

---

☰ Menu

---

1 MIN AGO   EDIT

# Plagiarism Detection Software

## Plagiarism Detection Software and Its Appropriate Use

Megan von Isenburg, Marilyn H. Oermann, and Valerie Howard

**NURSE AUTHOR & EDITOR, 2019, 29(1), 4**

Plagiarism by authors in manuscripts and by students in assignments is not a new problem, but it has been compounded by the ease of “copying and pasting” content from articles and other information sources on the Internet. Using content developed by others may be intentional but more often, this is an inadvertent error. With software tools that are currently available, editors, teachers, reviewers and others can more easily identify plagiarism and do not need to rely on their own abilities to recognize similarities to previously published text (Higgins, Lin, & Evans, 2016). Plagiarism detection software can be used to scan author manuscripts and student papers in a few minutes, matching what they have submitted to already

published work. This article examines plagiarism detection software and provides some recommendations for appropriate use by editors and authors.

## **WHAT IS PLAGIARISM DETECTION SOFTWARE?**

Plagiarism detection software assesses the similarity of content in papers with published literature and other types of information. The software compares the author's text against abstracts and citations in PubMed/MEDLINE; literally millions of journal articles, conference proceedings, and books from leading publishers (including Elsevier, Lippincott, Ovid, Sage, Springer, and Wiley Blackwell, among many others); and varied databases such as EBSCOHost, Gale InfoTrac, and ProQuest. In addition, plagiarism detection software searches the Internet for similar content. iThenticate, a leading software program, has its own web crawler that indexes more than "10 million web pages daily" (iThenticate, 2018). With plagiarism detection software, all types of documents can be checked: manuscripts, written assignments for courses, grants, theses and dissertations, other scholarly projects, and other types of reports.

With some plagiarism detection products, such as iThenticate, documents uploaded for assessment and the results are stored in their private and secure database, available only to the author who can delete the files from the system. Documents are not shared and cannot be used by others. However, this is not true for some other software programs, and authors are cautioned to investigate this before using.

Turnitin is a product designed to check the originality of student papers. It was made specifically for classroom use and for reviewing student work. Turnitin can be integrated within learning management systems for students to review their own papers before submitting them and teachers to then assess those papers for plagiarized content. In addition to searching databases and the Internet similarly to

iThenticate, Turnitin compares students' papers against millions of other student papers in its database (Turnitin, 2018).

## **USE IN ACADEMIC PUBLISHING**

Numerous health sciences and biomedical journals worldwide use software to detect plagiarism in manuscripts submitted for publication (Butler, 2010; Garner, 2012; Roach, Gospe, Ng, & Sahin, 2014). Despite this widespread adoption, there is not a standard for what similarity rate, or percentage of text matching that of another document, constitutes plagiarism. Similarity rates in manuscripts as low as 15% and as high as 80% have been used to operationally define plagiarism (Garner, 2012; Higgins et al., 2016; Kalnins, Halm, & Castillo, 2015; Park et al., 2017).

The prevalence of plagiarism has been measured in different settings. Research using plagiarism detection software has uncovered plagiarism rates of between 6% and 23% of scanned manuscripts (Butler, 2010). One study suggested that 3000 papers added to MEDLINE each year have strong similarity to papers already indexed in MEDLINE (Garner, 2011). Park et al. (2017) detected plagiarism (as defined as a 20% similarity index) in 9% of articles scanned, though this rate declined over a 5-year period. Despite the labor and economic costs of scanning articles with plagiarism detection software, one study of plagiarism in two radiology journals concluded that the process is a cost-effective and important step for avoiding the risks of publishing plagiarized articles (Kalnins et al., 2015).

Editors and journal staff using software to detect plagiarism should consider the role of text recycling or “unattributed rehashing” (Garner, 2012) from authors' prior manuscripts. Recent guidance from the publishing platform BioMed Central and the Committee on Publication Ethics (COPE) notes that some text recycling, also known as self-plagiarism, is justified. Those interpreting similarity rate reports from

plagiarism detection software should take into consideration the amount of text recycling, where it occurs in the article, and whether it is acknowledged (BioMed Central, n.d.). In some cases, text recycling is positive, as it can provide a standard way of describing the same methods that resulted in multiple publications (Moskovitz, 2017).

## **USE IN STUDENT PAPERS**

In an integrative review on plagiarism in nursing education, Lynch et al. (2017) argued that it is important to address plagiarism in the academic setting because of potential links to professional misconduct. They identified self-reported plagiarism rates among undergraduate nursing students as being between 38% to 60%, but caution that not all plagiarism is deliberate. Some accidental plagiarism can be due to poor organization, a lack of skills, and a lack of knowledge about good writing practices.

Similar to publishers, many graduate and professional academic programs have implemented plagiarism detection software as a means to detect plagiarism in assignments that are turned in for a grade (Dahl, 2007; Marshall, Taylor, Hothersall, & Perez-Martin, 2011; Marusic, Wager, Utrobicic, Rothstein, & Sambunjak, 2016; Walker, 2010); as a tool for teaching writing skills (Hampton, 2018); and as a method for providing formative feedback to students on writing assignments before the assignment is due (Whittle & Murdoch-Eaton, 2008). Student perceptions of the use of plagiarism detection software to screen their writing are reported as being fairly positive (Dahl, 2007; Whittle & Murdoch-Eaton, 2008), but teachers should be mindful to respect student privacy and should gain student approval if the software stores a copy of the student's work for future use (Brinkman, 2013).

In teaching about plagiarism and ethics in writing, one study of two master's level health programs found that software alone is not enough to reduce the rate of

plagiarism in assignments, but that adding lectures and modules on plagiarism resulted in lower plagiarism rates (Marshall et al., 2011). A systematic review of interventions to promote publication integrity found that while results varied across studies, plagiarism software, practical exercises, and other active learning techniques reduced plagiarism in some settings (Marusic et al., 2016).

## **HOW DOES IT WORK?**

After uploading the document at the product's website, the software highlights text that matches text in its database. This alerts authors to check this content in their own paper before submitting it, and editors and teachers to review the text that is similar to published work. The text that is highlighted includes content that has been cited properly by the author and thus is not plagiarized as well as text that is similar to published content. The highlighted text is used to generate a score that represents the percentage of similarity between the author's document and published content. Varied colors are used for the highlighting to represent different percentages of similar text.

## **HOW TO USE IT**

For responsible use of plagiarism detection software, editors and teachers cannot make a decision about text being plagiarized based only on the score generated by the software: they need to review every segment of highlighted text to determine if the content is too similar to published work or has been properly cited and can be used. Some similarity is acceptable, such as commonly used terms and phrases in nursing, lists of author names in the text (for example, when using the American Psychological Association reference style), and some text recycling. Decisions are needed by editors and publishers as to whether students' theses and dissertations submitted to ProQuest are considered a form of published content. Note that both the title page and reference list should be excluded from the document when checking for plagiarism.

If using the software as an author, it is worthwhile to check the document for plagiarism about half way through the paper. This may help identify some phrases and sentences that are too similar to published work, and need to be revised, before the entire paper is written.

## COMPARISON OF A FEW PRODUCTS

There are many plagiarism detection software programs including some that are free. Tables 1 and 2 provide a comparison of iThenticate and Turnitin with two free products.

Table 1. Comparison of iThenticate and Turnitin

Service	iThenticate	Turnitin Feedback Studio (formerly just Turnitin)
URL	<a href="http://www.ithenticate.com/">http://www.ithenticate.com/</a>	<a href="https://www.turnitin.com/">https://www.turnitin.com/</a>
Who are the users?	Editors, faculty, and researchers	Faculty and students in a classroom setting/online
What does it scan against?	Millions of journal articles, conference proceedings, books; databases such as PubMed/MEDLINE, EBSCOHost, ProQuest; more than 10 million web pages/day	929 million archived student papers; 67 billion web pages; millions of journal articles and subscription content sources; content with Crossref
What does user do?	Click on the 'Upload a file' link; choose file and upload, or 'cut and paste' text to be analyzed	Faculty use access code to create a class ID; students use class ID to create account and upload documents; users can also 'cut and paste' text to be analyzed
How are results displayed?	Identifies material in database that matches text (highlighted). User receives a Similarity Report with a % of similarity between the submission and content in the database. Since the highlighted text includes properly quoted and cited text (not necessarily plagiarized), it is extremely important to review all sections and verify that it has been properly referenced or summarized. This can help identify instances of unintentional plagiarism	Students and faculty receive a % similarity score. Faculty can view all similarity reports in the assignment 'inbox' page and insert feedback for students; multiple drafts can be uploaded for review. Be aware that the similarity index could be high since it is comparing with previous submission drafts from the same student. Need to review all highlighted text carefully as it includes properly quoted and cited text
Added Features	Organizational subscriptions can include a private, custom database for document-to-document comparison	Integrates with several learning management systems Instructors can leave voice and text comments in various forms
Privacy	Submissions are not added to database and are kept private for the user	Submissions are added to database by default (can be changed by faculty member)
Comments	<ul style="list-style-type: none"> <li>Partnership with ProQuest (for thesis and dissertations)</li> <li>Multiple web resources to assist authors and editors including blogs, papers and reports on plagiarism</li> <li>Can share documents and collaborate in groups</li> </ul>	<ul style="list-style-type: none"> <li>Partnership with Gradescope (automates grading of existing or new assignments)</li> <li>Turnitin has been 'reimagined for the modern classroom' and is now named 'Turnitin Feedback Studio'</li> <li>Many resources to assist faculty, including but not limited to, standardized email messages to faculty / students, tutorials, and videos</li> </ul>
Creator	Turnitin	iParadigms
Cost	Individual: \$100 per credit. One credit is good for a manuscript of 25,000 words or less. Organizations: Volume discounts are available. Check to see if iThenticate is available to you through your employer.	Not available to individuals. Must be purchased through an organizational contract. Turnitin does not provide pricing on its website.

Table 2. Comparison of Two Free Products, Grammarly and Helioblast

Service	Grammarly	Helioblast
URL	<a href="https://www.grammarly.com/plagiarism-checker">https://www.grammarly.com/plagiarism-checker</a>	<a href="https://helioblast.heliotext.com/">https://helioblast.heliotext.com/</a>
What does it scan against?	"Billions of web pages" Published academic content from Proquest's academic databases	MEDLINE abstracts and titles (not full text)
What does user do?	Paste text into large box on website	Paste text into large box on website. Text is limited to 1,000 words, and users are encouraged to scan only an abstract or paragraph
How are results displayed?	Report indicates whether there are plagiarism issues. No specific score is provided	Results show abstracts that are similar to query. A score is provided but not explained
Added Features	Report also indicates how many writing issues, such as grammar, spelling, diction, syntax, and tone, are detected. These issues are not disclosed without the user creating a paid account with Grammarly. If user has an account, Grammarly can be integrated into MS Word applications to scan for writing issues and/or plagiarism	Two analysis tools are available: - Experts: other authors who write on similar topics - Implicit keywords identify terms in similar papers that were not in the original Search engine could be customized to other databases, so presumably faculty or publishers could contract with Helioblast to scan their own databases or existing papers or publications
Privacy	Does not store user writing publicly or in any database	Not disclosed
Comments	The free version of Grammarly offers a high level assessment of whether there are issues with plagiarism but provides no details	Because the service is basically limited to scanning abstracts against other abstracts, this tool may miss many forms of text reuse that could occur later in a paper
Creator	Grammarly is a for-profit company founded in 2008 that uses artificial intelligence to identify grammar issues in writing	Kim Menier

## SUMMARY

Plagiarism remains an ongoing issue in publishing and in academia, and software programs are available to detect similarities to published content, previously submitted student papers, thesis, dissertations, and websites. These free or for-purchase software programs can assist new writers with identifying instances of 'unintentional' plagiarism, and help editors and faculty members to identify all types of plagiarism. While using any plagiarism detection software, the user must be vigilant about reviewing the detailed results. A similarity score or index does not, alone, indicate plagiarism as properly referenced materials may be documented as 'similar'. In addition, some plagiarism detection sites retain ownership of the draft

submissions in the database and use these submissions for future similarity checks. While the software programs can be extremely helpful in detecting and counseling authors about plagiarism, we recommend that editors, faculty, researchers, and students receive proper instruction on the use, ultimate retention of submitted documents, meaning of results, and consequences of using any plagiarism detection system before implementation.

## REFERENCES

1. BioMed Central. (n.d.). How to deal with text recycling. [http://media.biomedcentral.com/content/editorial/BMC-text-recycling-editorial\\_guidelines.pdf](http://media.biomedcentral.com/content/editorial/BMC-text-recycling-editorial_guidelines.pdf)
2. Brinkman, B. (2013). An analysis of student privacy rights in the use of plagiarism detection systems. *Science and Engineering Ethics*, 19(3), 1255-1266. doi:10.1007/s11948-012-9370-y
3. Butler, D. (2010). Journals step up plagiarism policing. *Nature*, 466(7303), 167. doi:10.1038/466167a
4. Dahl, S. (2007). The student perspective on using plagiarism detection software. *Active Learning in Higher Education*, 8(2), 173-191. doi:10.1177/1469787407074110
5. Garner, H. (2012). Science publishing: How to stop plagiarism. *Nature*, 481(7379), 21-23. doi:10.1038/481021a
6. Garner, H. R. (2011). Combating unethical publications with plagiarism detection services. *Urologic Oncology*, 29(1), 95-99.
7. Hampton, M. D. (2018). A 5-year evaluation of faculty use of practical within-course writing development strategies in a DNP Program. *Nurse Educator*. doi:10.1097/nne.0000000000000589



8. Higgins, J. R., Lin, F. C., & Evans, J. P. (2016). Plagiarism in submitted manuscripts: Incidence, characteristics and optimization of screening-case study in a major specialty medical journal. *Research Integrity and Peer Review, 1*, 13. doi:10.1186/s41073-016-0021-8
9. iThenticate. (2018). FAQs: Plagiarism software. <http://www.ithenticate.com/products/faqs> Accessed December 8, 2018.
10. Kalnins, A. U., Halm, K., & Castillo, M. (2015). Screening for self-plagiarism in a subspecialty-versus-general imaging journal using iThenticate. *AJNR American Journal of Neuroradiology, 36*, 1034-1038. doi:10.3174/ajnr.A4234
11. Lynch, J., Everett, B., Ramjan, L. M., Callins, R., Glew, P., & Salamonson, Y. (2017). Plagiarism in nursing education: An integrative review. *Journal of Clinical Nursing, 26*(19-20), 2845-2864. doi: 10.1111/jocn.13629
12. Marshall, T., Taylor, B., Hothersall, E., & Perez-Martin, L. (2011). Plagiarism: A case study of quality improvement in a taught postgraduate programme. *Medical Teacher, 33*(7), e375-381. doi:10.3109/0142159x.2011.579201
13. Marusic, A., Wager, E., Utrobicic, A., Rothstein, H. R., & Sambunjak, D. (2016). Interventions to prevent misconduct and promote integrity in research and publication. *Cochrane Database of Systematic Reviews, 4*, Mr000038. doi:10.1002/14651858.MR000038.pub2
14. Moskovitz, C. (2017). Text recycling in health sciences research literature: a rhetorical perspective. *Research Integrity and Peer Review, 2*, 1. doi:10.1186/s41073-017-0025-z
15. Park, S., Yang, S. H., Jung, E., Kim, Y. M., Baek, H. S., & Koo, Y. M. (2017). Similarity Analysis of Korean Medical Literature and Its Association with Efforts to Improve Research and Publication Ethics. *Journal of Korean Medical Science, 32*, 887-892. doi:10.3346/jkms.2017.32.6.887

16. Roach, E. S., Gospe, S., Ng, Y. T., & Sahin, M. (2014). Trust but verify: the introduction of plagiarism detection software. *Pediatric Neurology, 50*(4), 287. doi:10.1016/j.pediatrneurol.2013.12.021
17. Turnitin. (2018). Website. <https://www.turnitin.com/about/content>. Accessed December 9, 2018
18. Walker, J. (2010). Measuring plagiarism: Researching what students do, not what they say they do. *Studies in Higher Education, 35*(1), 41-59. doi:10.1080/03075070902912994
19. Whittle, S. R., & Murdoch-Eaton, D. G. (2008). Learning about plagiarism using Turnitin detection software. *Medical Education, 42*, 528. doi:10.1111/j.1365-2923.2008.03058.x

## ABOUT THE AUTHORS

Megan von Isenburg, MSLS, is Associate Dean for Library Services and Archives at Duke University School of Medicine, Durham, North Carolina, USA. [megan.vonisenburg@duke.edu](mailto:megan.vonisenburg@duke.edu)

Marilyn H. Oermann, PhD, RN, ANEF, FAAN, is Thelma M. Ingles Professor of Nursing, Duke University School of Nursing, Durham, North Carolina, USA. She is Editor of *Nurse Educator* and the *Journal of Nursing Care Quality*. [marilyn.oermann@duke.edu](mailto:marilyn.oermann@duke.edu)

Valerie Howard, EdD, MSN, RN, CNE, FAAN, is Associate Dean for Academic Affairs and Professor of Nursing at Duke University School of Nursing, Durham, North Carolina, USA. [valerie.howard@duke.edu](mailto:valerie.howard@duke.edu)

2019-29-1-4 von Isenburg et al