Utilization of NLP-Technology in Current Applications for Education and Research by Indonesian Student, Teacher, and Lecturer

M. Kharis
Universitas Negeri Surabaya
Universitas Negeri Malang

Kisyani Laksono
Universitas Negeri Surabaya

Suhartono
Universitas Negeri Surabaya

This study aims to reveal the use of software that works with Natural Language Processing (NLP) technology by students, teachers, and lecturers in Indonesia. A total of 166 respondents from 51 institutions in 12 provinces in Indonesia have participated in data collection through questionnaires and interviews. Questionnaire data were presented in percentage form, while interview data were analyzed using the purposive sampling technique. The analysis results show that most students are unfamiliar with NLP but have used the technology in various applications. Based on the data analysis, students and lecturers have commonly used Google translate and Grammarly. However, PoP and elicit applications useful for searching metadata for scientific papers are not widely known. VOS-Viewer and Nvivo as qualitative data processing applications have not been widely used. Thus, it is crucial to conduct training on using the applications mentioned.

Keywords: NLP technology, current application, education

INTRODUCTION

In addition to teaching, students, teachers, and lecturers in Indonesia are currently required to conduct research and publish their results in national and international journals. For their assignments and work, the application of Natural Language Processing (NLP) technology in the form of computer applications should be able to assist them in education, research, and publication. The critical question is, how are students, teachers, and lecturers mastered using NLP technology in their assignments and work.

NLP is a field of computer science that links computer and human-machine communication through human language analysis. This method is developed to recognize and process human language automatically or semi-automatically in computers, and this process is referred to as computational linguistics (CL). NLP processing is artificial intelligence designed to allow computers to understand, interpret, and manipulate human language. NLP technology is integrated into various disciplines, including

computer science and CL, to create communication between humans and computers (Goesderilidar, 2021). NLP technology has been realized in various computer applications for education and research. Google Translate, Google Search, speech recognition, Chatbot, and other Artificial Intelligence (AI) machine applications are some applications that use NLP and are related to the use of human language, both oral and written.

Some researchers argue that NLP and CL are used interchangeably, which refers to the same meaning. In short, it can be stated that CL is a combined science between computer science and linguistics. This science seeks to find a computational model for linguistics to answer the question of how human language (natural language) can be processed automatically by computers (Dwiastuti, 2019).

Thus, understanding and interaction of computers with users depends on how well the language is arranged at the level of morphology, syntax, semantics, phonetics, and grammar by humans to computers. The better the settings and input gave to the system in the computer, the better the computer understands human language. Thus, NLP techniques follow the pattern of human grammar because the NLP works and refers to techniques for manipulating and analyzing human language input.

The number of human language texts continues to grow, are interconnected with one another, and contain excellent knowledge. However, the larger the number of texts, the more difficult it is for humans to disseminate the text to find its knowledge. An NLP is expected to do this assignment as effectively and accurately as humans (Chowdhary, 2020).

With NLP technology, computers can read text, hear, speak, interpret, measure sentiment, and determine important parts of sentences. NLP is used for tokenization and parsing, lemmatization/stemming, part-of-speech tagging, language detection, and identification of semantic relationships. Thus, it can be stated that NLP breaks language into pieces of shorter sentence elements, then understands the relationships between these pieces, links them together, and works together to create meaning (Rosyadi et al., 2020). An example of a technology application that utilizes the role and function of NLP in its linguistic process is the development research by Kisyani et al. (2021). This research has developed a machine translation device that utilizes speech-to-text technology, which is named Takarir [7] and has been evaluated by Kharis et al. in their research (Kharis et al., 2021). Other applications using NLP technology that are discussed in this paper include Google Translate (GT), Grammarly, Nvivo, VoS-Viewer, Chatbot, Quillbot, Publish or Perish, and Elicit. Some of these applications can run on the desktop or online. Referring to the development of NLP application technology, students, teachers, and lecturers should use NLP-based applications to help their assignments so that they are faster and more precise.

LITERATURE REVIEW

Using NLP in Several Applications

The main goal of NLP technology is to make machines that can understand the meaning of human language and then provide appropriate responses, such as speech recognition, text-to-speech, voice commands, and machine translation (Constantin et al., 2020). The development of technology by utilizing the NLP work system has been carried out by Migunani et al., who developed a virtual teacher chatbot application. With this application, virtual teachers help students learn, especially outside school (Migunani & Aditama, 2020). With the NLP tools, building a chatbot from scratch has become a slightly more straightforward process (Sophia et al., 2020).

Another application is VOS-Viewer, which is an application that can generate visualizations to reveal research collaborations and research trends by connecting subjects and research objects. VOS-Viewer may also be used as a mapping application for bibliometric data analysis (Al Husaeni & Nandiyanto, 2021). This pattern of research publication by lecturers, students, and researchers can be illustrated using the visualization platform VOS-Viewer (Tsai, 2019), which works to find relationships between publications and can help researchers understand the following research and subsequent publications.

Pros and Cons: An Example

As a machine translation, Google translate continues to improve its translation quality with better grammar accuracy. Students, teachers, and lecturers with low foreign language skills can use the GT application to read and write texts in foreign languages. Tsai's research shows that EFL students are satisfied with using GT in their English writing, especially in finding the correct vocabulary and speeding up the process of writing English texts (Tsai, 2019).

As a machine, GT cannot translate perfectly. The translation result can be referred to as a pre-translation that still needs to be revised, including syntactic, morphological, semantic, and orthographic errors (Santoso, 2010). Meanwhile, Rahmannia & Triyono revealed that the GT translation results were often inappropriate or had several errors. Analysis errors occur in grammar and word selection that are not appropriate. From the analyzed data, the dominance of translation errors occurs in the shift of meaning (Rahmannia & Triyono, 2019).

Meanwhile, several research results show comparisons between the manual and technological approaches in qualitative data analysis. The debate about using Nvivo as a tool to analyze qualitative data has been conducted by Mattimoe et al. (2021). This shows that using Nvivo as a tool to analyze qualitative data has strengths and weaknesses. However, researchers in qualitative research should use computer applications to analyze qualitative data, and Nvivo is one of the applications that can be used at every stage of the analysis process (Kalpokas & Radivojevic, 2022) because Nvivo can generate, refine, and test a complex data from many data sources simultaneously (Dalkin et al., 2021). Referring to these great benefits of Nvivo, Soehardi conducted training on using Nvivo for young researchers (Soehardi et al., 2021).

METHOD

Participants

Research data were obtained from 166 respondents, which is consisting of undergraduate, master, and doctoral students, teachers, and lecturers. Respondents came from 12 provinces from 51 institutions in Indonesia.

Instruments

The research instrument was questionnaires and interviews via the link bit.um.ac.id/NLP_pend_pen, which was compiled in the form of a google form. Nine applications work with NLP technology. These are the top nine applications used in the field of education, teaching, research, and publications. Questions were answered in a descriptive questionnaire scaled questionnaire with intensity levels, namely *never, seldom, sometimes, often,* and *very frequently*. Interviews were conducted to obtain data about respondents' knowledge of other applications related to the questions.

Data Analysis

Questionnaire data were presented in percentage form, while interview data were analyzed using the purposive sampling technique. Researchers determine the data taken by unique characteristics that follow the research objectives to answer research problems.

RESULTS AND DISCUSSION

Demographics Respondent

Respondents who participated in this study came from various provinces and institutions, which are described in figure 1 below:

FIGURE 1
DEMOGRAPHICS OF RESPONDENTS

Respondents are academics, both in secondary, middle, and higher education. The following is a description of the characteristics of the respondents based on their educational and occupational backgrounds:

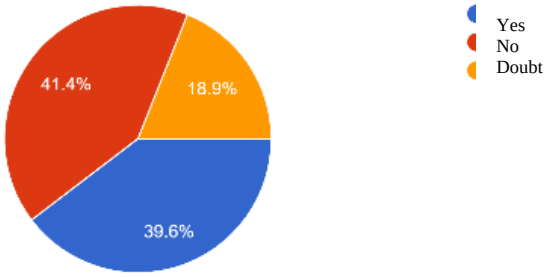
TABLE 1
CHARACTERISTICS OF RESPONDENTS

No	Characteristic of respondent	Amount
1.	Teacher	1
2.	Undergraduate students	55
3.	Undergraduate students; Teacher	1
4.	Master students; Teacher	3
5.	Master students	14
6.	Doctoral students; Lecturer at university	21
7.	Doctoral students	25
8.	Lecturer at university	46

Questionnaire About Use Application

Respondents filled out a questionnaire about the use of NLP-based applications. The first question is whether the respondent is familiar with the term NLP. The following diagram shows the respondents' answers to these questions:

FIGURE 2
KNOWING THE TERM NLP

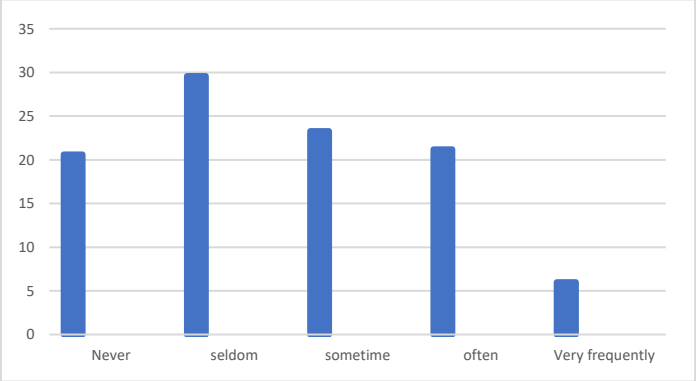


As many as 41.4% of respondents stated that they did not know the term NLP. However, they frequently use GT to translate daily conversation texts, vocabulary, short sentences, paragraphs, text from websites, standard texts, and literary texts. Most respondents (49.7%) use GT to translate standard texts, such as

scientific papers. Meanwhile, 45% use GT to translate the vocabulary, and another 38% translate paragraphs. Respondents also use applications other than GT, for example, Oxford Advanced Learner Dictionary, Linguee, Lexico, Deepl Translator, U-Dictionary, Alepsa, Bing, Baidu Fanyi, Papago, U-Dictionary, Pleco, Hanping Lite, Oldae, Siri, Cambridge Dictionary, Pons, Reverso, Dict. cc, Yandex, DingTalk.

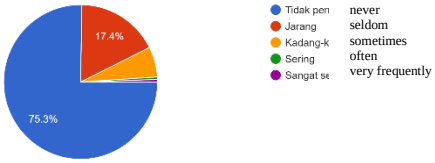
Grammarly is used for corrective feedback on English texts as an application that uses NLP technology. This study revealed that 21% of respondents had never used Grammarly. 72% of respondents who have used Grammarly check grammar and word choice in standard texts, for example, scientific papers, and check spelling and/ or punctuation. Meanwhile, only 18.9% checked the level of similarity, as shown in the following diagram:

FIGURE 3
USE OF GRAMMARLY



Besides Grammarly, respondents used U-dictionary Whitesmoke, pro writing aid, paraphrasing tool, Quillbot, enago, and languagetool.org to check their text. The next question is related to chatbot applications. The questionnaire's results show in the following data:

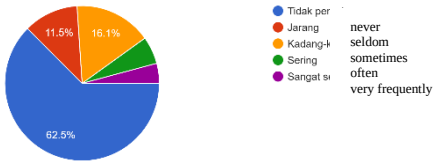
FIGURE 4
USE OF CHATBOT-BASED APPLICATIONS



As many as 75.5% of respondents said they had never used a chatbot. 12.5% of respondents who have used chatbot technology aim to contact Customer Service at an institution/company. Interestingly, 16.1% of respondents use chatbots for fun because they want to try the existing system. Only a tiny percentage (9%) of chatbot technology users use it to learn a learning material.

The following application is Quillbot, which can help users paraphrase sentences or paragraphs. The results of the questionnaire show the following data:

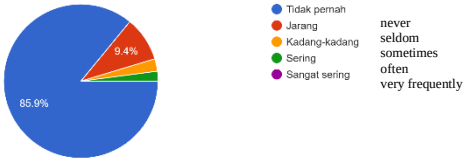
FIGURE 5
QUILLBOT APP USAGE



As many as 62% of respondents stated that they had never used the Quillbot online application. 31.8% of respondents have used Quillbot to paraphrase sentences, phrases, and or paragraphs, and 20.8% used Quillbot because they want to avoid plagiarism. A small percentage of respondents (15.1%) use Quillbot to find synonyms for a word, phrase, or expression in a paragraph.

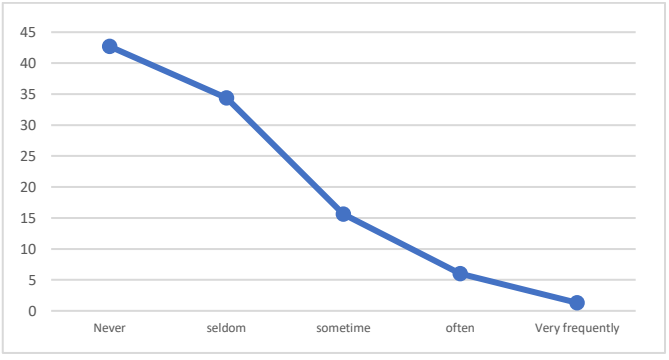
In addition to Quillbot, respondents use applications such as Smodin, Jasper AI, Spinbot, smallseotoll, paraphrasing tool, Spinner, and paraphrase. it, and Paraphrase Online Indonesia.

FIGURE 6
USAGE OF THE VOS-VIEWER APP



85.9% of respondents stated that they had never used the VOS-Viewer application, and 12% of respondents, who had used VOS-Viewer, used it to find novelty in research. Only 10.9% of respondents wanted to find gaps for new research themes, and a small percentage (7.8%) used VOS-Viewer to find the relationship between publications of research results. In addition to using the VOS-Viewer, several respondents used the Open Knowledge Map application, Google Scholar, Mendeley, and the Scopus website to gain knowledge about the novelty of the research.

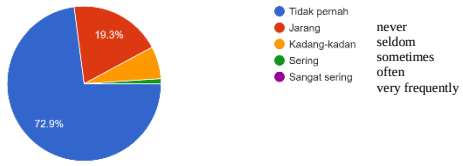
FIGURE 7
USE OF APP THAT WORKS WITH THE SPEECH RECOGNITION SYSTEM



42.7% of respondents stated that they had never used an application that worked with a speech recognition system. 27.6% of respondents use speech recognition applications to test the level of similarity in pronunciation/ pronunciation. Other 17.7% of respondents used it to send written messages by voice

when they did not want to write, and a small proportion of respondents (9.4%) used speech recognition applications as part of their research study. Some respondents use other applications with speech recognition systems, including Google Speaking, Google Assistant, Braina, Digital Oxford Dictionary, Siri, Google Translate, Pleco, Speech texter, and Google Speech.

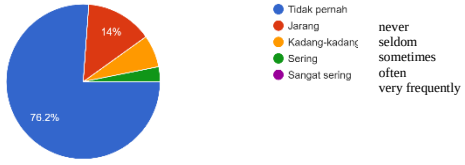
FIGURE 8
USE OF NVIVO DESKTOP APP



This study shows that 72.9% stated that they had never used Nvivo, and 20.2% used Nvivo to analyze qualitative data. As many as 13% of Nvivo users visualize some phenomenons, 8.8% compose a literature review, and 6.7% find sentiment in articles. Other applications used by respondents for this assignment are Winstep, RQDA, and Qualcoder.

The following application is Publish or Perish (PoP), which is an application that provides publication metadata. PoP is used in the process of doing a literature study on the selected topic. Therefore, PoP establishes a comparable study topics database (Al Husaeni & Nandiyanto, 2021).

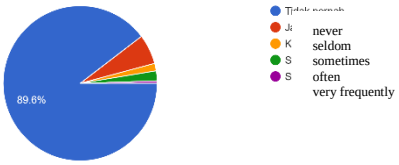
FIGURE 9
USE OF POP DESKTOP APP



76.2% stated that they had never used the PoP application. 16.1% of respondents use PoP to find metadata from sources/references only from reputable journals, 16.6% looked for source/reference metadata from any journal, and 13% looked at data on citation metrics and an author's H-Index. Other applications used by respondents are Mendeley, Google Scholar, elicit, and Biohunter.

The following application is Elicit, which is an application that provides publication metadata as well as a summary of the abstract. The author can take it as a reference source.

FIGURE 10
USE OF ELICIT



89.6% stated that they had never used the elicit online application. Only 9.8% of respondents used elicit to get metadata of sources/references from reputable journals and a summary/from an article. Several

respondents used similar applications: Researchrabbit AI, lens.org, google scholar, connected papers and DOAJ, and Mendeley Reference Manager.

Respondents' Suggestions and Desires

In addition to the questionnaire, respondents gave suggestions regarding using applications with NLP technology discussed in this paper. The following are the comments and suggestions of respondents:

- NLP-based applications are beneficial for me in the field of education and research.
- The applications mentioned in this study are critical to support the completion of our work. Thus, students should be able to use the application as a tool to facilitate completing the work.
- There are many applications that I do not know or do not understand, so it is necessary to conduct training on using the applications mentioned in this study, for example, through online webinars.
- We want a software/application that includes translation, grammar testing, plagiarism testing, and reference searching. The ease of using one application with the four functions above will be able to help researchers to work more effectively.

CONCLUSIONS

The applications discussed in this paper have many advantages. GT and Grammarly have been commonly used among students and lecturers. However, PoP and elicit applications useful for searching metadata for scientific papers are not widely known. VOS-Viewer and Nvivo as qualitative data processing applications have not been widely used. Students, teachers, and lecturers should improve their qualifications by increasing their skills in using NLP-based applications to assist them in their assignments and work for learning, teaching, research, and publications.

ACKNOWLEDGMENTS

The DRTPM funded this research. The authors appreciate the DRTPM and LPPM Unesa in supporting this research. Furthermore, the authors thank colleagues for their participation and support in helping and composing this paper and all respondents who provided data for this study.

REFERENCES

Al Husaeni, D.F., & Nandiyanto, A.B.D. (2021). Bibliometric Using Vosviewer with Publish or Perish (using Google Scholar data): From Step-by-step Processing for Users to the Practical Examples in the Analysis of Digital Learning Articles in Pre and Post Covid-19 Pandemic. *ASEAN Journal of Science and Engineering*, *2*(1), 19–46. https://doi.org/10.17509/ajse.v2i1.37368

Chowdhary, K.R. (2020). Natural Language Processing. In K.R. Chowdhary (Ed.), *Fundamentals of Artificial Intelligence* (pp. 603–649). Springer India. https://doi.org/10.1007/978-81-322-3972-7_19

Constantin, Y., Darusalam, U., & Nathasia, N.D. (2020). Aplikasi Personal Assistant Berbasis Voice Command Pada Sistem Operasi Android Dengan NLP. *JOINTECS (Journal of Information Technology and Computer Science)*, *5*(2), 121–128. https://doi.org/10.31328/jointecs.v5i2.1246

Dalkin, S., Forster, N., Hodgson, P., Lhussier, M., & Carr, S.M. (2021). Using computer assisted qualitative data analysis software (CAQDAS; NVivo) to assist in the complex process of realist theory generation, refinement, and testing. *International Journal of Social Research Methodology*, *24*(1), 123–134. https://doi.org/10.1080/13645579.2020.1803528

Dwiastuti, M. (2019). Pengenalan Computational Linguistics. *Meisyarah Dwiastuti*. Retrieved from http://meisyarahd.github.io/blog/2019/11/03/intro_to_computational_linguistics/

Goesderilidar, G. (2021). Prospek Pengembangan Komputasional Linguistik Menggunakan Pemrosesan Bahasa Alamiah (Natural Language Processing /NLP). *IndraTech*, *2*(2), 81–85. https://doi.org/10.56005/jit.v2i2.84

Kalpokas, N., & Radivojevic, I. (2022). Bridging the Gap Between Methodology and Qualitative Data Analysis Software: A Practical Guide for Educators and Qualitative Researchers. *Sociological Research Online*, *27*(2), 313–341. https://doi.org/10.1177/13607804211003579

Kharis, M., Kisyani, K., Suhartono, S., & Yuniseffendri, Y. (2021). Takarir: A new simultaneous translator voice to text to promote bi/multilinguality. *Journal of Language and Linguistic Studies*, *17*(3), 1175–1183. https://doi.org/10.52462/jlls.83

Kisyani, K., Sumarno, A., Yustanti, W., & Suhartono. (2020). *Pengembangan Takarir dalam Masa Pandemi Covid-19* [Laporan Akhir Penelitian Guru Besar]. Universitas Negeri Surabaya.

Mattimoe, R., Hayden, M.T., Murphy, B., & Ballantine, J. (2021). Approaches to Analysis of Qualitative Research Data: A Reflection on the Manual and Technological Approaches. *Accounting, Finance, & Governance Review*, *27*(1), 15.

McAllister, J.T., Lennertz, L., & Atencio Mojica, Z. (2021). Mapping A Discipline: A Guide to Using VOSviewer for Bibliometric and Visual Analysis. *Science & Technology Libraries*, *0*(0), 1–30. https://doi.org/10.1080/0194262X.2021.1991547

Migunani, & Aditama, K. (2020). Pemanfaatan Natural Language Processing dan Pattern Matching dalam Pembelajaran melalui Guru Virtual. *Elkom: Jurnal Elektronika Dan Komputer*, *13*(1), 121–133.

Rahmannia, M., & Triyono, S. (2019). A Study of Google Translate Translations: An Error Analysis of Indonesian-to-English Texts. *International Journal of Linguistics, Literature, and Translation (IJLLT)*, *2*(3). https://doi.org/10.32996/ijllt.2019.2.3.22

Rosyadi, H.E., Amrullah, F., Marcus, R.D., & Affandi, R.R. (2020). Rancang Bangun Chatbot Informasi Lowongan Pekerjaan Berbasis Whatsapp dengan Metode NLP (Natural Language Processing). *Briliant: Jurnal Riset dan Konseptual*, *5*(3), 619–626. https://doi.org/10.28926/briliant.v5i3.487

Santoso, I. (2010). Analisis Kesalahan Kebahasaan Hasil Terjemahan Google Translate Teks Bahasa Indonesia Ke Dalam Bahasa Jerman. *Seminar Internasional Jurusan Pendidikan Bahasa Jerman*, *1*.

Soehardi, F., Putri, L.D., & Dinata, M. (2021). NVivo Software Training for Young Researchers: *Mattawang: Jurnal Pengabdian Masyarakat*, *2*(1), 8–13. https://doi.org/10.35877/454RI.mattawang265

Sophia, J.J., Kumar, D.A., Arutselvan, M., & Ram, S.B. (2020). *A Survey on Chatbot Implementation in Health Care using NLTK*, *5*. https://doi.org/10.29207/resti.v4i4.2107

Tsai, S.-C. (2019). Using google translate in EFL drafts: A preliminary investigation. *Computer Assisted Language Learning*, *32*(5–6), 510–526. https://doi.org/10.1080/09588221.2018.1527361