

Are the More Digitally Literate the Better Researching? The Case of English Department Students

Nurani Virginia
UIN Sunan Ampel Surabaya, Indonesia
nuranivirginia9@gmail.com

Aldianto Aldianto
UIN Sunan Ampel Surabaya, Indonesia
aldian041988@gmail.com

Alvin Uways Al-Qorni
The University of Melbourne, Australia
aalqorni@student.unimelb.edu.au

Abstract. Digital literacy has been seen as a key skill in academic research by university students in terms of recognizing, evaluating, and using credible scholarly resources. While previous studies have focused on digital literacy in the context of e-learning and information search, little research has been done on the connection between digital literacy and literature mining skills for academic research purposes. This study attempted to investigate the link between digital literacy and literature mining skills among university students studying in an English Language Education Department. Quantitative correlational research methodology was applied involving 34 students who were taking the Academic Article Writing course. Data were gathered through two structured questionnaires assessing digital literacy and literature mining skills of the participants. The tools used to collect data were found to be valid and reliable after item validity tests and calculation of Cronbach's Alpha scores. Pearson correlation test was conducted in SPSS version 25. The finding showed that there was a significant positive correlation between the digital literacy and literature mining skills of the participants ($r = 0.583$, $p < .001$). The findings show that those who have higher digital literacy skills also have high literature mining skills in conducting research in academic work. But from the fact that the relationship between the two variables is moderate, one can conclude that having high digital literacy skills alone cannot help one develop good research skills. This implies that critical thinking, evaluation of sources, information synthesis, and literature search should be done effectively.

Keywords: Academic Article Writing, correlation study, digital literacy, literature mining, research literacy instruction.

1. INTRODUCTION

Rapid evolution in technology has had a profound impact on the way knowledge creation, retrieval, evaluation, and dissemination is carried out in higher learning institutions. Besides having to retrieve information via digital means, university students are supposed to evaluate the authenticity of such information. As such, digital literacy is now a necessary skill when carrying out academic research in the twenty-first century. Digital information should be navigated in order to find credible sources to back up academic arguments.

Moreover, digital literacy is not limited to the technical ability of handling digital devices. The term comprises of the knowledge, skills, and critical stance which one needs to be able to search for, analyze, organize, generate and share information in the context of digital technologies (Vuorikari et al., 2022). The same applies to the emphasis of UNESCO (2023) on the need to be critically engaged with digital information and to demonstrate responsibility in the digital environment. In the realm of higher education, this skill is especially significant since the students should be able to differentiate valid scholarly information from the information found online.

Prior researchers have assessed the significance of digital literacy within the context of educational studies. Atoy et al. (2020) have studied digital literacy from the perspective of online information search practices by university students. Haryanto et al. (2022) assessed its importance regarding the critical thinking of the students, while the contribution of digital literacy to academic performance was studied by Muhammadiyah et al. (2021). With regard to English language education, Mudra (2020) and Sanova et al. (2022) have assessed its significance in the context of language education and use of digital tools for learning. From the above discussion, it can be concluded that digital literacy has some significance in the academic performance of the students. Nonetheless, the significance of digital literacy in a particular research practice needs more attention.

The second competency that is important for conducting academic research is literature mining. In this research, literature mining means a set of actions related to the search for, identification, selection, organization, evaluation, and synthesis of scientific literature concerning a research issue. Thus, the above-mentioned definition allows us to distinguish literature mining as a student research competency from computational text mining – a process that consists in the application of computational techniques for the extraction of patterns, relations, or information from the textual material. Literature mining needs such competencies as the formulation of adequate search requests, finding of relevant keywords, the use of academic databases, source evaluation, comparison of research results, and integration of evidence into a literature review.

Previous studies about text mining have contributed greatly to understanding the process of information retrieval, organization, and summarization. Some examples include the discussion about text mining techniques and applications by Gupta and Lehal (2009) such as information extraction, categorization, clustering, and summarization. Other researchers such as Tao et al. (2020) and Xie et al. (2020) studied the application of text mining and information extraction in research-related settings. Yim (2017) also studied the application of text mining in facilitating collaborative writing in academic setting. Even though these studies have provided valuable insight in the organization and

extraction of information from texts, computational text mining is not the same as the literature mining skills that are studied in the current research project.

Information literacy studies offer an additional perspective on the problem of students' engagement with scholarly resources. Specifically, Tachie-Donkor & Ezema (2023) examined information literacy competencies in regard to information-seeking behavior and lifelong learning of university students. In turn, their research underlines the importance of information literacy as it pertains to the identification and utilization of information. Such a perspective is significant since the search for scholarly publications goes beyond just inserting keywords into the search engine. It also implies the need to understand the information needs and evaluate the retrieved information in terms of its applicability to a specific purpose. However, information-seeking behavior and literature mining are not synonymous constructs. Indeed, the latter, as applied in this study, includes the organization and synthesis of scholarly literature findings. It is also possible to evaluate the association between digital literacy and academic success by assessing student achievement more broadly. According to Ardhiani et al. (2023), there exists a positive correlation between digital literacy and academic performance of university students based on a meta-analysis of the research on this topic. It suggests that digital literacy is relevant for the academic success of the students. At the same time, academic performance is a very broad concept that can involve various skills and factors that influence learning processes. Academic performance does not necessarily mean that a student can develop a good search of literature or critically evaluate the information sources and previous researches in academic work.

Further evidence of source evaluation practices is provided by research on online source evaluation. Liu (2025) conducted a study investigating how Chinese EFL undergraduates evaluate sources of information when searching online, concentrating on their evaluation behavior and criteria used. This field of investigation seems especially valuable for teaching English language, since EFL undergraduates should find and evaluate research conducted in various academic environments. Source evaluation is one of the components of efficient literature search process. Nevertheless, while conducting research about source evaluation processes, one cannot conclude that students with a higher level of digital literacy will have better literature mining skills.

As a whole, earlier research works have confirmed the relevance of the above-mentioned concepts of digital literacy, information literacy, and source evaluation for engagement in digital information by the students. However, there are still a number of questions that require further exploration. First, while research works dedicated to digital literacy mainly consider the concept in relation to general learning outcomes, academic success, or online information searches of the students, the development of their literature mining abilities has not been fully addressed. Secondly, although earlier research works dealing with the study of information literacy and source evaluation provided certain information regarding ways in which students find and evaluate information, the connection between their overall digital literacy and literature mining skills remains to be clarified.

With this in mind, the current research aims to explore the connection between digital literacy and literature mining ability among the undergraduate students of the English Language Education Department in a university in Indonesia. In particular, the

current research will attempt to explore whether students who have a higher digital literacy score are more likely to be high in literature mining ability as well. Employing the quantitative correlation research design, the current research provides localized evidence on the link between digital skills and scholarly literature engagement among the students. It should be pointed out that in studying the two variables together, the current research acknowledges that digital literacy is just one possible aspect of research competency.

Accordingly, this study addresses the following research question: Is there a significant relationship between undergraduate students' digital literacy and their literature mining skills for academic research?

2. METHOD

The research utilized the quantitative correlational research design in order to investigate the correlation between students' digital literacy and literature mining competences in academic writing. It was deemed fit to use correlational design in the current study because the aim was to find out the presence or absence of any correlation between the two factors rather than finding out the cause-and-effect relationship between the two variables. Quantitative research refers to gathering and analysis of numeric data while correlational design seeks to discover the nature and degree of the relationship between variables. The research was carried out at the English Language Education Department of UIN Sunan Ampel Surabaya. The population of the research was comprised of 74 sixth semester students taking the Academic Article Writing subject in three classes; A, B, and C. The samples were drawn through cluster random sampling in which the clusters referred to the classes. Through the process of cluster sampling, class C was chosen as the research sample which was composed of 34 students.

Data were gathered through two questionnaires, both of which were created using the Google Forms platform. The questionnaires were measuring the digital literacy and literature mining skills of the respondents through 45 close-ended statements taken from other works on the same subject. All responses were measured using a five-point Likert scale with choices of rating ranging from 1 (Strongly Disagree) to 5 (Strongly Agree). While the first set of questions measured the students' capacity to access, evaluate, organize, and use the digital information for their studies, the second set measured the students' ability to search for, retrieve, evaluate, organize, and synthesize the scholarly literature used in their research. Before the main data gathering, validity and reliability tests were performed using the IBM SPSS Statistics Version 25 software program. For item validity, the Pearson Product-Moment correlation was utilized and its critical value at the 5% significance level is $r = 0.339$ ($N=34$). Items that had correlation coefficients greater than the critical value were found to be valid. The reliability test was done using Cronbach's Alpha that evaluates the internal consistency of the questionnaire items. If the coefficient was greater than 0.60, then the instrument had sufficient reliability for research purposes.

Prior to the analysis of the research hypothesis, the normality of the data set was tested using the Kolmogorov-Smirnov normality test. A value of the significance level of more than 0.05 implied that the data set passed the normality assumption. Having ensured that the data fulfilled the statistical assumptions, the relationship between digital literacy

and literature mining skills was evaluated using the Pearson Product-Moment correlation coefficient in IBM SPSS Statistics Version 25.

Based on the research objective, the hypotheses were formulated as follows:

H₀: There is no significant relationship between students' digital literacy and literature mining skills for academic research.

H₁: There is a significant positive relationship between students' digital literacy and literature mining skills for academic research.

In order to know the relationship between digital literacy in research and literature mining in research, the researcher utilized the program SPSS 25 to analyze the data. In addition, the researchers used a category value from (James D. Evans) in order to find out whether the variables are correlated or not. The category values are as follows in Table 1.

Table 1. Interpretation of the two variables.

<i>Scale</i>	<i>Category</i>
0.000 – 0.199	Very low
0.200 – 0.399	Low
0.400 – 0.599	Neutral
0.600 – 0.799	High
0.800 – 1.000	Very high

3. RESULT

Before proceeding with the main analysis, the research instruments were subjected to validation and reliability testing to ensure that the questionnaires were able to measure the relevant variables. Validity of the questionnaires was done through the Pearson Product-Moment correlation test. A variable was said to be valid if it had a correlation coefficient greater than the critical value of $r = 0.339$ at the 5% significance level ($N = 34$). It was found out that there were 17 variables measuring digital literacy that satisfied the validity criteria. On the other hand, 24 out of 26 variables measuring literature mining satisfied the validity criteria while two items (Items 7 and 16) did not satisfy the validity criteria.

The reliability of the instruments was assessed using Cronbach's Alpha to determine the internal consistency of the questionnaire items. The results of the digital literacy questionnaire are presented in Table 2.

Table 2. Reliability of the Digital Literacy Questionnaire.

<i>Reliability Statistics</i>	
Cronbach's Alpha	0.880
Number of Items	17

As shown in Table 2, the digital literacy questionnaire obtained a Cronbach's Alpha coefficient of 0.880, indicating high internal consistency. Therefore, the instrument was considered reliable for data collection. Meanwhile, the reliability of the literature mining questionnaire is presented in Table 3.

Table 3. Reliability of the Literature Mining Questionnaire.

<i>Reliability Statistics</i>	
Cronbach's Alpha	0.918
Number of Items	24

Table 3 shows that the literature mining questionnaire achieved a Cronbach's Alpha coefficient of 0.918, which exceeded the acceptable threshold of 0.60. Therefore, the instrument demonstrated satisfactory reliability and was considered appropriate for measuring students' literature mining skills.

3.1 Students' Digital Literacy

The results of students' digital literacy are presented in Table 4.

Table 4. Results of students' digital literacy.

<i>No.</i>	<i>Statement</i>	<i>Percentage</i>
1	I can quickly obtain the sources I need through various websites.	94.5%
2	I can access various resources from online news, reports, websites, blogs, and social media.	97.5%
3	I can determine whether the information I obtain is relevant to my research needs.	87.2%
4	I understand the reliability of the sources I find.	87.1%
5	I can evaluate sources based on their relevance to my research topic.	84.6%
6	I can identify the credibility and trustworthiness of the sources I find.	61.6%
7	I can synthesize information from various sources using my own words.	87.2%
8	I can use digital media to create, present, and share information.	92.3%
9	I can integrate information from multiple sources into my research project.	74.4%
10	I can connect information from different sources.	82.2%
11	I can distinguish between reputable and non-reputable academic sources.	87.2%
12	I can draw conclusions from multiple sources for research purposes.	99.7%
13	I can correctly cite academic sources.	84.6%
14	I can use reference management systems to organize citations.	82.1%
15	I can retrieve citations from a reference management system.	84.6%
16	I always provide references in my academic writing.	82.1%
17	I can synthesize information from various sources into academic writing.	87.2%

As seen in Table 4, the respondents mostly had a very high level of digital literacy. It was noted that most of the students could access information from different online sources, select appropriate information, incorporate sources into academic writing, and follow appropriate citing practices. The largest number (99.7%) was registered in students' capability to reach conclusions based on information drawn from different sources in research. However, some other areas needed improvement. In particular, 61.6% of the

respondents felt confident in evaluation of credibility of information sources, while 74.4% were able to incorporate sources into research projects. Thus, it can be assumed that although the students have good digital skills, there is still a need to improve information evaluation and synthesis skills.

3.2 Students' Literature Mining Skills

The results of students' literature mining skills are presented in Table 5.

Table 5. The results of students' literature mining skills.

<i>No.</i>	<i>Statement</i>	<i>Percentage / Response</i>
1	I always establish criteria to obtain relevant information.	92.3%
2	I outline my research before searching for literature.	79.5%
3	I use various research tools to collect academic sources.	92.3%
4	Research tools currently used	Google Scholar, Open Knowledge Map, Connected Papers, WorldCat, EndNote, Taylor & Francis Online, Publish or Perish, ScienceDirect, Sci-Hub, ERIC, Sage
5	Research tools I would like to learn	Connected Papers, Garuda, Google Scholar, IJET, Publish or Perish, Scopus, Sage, Zotero, ResearchGate
6	I use additional keywords or modifiers to refine my searches.	89.2%
7	I consult peers or lecturers when developing search keywords.	89.8%
8	I keep up with current terminology in my research field.	71.8%
9	I analyze research topics before searching for literature.	82.0%
10	I clearly define the research topic before searching for sources.	89.7%
11	I revise my research questions when necessary.	82.0%
12	I check plagiarism after completing my academic writing.	63.1%
13	I conduct manual annotation when reviewing literature.	58.9%
14	I use annotation software when reviewing literature.	76.9%
15	Annotation software used	Mendeley, Zotero, ReadCube Papers
16	I ensure that the sources I use are up to date.	84.6%
17	I consider the target audience and purpose of each source.	81.3%

18	I critically evaluate the quality of information.	71.8%
19	I identify the strengths and weaknesses of previous studies.	71.8%
20	I thoroughly proofread my writing before submission.	89.7%
21	I prepare a search plan before collecting information.	66.7%
22	I create a systematic literature search strategy.	39.0%
23	I develop an outline to maintain consistency throughout my research.	64.3%
24	I identify common themes among previous studies.	71.8%
25	I evaluate source reliability based on evidence, citations, and references.	69.3%
26	I identify and evaluate the essential components of research articles.	84.6%

In general, the participants possessed acceptable literature mining abilities. Majority of the students claimed that they utilized different academic databases and digital tools such as Google Scholar, Open Knowledge Map, Connected Papers, WorldCat, ERIC, ScienceDirect, and Taylor & Francis Online for identifying academic literature. Reference manager tools such as Mendeley, Zotero, and ReadCube Papers were also widely used by the participants for organizing citations and references. However, there were some shortcomings. Only 39.0% of the participants said that they developed a systematic strategy prior to literature gathering, while 58.9% annotated scholarly literature manually. Besides, only 63.1% verified the absence of plagiarism after finishing academic writing. The participants also lacked ability to critically appraise source credibility, assess evidence critically, and synthesize previous research. This shows that literature mining is not just a technical skill, but it also involves critical analysis and higher order thinking.

3.3 Correlation between Digital Literacy and Literature Mining Skills

Pearson Product-Moment correlation analysis was conducted to examine the relationship between students' digital literacy and literature mining skills. The results are presented in Table 6.

Table 6. Correlation between Digital Literacy and Literature Mining Skills.

<i>Variables</i>	<i>Digital Literacy</i>	<i>Literature Mining</i>
Digital Literacy	1	0.583**
Literature Mining	0.583**	1
Sig. (2-tailed)	—	0.000
N	34	34

Note. ** Correlation is significant at the 0.01 level (2-tailed).

The analysis revealed a moderate positive correlation between digital literacy and literature mining skills ($r = 0.583$, $p < .001$). This finding indicates that students with higher levels of digital literacy tended to demonstrate better literature mining skills for

academic research. Therefore, the null hypothesis (H_0) was rejected, while the alternative hypothesis (H_1) was accepted.

Moreover, the results obtained from the survey revealed that students make use of digital research resources like Google Scholar, Open Knowledge Map, Connected Papers, Publish or Perish, ResearchGate, and ScienceDirect for conducting searches. The software programs most commonly used for citation management were Mendeley and Zotero. This clearly indicates that the students understand the significance of using digital tools in order to conduct searches and writing literature. However, based on the descriptive statistics, it can be concluded that students need further guidance in assessing the sources, formulating search strategies, and critical analysis of research evidence before using it in their academic work. The purpose of the current study was to investigate the correlation between digital literacy and the ability to conduct literature mining among undergraduate students from the English Language Education Department. It was found out that there was a statistically significant positive correlation between the above-stated variables ($r = 0.583$, $p < .001$). Thus, the null hypothesis (H_0) that states that there is a significant positive correlation between students' digital literacy and literature mining skills in academic research was accepted.

These results align with prior research focusing on the significance of digital literacy within higher education. As mentioned by Atoy et al. (2020), the higher students' level of digital literacy, the better they can find, evaluate, and use the online content. Also, Haryanto et al. (2022) noted that digital literacy has a positive effect on students' critical thinking and academic achievements. Furthermore, in the field of English language education, Mudra (2020) also claimed that digital literacy enables students' self-directed learning and makes it easier for them to access and manage digital educational materials. The current study contributes to the existing research by revealing that digital literacy positively relates to literature mining abilities of EFL university students in Indonesia.

In addition to this, the results obtained from the descriptive data indicated that most students had developed the necessary digital literacy skills. It was evident that the majority of the respondents were able to gain access to different sources of academic knowledge, locate relevant information, and incorporate scholarly citations in their academic assignments. They had also shown proficiency in the use of different digital tools for research, such as Google Scholar, ERIC, ScienceDirect, Connected Papers, Open Knowledge Map, Mendeley, and Zotero.

Nevertheless, there are several issues that should be addressed in addition to the positive findings described above. As indicated by the moderate correlation coefficient ($r = 0.583$), digital literacy does not totally contribute to explaining students' ability to conduct literature mining. Some of the questions on the questionnaire made it clear that most of the participants still had problems with evaluating the credibility of information sources, conducting a systematic search of literature, analyzing results of multiple studies, and critically interpreting research evidence. Such findings show the gap that exists between students' capacity to access digital information and higher-level cognitive skills necessary to evaluate and synthesize scholarly knowledge. This may be explained by the difference between digital literacy and critical research skills. Digital literacy helps the student access and organize information, but effective literature review requires both digital literacy and critical thinking skills and analytical abilities. There is no problem in

having great skills related to the use of digital devices but lacking the ability to evaluate sources, analyze conflicting results, find the gaps in research, and write literature review.

Furthermore, the results validate the claim that literature mining is not just a technical search process but a comprehensive research competency that entails identifying, evaluating, organizing, synthesizing, and interpreting scholarly literature. Consequently, universities need to incorporate digital literacy skills training into research methodology, academic writing, information evaluation, and literature reviews to better equip students in conducting research. In general, this research adds to the existing literature on digital literacy in proving that there is a positive relationship between digital literacy and literature mining skills among Indonesian EFL university students. Nevertheless, the results also confirm that enhancing research competence among students also requires critical thinking and information evaluation skills besides digital skills.

4. CONCLUSION

This study suggests that students with high digital literacy tend to have good skills in literature mining in academic research. This shows that digital literacy significantly contributes to the ability of students to manage and use academic information through the whole process of research. Nonetheless, digital literacy itself is not enough for developing full research competence. There are some difficulties related to assessing the credibility of sources of information, literature search strategy, synthesis of previous studies, and critical interpretation of research evidence. Thus, this study shows that such higher-level skills as critical thinking and evaluation of information are necessary for academic research along with digital literacy. This study recommends universities to include digital literacy along with research methodology, academic writing, information evaluation, and literature review in their curriculums. This approach will make it possible for students to conduct academic research of a high level. It is also recommended for future researchers to use larger and more varied sample from different universities and disciplines. Mixed methods and qualitative methods could also be used.

REFERENCES

- Allahyari, M., Pouriyeh, S., Kochut, K., Safaei, S., Ariaifar, K., Ruiz, C., & Arabnia, H. R. (2017). A brief survey of text mining: Classification, clustering and extraction techniques. *arXiv Preprint arXiv:1707.02919*.
- Ardhiani, O., Hadjam, M. N. R., & Fitriani, D. R. (2023). Digital literacy and student academic performance in universities: A meta-analysis. *Journal of Psychology and Instruction*, 7(3). <https://ejournal.undiksha.ac.id/index.php/JoPaI/article/view/68191>
- Atoy, M. B., Garcia, F. R. O., Cadungog, R. R., Cua, J. D. O., Mangunay, S. C., & de Guzman, A. B. (2020). Linking digital literacy and online information searching strategies of Philippine university students. *Journal of Librarianship and Information Science*, 52(4), 1015–1030. <https://doi.org/10.1177/0961000619871618>
- Evans, J. D. (1996). *Straightforward statistics for the behavioral sciences*. Brooks/Cole Publishing.

- Fraenkel, J. R., Wallen, N. E., & Hyun, H. H. (2019). *How to design and evaluate research in education* (10th ed.). McGraw-Hill Education.
- Gilster, P. (1997). *Digital literacy*. John Wiley & Sons.
- Gupta, V., & Lehal, G. S. (2009). A survey of text mining techniques and applications. *Journal of Emerging Technologies in Web Intelligence*, 1(1), 60–76. <https://doi.org/10.4304/jetwi.1.1.60-76>
- Haryanto, E., Widiastuti, I., & Sulisty, G. H. (2022). Digital literacy and critical thinking among Indonesian university students. *International Journal of Instruction*, 15(2), 789–806.
- Liu, J. (2025). Investigating EFL undergraduates' processes of source evaluation during web search: Behaviours and criteria. *SAGE Open*, 15(3). <https://doi.org/10.1177/21582440251334873>
- Mudra, H. (2020). Digital literacy among young learners: How do EFL teachers and learners view its benefits and barriers? *Teaching English with Technology*, 20(3), 3–24.
- Muhammadiyah, M., Salija, K., & Jabu, B. (2021). Digital literacy and students' academic achievement in higher education. *Journal of Language Teaching and Research*, 12(6), 1004–1012.
- Ng, W. (2012). Can we teach digital natives digital literacy? *Computers & Education*, 59(3), 1065–1078. <https://doi.org/10.1016/j.compedu.2012.04.016>
- Nuryadi, Astuti, T. D., Utami, E. S., & Budiantara, M. (2017). *Dasar-dasar statistik penelitian*. Sibuku Media.
- OECD. (2021). *21st-century readers: Developing literacy skills in a digital world*. OECD Publishing. <https://doi.org/10.1787/a83d84cb-en>
- Sanova, S., Syafryadin, S., & Nurkamto, J. (2022). Digital literacy in English language learning: Challenges and opportunities in Indonesian higher education. *Studies in English Language and Education*, 9(2), 851–867.
- Tachie-Donkor, G., & Ezema, I. J. (2023). Effect of information literacy skills on university students' information seeking behaviour and lifelong learning. *Heliyon*, 9(8), Article e18427. <https://doi.org/10.1016/j.heliyon.2023.e18427>
- Tao, D., Wang, T., Wang, T., Zhang, T., & Qu, X. (2020). A systematic review of literature mining techniques for evidence synthesis. *Journal of Informetrics*, 14(4), 101091.
- UNESCO. (2023). *Global education monitoring report 2023: Technology in education—A tool on whose terms?* UNESCO Publishing.
- Vuorikari, R., Kluzer, S., & Punie, Y. (2022). *DigComp 2.2: The Digital Competence Framework for Citizens*. Publications Office of the European Union. <https://doi.org/10.2760/115376>
- Xie, Y., Zhang, Y., & Cheng, G. (2020). Literature mining techniques for scientific knowledge discovery: A review. *Information Processing & Management*, 57(6), 102367.