European Science Editing

Original Article

Using text-matching software in educational science research: Research results from 18 universities in Vietnam

Thuan Van Pham[1], Hien Le Thi Thu[1]✉, Thanh Thi Nghiem[1], Hang Do Thi Thu[2], Can Nguyen Duc[1], Huy Nguyen Hoa[3]

[1]VNU University of Education, Vietnam National University, Hanoi, Vietnam
hienltt.1978@gmail.com
orcid.org/0000-0003-0805-5084
orcid.org/0000-0001-8847-6369
orcid.org/0000-0002-4516-9630
orcid.org/0009-0002-6672-2949

[2]Hanoi Metropolitan University, Hanoii, Vietnam
orcid.org/0009-0008-2268-6425

[3]VNU Center for Education Accreditation, Vietnam National University, Hanoi, Vietnam
orcid.org/0000-0002-3905-8088

European Science Editing

Abstract

Background: Plagiarism by researchers and college students in Vietnam has become a major concern for publishers. Many cases of master's theses of graduate students in Vietnam being cancelled or their diplomas being revoked for plagiarism are recorded, and some scientists also have been warned or criticized for plagiarism or self-plagiarism.

Objectives: The purpose of this study was to analyse the use in educational research of 13 popular text-matching software packages at universities in Vietnam.

Methods: The study was based on semistructured interviews of 104 researchers from 18 universities in Vietnam with reference to the use of text-matching software by the researchers.

Results: The three most commonly used text-matching software packages were Turnitin, DoIT, and iThenticate. Three-fourths of the 18 universities employ text-matching software and although 17 out of 104 researchers were unfamiliar with such software. Universities in Vietnam primarily require plagiarism checks for master's theses (79 out of 104 responses) and doctoral dissertations (72 out of 104 responses). Out of 104 participants, 32 use them for graduate theses or project reports, and 45 use them for research papers and project reports.

Conclusions: Many universities in Vietnam are yet to specify the requirements for use of text-matching software, and most researchers and students use it only when prompted by publishers or institutions. Researchers in educational science typically lack the financial resources and the requisite skills for using text-matching software.

Introduction

Plagiarism in research remains a significant concern worldwide, a challenge to publishers and journal editors. One common facet of the problem is improper copying and misattribution of references in works published or submitted for publication by researchers – a violation of ethics in which some researchers exploit the works of others without proper acknowledgement or without correct citations. Such failures in giving credit where it is due lead to misunderstandings and constitute unfair practices within the research community, ultimately affecting the legitimacy and reliability of research outcomes in the realm of scientific publishing.[1] Plagiarism also obscures transparency in publishing the findings of research. Some researchers may prioritize positive results while disregarding unfavourable results or include statistically insignificant results, compromising the credibility and accuracy of research findings. In recent years, unethical and dishonest practices in academia, including data fabrication, plagiarism, and deception, have emerged as serious concerns in research.

Plagiarism can be found in any scientific journal despite the use of tools to detect plagiarism. The absence of widely accepted definitions of misconduct in research and sole reliance on similarity checks prevent journal editors from effectively addressing complex cases of scientific information recycling, wasteful publishing, or any form of intellectual theft. Some studies have shown that in non-English-speaking countries, writing and copying texts pose significant challenges in university education. In some countries, poor management of research, lack of instruction in ethics related to publishing, and limited access to high-quality research sources contribute to plagiarism becoming a cross-cultural, cross-disciplinary, and widespread phenomenon.[2]

Plagiarism including self-plagiarism is currently a serious challenge for university management in Vietnam in general: some social science and humanities researchers in educational institutions in Vietnam have violated academic integrity by resorting to plagiarism in their publications,[3] and as many as 91.7% of graduate theses were identified as involving plagiarism at some universities in Vietnam that did not use any text-matching software; even in universities that did use such software, the proportion was as high as 61.7%.[4]

What is plagiarism in scientific research?

In this digital age, with continuing development of information technology, textual plagiarism has become one of the pressing concerns in higher education and research. Plagiarism has serious legal and ethical consequences for society, undermining its value system;[2] is considered a violation of publishing ethics; and threatens the credibility of research and academic writing.[3] Also known as 'literary theft', plagiarism refers to the act of appropriating (ideas, passages, etc.) from the work of another author. Literary theft can be seen as someone taking the creations of others and using them as their own (thoughts, writings, inventions, etc.) or copying (literary works, ideas, etc.) improperly or without acknowledgement, treating the thoughts, work, etc. of others as their own.[4] If instances of plagiarism are discovered, the evaluation of such research work – and punishment to its authors – will affect the reputation and credibility of those involved in the assessment and review and of funding entities supporting the publications based on that research.

Plagiarism can be understood as the act of stealing someone else's writing or ideas to develop them as one's own, but it can also be seen as using preexisting works of others without appropriately acknowledging the

source and presenting them as one's own. In science, plagiarism can be regarded as the appropriation by one person of another's products, including academic texts, research methods, graphics, and ideas, and it is considered unethical conduct.

The research reported in this article aims to analyse the frequency of use, types of text-matching software used, and regulations related to plagiarism in 18 universities in Vietnam and the skills of researchers in those universities in using text-matching software in research on educational science.

Text-matching software in scientific research
Although plagiarism in academia is not new, rapid and continuing advancements in information technology have provided convenient and immediate access to vast amounts of information and to the texts in which that information is embedded, making plagiarism easier than ever before. At the same time, information technology also provides tools for detecting academic plagiarism.[5]

Checking for plagiarism manually requires significant effort and excellent memory from the person engaged in such checking. Additionally, the manual approach is so time-consuming – it requires reading and searching through multiple documents – as to make it impractical,[6] because it can hold up the publication of the manuscript in question – sometimes for so long that the findings being reported cease to be relevant. These considerations make it impractical to attempt to detect plagiarism using manual methods.[7]

At the same time, the remarkable advancements in science and technology, particularly in information technology, have made plagiarism easier.[8] Fortunately, text-matching software has kept pace with these developments and offers the best way to curb plagiarism.[6] In recent years, several online tools have emerged that can accurately identify potential

plagiarism in scientific publications, making it easier for research institutions, students, and researchers to manage and check for plagiarism in their research outputs.[9] However, users must exercise caution and possess the knowledge to interpret the results of the software's analysis, as common technical terms or mathematical formulas found in multiple instances across different papers may not necessarily be considered plagiarism.[10]

The use of text-matching software has both positive and negative aspects.[2] It is important to note that software cannot determine plagiarism but can only identify certain similarities in texts that may constitute plagiarism: whether to regard a particular instance or work as plagiarism is a decision that must be taken by users based on their analytical and evaluative abilities.[11] Text-matching software has helped uncover basic forms of text recycling in journals; however, relying solely on text-matching software is not effective in preventing more complex forms of plagiarism. These limitations include the inability to detect interpretive and manual revisions that do not involve outright copying of entire text (e.g., reordering sentences or words or altering some words and their sequence while maintaining the original content and ideas).[12]

Methods

We used a mixed-method approach combining quantitative analysis (based on survey responses from a representative sample of lecturers, research students, and doctoral students in educational science from universities in Vietnam) and qualitative analysis of the regulations on plagiarism (in selected teacher-training universities). The approach involved analysing the relevant regulations and selected representative journals published by teacher-training universities to assess the current usage of text-matching

software in educational science research at those universities.

Samples for research

Text-matching software. We chose 13 widely used text-matching software programs (listed later) used in universities worldwide to investigate the usage of text-matching software in research in universities in Vietnam. The software programs were CrossCheck,[13] DoIT,[14] DupliChecker,[15] iThenticate (www.ithenticate.com),[16] Paper Rater,[17] PlagAware,[18] Plagiarism Check,[19] Plagiarisma,[20] PlagScan,[21] Quetext,[22] Urkund,[23] Turnitin,[24] and Viper Plagiarism Checker.[25]

Survey sample. The survey involved 104 lecturers, graduate students, and researchers from 18 universities in Vietnam specializing in educational research (Table 1). The participants were selected to represent the academic community in educational science and research (Table 2).

Case studies. We chose the following seven key pedagogical universities from the list of 18 to analyse the use of text-matching software: Thai Nguyen University of Education, Hanoi National University of Education, Vinh University, Hue University's University of Education, Hanoi Pedagogical University (two journals), Ho Chi Minh City University of Education, University of Science and Education – The University of Da Nang.

Survey questionnaire. The survey focused on the following topics: whether the university publishes any journals and whether those journals conduct plagiarism checks, whether the university uses text-matching

Table 1. Vietnamese universities and the number of researchers that participated in the survey

University	Number of Participants
Can Tho University	2
Dong Thap University	10
Thai Nguyen University of Agriculture and Forestry	4
Thai Nguyen University of Education	5
Hanoi National University of Education	4
Tay Nguyen University	8
Ton Duc Thang University	1
Vinh University	2
An Giang University	3
University of Education of Vietnam National University	25
Hung Vuong University	5
University of Science and Education – The University of Da Nang	11
Sai Gon University	1
Hue University's College of Education	7
Hanoi Pedagogical University 2	13
Tay Bac University	1
Pham Van Dong University	1
Ho Chi Minh City University of Education	1

Table 2. Status, age, and sex of the participants

Status	Number of Respondents	Age Group (Years)			Sex	
		<30	30–45	>45	Men	Women
Lecturer	82	0	66	16	43	39
Master's student	17	12	5	0	11	6
PhD student	5	0	5	0	4	1

software and requires plagiarism checks for scientific works, researchers' knowledge of text-matching software, proficiency in using text-matching software, and difficulties encountered when using text-matching software.

Data collection and analysis. The survey was conducted using Google Forms and sent to participating researchers. In addition, case studies were conducted on the seven selected universities based on the publications available on the universities' websites to provide insights into the use of text-matching software by those universities.

Results

Text-matching software

Of the 104 respondents, 92% said that their university publishes a journal or journals, which offer an outlet for researchers in those universities to publish their research. Among the survey participants, 59% answered that none of the journals stipulates a cap on similarity percentages of the text of the submitted articles, whereas 49% of the respondents answered that the journals do stipulate such a cap or limit (Table 3).

To further confirm that journals affiliated with universities do not require authors to check for plagiarism before submitting their manuscripts for publication, the study accessed the websites of scientific journals from the seven universities (Table 4). Whereas all of them had drawn up specific guidelines for submission

of manuscripts, none had any stated policy on plagiarism. Of the seven, two accepted submissions online, namely, Hanoi National University of Education and Ho Chi Minh City University of Education (both published journals with the same title, *Journal of Science*).

The absence of any policies on plagiarism contributes to greater plagiarism in universities and makes it difficult for them to improve the quality of their journals. This observation is consistent with that of Vuong Quan Hoang on violations of research ethics in Vietnam.[3]

Use of text-matching software

Of the 104 participants, 35 said that their university uses software to detect plagiarism; it is clear that the use of text-matching software to check theses, dissertations, and research topics is yet to be widely adopted. Although the universities are aware of the problem of plagiarism in research, the use of text-matching software is not mandatory in most universities in Vietnam.

To establish a basis for ascertaining the use of text-matching software in universities in Vietnam, we accessed the websites of seven key teacher-training universities to analyse their requirements related to checking for plagiarism (Table 5).

Of the seven universities, five check for plagiarism in the dissertations and theses submitted by their students, use text-matching software for the purpose, and stipulate that similarity in no case must exceed 20%. In particular, Ho Chi Minh City Pedagogical

Table 3. Regulations concerning plagiarism

Question	Yes		No	
	Number of Responses	%	Number of Responses	%
Does the university where you work and study have an affiliated scientific journal?	96	92.3	8	7.7
Do the submission guidelines of the faculty-affiliated journals at the university that trains teachers specify plagiarism checks through software?	43	41	61	59

Table 4. Journals and universities selected for case studies

Journal	ISSN	University	Submission Online/ Offline	Website
Journal of Science	0868-3719	Hanoi National University of Education	Online	http://stdb.hnue.edu.vn/portal/index.php?tabid=1019
HPU2 Journal of Sciences	2815-3567	Hanoi Pedagogical University 2	Offline	https://sj.hpu2.edu.vn/index.php/journal
TNU Journal of Science and Technology	1859-2171	Thai Nguyen University Of Education	Offline	http://jst.tnu.edu.vn/jst
Vinh University Journal of Science	1859-2228	Vinh University	Offline	https://tckh.vujs.vn/
Journal of Science	1859-1612	Hue University of Education	Offline	https://jst-ud.vn/jst-ud
Journal of Science and Technology	1859-1531	University of Science and Education – The University of Da Nang	Offline	http://ued.udn.vn/page/tap-chi-khoa-hoc.html
Journal of Science	1859-1272	Ho Chi Minh City Pedagogical University	Offline	http://tapchikhgdkt.hcmute.edu.vn/

Table 5. Policies and software used for detecting plagiarism in theses and dissertations in educational science at seven key teacher-training universities in Vietnam

University	Plagiarism Policy	Use of Text-Matching Software	Website
Hanoi National University of Education	Yes	Turnitin	http://stdb.hnue.edu.vn/portal/index.php?tabid=1019
Hanoi Pedagogical University (two journals)	Yes	Turnitin	http://www.hpu2.edu.vn/en/
Thai Nguyen University of Education	Yes	Turnitin	https://tnue.edu.vn/phan-mem-quan-tri
Vinh University	Yes	DoIT	http://.vinhuni.edu.vn/
Hue University of Education	No	No	http://tckhgd.huce.vn/
University of Science and Education – The University of Da Nang	No	No	http://ued.udn.vn/page/tap-chi-khoa-hoc.html
Ho Chi Minh City Pedagogical University	Yes	iThenticate	http://tapchikhgdkt.hcmute.edu.vn/

University has laid down clear regulations on research ethics, types of plagiarism, acceptable level (percentage) of similarity, and consequences for violators. Among the five, three use Turnitin, one uses DoIT, and one uses iThenticate. These universities have bought licenses to allow users access to the software package without any additional fees. However, two universities, namely, Hue University of Education and University of Science and
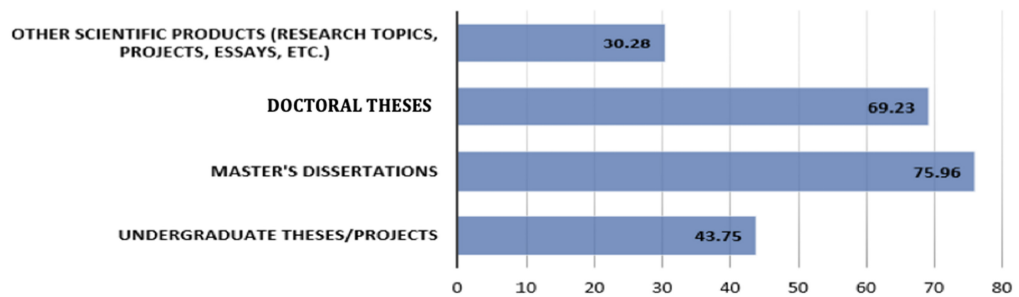
Figure 1. Plagiarism checks by universities on different types of theses and scientific products.

Education – The University of Da Nang, neither use text-matching software nor have any regulations on checking text similarity.

Plagiarism checks are used for different types of text (Figure 1): nearly 76% of the respondents indicated that such checks are conducted on master's theses, whereas the corresponding proportion was only about 30% for other texts such as project reports and essays.

Use of text-matching software among researchers

The survey asked the participants about several aspects related to plagiarism and the use of text-matching software.

Researchers' understanding of plagiarism: The survey showed that the participants' understanding of the scope of plagiarism was inadequate (Figure 2). Almost everyone (97%) believed that plagiarism is limited to copying a paragraph from someone else's text, and most (93%) understood that plagiarism

extends to copying someone else's ideas too; on the other hand, only 18% were aware that plagiarism may apply to recycling one's own text and only 21% were aware that claiming somebody else's work as one's own also constitutes plagiarism. This lack of comprehensive understanding contributes to plagiarism in scientific research.

The responses also showed that for researchers, plagiarism is not a particularly serious issue and they address it only when compelled to: a little over 95% of the respondents admitted to using plagiarism checking software only when it was mandated by the university and approximately 88% when it was mandated by the publishers (Figure 3). On the other hand, only 26% said that they use such software voluntarily. To be fair, 97% of the respondents also said that insufficient budget for buying such software was a serious hurdle (Figure 3), which they overcame by using software provided free of charge by academic institutions.
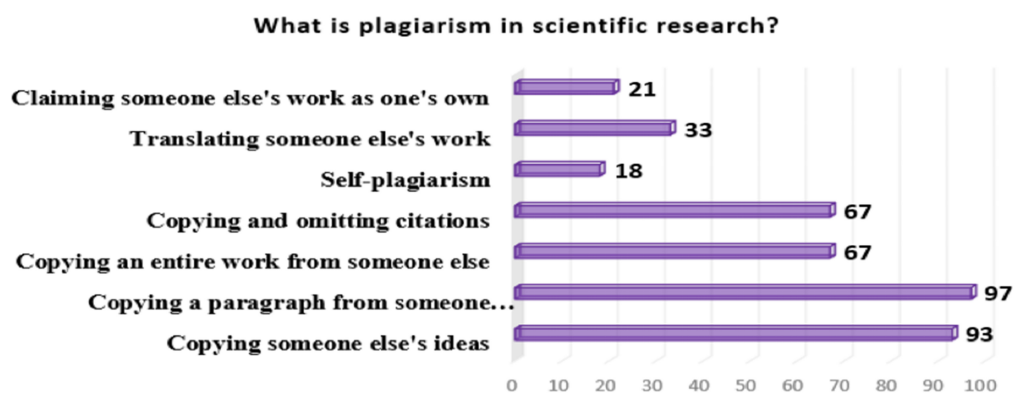

Figure 2. Researchers' understanding of plagiarism in research.
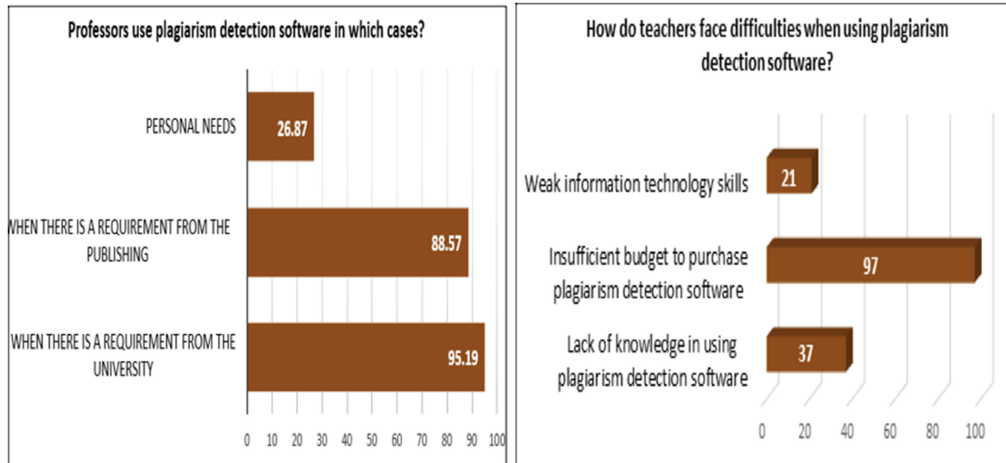
Figure 3. Researchers' awareness of plagiarism checking and challenges to using text-matching software.

The respondents also showed insufficient awareness of various text-matching software: only about 60% were familiar with Turnitin and only 50% with DoIT (Figure 4).

The use of a given software matched the awareness of that software: of the 14 software packages listed, only Turnitin and DoIT were frequently used, and even those were used by slightly fewer than 50% of the respondents (Figure 5).

Turnitin is bought and provided by universities, granting access to lecturers, researchers, and students (as evident from Table 5). The users therefore incur no cost, and universities insist that the software be used for plagiarism checks before the results of any research are published. The second most popular text-matching software is DoIT, an indigenous product enriched with significant amounts of text in Vietnamese language. Of
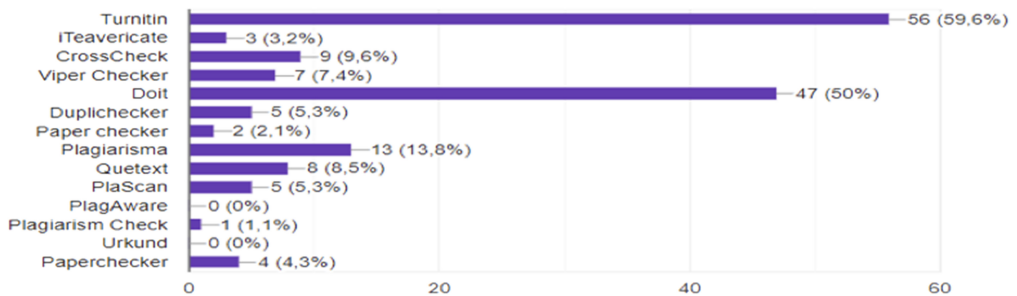

Figure 4. Level of awareness of text-matching software among educational researchers.
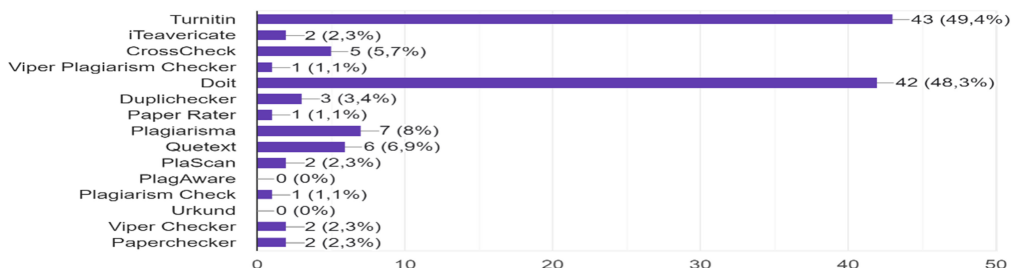

Figure 5. Frequency of text-matching software usage by educational science researchers in Vietnam.

the remaining 12 software packages, 10 are used only infrequently and 2, PlagAware and Urkund, are unknown and never used by Vietnamese researchers.

Discussion

The survey results show that plagiarism still exists and occurs at some universities that do not require their students to use software to check for similarity.

Text-matching software is a useful tool in detecting plagiarism and ensuring transparency and quality in scientific publications. Promoting the use of such software in academic publishing in Vietnam will play a crucial role in enhancing the quality and credibility of the country's research. However, such use is yet in the development stage and is neither widespread nor mandatory. Three possible reasons explain this state of affairs.

- **Lack of awareness**. Some scientific journals and researchers in Vietnam continue to be unaware of publishing ethics in general and of the importance of using text-matching software in particular.
- **Technical limitations**. Lack of funds as well as of technical resources and ability to use such resources, namely, text-matching software, constrain the routine use of software to curb plagiarism.
- **Scant regard for intellectual property**. Some researchers and authors do not yet have a full understanding of intellectual property rights and the consequences of plagiarism and, therefore, do not regard the use of text-matching software as necessary.

Our qualitative analysis aligns with the quantitative results of the survey. Many universities in Vietnam currently lack regulations on research ethics and plagiarism, criteria for plagiarism, procedures for checking plagiarism, and consequences of committing

plagiarism. The reasons for such lack are many.

1. Some universities and lecturers have only limited awareness of the importance of using text-matching software to prevent improper copying and to ensure integrity in research. Therefore, they do not consider plagiarism and inaccurate citation as important.

2. Routine use of text-matching software entails considerable technical and financial resources. Universities may find it difficult to invest in and deploy such software, especially universities with large numbers of academic staff and students or universities with limited resources.

3. Lack of clear regulations and policies from universities on the use of text-matching software can lead to inconsistencies and disparities in its application. The absence of regulations and policies may result in different practices among universities in using text-matching software.

4. Researchers' technological readiness in adopting and using text-matching software may be limited because of inadequate capabilities in using such tools.

5. Academic staff and students in universities in Vietnam are yet to integrate research ethics into their work and often do not take personal responsibility for their scientific products. This problem has been addressed by Samuel Bruton and Dan Childers regarding contrasting perspectives among scientists on plagiarism: one perspective suggests that it is the authors' responsibility, not that of institutions and schools, to ensure that their works are free of plagiarism whereas another

perspective is that both the parties, namely researchers (authors) and the institutions that employ them, including schools in the case of authors who are students, are accountable for all aspects of scientific publishing.

Plagiarism is a serious violation of the ethics of research, and text-matching software is a crucial tool for researchers, publishers, and research management agencies to deal with such violation because the tool enables quick and accurate identification of concerns related to plagiarism and thus promotes academic integrity. Researchers need to learn how to use such software to check for text similarities as well as to understand different forms of plagiarism. Several universities in Vietnam have recognized the importance of using text-matching software in scientific research and have incorporated requirements for plagiarism checks in theses, dissertations, and final reports of research projects and publications arising out of the projects. However, mandatory checking for plagiarism by universities remains relatively low, and a significant number of institutions are yet to acquire licensed copies of text-matching software. Consequently, researchers (including lecturers, research fellows, graduate students, and undergraduates) in educational science in universities in Vietnam have limited proficiency in using the software. They are primarily familiar with Turnitin (provided by their institutions) and the domestically produced DoiT, but are only marginally aware of other software packages available worldwide.

References

1. Velip PK. Ethical issues in research writing. *IJTSRD*. 2018;Volume-2():2429-2432. [CrossRef]
2. Marjanović Jakovljević M, Tomašević V, Živković D. *Anti-plagiarism Software: Usage, Effectiveness and Issues*; 2015:119-122. [CrossRef]
3. Masic I. *Plagiarism and How to Avoid It 19* . 2 Definition of the Plagiarism Problem. 2020; 163-177.
4. Kashkur M, Parshutin S, Borisov A. Research into plagiarism cases and plagiarism detection methods. *Sci J Riga Tech Univ Comput Sci*. 2011;42(1):138-143. [CrossRef]
5. Foltýnek T, Meuschke N, Gipp B. Academic plagiarism detection: A systematic literature review. *ACM Comput Surv*. 2020;52(6):1-42. [CrossRef]
6. Shankar P, RC. Anti-plagiarism [software] – *A tool to ensure quality research output*. *J Chem Inf Model*. 2013;53:1-10.
7. Bretag T, Mahmud S. A model for determining student plagiarism: electronic detection and academic judgement. *J Univ Teach Learn Pract*. 2009;6(1):57-69. [CrossRef].
8. Heckler NC, Rice M, Hobson Bryan CH. Turnitin systems: A deterrent to Plagiarism in College Classrooms. *J Res Technol Educ*. 2013;45(3):229-248. [CrossRef].
9. Lykkesfeldt J. Strategies for using plagiarism software in the screening of incoming journal manuscripts: recommendations based on a recent literature survey. *Basic Clin Pharmacol Toxicol*. 2016;119(2):161-164. [CrossRef]
10. Birkić T, Celjak D, Marko Cundeković SR. *Analysis of Software for Plagiarism Detection in Science and Education*; 2016.
11. Foltýnek T, Dlabolová D, Anohina-Naumeca A, Razı S, Kravjar J, Kamzola L, et al. Testing of support tools for plagiarism detection. *Int J Educ Technol High Educ*. 2020;17(1):1-38. [CrossRef]
12. Gasparyan AY, Nurmashev B, Seksenbayev B, Trukhachev VI, Kostyukova EI, Kitas GD. Plagiarism in the context of education and evolving detection strategies. *J Korean Med Sci*. 2017;32(8):1220-1227. [CrossRef]
13. Crossref.org. Available at: https://www.crossref.org/services/similarity-check/.
14. Doit.uet.vnu.edu.vn. Available at: http://doit.uet.vnu.edu.vn/.
15. duplichecker.com. Available at: http://www.duplichecker.com.
16. iThenticate.com. Available at: http://www.ithenticate.com/.

17. Paperrater.com. Available at: https://www.paperrater.com/.

18. Plagaware.com. Available at: https://www.plagaware.com/.

19. plagiarismcheck.org. Available at: https://plagiarismcheck.org/.

20. Plagiarisma.net. Available at: http://plagiarisma.net.

21. Plagscan.com. Available at: https://www.plagscan.com.

22. Quetext.com. Available at: https://www.quetext.com.

23. Urkund.com. Available at: https://www.urkund.com/.

24. Turnitin.com. Available at: www.turnitin.com.

25. Scanmyessay.com. Available at: https://www.scanmyessay.com/.

ease/publications

ese/ European Science Editing

European Science Editing is an official publication of EASE. It is an open access peer-reviewed journal that publishes original research, review and commentary on all aspects of scientific, scholarly editing and publishing.

https://ese.arphahub.com/
https://www.ease.org.uk
https://twitter.com/Eur_Sci_Ed
https://www.linkedin.com/company/easeeditors/