



Mapping Strategy for Student Theses in the Field of Language: Research Roadmap with Bibliometric Analysis

Tessa Simahate¹, Harly Christy M. Siagian¹

¹UPT. Perpustakaan, Universitas Negeri Medan, Medan, Indonesia
echa@unimed.ac.id

Abstract. This study aims to evaluate the correlation between the thesis topics of students from the Faculty of Language and Arts at Universitas Negeri Medan from 2017 to 2023 and the university's vision, mission, and research roadmap. Using qualitative and quantitative analysis methods, the thesis topics are analyzed and classified into several main categories: innovation in education, the utilization of technology in education, the preservation and development of local arts and culture, and the improvement of community welfare. The analysis results show that the majority of student research topics align with Universitas Negeri Medan vision to become an excellent university in education, globally competitive, and culturally characterized. Furthermore, these topics support Universitas Negeri Medan mission in developing quality education, utilizing technology, preserving local culture, and improving community welfare. This study indicates that Universitas Negeri Medan successfully directs student research according to the priorities set in the research roadmap, significantly contributing to the achievement of the university's strategic goals. The findings of this study provide a scientific foundation for Universitas Negeri Medan to continue encouraging student research that is relevant to societal needs and the advancement of knowledge.

Keywords: Thesis topic analysis, Student research, Bibliometric analysis, Research roadmap

1 Introduction

Academic research is a fundamental pillar in the development of science and technology. One of the outputs of academic research produced by students is the thesis. A thesis is a form of scientific work that students must compile as one of the graduation requirements in higher education. Student theses often reflect trends, focuses, and evolutions in specific studies, such as in the field of language. The process provides valuable insights related to the direction and development of research conducted by students, thus necessitating the mapping of theses.

The strategy of mapping student theses in the field of language through bibliometric analysis is an innovative and systematic approach to evaluating and understanding research trends. Bibliometric analysis involves the collection and analysis of academic

publication data to identify patterns, relationships, and developments in scientific literature [1]. This method enables the identification of the most popular research topics, frequently used methods, and significant contributions in the field. Bibliometric analysis is an effective tool for exploring patterns of scientific communication and research trends [2]. Additionally, this bibliometric analysis can provide insights into the strengths and weaknesses of research in a specific field and assist in planning future research strategies. As stated by Van Eck & Waltman [3], “bibliometric mapping provides a powerful tool for analyzing the structure and development of scientific research”.

This approach not only aids in understanding the current research landscape but also in formulating a research roadmap for the future. Through the process of identifying research gaps, emerging trends, and potential academic collaborations, bibliometric analysis can provide guidance for students and researchers in designing relevant and impactful studies. It offers valuable insights into the dynamics and scientific developments that can be utilized for strategic planning in research [4].

This mapping strategy aligns with the vision, mission, and goals of Universitas Negeri Medan to promote academic excellence and make substantial contributions to science. Universitas Negeri Medan vision focuses on achieving excellence as a university in the fields of education, industrial engineering, and culture [5]. Universitas Negeri Medan mission is to provide education, conduct research, engage in community service, and collaborate with various domestic and international agencies. The aim is to develop Universitas Negeri Medan into a leading teaching and research institution, to cultivate creativity in industrial engineering and technology, to foster a scientific and ethnic culture, to promote entrepreneurship, and to nurture a healthy organizational climate and academic environment. Furthermore, Universitas Negeri Medan goals are to produce graduates who are excellent, professional, and characterful, with intellectual intelligence, entrepreneurial skills, and a national perspective. Universitas Negeri Medan also strives to produce, develop, and disseminate knowledge, technology, and/or arts to fulfill the tridharma (three pillars) of higher education and to generate scientific works that establish Universitas Negeri Medan as a center for innovation in learning and research. By integrating the university's vision, mission, and goals into the research mapping strategy, it is expected to create a positive synergy between academic efforts and institutional aspirations, thereby fostering significant progress in the fields of Language and Education in general.

By identifying research trends and opportunities, universities can more effectively allocate resources and support to students and faculty, enhancing the quality of research. This approach also fosters increased collaboration between faculties and researchers, resulting in scholarly works with broad impact and supporting the achievement of the university's strategic goals.

In this study, we will outline the strategy for mapping student theses in the field of language using bibliometric analysis. The research aims to determine the extent of the distribution of both emerging and underdeveloped research topics in the field of language, to create a map of academic fields for student research based on groups of disciplines using the Dewey Decimal Classification (DDC), and to present a research roadmap for undergraduate students at the Faculty of Language and Arts at Universitas

Negeri Medan over the last seven years. Consequently, this study seeks to provide a comprehensive guide for students, faculty, and researchers in directing their research according to existing scientific trends and needs, as well as in support of the university's vision, mission, and objectives.

2 Method

This study employs a bibliometric analysis approach to examine thesis data previously published, focusing on content analysis to uncover the main topics in bibliometric research on Language. The thesis data was collected from university libraries and academic repositories that include language student theses. The study selected all theses from Faculty of Language and Arts graduates from the years 2017 – 2023 as research objects. The sampling was conducted using purposive sampling method, limited to majors in Indonesian Language Education, English Language Education, German Language Education, and French Language Education from the years 2017 – 2023, comprising 2520 theses in total. The purposive sampling technique is a non-random sampling method where the researcher determines sample selection based on specific considerations in line with the aims and objectives of the study, ensuring that the samples indeed possess the data desired by the researcher for the success of the study [6].

The analysis was conducted using bibliometric software tools such as VOSviewer to identify the main topics, research trends, and research gaps, and OpenRefine to clean, group thesis data, and convert file formats for use in VOSviewer. VOSviewer is a computer application that can be used to view bibliometric maps [7]. It is employed to present maps of keywords or co-words and keyword density visualizations (density visualization) that appear in the student thesis documents of the Faculty of Language and Arts at Universitas Negeri Medan for the years 2017 – 2023. The data was analyzed quantitatively to produce a comprehensive research map. The analytical steps included data collection from university libraries and academic repositories, identification of key keywords, analysis of frequency and distribution of publications, and the use of visualization techniques to present findings clearly.

3 Result and Discussion

This study identified the number of research topics in the theses of Faculty of Language and Arts students at Universitas Negeri Medan over the period from 2017 to 2023, encompassing approximately 2520 research topic titles. This information is vividly and clearly depicted from the processed research data regarding the distribution of research topics in the theses of Faculty of Language and Arts students at Universitas Negeri Medan, graduates from 2017 to 2023, as can be seen in Table 1 and Figure 1.

Table 1. Distribution of Thesis Research Topics for Faculty of Language and Arts Graduates 2017-2023

No.	Departement	Year							Total
		2017	2018	2019	2020	2021	2022	2023	
1	English Language Education	136	150	186	79	174	170	19	914
2	Indonesian Language Education	157	145	182	132	129	134	15	894
3	German Language Education	31	35	65	70	70	35	40	346
4	French Language Education	42	43	49	57	78	59	38	366
Total		366	373	482	338	451	398	112	2520
Percentage (%)		14.52 %	14.80 %	19.13 %	13.41 %	17.90 %	15.79 %	4.44 %	100 %

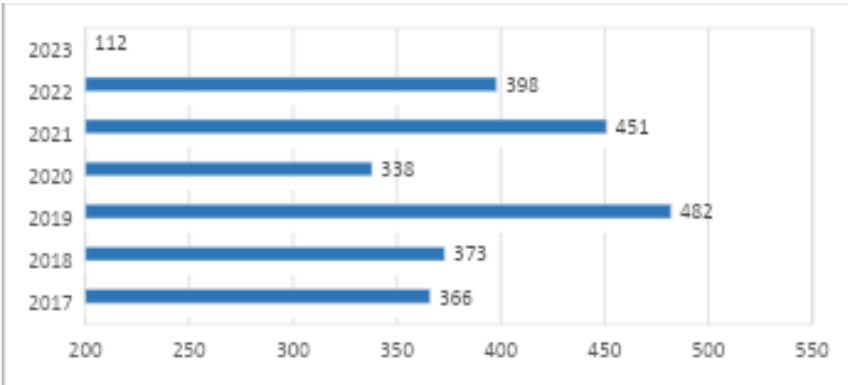


Fig 1. Graph Distribution of Thesis Research Topics for Faculty of Language and Arts Graduates 2017-2023

The information obtained from the illustrations in Table 1 and Figure 1 indicates that the year with the highest number of thesis research topics was 2019, with 482 titles (19.13%). The second highest number of research topics occurred in 2021, with 451 research topics (17.90%), while the year with the fewest research topics was 2023, accounting for 112 research topics (4.44%).

The distribution of research topics based on study programs. The information on the distribution of research topics is illustrated by the distribution of thesis research at the Faculty of Language and Arts of Universitas Negeri Medan from 2017-2023 across various study programs, as depicted in the illustrations in Table 2 and Figure 2.

Table 2. Distribution of Thesis Research Numbers by Program of Study 2017-2023

No.	Department	Volume	Percentage %
1.	English Langugae Education	914	36%
2.	Indonesian Languge Education	894	35%
3.	German Langugae Education	346	14%
4.	French Langugae Education	366	15%
Total		2520	100%

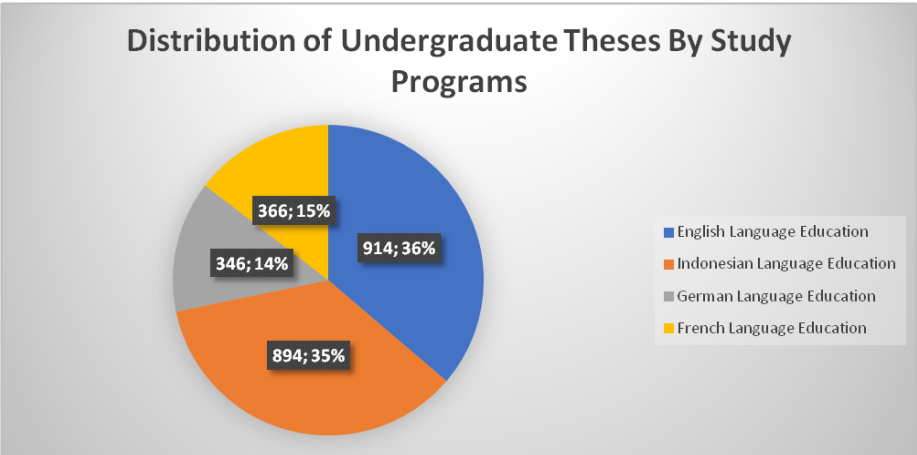


Fig 2. Graph of Distribution of ThesisNumbers for Faculty of Language and Arts Graduates from 2017 to 2021 by Study Programs

The information derived from Table 2 and Figure 2 reveals that the distribution of thesis research at the Faculty of Language and Arts of Universitas Negeri Medan from 2017 to 2021, based on study programs, shows that the English Language Education program conducted the most research, with 914 studies (36%). This is followed by the Indonesian Language Education program with 894 studies (35%). The French Language Education program follows with 266 studies (15%). However, the German Language Education program had the fewest studies, totaling only 346 (14%).

Distribution of Fields of Study in Student Research at Faculty of Language and Arts Universitas Negeri Medan Based on DDC. Table 3 displays the categorization of research topics in the theses of students at the Faculty of Language and Arts according to the DDC system. Research in the theses of the Indonesian Language Education program from 2017 to 2023 falls under the following academic categories: Education (370), Language (400), and Rhetoric (808).

Table 3. Distribution of the Number of Fields of Study in Theses from the Indonesian Language Education Program at Faculty of Language and Arts Based on DDC the Most Researched Fields

No .	Class	Bidang Ilmu	Volume	Percentage%
1.	808.066	Professional, technical, expository literature	116	13%
2.	371.3	Methods of instruction and study	82	9%
3.	401.41	Discourse analysis	54	6%
4.	808.107 12	Rhetoric of poetry - study and teaching	48	5%
5.	410.071 2	Linguistic : secondary education	46	5%
6.	808.307 12	Professional, technical, expository literature	46	5%
7.	371.33	Rhetoric of fiction : secondary education	46	5%
8.	808.802 2	Collections of literary texts from more than two literatures	29	3%
9.	801.95	Teaching aids, equipment, materials	28	3%
10.	418.4	Moral, ethical, character education	22	2%
Jumlah			517	52%

From the data provided, it can be explained that according to the DDC, the data shown represents the 10 most researched fields of study out of a total of 90 fields in the theses of the Indonesian Language Education program at the Faculty of Language and Arts of Universitas Negeri Medan from 2017-2023. From the table, it is apparent that the research topic with the highest number of theses is in the category of Professional, technical, expository literature, with 116 studies (13%), and the category with the fewest is Moral, ethical, character education, with 22 studies (2%).

Table 4 illustrates the categorization of research topics in the theses of students from the English Language Education program at the Faculty of Language and Arts for the years 2017-2023, using the DDC system. The research topics fall under the academic categories of Education (370), English Language (400), and Rhetoric (808).

Table 4. Distribution of the Number of Fields of Study in Theses from the English Language Education Program at Faculty of Language and Arts Based on the Most Researched DDC

No.	Class	Fields	Volume	Percentage%
-----	-------	--------	--------	-------------

1.	428.4	Reading	145	16%
2.	371.33	Teaching aids, equipment, materials	111	12%
3.	808.066	Professional, technical, expository literature	108	12%
4.	428	Standard English usage	86	9%
5.	371.3	Methods of instruction and study	76	8%
6.	420.71	English Language-Study and Teaching	51	6%
7.	808.802 2	Description (Rhetoric)	42	5%
8.	420.712	English Language-Study and Teaching (Primary)	34	4%
9.	371.26013	Validity and reliability of tests	31	3%
10.	808.802 3	Narration	27	3%
Total			711	78%

Based on the DDC, there is a distribution across 66 fields of study in the theses of students from the English Language Education program at the Faculty of Language and Arts of Universitas Negeri Medan, from 2017 to 2023. From the provided table, it is observed that the most researched topic is Reading, with 145 studies (16%), and the least researched topic is Narration, with 27 studies (3%).

Furthermore, Table 5 categorizes the research topics in the theses of students from the German Language Education program, also using the DDC. The research in this program from 2017 to 2023 falls under the academic categories of Teaching aids, equipment, materials (371); German and related languages (430); and Comic strip (741). According to the data in Table 5, there is a distribution across 111 fields of study in the German Language Education program's theses. The topic with the most research is Teaching aids, equipment, materials, with 43 studies (12%), and the topic with the fewest studies is Translating and interpreting, with 9 studies (3%).

Table 5. Distribution of the Number of Fields of Study in Theses from the German Language Education Program at Faculty of Language and Arts Based on the Most Researched DDC

No.	Class	Field	Volume	Percentage %
1.	371.33	Teaching aids, equipment, materials	43	12%
2.	371,334	Computer sciences	19	5%
3.	430.71	German and related languages	19	5%
4.	371.33523	Motion pictures and video recordings	16	5%

5.	430.071 2	German language-study and teaching (secondary)	16	5%
6.	741.56	Cartoons, caricatures, comic strip	13	4%
7.	438.071 2	Standard German usage (prescriptive linguistic)	12	3%
8.	371.335 2	Slides and filmstrips, motion picture and video, pictures	10	3%
9.	006.76	Programming	9	3%
10.	430,803	Translating and interpreting	9	3%
Total			166	48%

Table 6 displays the categorization of research topics in the theses of students from the French Language Education program at the Faculty of Language and Arts, based on the DDC from 2017 to 2023. The research within these theses falls under the academic categories of Teaching aids, equipment, materials (371); French and Related Roman Languages (445); and French Writing (808).

Table 6. Distribution of the Number of Fields of Study in Theses from the French Language Education Program at Faculty of Language and Arts Based on the Most Researched DDC

No.	Class	Field	Volume	Percentage %
1.	371.33	Teaching aids, equipment, materials	52	14%
2.	445	Grammar of standard French	47	13%
3.	440,071	French languages - study and teaching	22	6%
4.	445.6	Grammar: Spanish, Portuguese, Galician	22	6%
5.	440.071 1	French languages - study and teaching (Higher education)	21	6%
6.	808,066	Professional, technical, expository literature	12	3%
7.	440,014	French languages - semantic	10	3%
8.	448.040 41	French languages - translating	9	2%
9.	440.014 1	French languages - semantic	8	2%
10.	441.5	Writing systems, phonology, phonetics of standard French : prepositions	8	2%
Total			211	58%

Based on the data provided and using the DDC, the distribution of academic fields in the theses from the French Language Education program at the Faculty of Language and Arts of Universitas Negeri Medan from 2017 to 2023 is outlined. The analysis shows that the most researched topic is Teaching aids, equipment, materials, with 52 studies (14%). The least researched topic is Writing systems, phonology, phonetics of standard French: prepositions, with 8 studies (2%).

Network Visualization of Keyword and Subject Map. This research successfully presents the analytical results obtained from VosViewer Version 1.6.17, specifically the network visualization of Subject Headings in the theses of each study program at the Faculty of Language and Arts of Universitas Negeri Medan, spanning the graduation years 2017-2023. The study reports the identification of subject heading clusters across various language education programs: 12 clusters for the Indonesian Language Education program, 11 clusters for the English Language Education program, 8 clusters for the German Language Education program, and 7 clusters for the French Language Education program. The distribution of trends in research topic linkages is detailed in Table 7:

Table 7. Distribution Trend of Research Topic Links from Indonesian Language Education Graduates

Keywords and Subjects	Cluster	Occurance	Total of Network Link	Link	Research Year
Kemampuan menulis	5	166	387	70	2019
Writing	8	110	275	60	2019
Writing Skills	10	69	154	36	2020
Teaching-Aids and Device	4	37	52	33	2019
Explanation (Linguistic)	2	33	69	22	2019
Teks negosiasi	12	30	65	24	2019
Kurikulum 2013	3	25	57	29	2018
Reading	11	25	34	18	2018
Critism, contextual	7	22	33	18	2020
Moral education	1	24	33	12	2020
Morfologi	6	5	13	10	2018

Source: Theses from the Indonesian Language Education graduates from 2017-2023 and LCSH (Library of Congress, 2000)

Table 8. Distribution Trend of Research Topic Links from English Language Education Graduates

Keywords and Subjects	Cluster	Occurance	Total of Network Link	Link	Research Year
Research and Development (R&D)	2	124	317	43	2019
Writing	4	140	276	57	2020

Reading	3	106	238	53	2018
Learning strategies	1	82	114	54	2020
English language-study and teaching	5	80	93	31	2020
Reading comprehension	6	52	92	43	2019
Procedure text	11	36	84	26	2019
Students	9	17	49	23	2021
Translating and interpreting	10	19	26	11	2019
Vocabulary	8	13	20	19	2018
Writing descriptive text	7	6	18	10	2021

Source: Theses from the English Language Education graduates from 2017-2023 and LCSH (Library of Congress, 2000)

Table 9. Distribution Trend of Research Topic Links from German Language Education Graduates

Keywords and subjects	Cluster	Occurance	Total of Network Link	Link	Research Year
Teaching-aids and devices	6	110	145	20	2020
Die erstellung	2	53	81	23	2020
German language	8	23	44	9	2021
Comic books, strips, etc.	3	20	37	9	2020
Translating and interpreting	1	19	24	8	2021
German language-study and teaching	5	19	12	9	2021
Web sites-design	7	11	12	5	2020
Vocabulary	4	5	3	1	2018

Source: Theses from the German Language Education graduates from 2017-2023 and LCSH (Library of Congress, 2000)

Table 10. Distribution Trend of Research Topic Links from French Language Education Graduates

Keywords and subjects	Cluster	Occurance	Total of Network Link	Link	Research Year
Teaching-Aids and Device	2	85	101	13	2021

2018; Reading is in cluster 11 with 34 links researched in 2018; Criticism, Contextual is in cluster 7 with 33 links researched in 2020; Moral education is in cluster 1 with 33 links researched in 2020; and Morfologi is in cluster 6 with 13 links researched in 2018.

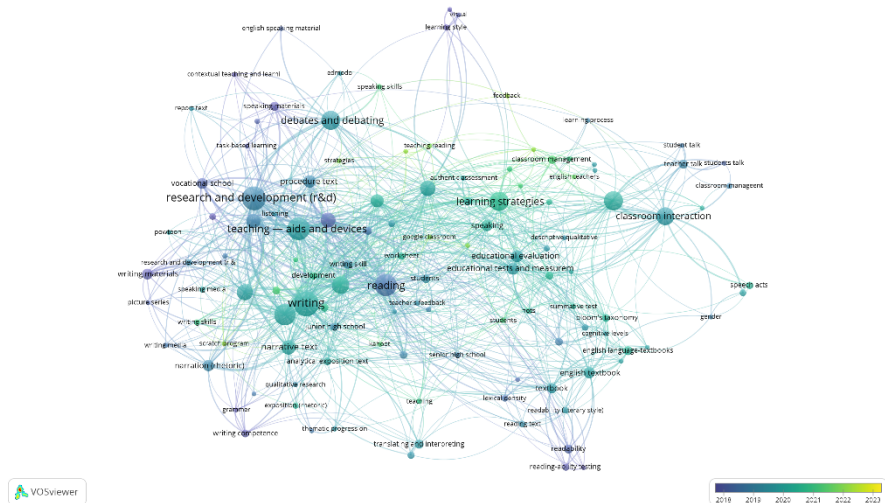


Fig 4. Visualization of the Spread of Keywords and Subject Headings in the English Language Education Program

Source: Data processing results

Table 8 and Figure 4 depict the distribution of research trends over the years in the theses of the English Language Education program as follows: "Research and Development" is found in cluster 2 with 317 links, researched in 2019; "Writing" is in cluster 4 with 276 links, researched in 2020; "Reading" appears in cluster 3 with 238 links, researched in 2018; "Learning Strategies" is in cluster 1 with 114 links, researched in 2020; "English Language-Study and Teaching" is in cluster 5 with 93 links, researched in 2020; "Reading Comprehension" is in cluster 6 with 52 links, researched in 2019; "Procedure Text" is in cluster 11 with 84 links, researched in 2019; "Students" are in cluster 9 with 49 links, researched in 2021; "Translating and Interpreting" is in cluster 10 with 26 links, researched in 2017; "Vocabulary" is in cluster 8 with 2 links, researched in 2019; and "Writing Descriptive Text" is in cluster 7 with 7 links, researched in 2021.

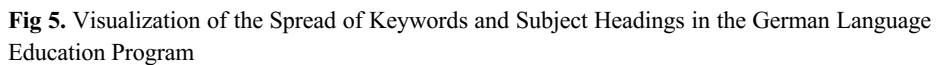


Table 9 and Figure 5 depict the distribution of research years in the theses of the German Language Education program as follows: "Teaching aids and devices" are found in cluster 6 with 145 links, researched in 2020; "Die Erstellung" is in cluster 2 with 81 links, also researched in 2020; "German Language" appears in cluster 8 with 44 links, researched in 2020; "Comic books, strips, etc." are in cluster 3 with 81 links, researched in 2020; "Translating and Interpreting" is in cluster 1 with 24 links, researched in 2021; "German Language-Study and Teaching" is in cluster 5 with 12 links, researched in 2021; "Web Sites-Design" is in cluster 7 with 12 links, researched in 2020; "Vocabulary" is in cluster 4 with 3 links, researched in 2018.

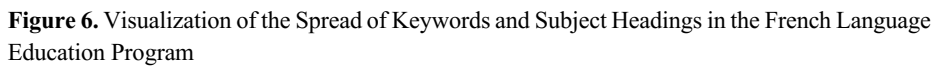
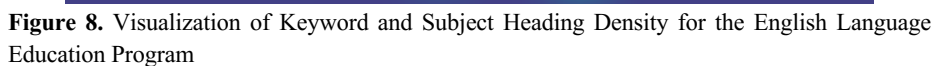


Table 10 and Figure 6 illustrate the distribution of research trends over the years in the theses of the French Language Education program as follows: "Teaching aids and devices" are found in cluster 2 with 101 links, researched in 2021; "Medias d'apprentissage" is in cluster 5 with 46 links, also researched in 2021; "Le Roman" appears in cluster 4 with 20 links, researched in 2018; "Analyse" is in cluster 1 with 11 links, researched in 2020; "Translating and Interpreting" is in cluster 3 with 11 links, researched in 2019; "Speech Acts (Linguistics)" is in cluster 7 with 3 links, researched in 2019; and "Politeness (Linguistics)" is in cluster 6 with 3 links, researched in 2020.

Density Visualization. This research has successfully presented information from the analysis conducted using the VosViewer program, which included a visualization of the density of Subject Headings in the theses from the Faculty of Language and Arts at Universitas Negeri Medan, covering the graduates from 2017 to 2023. This can be viewed in Figure 7, 8,9, 10.



Source: Data Processing Results



Source: Data Processing Results

4 Conclusion

Based on the bibliometric analysis of the mapping of student theses in the Language field, it can be concluded that this research provides a comprehensive overview of the distribution of research topics in the theses of students from the Faculty of Language and Arts at Universitas Negeri Medan during the period 2017-2023. The analysis shows a variety of research trends across different study programs and fields, which can serve as a basis for the development of academic policies and future research. This study also reveals that the majority of thesis topics from the Faculty of Language and Arts at Universitas Negeri Medan are significantly linked to the university's vision, mission, and research roadmap. The thesis topics generally align with Universitas Negeri Medan's goal to be an institution excelling in education, globally competitive, and culturally distinctive. Additionally, student research supports the university's mission in developing high-quality education, utilizing technology, preserving local culture, and enhancing community welfare. These findings indicate that Universitas Negeri Medan has successfully directed student research in line with its strategic priorities, thus significantly contributing to achieving the university's objectives.

References

1. Herawati, P., Utami, S. B., & Karlina, N. (2022). Analisis Bibliometrik: Perkembangan Penelitian Dan Publikasi Mengenai Koordinasi Program Menggunakan Vosviewer. *Jurnal Pustaka Budaya*, 9(1), 1–8. <https://doi.org/10.31849/pb.v9i1.8599>
2. Ellegaard, O., & Wallin, J. A. (2015). The bibliometric analysis of scholarly production: How great is the impact? *Scientometrics*, 105(3), 1809–1831. <https://doi.org/10.1007/s11192-015-1645-z>
3. Van Eck, N. J., & Waltman, L. (2010). Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics*, 84(2), 523–538. <https://doi.org/10.1007/s11192-009-0146-3>
4. Van Raan, A. F. J. (2005). Fatal attraction: Conceptual and methodological problems in the ranking of universities by bibliometric methods. *Scientometrics*, 62(1), 133–143. <https://doi.org/10.1007/s11192-005-0008-6>
5. Universitas Negeri Medan. (2022). *Visi-Misi Unimed*. <https://www.unimed.ac.id/visi-misi/>
6. Muljono, P., Sujana, J., & Prabowo, B. (2009). *Metodologi Penelitian dan Laporan Kearsipan*. Universitas Terbuka.
7. Tupan. (2016). Peta Perkembangan Penelitian Pemanfaatan Repositori Institusi Menuju Open Access: Studi Bibliometrik dengan VOSViewer. *Khazanah Al-Hikmah: Jurnal Ilmu Perpustakaan, Informasi, Dan Kearsipan*, 4(2), 104–117. <https://doi.org/10.24252/kah.v4i2a1>

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

