing in the process of healthcare innovation, competitive advantage, transfer of learning and collaborative work [7].

This work will focus on supporting the group of palliative care specialists [8] who treat patients with breast cancer in health organizations in Panama city by providing updated knowledge that will be extracted from the medical records of patients for making decisions more effectively.

In Latin America and the Caribbean, 27% of new cancer cases and 15% of cancer deaths are due to breast cancer [9,10]. In North America, 30% of new cases and 15% of cancer deaths in women are related to breast cancer [9,11].

Breast cancer is the most common cancer in women in both developed and developing countries [12,13]. Is the most frequently diagnosed cancer [14,15] and the leading cause of cancer death in females worldwide, accounting for 23% (1.38 million) of the total new cancer cases and 14% (458,400) of the total cancer deaths in 2008. About half the breast cancer cases and 60% of the deaths are estimated to occur in economically developing countries[16].

The incidence of breast cancer [17] is increasing in the developing world due to increased life expectancy, increased urbanization, aggressive screening programs, and the adoption of lifestyles associated with westernization. In Panama, 21% of new cancer cases and 15% of cancer related deaths in women are due to breast cancer [18,19]. The trend of recent years shows that both the incidence and mortality from breast cancer is still increasing.

The clinical data of patients in the palliative care unit are stored as text medical documents. Medical specialists need specialized tools for the retrieval and analysis of such data. There is no technological model based on innovative sources that provide substantial patient information or that can help doctors to develop research or knowledge management and allow the development of procedures and techniques necessary to improve the quality of patients in palliative care centers in the country.

This manuscript presents the development and implementation of a framework for extracting knowledge, based on the development of ontology to structure information and text mining using algorithms that are applied to find patterns to elicit information from medical records of patients with breast cancer.

Capabilities Framework

Palliative Care

Classically, the attention of the person with terminal cancer has become the reason for palliative care [20]. Palliative care is defined as an approach to improve the quality of life of patients and their families facing the problem associated with a mortal illness, through the prevention and relief of suffering by means of early identification, assessment and treatment of pain and other physical, psychosocial and spiritual problems [21].

Technological Solution

The development of efficient information technology solutions that enable an effective way to control the cost of healthcare and innovative e-health systems have become an important area of development.

Our goal was to develop and implement a framework technology approach oriented e-health, with information available for scientific research taking as references the methods presented in [22-24]. The following capabilities are included in the technological solution:

- ICT infrastructure for monitoring patients: visualization the patient's record in electronic form and thereby with updated information.
- Information retrieval and knowledge management: with the data available automatically, recovery is fast and easy, managing the knowledge of those aspects in the data invisible to the naked eye.
- Data stored on a reliable structure database: the storage will be through a database, which is a reliable and easily accessible structure.
- Streamline decision-making: with quick access to information and knowledge management, decisions may be taken in an agile and accurate manner.
- Statistical analysis: accurate data sets will help specialists to perform related research in breast cancer.
- Evaluate the health status of the patient: to assess the health status of the patient using the collected information to perform a more precise diagnosis.

- Recognition of unknown patterns: with the implementation of algorithms for text mining models, find patterns of information that provide diagnostic support to the specialists.
- Integrated systems for e-health: interoperability with other systems or information models of health knowledge.
- Effectiveness, efficiency and robustness in information: for specialists information is vital and should be accurate.
- Privacy, security and trust information: extreme confidentiality will be ensured in the patient data as part of the security of the information that is required in these cases.

Major changes in the demographic and social structure apply pressure on the sustainability of health and social care, and thus on the quality of life of patients with cancer.

In this sense the proposed framework is novel in technological context, using information structuring tools such as ontologies, as well as techniques and algorithms for text mining.

The social impact to the Panamanian population will be noticeable in the extraction of knowledge inclusive regarding the breast cancer and as a consequence new research could be conducted in regards of extracted information.

Framework Components

The general structure of the proposed Framework [25-27], is based on elements that act as components to facilitate integration with functional features. Each component represents one or more features or functions, while the components can be used independently if necessary. These components can be seen in Figure 1.
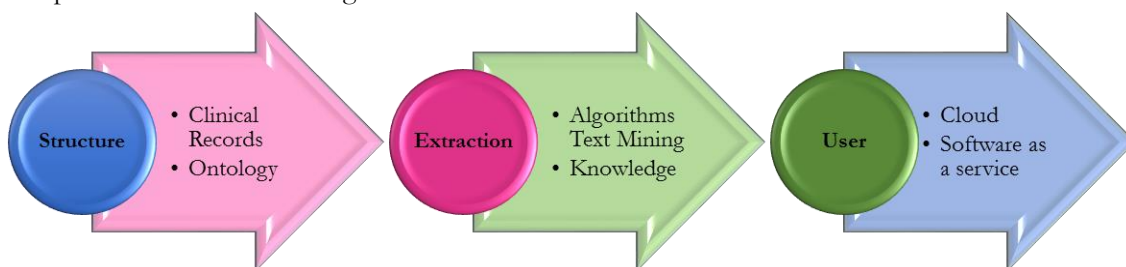

Figure 1. Framework Elements

Below are the details of each of these components:

- **Structure Component:** is considered the initial component where the data of the patients are structured.

  This component includes the development of ontology. As mentioned in scientific literature [28-30], ontologies are prominent in a knowledge sharing place. An ontology is a set of definitions with a central role in understanding and viewing the world [31,32].

  Ontologies can be used to structure information. Recently many information retrieval systems made use of ontology to help users better understand their information needs [33,34].

  According to the IEEE (Institute of Electrical and Electronics Engineers), interoperability is: the "ability of a system or a product to work with other systems or products without special effort on the part of the customer. Interoperability is made possible by the implementation of standards." [35,36]

- **Extraction Component:** a series of text mining algorithms are applied on structured data to find patterns of information.

  Text mining refers to the process of deducing insights from the text information. Knowledge Discovery in Databases (KDD) is the most common area of knowledge discovery and its origins are in areas such as databases, statistics and machine learning [37].

  If that knowledge is obtained in text form, this process is called text mining, and data source can be electronic documents [38-40]. This component is important because it generates

knowledge, which is a resource that becomes a subject with enormous potential to change the world due to advances in new information technologies.

Understanding the information that health professionals need is vital to the process of designing a system of information management [41]. The term knowledge management was first introduced in Europe Management Conference 1986 [42].

The outcome of this component is to acquire new knowledge obtained from medical records of patients.

- **User Component**: allows the management of the acquired knowledge and gives different access to the knowledge to different users of the Framework (such as medical specialists, researchers, administrators, etc.).

  The use of software such as the SaaS layer service of Cloud Computing give easy access to the knowledge and this item is vital as the data can be stored on any server in the cloud. This component can emerge as a new paradigm to accommodate and provide services-through the internet [43]. This technology has the potential to improve collaboration, agility, scale and availability of computer resources, besides reducing costs through optimized resource efficiency [44,45].

  The SaaS layer is the top layer of the cloud [46,47] and is the layer that most consumers use it. Built on top of both infrastructure as a service (IaaS), platform as a service (PaaS) [48], SaaS provides applications, programs, software and web tools to the public for free or for a price. Accessible via computer, tablet or smartphone, the layer of SaaS Cloud encompasses the largest and most accessible of cloud computing layer. However, a special attention need to be paid to confidentiality and HIPPA accessibility while data should be deidentified before storing them on the internet.

The three services that will be implemented are as follows: storage, visualization and interoperability.

- The information is stored in various storage media, such as high school servers, compact discs (CD's), digital versatile discs (DVD's), universal serial bus or any other means.

- In terms of presentation of the results these must be obtained in real time, easy to read and interpret, and accessible. The devices used to show the results of the Framework will be computers, tablets, smartphones or if it will be preferred, printed document.

- A service for exchanging information will also be available to allow the use of data by other specialists or researchers.

Framework Arguments

This Framework does not indicate that information will be entered by medical specialists, but they will be able to define clinical concepts that wish to manage and required information that will be recorded for each type of activity.

The incidence of breast cancer in Panama is high [49,50]. This type of cancer is of interest because of the large number of new cases, especially in young women are diagnoses lately [51].

Despite the fact that the Framework does not provide any cure, it will be helpful for medical specialists to find valuable information to help improve diagnosis, treatment, and palliative care. This Framework once developed can be applied to any other clinical data.

The proposed solution will define some elements, which were selected considering the desirable characteristics of knowledge to be achieved extracted from text documents (medical records) on patients with breast cancer. The selected items were:

- Modeling text documents through an ontology.
- Extraction of patterns through text mining.
- Other technologies on which to develop.

The use of ontologies in the field of information technology is very common and is used in various areas to achieve different goals. The criteria for choice of methodology to develop the ontology in the particular case of the proposed framework are as follows:

- **Selection of ontological methodology:** Ontology defines a common vocabulary for researchers who need to share information in a domain. It contains definitions of basic concepts and their relationships that can be interpreted by a machine.

  Before proceeding with the application of ontology, a comparison of different methodologies to establish what is the optimal development was conducted. For the design of any ontology is necessary to have a specific methodology. Many proposals exists in the scientific literature. For example the CYC methodology published by Lenat and Guha since 1990 [52], which reported some general steps for building ontologies. In 1995, the methodology of Uschold and King [53] that recreate a series of steps that allow to capture and specify the knowledge existent in a specific domain; this was not considered because its base on ontology that emerges from business and healthcare is considerable different by business. In parallel with the previous one, the methodology of Gruninger and Fox [54] arises, based also on ontologies for business type activities.

  Methontology was developed within the Ontology Engineering Group at the Polytechnic University of Madrid. The Methontology methodology was used because it was considered appropriate for the following reasons:
  - It has been used in the construction of ontologies in different fields. [55,56]
  - ODE and WebODE [57] are tools to support this methodology, however any other tool can be used to support. For example see Protégé [58].
  - Has already been applied to study patients with cancer [59].
  - It is rooted in the activities identified by the software development process proposed by the IEEE organization and other knowledge engineering methodologies [56].
  - Create ontologies from scratch or build on the reuse of existing ones.
  - Includes the identification of the development process of ontology (timing, control, quality assurance, knowledge acquisition).
  - Proposes the most accurate description of each activity performed.
  - Recommended for building ontologies by the Foundation for Intelligent Physical Agents (FIPA) [56], which promotes interoperability between applications based on agents Methodology.
  - It is the most comprehensive, documented and for being one of the most used compared with existing ones.
  - Proposes a lifecycle ontology building based on evolutionary prototyping, because it allows you to add, change and remove terms in each new version.

  For all these reasons it is considered that Methontology fits well to our objectives, allowing us to create an ontology which then serves as the basis of input to the process of text mining. The main characteristics of some existing ontologies are presented in Table 1.

Table 1. Brief comparison of some ontologies

Methodology	Year	Creator	Application	Process	Language	Re-use	Ontology
Cyc	1990	Lenat&Guha	Natural language processing	x	CycL	X	Cyc
Grüninger y Fox	1995	Grüninger y Fox	Enterprise	x	C++ Prolog	X	TOVE
Uschold y King	1995	Uschold y King	Enterprise	x	Ontolingua	X	Enterprise
Methontology	1998	Fernández-López	Domain	x	Ode WebOde Protegé OntoEdit Kaon	X	Medical Law Microbiology Education

- **Text Mining**: The classic text mining technique is divided in stages, specifically three that are pre-processing, representation and discovery [60]. In the pre-processing data are taken from the

ontology and are prepared before being processed. It can be treated as a filtering phase in which all texts are unified.

The objective of this stage is that the texts are translated as structured representations that facilitate analysis. Pre-processed texts arrive at the stage of representation . This step varies with the type of pre-processing selected. The result obtained in this stage will determine the discovery algorithm to be used.

At the stage of discovery algorithms, based on a structured representation of information, are able to find repeating patterns in text documents apply. All stages interact at a high level [61,62]. In the development process, each stage has a direct impact on the following ones. Figure 2 shows the overall model of the proposed Framework.
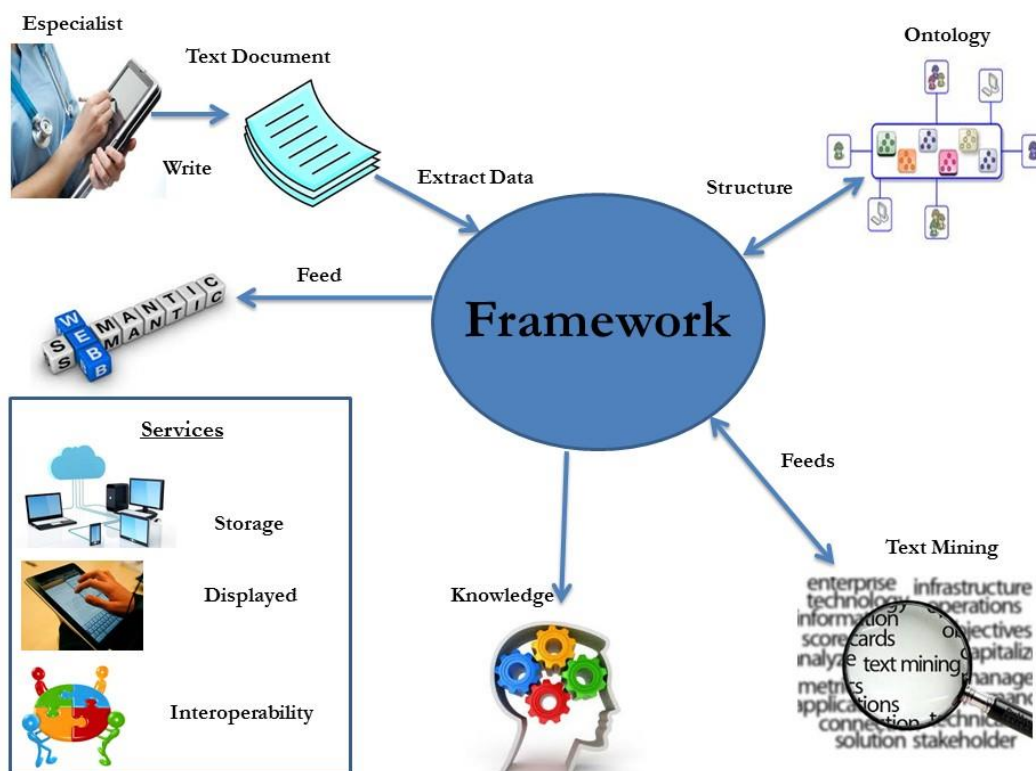

Figure 2. The overall model of the Framework

Like other fields of application, healthcare can benefit from the advantages of information technology [63]. Our main objective is to provide specialists and researchers with the processing of information. Specialists in palliative care have a way to work actively to manage the information enabling it to obtain relevant knowledge about patients with breast cancer.

Concluding Remarks

The development of the Framework is useful for acquiring knowledge about palliative care for patients with breast cancer in Panama.

The knowledge base that will manage the Framework may be used for future research in the field of palliative care of patients with breast cancer in Panama.

Developing ontology provides a basis for the development of other domain ontologies of breast cancer and can be reused by other researchers.

This Framework could be a model other types of cancers. Furthermore, the Framework will enable interoperability with other local or international similar platforms.

Conflict of Interest

The authors declare that they have no conflict of interest.

References

1. Saldaña J, Vargas-Lombardo M. eHealth Management Platform for Screening and Prediction of Down ' s Syndrome in the Republic of Panama. E-Health Telecommunication Systems and Networks 2014;3(3):33-42.
2. Fonseca J, Mora A, Barroso P. The web and the new generation of medical information systems. In: Taktar AFG, Fisher AC (Eds). Outcome Predict Cancer. Elsevier; Netherland 2007, pp. 391-
3. Frisse ME. Health information technology: One step at a time. Health Aff 2009;28(2):w379-w384.
4. Blaya JA, Fraser HSF, Holt B. E-health technologies show promise in developing countries. Health Aff 2010;29(2):244-251.
5. Clancey WJ. Situated cognition: On human knowledge and computer representations. Cambridge University Press; 1997.
6. Ardashkin IB, Popova A V. Knowledge Management as a Form of Student Initiative and a Tool to Increase Education Efficiency. Procedia-Social Behav Sci 2015;166:270-276.
7. Terzieva M. Project Knowledge Management: How Organizations Learn from Experience. Procedia Technol 2014;16:1086-1095.
8. Care P. Discussing palliative care with patients. Ann Intern Med 1999;130:744-749.
9. Justo N, Wilking N, Jönsson B, Luciani S, Cazap E. A review of breast cancer care and outcomes in Latin America. Oncologist 2013;46:248-256.
10. Galaviz KI, Harden SM, Smith E, Blackman KC, Berrey LM, Mama SK, et al. Physical activity promotion in Latin American populations: a systematic review on issues of internal and external validity. Int J Behav Nutr Phys Act [Internet] 2014;11(1):77. Available from: http://www.ijbnpa.org/content/11/1/77
11. Amadou A, Torres-Mejía G. Breast cancer in Latin America : global burden, patterns, and risk factors. Salud Publica Mex 2014;56(5):547-54.
12. Font A, Rodríguez E. Eficacia de la terapia de grupo en cáncer de mama: variaciones en la calidad de vida. Psicooncología 2004;1:67-86.
13. Hemminki K, Li X, Czene K. Cancer risks in first-generation immigrants to Sweden. Int J Cancer 2002;99:218-228.
14. Siziopikou KP. Ductal carcinoma in situ of the breast: current concepts and future directions. Arch Pathol Lab Med 2013;137:462-466.
15. Bhikoo R, Srinivasa S, Yu TC, Moss D, Hill AG. Systematic review of breast cancer biology in developing countries (part 1): Africa, the Middle East, Eastern Europe, Mexico, the Caribbean and South America. Cancers (Basel) 2011;3:2358-2381.
16. Parkin DM, Pisani P, Ferlay J. Global cancer statistics. CA Cancer J Clin 1999;49(2):133-164.
17. Kruijshaar ME, Barendregt JJ, van de Poll-Franse L V. Estimating the prevalence of breast cancer using a disease model: data problems and trends. Popul Health Metr 2003;1:5.
18. Panama Ministry of Health. Informe del Cáncer de Mama en Panamá [Internet]. 2014 [cited 2015 May 28]. Available from: http://190.34.154.93/rncp/sites/all/files/ca mama 2010_0.pdf
19. Panama Ministry of Health. Informe del Cáncer en Panamá [Internet]. 2015 [cited 2015 May 28]. Available from: http://190.34.154.93/rncp/sites/all/files/capma_2011.pdf
20. Reis A, Pedrosa A, Dourado M, Reis C. Information and Communication Technologies in Long-term and Palliative Care. Procedia Technol 2013;9:1303-1312.
21. Temel JS, Greer JA, Muzikansky A, Gallagher ER, Admane S, Jackson VA, et al. Early palliative care for patients with metastatic non-small-cell lung cancer. N Engl J Med 2010;363:733-742.

22. Lucas H. New technology and illness self-management: Potential relevance for resource-poor populations in Asia. Soc Sci Med 2014;1-9.
23. Ludwig W, Wolf KH, Duwenkamp C, Gusew N, Hellrung N, Marschollek M, et al. Health-enabling technologies for the elderly - An overview of services based on a literature review. Comput Methods Programs Biomed 2012;106(2):70-78.
24. Calvo GS, Gómez-Suárez C, Soriano JB, Zamora E, Gónzalez-Gamarra A, González-Béjar M, et al. A home telehealth program for patients with severe COPD: The PROMETE study. Respir Med 2014;108:453-462.
25. Rui L, Maode D. A Research on E - learning Resources Construction Based on Semantic Web. Phys Procedia 2012;25:1715-1719.
26. Lieto A. A Computational Framework for Concept Representation in Cognitive Systems and Architectures: Concepts as Heterogeneous Proxytypes. Procedia Comput Sci 2014;41:6-14.
27. Erickson P, Willke R, Burke L. A Concept Taxonomy and an Instrument Hierarchy: Tools for Establishing and Evaluating the Conceptual Framework of a Patient-Reported Outcome (PRO) Instrument as Applied to Product Labeling Claims. Value Heal 2009;12(8):1158-1167.
28. Choukri D. A New Distributed Expert System to Ontology Evaluation. Procedia Comput Sci 2014;37:48-55.
29. Bilgin G, Dikmen I, Birgonul MT. Ontology Evaluation: An Example of Delay Analysis. Procedia Eng 2014;85:61-68.
30. Miled AB. Reusing knowledge based on Ontology and Organizational Model. Procedia Comput Sci 2014;35:766-775.
31. Gruber TR. A translation approach to portable ontology specifications. Knowl Acquis 1993;5(2):199-220.
32. Gavrilova TA, Leshcheva IA. Ontology Design and Individual Cognitive Peculiarities: A Pilot Study. Expert Syst Appl 2015;42(8):3883-3892.
33. Zidi A, Bouhana A, Abed M, Fekih A. An Ontology-based Personalized Retrieval Model Using Case Base Reasoning. Procedia Comput Sci 2014;35:213-222.
34. Khan S, Safyan M. Semantic matching in hierarchical ontologies. J King Saud Univ Inf Sci 2014;26(3):247-257.
35. Jardim SVB. The Electronic Health Record and its Contribution to Healthcare Information Systems Interoperability. Procedia Technol 2013;9:940-948.
36. Benabdelhafid MS, Boufaida M. Toward a Better Interoperability of Enterprise Information Systems: A CPNs and Timed CPNs-based Web Service Interoperability Verification in a Choreography. Procedia Technol 2014;16:269-278.
37. Luther S, Berndt D, Finch D, Richardson M, Hickling E, Hickam D. Using statistical text mining to supplement the development of an ontology. J Biomed Inform 2011;44:S86-93.
38. Spasić I, Livsey J, Keane JA, Nenadić G. Text mining of cancer-related information: Review of current status and future directions. Int J Med Inform 2014;83(9):605-623.
39. Pletscher-Frankild S, Pallejà A, Tsafou K, Binder JX, Jensen LJ. DISEASES: Text mining and data integration of disease–gene associations. Methods. 2015;74:83-89.
40. Zhu F, Patumcharoenpol P, Zhang C, Yang Y, Chan J, Meechai A, et al. Biomedical text mining and its applications in cancer research. J Biomed Inform 2013;46(2):200-211.
41. Revere D, Turner AM, Madhavan A, Rambo N, Bugni PF, Kimball A, et al. Understanding the information needs of public health practitioners: a literature review to inform design of an interactive digital knowledge management system. J Biomed Inform 2007;40(4):410-421.
42. Allameh SM, Zare SM, Davoodi SMR. Examining the impact of KM enablers on knowledge management processes. Procedia Comput Sci 2011;3:1211-1223.
43. Avram M-G. Advantages and challenges of adopting cloud computing from an enterprise perspective. Procedia Technol 2014;12:529-534.
44. Shahzad F. State-of-the-art Survey on Cloud Computing Security Challenges, Approaches and Solutions. Procedia Comput Sci 2014;37:357-362.
45. Zrnec A. Software engineering in the cloud for reducing the application time-to-market. Elektroteh Vestn Electrotech Soc Slov Ljubljana 2011;78(3):123-127.

46. Weinhardt C, Anandasivam A, Blau B, Borissov N, Meinl T, Michalk W, et al. Cloud Computing – A Classification, Business Models, and Research Directions. Bus Inf Syst Eng 2009;1:391-399.

47. Hoogvliet MT. SaaS Interface Design. Design [Internet]. 2008; Available from: http://catalystresources.com/pdf/whitepaper_maartenhoogvliet_saasinterfacedesign.pdf

48. Andreadis G, Fourtounis G, Bouzakis K-D. Collaborative design in the era of cloud computing. Adv Eng Softw 2015;81:66-72.

49. Cazap E, Buzaid AC, Garbino C, De La Garza J, Orlandi FJ, Schwartsmann G, et al. Breast cancer in Latin America: Results of the Latin American and Caribbean Society of Medical Oncology/Breast Cancer Research Foundation expert survey. Cancer 2008;113:2359-2365.

50. Forouzanfar MH, Foreman KJ, Delossantos AM, Lozano R, Lopez AD, Murray CJL, et al. Breast and cervical cancer in 187 countries between 1980 and 2010: A systematic analysis. Lancet 2011;378(9801):1461-1484.

51. Luciani S, Cabanes A, Prieto-Lara E, Gawryszewski V. Cervical and female breast cancers in the Americas: current situation and opportunities for action. Bulletin of the World Health Organization 2013;604-609.

52. Sowa JF. Building large knowledge-based systems: Representation and inference in the cyc project. Artif Intell 1993;61:95-104.

53. Uschold M, King M. Towards a Methodology for Building Ontologies. Methodology 1995;80:275-280.

54. Grüninger M, Fox MS. Methodology for the Design and Evaluation of Ontologies. 1995;

55. Fernandez-Lopez M, Fernandez-Lopez M, Gomez-Perez a, Gomez-Perez a, Sierra JP, Sierra JP, et al. Building a Chemical Ontology Using Methonology and the Ontology Design Environment. IEEE Intell Syst 1999;14:37-46.

56. Corcho O, Fernández-López M, Gómez-Pérez A, López-Cima A. Building legal ontologies with METHONTOLOGY and WebODE. Lect Notes Comput Sci 2005;3369:142-157.

57. Arpírez JC, Corcho O. WEBODE in a Nutshell. AI Mag 2003;24(3):37-47.

58. Horridge M, Knublauch H, Rector A, Stevens R, Wroe C. Protege Tutorial, Uni Manchester. World. 2004; p. 117.

59. El-Subaihi JA, Anton F, Mioc D, Andersen SV. Child Cancer Follow-up Ontology and Information System. In: Rabase A, Brebbia CA, Bia A. (Eds) Data Management and Security : Applications in Medicine, Science and Engineering. WIT Press, 2013, 12 pages.

60. Missikoff M, Velardi P, Fabriani P. Text mining techniques to automatically enrich a domain ontology. Appl Intell 2003;18:323-340.

61. Hotho A, Nürnberger A, Paaß G. A Brief Survey of Text Mining. LDV Forum - Gld J Comput Linguist Lang Technol 2005;20:19-62.

62. Berry MW, Castellanos M. Survey of text mining. Comput Rev 2007;45:548.

63. Vargas-Lombardo M, Jipsion A, Vejarano R, Camargo I, Álvarez H, Mora EV, et al. Technologies to better serve the millions of diabetic patients: A holistic, interactive and persuasive ICT model to facilitate self care, in extremely poor rural zones of central America. J Med Syst 2012;36:595-600.