

Ontology Based Decision Support System for Youth Counseling

Ibukun T. Afolabi , Gbolahan E. Olujinmi and Raphael O. Nwokoye

Abstract— The amount of data residing in social media currently untapped is certainly limitless as millions of people are constantly posting one or more messages to public forums on the internet. Social media, for example, Twitter with over 320 million monthly active users has proven to be a fertile ground for harvesting opinions from multiple people. This project develops an ontology based decision support system that analyses social media content for the purpose of youth counseling.

For the development of this project, Java programming language and My Structured Query Language (MYSQL) was used to design the interface while Java programming Language, Stanford Natural Language Processor (NLP), Jena Application Program Interface (API) and Protégé Owl was used to develop the text analysis component of the system.

This project has been able to create a system that bridges gap between parents and their wards, as teenagers spend a lot of time online. It serves as a tool for parents to monitor the interactions of their wards with other social media users as well as understand the mental state of their children. It will also help youth become an effective member of the society.

Index Terms— Ontology, Decision Support System, Text analysis, Social Media, Youth, counseling

I. INTRODUCTION

The term “Social media” refers to the wide range of internet-based and mobile services which allow various users to take part in online exchanges, contribute user-created content, or join online communities (Dewing, 2010). Social Media can also be defined as the set of web-based broadcast technologies that enable the democratization of content, giving users the ability to emerge from consumers of content to publishers Nei *et al.* (2014). It is an online technology tool that enables people to speak. Other people use social media to share data, text, audio, video, images, podcasts, and different transmission communication. Engaging in numerous sorts of social media could be a routine activity which according to statistics, teach adolescents by enhancing communication,

Manuscript received July 1, 2017; revised July 23, 2017. This work was supported by Covenant University. I.T. Afolabi, Olujinmi G.E. and Nwokoye R.O. thanks Covenant University for their Sponsor and financial support towards the success of this publication.

I.T. Afolabi is a faculty in the department of Computer and Information sciences, Covenant University, Ota, Nigeria(corresponding author to provide phone: +234- 08021247616; e-mail:ibukun.fatudimu@covenantuniversity.edu.ng).

Olujinmi G.E. is in the department of Computer and Information sciences, Covenant University, Ota, Nigeria (e-mail:gbolahan.olujinmi@stu.cu.edu.ng)

Nwokoye R.O. is in the department of Computer and Information sciences, Covenant University, Ota, Nigeria (e-mail:raphael.nwokoye@stu.cu.edu.ng)

social connections and even practical skills. Social media has led to series of questions regarding its effect on the society, while it is already an accepted fact that social media affects people’s living style.

A lot of issues tend to arise when unlimited access is granted to people without thorough screening, as it is on various Social Media platforms. People, mostly youths, tend to take advantage of the fact that most of the contents and data they upload on social media platforms, be it good or harmful, cannot be sanctioned and condemned by law enforcement agencies set up against public misconducts.

Decision Support Systems are interactive, computer-based systems that help user in judgement and choice activities. Decision Support Systems provide data storage and retrieval but enhance the traditional information access and retrieval function with support for model building and model-based reasoning (Wu and Chan, 2009).

Decision Support Systems has been used in various areas such as management and planning in business, the health care sector, the military, and also areas where management will encounter complex decision situations. Decision support systems are mainly used for tactical and strategic decisions which are faced in situations where the time taken to think through and model the challenges is required to be short.

Youth Counselling is a learning-oriented process that takes place usually in an interactive relationship, sorely with the goal of helping a youth learn more about him or herself, and then use such understanding to enable the person become an effective member of the society (Ellis *et al.*, 2001). It is a means through which a youth can be shown care and concern by a helper or counsellor, which is usually an older person. This is done to bring about change in the life of the youth (Ellis *et al.*, 2001).

Both the concepts of Youth Counselling and Decision Support Systems go hand-in-hand, as Youth Counselling makes use of the Decision Support Systems to carefully guide youths. The application of the Decision Support System (DSS) helps to drastically minimize challenges and risks in the area of strategic choice making as regards Youth Counselling.

This paper therefore introduces an ontology based decision support module or plugin into a social media website in order to monitor youth behaviour on the social media. This is not trivial as there is an urgent need to keep tabs on and track the behaviour of youths on social media platforms, as it helps in curbing fast arising issues such as terrorism and many others. The system is intended to be used to infer the mental state of a user/ward, thus, aiding parents in monitoring their wards.

II. LITERATURE REVIEW

Decision support systems are developed to aid their users in making more realistic choices which are based on whatever parameters and criteria as regarding a certain decision that is meant to be taken. Various Decision Support Systems have been in existence for a number of years, each one having its own unique niche and areas they intend to help their users with, through the specific and particular services provided on such platforms. The list is endless, some of which include the following; Clinical Decision Support System Archimedes IndiGO (Parsania, 1992); SAP (Systems, Applications, and Products in data processing), (March and Garvin, 1996); The Analytica system which is a visual software environment that was developed by Lumina Decision Systems for building, analyzing, exploring and sharing quantitative decision models (Bhargava *et al.*, 1999); Paramount Decisions which is based on Choosing by Advantages (CBA) (Suhr, 1999); The CoBRA software currently provides a wide range of decision-making tools, databases, interactive maps, checklists, incident reporting capabilities and standard response protocols for hundreds of fire and emergency response, defence and law enforcement issues (<https://opecosystems.com/shop/item/cobra-incident-management-system>); (Liedgren *et al.*, 2016) in a review of literature discovered that decision support systems have had modest use in social work or youth counselling. They concluded that decision support systems focusing on a typical rather than typical cases, are perceived as the most useful among experienced staff.

Text analysis is an extension of knowledge mining that tries to search out patterns from giant unstructured sources. That is, it usually referred to as the method of extracting information from unstructured texts. There are several decision support systems based on text analysis which includes the following to mention a few; A hybrid system for sentiment analysis in twitter messages twitter (Wilson *et al.*, 2009). This system adopts a hybrid rule-based, lexicon-based and machine learning techniques. Konig & Brill, 2006 proposed a hybrid classifier approach that exploits human reasoning over mechanically learnt text patterns to match machine learning which proved highly effective and was more effective than other classifiers that were evaluated. The AZFinText System was developed for the textual analysis of stock market prediction using breaking financial News. This structure uses an analytical machine learning method for financial news articles analysis via numerous diverse textual signs such as Bag of words, Named Entities and Noun phrases. (Schumaker & Chen, 2006). In addition to the above, the following are some of the text analysis based decision support systems; SAS Text Analytics System and Megaputer Polyanalyst (Qingu & Richard, 2008), PASW (Predictive analytics Software) modeller which makes it straightforward to get insights in your information with an easy graphical interface (<http://www.spss.com.hk/software/modeler/>).

As a result the review of existing decision support systems, it is clear and obvious that each one aims at improving the quality of the decisions taken by its users or subscribers, while also making the whole process of decision making less stressful for the decision makers and far easier, following a guided set of parameters and criteria set in place by each

system. Each decision support system is unique in its own way based on the varying methods of operation used in carrying out the decision making process. Also, from the review, it was discovered that text analysis is not a new topic to scientists as it has been applied in various areas e.g. career counselling or just counselling in general. It has also drawn an incredible amount of attention over the recent years because of the future possible application of text analysis either by individuals, corporations etc. A text analysis system can only be complete if it is able to identify small as well as large-scale patterns, understand context and, from content, to deduce the meaning, and finally build models. Currently, most text analysis based decision support systems do not do much semantic analysis particularly employing the use of ontologies for semantic analysis. This paper therefore aims at describing the process used in developing an ontology based decision support system for youth counselling.

III. METHODOLOGY

The methodology used in this research is based the combination of software development process for the social media site and ontology based text analysis process for the counselling plugin of the social media site. For the social media site developed, the functional requirement includes registration, sign up, login and logout. Users are also able to upload/post feeds or messages, view and update his or her profile, add and accept new friends, send instant chats to his or her friends privately and share e images and pictures on the site and much more. In developing the social media site, the Use case diagram, class diagram, sequence diagram, activity diagram and so on as used for the system analysis and it was implemented using the combination of HTML, CSS and Java programming language.

The decision support module consists of the slang ontology and the programming module used to analyze the social media text. The slang ontology was developed using the steps described in (Noy & McGuinness, 2000) which is summarized as follows;

Determine the domain and scope of the ontology: This was done by simply determining what domain the ontology will cover as well as users of the ontology. For this project the domain of the ontology was social media text particularly the uniqueness of its contents which includes slangs and acronyms.

Select textual source to use for the ontology: This is the process of discovering what type of words the ontology is made up of. It includes finding the right documents that those words will be gotten from. For example, the words in this slang ontology were gotten from different websites that have the different teenage slangs and present social media slangs. For this research, documents were gathered online and from social media sites.

Enumerate important terms in the ontology: That is the explanation of the various important terms, statements so that users can relate to it and understand.

Build the ontology preferably using the following steps

- Extract relevant terms for text
- Define the concepts, taxonomies, relations, attributes, function, instances, etc.
- Design an application interface or platform that the ontology can be accessed and queried.

Figure 1 describes the activity workflow of creating the ontology. The first step determines the domain and scope of the ontology. The source of the slang ontology are various teenager slang documents as well as different abbreviations used by teenagers in social media. Writing down the important slangs to be input in the ontology form the third step. The fourth step is to extract the relevant term from text and it was achieved by using the information extraction tools called Analog X keyword extractor. The relevant terms were chosen from the document corpus using the various information extraction processes such as stop word removal and stemming. In the stop word removal, common words with no meaning were eliminated. In the stemming process words that are related to each other or syntactic-similar were identified. After these text pre-processing steps, the relevant concepts and properties are then selected from the extracted keywords and organized into categories. This is known as the top down development process. To implement the slang ontology, protégé owl was used. The graphic user interface that was used to access and query the ontology was developed using the java NetBeans.

The text analysis based DSS module also performed the following text pre-processing steps on the extracted comments from users of the social media site.

Stop word Removal: This is the elimination of common words with no meaning as well as any related information to the task (Gupta and Lehal, 2009)

Stemming: This is the process of identify those words that are related to each other or that are syntactic-similar. This is done to obtain the stem or radix of each word, which emphasize its semantics. (Gupta and Lehal, 2009)

TF/IDF: Inverse Document Frequency- an Inverse Document Frequency (IDF) is a statistical weight that intends to reflect how important a word is to a document. IDF feature is combined which reduces the weight of terms that occur often in the document set and increases the weight of terms that occur rarely. (Vijayarani *et al.*,2014)

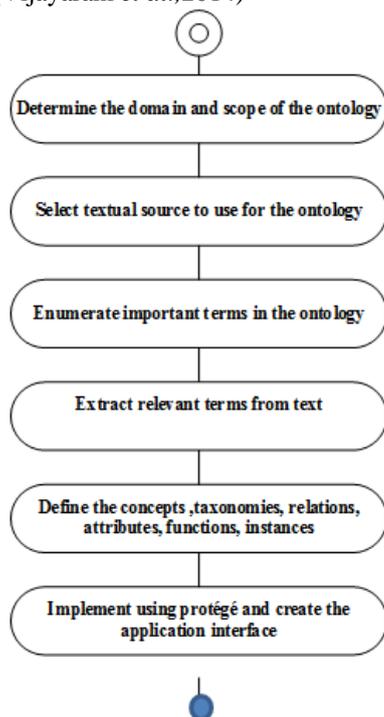


Figure 1: activity workflow of creating the ontology

For the text analysis plugin, posts and messages on the site are accessible to a text analyser which uses the Jena API to interface with the Protégé owl. Upon activated, the text analyser retrieves all the text of a particular user and runs it through a programming module that compares it with the contents of the ontology to determine the percentage positivity or negativity based on the developed slang ontology. Both the social media site and decision support system module which is also known as the text analysis plugin is implemented using the Java Programming language.

IV. ONTOLOGY BASED DSS FOR YOUTH COUNSELLING DESCRIPTION

The slang Ontology: The slang ontology can be defined as a data model that represents slangs as a set of concepts within a domain and the relationship between these concepts. It is also a form of slang management that captures the slangs within a society as a model. A slang ontology is a set of concepts-such as various slangs that are in some specific way in other to create an agreed-upon vocabulary for exchanging information. The slang ontology developed for this research is a semantic representation of information of the various slangs used by teenagers in the 21st century. The main users of this ontology are Text analytic applications: software application that needs to leverage on different slangs as well as their interpretations and Domain Experts: This is the person responsible for the update of the concepts and relationships in the slang ontology as well as its maintenance Structuring and representation of knowledge is the main concern of knowledge engineering (Sowa, 2000). The slang ontology will serve as a strong backbone or a central ingredient to be used by the domain expert to build a robust text analysis platform. Slang ontology is required to be extensible without having to change its model and assumption. This is because it will serve as a basis for the addition of new concepts to the ontology.

The slang ontology contains concepts, object axioms and 1660 individuals and it is described in detail as follows;

Positive slang: the concepts positive slang describes all the slang that is used as a sign of joy or happiness. It also includes those used as signs of well wishes etc.

Negative slangs: The concept negative slang describes all the slang that is used as a sign of depression, sexual or ill fare etc.

Neutral slang: these are slangs that are neither used for well wishes or ill fare.

Interpretations: These are the meaning of the various slangs and acronyms

The description of the main constructs used to represent the Slang ontology is described thus;

- **rdfs:label:** Used to store the general or popularly known name to refer to the concept or individual. For example, the individual LOL is labeled as Laugh Out Loud.
- **rdfs:DefinedBy :**Used to store the definition of the term. For example, the individual LOL can be defined as expressing great joy or making fun of someone.
- **rdfs:Comment:** Captures the further explanation on the term or individual. For example, the

individual LOL can be further explained as used in situations where there is a need to remove tension.

- owl:sameAs: Used to equate similar individuals.

The ontology based- DSS contains the following modules, their functions and how they work together in supporting counselors in advising youths.

Registration and Login Module: This module provides the non-existing user and incoming visitors an interactive interface in order to become a member and a registered user on the platform. The module takes in the necessary inputs for registration and implements the necessary access control and security features. The Login button provides a link to the login servlet which establishes a connection with the database so as to authenticate the users' credentials that have been inputted.

Live Feeds, Notification and Messages Module: This module provides the user an attractive display of the platform's content which has to do with the live posts and feeds to be uploaded by a user. It shows the posts of other users, friends and further allows users to search for other users on the platform and add them as friends. It displays all the notifications a user has such as the posts sent to him or messages sent to him which were on the livefeeds section and not the messaging section itself. It shows the messages and the user who sent it. It allows a user to receive and reply messages sent to them by other users on the platform. It shows the user and the time at which the message was sent and allows for easy reply to the sender. Figure 2 is the live feeds page (Landing page) and is the first page the user encounters upon login into the system.

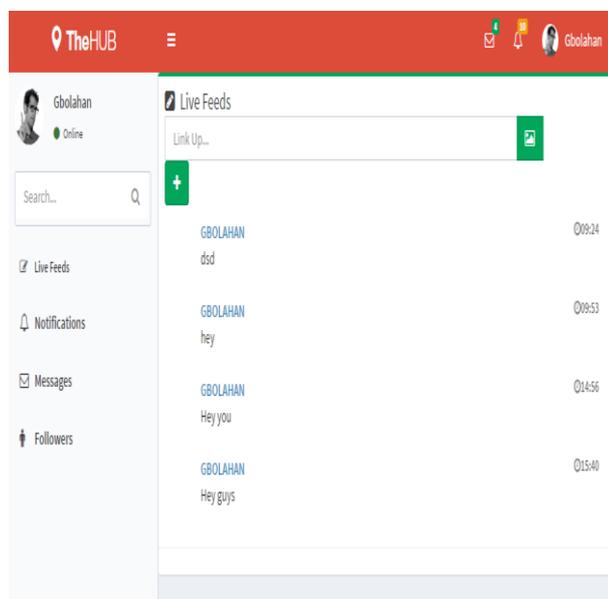


Figure 2: Screenshot showing the landing page which is the Live Feeds page

Figure 3 is the page that displays all the notifications a user has such as the posts sent to him or messages sent to him which were on the livefeeds section and not the messaging section itself. It shows the messages and the user who sent it as well.

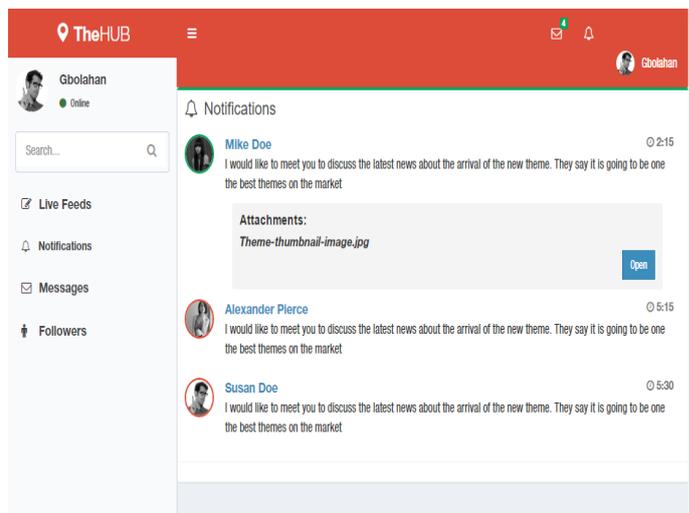


Figure 3: Screenshot showing the Notification page

Figure 4 shows the page that allows a user to receive and reply messages sent to them by other users on the platform. It shows the user and the time at which the message was sent and allows for easy reply to the sender.

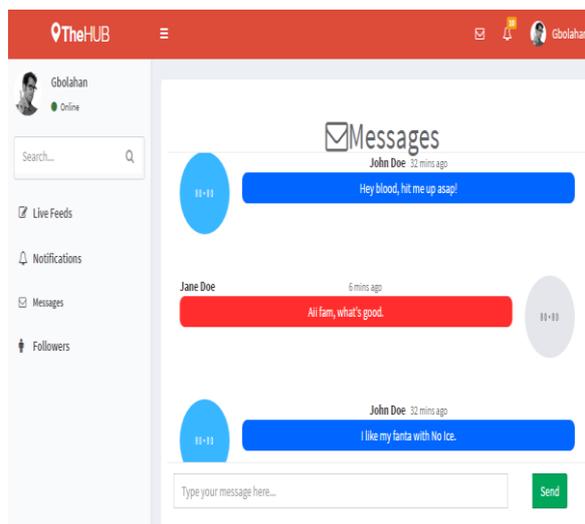
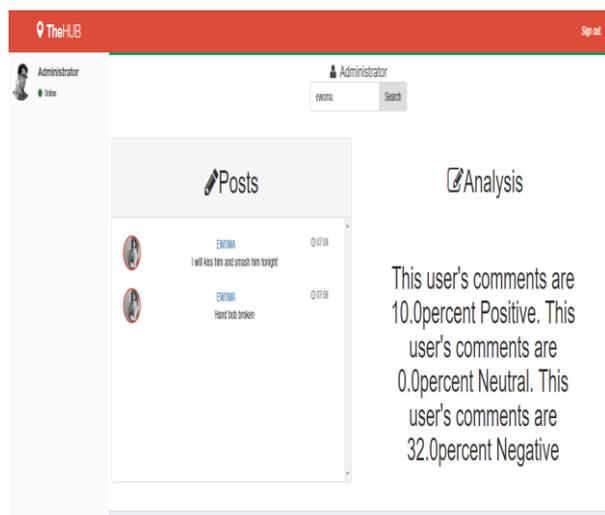


Figure 4: Screenshot showing the Messages Page

The Text analysis module: this module retrieves the messages or textual content stored by Live feeds, notification and message module with a unique identification, i.e. the textual messages belonging to each individual is stored uniquely in a database and combined. It then passes it through and a text preprocessing phase which performs stopword removal, stemming, TF-IDF and the result of this is the extracted keywords which represent the focus of the individual textual content. With the use of the JENA API, these keywords are compared with the contents of the ontology to analyze them semantically and determine the degree of negativity or positivity of the individual textual content. The result of this comparison is then stored and a function is written to determine the percentage positivity or negativity of the individuals textual content. The result of this is displayed in Figure 5. This text analysis module is only accessible to special category of users who are the counselors. Upon logging in and accessing the text analysis module, these categories of users they are able to run the

analysis, either for a particularly queried individual or for several individuals that they are interested in counseling.



V. CONCLUSION AND FURTHER STUDY

In conclusion, the developed system has provided a semantic approach to analysing social media text. In particular it has been able to address one of the concerns in youth counselling which is to focus on a typical rather than typical case by providing a decision support system that is able to single out a particular youth and identify the problem in particular.

Also the ontology based decision support system described in this project can be of tremendous benefit to schools, parents, researchers and anybody trying to guide youths unto the right path. It will also help to reduce online social misconducts, cybercrime and violence. For further work, the system can be improved by Creation of location indication/real-time mapping: This feature would allow for real-time mapping of an individual on the platform as regards his or her time of posting feeds on the system. This would allow a for easy locating in cases such as suicide or even kidnapping, it will further help in the creating and fostering relations of mini social communities or groups with the system. Also the ontology contents can be improved upon to take care of more concepts, individuals and relationships.

REFERENCES

- [1] Bhargava, H. K., Sridhar, S., and Herrick, C. (1999). Beyond spreadsheets: tools for building decision support systems. *Computer*, 32(3), 31–39.
- [2] Dewing, M. (2010). *Federal government policy on arts and culture*. Ottawa: Library of Parliament. (<https://lop.parl.ca/Content/LOP/ResearchPublications/2008-41-e.pdf>)
- [3] Ellis, A., Gordon, J., Neenan, M., and Palmer, S. (2001). *Stress counselling: A rational emotive behaviour approach*. Sage.
- [4] Gupta V. and Lehal G. S., “A Survey of Text Mining Techniques and Applications”, *Journal of Emerging Technologies in Web Intelligence*, vol. 1, no. 1, (2009).
- [5] Konig A. C. and Brill E. (2006) “Reducing the human overhead in text categorization” **KDD '06 Proceedings**

- [6] Liedgren P., Elvhage G., Ehrenberg A. & Kullberg C (2015) The Use of Decision Support Systems in Social Work: A Scoping Study Literature Review in *Journal of Evidence-Informed Social Work* Vol. 13 , Iss. 1, 2016
- [7] March, A., and Garvin, D. (1996). *SAP America. Boston: Harvard Business School.*
- [7] Nei, L. C., Cherng, L. Y., and Singh, M. M. (2014). A Case Study on Clickjacking Attack and Location Leakage. *International Journal of Scientific and Engineering Research*, 5(7), 190-199.
- [8] Noy N.F. and McGuinness D.L., "Ontology Development 101: A Guide to Creating Your First Ontology". http://protege.stanford.edu/publications/ontology_development/ontology101.pdf , 2000
- [9] Parsania, V. S., Kamani, K., Kamani, G. J., and Jani, N. N. (1992). *Reviewing and Modeling Clinical Decision Support System.*
- [10] Richard S. Segall, Qingyu Zhang , Mei Cao “ *Web-Based Text Mining of Hotel Customer Comments Using SAS® Text Miner and Megaputer Polyanalyst®* “<http://www.docstoc.com/docs/43446814/WEB-BASED-TEXTMINING-OF-HOTEL-CUSTOMER-COMMENTS-USING-SASTEXT>
- [11] Schumaker R.P. and Chen H. (2009)Textual analysis of stock market prediction using breaking financial news: The AZFin text system in *ACM Transactions on Information Systems (TOIS)*, Volume 27 Issue 2, No. 12
- [12] Shur J.(1999) “The Choosing By Advanatges Decision System” Published by quorum Books, United States of America.
- [13] Vijayarani S, "Preprocessing Techniques for Text Mining - An Overview", *International Journal of Computer Science & Communication Networks*, vol. 5, no. 1, pp. 7-16, 2014
- [14] Wilson T., Wiebe J. and Hoffmann P (2009) “Recognizing Contextual Polarity: An Exploration of Features for Phrase-Level Sentiment Analysis” in *Computational Linguistics*, Volume 35, Issue 3 Pp.399-433
- [15] Wu, Y., and Chan, C. W. (2009). A data analysis decision support system for the carbon dioxide capture process. *Expert Systems with Applications*, 36(6), 9949–9960.