



USABILITY OF A FILIPINO LANGUAGE TOOLS WEBSITE

Ria A. Sagum, MCS

Department of Computer Science, College of Computer and Information Sciences

Polytechnic University of the Philippines, Manila, Philippines

Email: rasagum@pup.edu.ph

Abstract

Natural language processing (NLP) assists in learning a natural language in the educational setting by developing scientific approaches using computer and internet for improvement of learning. Language tools are necessary to improve the output of different researches in this field. Different natural languages has different characteristics and rules that needs to be considered in developing their tools for language processing. The researcher introduced an NLP website that contains different language tools for Filipino language that may be used by students, specifically from Polytechnic University of the Philippines, to make it easier for them to develop the NLP system that necessitate to use these tools. This website was evaluated for its usability, the users who answered the survey agreed that the website is useful in their work in the field.

Keywords – Natural Language Processing, NLP Tools, Filipino Language Tools

I. INTRODUCTION AND ITS BACKGROUND

People communicate in different ways: through speaking and listening, making gestures, using specialized hand signals (such as when driving or directing traffic), using sign languages for the deaf, or through various forms of text. By text we mean words that are written or printed on a flat surface (paper, card, street signs and so on) or displayed on a screen or electronic device in order to be read by their intended recipient (or by whoever happens to be passing by). [1]

The field of Natural Language Processing (NLP) aims to convert human language into a formal representation that is easy for computers to manipulate. Current end applications include information extraction, machine translation, summarization, search and human-computer interfaces. [2]

Natural language process is an effective process to assist students in the process of scientific learning. Implementing NLP in the educational setting not only helps in developing effective language process, but it is also significant for enhancing the academic performance. The NLP techniques follow the approach of the natural process of language acquisition integrated with the scientific approach of using computer programs. [3]

NLP is a major factor associated with the branch of science, which focus on the development and improvement in the process of learning. It provides theoretical grounds to assist in the process of developing techniques and effective approaches for providing assistance in the scientific learning by utilizing the effective theories and approaches. NLP can be effectively applied in the education for promoting the

language learning and enhancing the academic performance of the students. In education setting, NLP, can assists in developing effective process of learning by developing scientific approaches by using computer and internet. In order to provide assistance, there exists different tools and approaches in enhancing language processing to ensure that students can easily develop understanding and use of a particular natural language. [4]

Different applications can be developed with the knowledge of NLP. The undergraduate/graduate students in different universities in the Philippines are now pursuing this field using the country's language, Filipino. However, the researcher observed that much of their times are being spent in developing language processing tools that is needed before they can proceed with the main process of their work. Different researchers in Filipino language tools has been made but were not available anywhere.

With this, the researcher introduced an NLP website wherein students, specifically students of Polytechnic University of the Philippines, can make use of the different NLP Tools developed by the NLP-Special Interest Group of the university. These language tools are mostly for Filipino language.

II. RELATED WORKS

Natural Language Processing (NLP) is an area of research and application that explores how computers can be used to understand and manipulate natural language text or speech to do useful things. NLP researchers aim to gather knowledge on how human beings understand and use language so that appropriate tools and techniques can be developed to make computer systems understand and

manipulate natural languages to perform the desired tasks. The foundations of NLP lie in a number of disciplines, viz. computer and information sciences, linguistics, mathematics, electrical and electronic engineering, artificial intelligence and robotics, psychology, etc. Applications of NLP include a number of fields of studies, such as machine translation, natural language text processing and summarization, user interfaces, multilingual and cross language information retrieval (CLIR), speech recognition, artificial intelligence and expert systems, and so on. [5]

Natural Language Processing and Programming Languages are both established areas in the field of Computer Science, each of them with a long research tradition. Although they are both centered around a common theme – “languages” – over the years, there has been only little interaction (if any) between them. The paper by Mihalcea, Rada et al tried to address the gap by proposing a system that attempts to convert natural language text into computer programs. While the researchers overview the features of a natural language programming system that attempts to tackle both the descriptive and procedural programming paradigms, in this paper researchers focused on the aspects related to procedural programming. Starting with an English text, they showed how a natural language programming system can automatically identify steps, loops, and comments, and convert them into a program skeleton that can be used as a starting point for writing a computer program, expected to be particularly useful for those who begin learning how to program. [7] Before the process of any application in NLP, like information extraction, machine translation, summarization and searching will be developed, there is a need to prepare the data to be used in their application. There is a need to do what is often referred to as “Pre-processing” activities in any NLP based application. These pre-processing activities includes, tokenization, named entity recognition, part-of-speech tagging, and stemming to name a few. Because of the need to include these process in NLP program development there are some individuals who published their work and share their tools online and there are universities who also shares their work to help researchers in this field lessen their time in developing their work. The NLP Groups’ web site of Stanford University is one of the examples.

The Natural Language Processing Group at Stanford University is a team of faculty, postdocs, programmers and students who work together on algorithms that allow computers to process and understand human languages. Their work ranges from basic research in computational linguistics to key applications in human language technology, and covers areas such as sentence understanding, automatic question answering, machine translation, syntactic parsing and tagging, sentiment analysis, and models of text and visual scenes, as well as applications

of natural language processing to the digital humanities and computational social sciences. [8] With this website different NLP researchers were able to download different language tools for English (although it seems that some of their tools works in some other natural language) and made use of this to do the process they would like to include in their application. This allow researchers to focus the development of their NLP systems on their main contribution in the field rather than dividing their time with the development of pre-processing activities.

III. METHODOLOGY

The website was designed to suit the needs of the would be users of the tools, keeping in mind that the tools must be portable enough to be downloaded and ready to be used as a method in their system development.

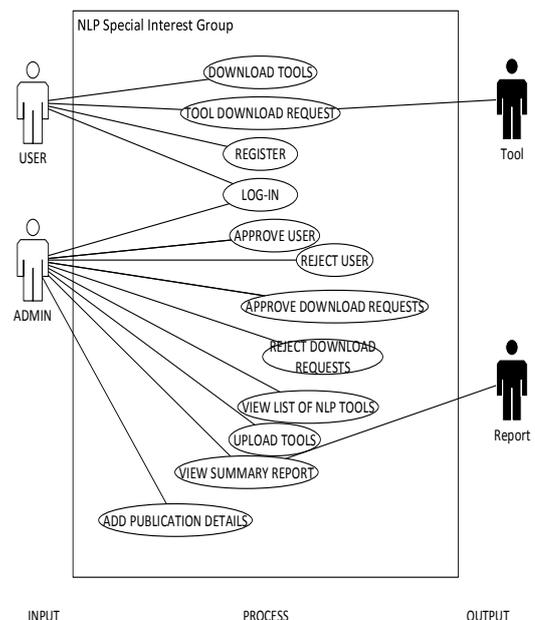


Figure 1. Use-case Diagram

Figure 1 shows the use-case diagram for NLPSIG. Follows the Use-Case Realizations.

3.1 Use-Case Realizations

3.1.1 User Register

- Brief Description
 - In this use case the system allows user to register before going directly to the website.
- Actors
 - Users
- Basic Flow of Events:



- This use case begins with the system requiring the users to register.
- End of use case.

3.1.2 *Approve or Reject User Registration*

- Brief Description
 - In this use case, the system allows the Admin to decide whether to accept or reject the user.
- Actors
 - Admin
- Basic Flow of Events:
 - This use case begins when user registers.
 - The admin verifies if the user is legible to visit the website. If so, the admin will accept the registration.
 - The admin will reject if otherwise.
 - End of use case.

3.1.3 *Accept or Reject Tool Download*

- Brief Description
 - In this use case the user requests for a tool to be downloaded.
 - The admin is to decide whether to accept or reject the request.
- Actors
 - Users
- Basic Flow of Events:
 - The system begins when a user requests for a tool download.
 - The admin will approve or deny the request.
 - If approved, the user can then download the tool.
 - If not, user may or may not request again.
 - End of use case.

3.1.4 *Add Tools*

- Brief Description
 - The purpose of this use-case is for the system to add or update the existing tools.
- Actors
 - Admin
- Basic Flow of Events:
 - The system begins when an admin adds tools.
 - End of use-case.

3.1.5 *Add Publication Detail*

- Brief Description
 - The purpose of this use-case is for the system to add or update the existing tools.
- Actors

- Admin

- Basic Flow of Events:

- The system begins when an admin adds publication details.
- End of use-case.

IV. RESULTS AND DISCUSSION

After the development of the website the researcher tested for the usability of this with the use of SUMI. Software Usability Measurement Inventory (SUMI) is a solution to the recurring problem of measuring users' perception of the usability of software. It provides a valid and reliable method for the comparison of products and different versions of the same products, as well as providing diagnostic information for future developments. SUMI provides an objective way of assessing user satisfaction with the software. The following are some example questions provided in SUMI.

- The way the tool information is presented is clear and understandable.
- There is enough information on the screen when it is needed.
- I think this tool is consistent.
- I can understand and act on the information provided by this tool.
- This tool is the norm when I want to do something which is not standard.
- There are fewer instructions to be read before you can use the tool.
- Tasks can be performed in a straightforward manner using this tool.

The questionnaire was designed to measure the efficiencies, simplicity, helpfulness and control of a product. During its development, the questionnaire was standardized as a measurement tool for some of the user orientated requirements expressed in the European Directive on Minimum Health and Safety Requirements for Work with Display Screen Equipment (90/270/EEC). [9]

SUMI was also mentioned in the ISO 9241 standard as a recognized method of testing user satisfaction [10].

For the analysis purpose, certain value was assigned to each scale as shown in Table 1. The average is gathered from the summation of input from the respondents. If the result is between 1.00 and 1.99, it means that users were satisfied and agree on the usability of the website. If the result is between 2.00 and 3.00, it shows that the users were



not sure (undecided) otherwise users totally disagree that the website is usable. Table 2 shows the details of the result.

Table 1: Attribute in Evaluation Form

Scale	Agree	Undecided	Disagree
Value	1.00-1.99	2.00-2.99	3.00-4.00

Table 2 Average Result for every question

Question	Result (Average)
1. This software responds fast enough to inputs.	1.75
2. Learning to navigate around the website is hassle-free to the users	1.92
3. The way that system information is presented is clear and understandable.	1.73
4. The information is helpful	1.58
5. The tool is consistent	1.92
6. Less to Read	1.67
7. Organization of the menu is logical	1.77
8. Easy to make the tool do exactly what we want	1.87
9. Accurate Result	1.80
10. There is enough Information	2.33
11. The speed of the website is fast enough.	1.91
12. There are too many steps required to get something to work.	2.00
13. Error prevention messages are not adequate.	2.48
14. I can understand and act on the information provided by the software.	1.81
15. I would recommend this software to my colleagues.	1.97
16. I know what I'm doing when using the website.	1.70
17. The software has very attractive presentation.	2.03
18. The website met my expectations.	1.80
19. Getting data files in and out of the system is easy.	1.60
20. I did not experience any difficulties when I use the website.	1.98
General Average	1.88

The questions were divided into three different choices: Agree, Disagree, and Undecided. Out of the 20 questions, the users agreed that the response time of the website is fast enough, that navigating around the website was hassle-free for them, that they understand the information presented in the website, that the information is helpful, that the tool is consistent, that the information in the website is less to read, that the organization of the menu is logical, that the speed of the website is fast enough for them, that they would recommend the website to their colleagues, and that they did not experience when they used the website. The users were undecided if the website gave enough information and on the attractiveness of the presentation of the website.

A summary of the result as perceived by the users as to the website's efficiency, simplicity, helpfulness and control of the product was presented in Table 3.

Table 3: Summary of the Usability of the website as perceived by the users

Criteria	Result	Verbal Interpretation
Efficiency	1.82	Agree
Simplicity	1.87	Agree
Helpfulness	1.87	Agree
Control of a product	1.93	Agree
General Average	1.87	Agree

The answers to the questions were classified to belong to the criteria Efficiency (5,9,18,19,20), Simplicity (2,3,6,7,12,17,20), Helpfulness (4,10, 15, 19), and Control of the product (1,8,11,13,15,16,20). As seen in Table 3, the users agreed that the website passed all the criteria for usability. But it must be noted that according to the average per each questions there exists the users' answers of undecided to the website's functionality and the information being provided by it.

Overall, the users were satisfied with their experience navigating through the website and they consider that the website is useful and helpful to them.

V. CONCLUSION AND RECOMMENDATIONS

The purpose of developing an NLP Website is to create an easier way for students, or anyone interested in NLP, to proceed with their NLP application development without worrying about the pre-processing activities. The findings and the evaluation result of the NLP Website provides important information to the researcher in terms of usability and satisfaction of the user when they use the website for learning about NLP. In conclusion, SUMI analysis shows that the NLP Website provides a good interface and being useful and fast enough for the users.

Based on the results gathered, the researcher was able to prove that the system provides a fast and reliable website.

Other researchers may expand the system further by

- Providing a graph in the PDF report to represent number of downloads per week,
- Making the PDF compatible to other browsers other than Mozilla Firefox,
- Making sure that the files in the download page are secured,
- Connecting the user's e-mail to the website for them to be notified if their requests were approved or denied.



ACKNOWLEDGEMENT

The researcher would like to acknowledge the help of the NLP tools group in providing their assistance in implementing the research. The CCIS Family in supporting the research activities of their faculty members. To her Family who is always her strength and inspirations in life and to GOD almighty for giving her the wisdom to do a research.

REFERENCES

- [1] R. Kibble (2013) "Introduction to Natural Language Processing"
http://www.londoninternational.ac.uk/sites/default/files/computing-samples/co3354_ch1-3.pdf
- [2] Dr. Khaled M. Alhawiti (2014) "Natural Language Processing and its Use in Education" (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 5, No. 12, 2014
- [3] Nadkarni, P. M., Ohno-Machado, L., & Chapman, W. W. (2011). Natural language processing: an introduction. Journal of the American Medical Informatics Association, 18(5), 544-551
- [4] Chowdhury, Gobinda G., Natural Language Processing
- [5] Mihalcea, Rada et al, NLP (Natural Language Processing) for NLP (Natural Language Programming), 2006
- [6] Zhang, N. (2001). Stanford NLP Tools
http://nlp.stanford.edu/courses/cs224n/2001/nzhang/NancyZhang_CS224nFinalProject.pdf
- [7] <https://www.alsintl.com/resources/languages/Filipino>
- [8] Bird, S. et. al. (2014). Natural Language Processing with Python
<http://7.ExtractingInformationfromText.html>
- [9] Manor, Z. et. al. (2012). The Evaluation of WebCost Using Software Usability Measurement Inventory.
- [10] Tanja, R. and Borka, J.B. A Case Study of Usability Testing – the SUMI Evaluation Approach of the EducaNext Portal. WSEAS Transactions on Information Science & Applications. 2(5), February 2008