Scholarly Communication and Plagiarism Check

K. Manjula,
Assistant Librarian,
Mysore University Graduate Library, Mysore.
E.mail: manju_bele@yahoo.in

Jayashri Jagadish,
Assistant Librarian,
Mysore University Library, Mysore.
E.mail: muljayashree@gmail.com

Dr. C. P. Ramasesh
Librarian,
Mysore University Library, Mysore.
E.mail: cpramasesh@gmail.com

Abstract

Paper depicts the trends in ensuring the originality of research output. Explains in detail the role of the popular anti-plagiarism software in avoiding the plagiarism in scholarly communications. The software packages: Turnitin and iThenticate are cited predominantly as examples for detecting the percentage of similar contents. Also dealt with the aspect of improving the quality of research submissions with the help of anti-plagiarism tools. Authors have narrated the experience of adopting the plagiarism check tools at the University of Mysore.

Key Words: Anti-plagiarism Software; Plagiarism- Overview; Turnitin; iThenticate.

Recently, scholars across the world have taken several measures to improve the quality of research output. The Institute for Scientific Information, Philadelphia, is bringing out several publications to measure the quality of publications as well as to capture the performance of the research institutions and individual research scholars by means of indicating impact factor and h-index. However, M/s. iParadigms, California, USA has introduced many software packages that can detect percentage of already published contents in the research submissions. By this one can improve upon the originality of the research submissions.

Plagiarism is usually explained in categorical terms: borrowing the ideas of others and presenting the same as if our own without giving credit to the authors. In other words, stealing someone's ideas and later on telling lies or defending act of literary theft. Plagiarism is simply an act of intellectual dishonesty.

Whenever a scientist or author borrows or paraphrases ideas belonging to others, he is expected to put the ideas within quotation marks and render complete bibliographic details with names of author/s. Even changing keywords and retaining the ideas without giving credit to authors is also treated as plagiarism. iThenticate is one of the most popular anti-plagiarism software being used to scan annually 2.3 million research papers and is highlighting to note that more than 10 million matches identified from the previously published content over the past 18 months. On account of this trend and with the intention of improving the original contribution in research output, a large number of publishers of journals now-a-days insist on plagiarism check before accepting the papers for publication. Elizabeth Wager says, “Promotion, appointments, and academic careers are relying on publication and in a way good for the publishers to open up some opportunities, I think if the pressure is too high, it will create an atmosphere in which the temptation to commit research misconduct is increased.” Hence, it is necessary to cite the source whenever authors quote ideas or sentences of others or paraphrases works of others. When the sources are cited, there should be uniformity in rendering references, depicting original author, title, edition, place of publication, publisher, year of publication and pagination. It is also important fact that a standard format has to be followed for rendering bibliography and references.
Reasons for plagiarism are many. In the past, ideas originated from the brain. However, of late, it comes from Google! The prominent reasons for plagiarism are

1. Lack of research skills.
2. Lack of command over the language.
3. Temptation to publish soon for career advancement.
4. Desire to get recognition soon.
5. Lack of time for understanding basic concepts and generation of new ideas.

Complaints against plagiarism and measures to avoid plagiarism was there even earlier days. Attempts were made to verify plagiarized contents or similar contents in the manuscripts manually. In fact, it was a tedious task, requiring enormous reading and memory. Now-a-days, in the digital era, the anti-plagiarism software packages scans the web information resources, online databases, electronic books, journals, theses, reports, monographs and proceedings of workshops and conferences and even the digital repositories of the institutions as well.

Anti-plagiarism software was first used at the Harvard, USA to check the plagiarized content in the student submissions of project works and assignments. Now, a large segment of academic institutions in the United States are relying upon the tool to avoid plagiarized content in research papers and reports. Working of the tool is explained here. Initially, the manuscript to be submitted for publication or doctoral degree will be scanned. In the next step, the software scans the world wide web, archives and databases will also be scanned to detect the similarity in content, word by word. Then it is likely you will see some red in your results. Common keywords and phrases may trigger in red. If there are complete sentences that are not original, directly taken from the already published sources, then the software will identify the original source of plagiarized content that was copied.

In the further steps, when you click on any red line, you can see original source for your reference and record; the tool allows you to verify paragraphs side by side.

Finally, you will be given an exact percentage that tells you to what extent your research paper/thesis is original. Use of anti-plagiarism software is not to be mistaken for fault finding mission. In fact, it is a fact finding exercise, aiming to identify same and similar content in already published sources and thereby avoiding it, with the intention of improving the originality of research contribution.

Some anti-plagiarism software packages that are commonly seen are

1. iThenticate, California, USA
2. Turnitin, California, USA
3. WriteCheck, California, USA
4. Viper, England
5. PlagAware, Ulm, Germany
6. PlagScan, Germany
7. Urkund, Sweeden
8. Docoloc, Germany
9. Plagiarism Checker X, New York, USA
10. Plag Tracker, Ukranian
11. Nitya D' Arch, Kottayam / Cochin, Kerala
Anti-plagiarism Software from Kottayam: NityaDY Arch scans Indian languages. Large majority of the customers for this commercial software is the researchers who have drafted research papers and theses in Malayalam. On the other hand 'Viper' is an open source software from England. The coverage of the software is about 10 billion pages. Whereas, iThenticate covers over 40 billion pages of information before it throws away similar content borrowed from published sources.

Here below, a sample page depicting the plagiarized content or similar content is shown. There is provision to compare the manuscript and the published source side-by-side. The author can rectify and eliminate the plagiarized content or can render citation mentioning the original source and further, giving credit to the author.

The software scans quite a number of sources covering more than 40 billion pages of information before depicting the degree of similar content. Information sources made available on the World Wide Web, electronic books and journals, doctoral theses, proceedings of the seminars, conferences and workshops, institutional repositories of various academic and research organizations and their digital libraries hosted on the web.

1. Internet resources
2. Cross-Check database: Journals
3. Proceedings of Conferences and Seminars, Research Reports, Theses
4. Online Publications: Books, Journals and Periodicals

5. Institutional Repositories

The research submissions while brought for scanning, usually the entire theses has to be in single word or PDF file eliminating preliminary pages like declaration, certificate, acknowledgement, table of contents, list of charts, graphs and tables. Further, the appendices, comprising of glossary, questionnaires, indexes have to be removed while uploading for scanning. Finally, the similarity report is generated after excluding the following content from the research submissions.

1. Software facilitates exclusion of text that are rendered within quotes.
2. When the text or paragraphs carry references in the theses, the cited text gets excluded.
3. Make provision for exclusion of phrases and key-words used in the title of the theses. These phrases often appear in the text of the theses here and there frequently. Hence, this has to be exercised in the folder option given for exclusion.
4. There is common option to exclude abstract, methods and materials (methodology) and bibliography.
5. Provision is made for exclusion of small matches comprising of up to 10 (standard) or more words that appear in the text of the theses.
6. In the software, provision is made to exclude small sources comprising of up to 10 or more words.
7. A standard format has to be followed for rendering references/bibliography. Only in such cases, the paragraphs taken from various sources get excluded.
8. Uniformity and consistency should be there while rendering bibliographic references. The references and bibliography should not be copy and pasted because, neither the references and bibliography get excluded nor the cited text get excluded. Therefore, the text or the references should not be copied and pasted from Google.

Plagiarism in research reporting is a punishable offence. However, on account of the growing population of scientists and researchers and the pressure to publish, an increasing number of scientists and researches have taken to cutting corners, resulting in falsified research, fraudulent data, paraphrasing, duplication and blatant plagiarism. The undesirable attempts are classified under 'research misconduct'. This research misconduct creates a ripple effect of costly damages for researchers, organizations and the general public as well, including the impact on patients to a greater extent. Organizations may miss financial support on account of the frauds and misconduct and their undesirable impact on the public and society. Repurcussions of plagiarism are enlisted here below.

1. Immediate failure of research
2. Value of the research is questionable
3. Retraction of research paper
4. Black listing of scientist and institutions
5. Punishable offence / fine can be imposed in court of law

It is essential to overcome plagiarism; avoiding plagiarized content in research reporting is an essential requisite. The software supports fighting against plagiarism by projecting similar ideas drawn from published sources. It is also the fact that there is high risk of being blacklisted by the search engines. The boundary between plagiarism and quality research work is unclear. Popular publishers of scholarly journals fix tolerance limit or restrict the percentage of similar content. Some of the reputed publishers are rejecting 23% of the submissions on account of the plagiarism or self-plagiarism to an enormous extent. In case of research papers of journals, very little of similar content is expected, especially in the paragraphs covering introduction/background of the topic, definitions, objectives and methodology. The rest of the content will be usually nothing but interpretation and inferences drawn from the original data. Of course, findings of the research, discussions, recommendations and conclusions, there cannot be similar content unless the author cites similar studies conducted previously.

However, drafting doctoral thesis involves survey of literature covering a period of 2 to 5 decades. Further, many of the definitions and important studies conducted until that period in the discipline have to be discussed
in detail. Sometimes, about the organization selected for the study has to be explained, covering the aspects of establishment and developments, including products and services offered. Discussing such of these issues, one has to rely upon previously published sources, and hence the percentage of similar content will obviously increase. Percentage of similar content increases in case of

1. Descriptive research
2. Tracing origin and historical background
3. Narrating developmental aspects
4. Projecting views of authors
5. Theoretical subjects like philosophy, religion, sociology, literature, and so forth, where the research relies much on already published literature and old classics.

In most of the above cases, the percentage of similar content will be high. Hence, the tolerance limit may be allowed to an higher extent in social science and humanities than science and technology.

The University Grants Commission introduced the anti-plagiarism tools under Shodhganga Project. ‘Shodhganga’ is the digital repository of the Indian electronic theses and dissertations. UGC-INFLIBET Centre has created the website of Shodhganga to host electronic theses and dissertations of Indian universities on open access platform so that, academicians and research scholars can access the theses for reference and downloading. The database also helps in avoiding duplication of research work at the national level.

The University of Mysore has made the scanning of doctoral theses with anti-plagiarism software mandatory way back in June 2014 to monitor the originality of research output. Several workshops/orientation programmes have also been conducted to create awareness among the research scholars and to provide guidelines for plagiarism check before submission of theses. The workshops/orientation programmes were conducted in the P. G. Departments, University Library and affiliated institutions as well. The guidelines pertaining to submission of theses for plagiarism check was circulated among the P. G. Departments and also hosted on the University Web Site for reference.

The scanning of research manuscripts and doctoral submissions helps in detecting the plagiarized content if any, and finally projects the percentage of similar contents taken from already published sources. It is an opportunity for the research scholars and research supervisors to rectify the contents of their research reports and ensure the originality of research contributions. The merits and usefulness of the tool are predominantly to

1. Improve originality of research contribution.
2. Correct the statements/quotes in case of inaccuracies or omissions.
3. Rectify the citations when they are incomplete
4. Avoid plagiarism while referring to similar content depicted
5. Trace the original information sources, contents and authors
6. Reduce percentage of similar content in case it is very high.

Therefore, it is necessary to check scholarly communications with anti-plagiarism tools. If you purchase a car, you have to verify the titles. If you purchase a house, you have to go through a title deed. Similarly, we need to verify the intellectual property before we try to publish and sell it.

References


Seglen, P O (1997), Citations and journal impact factors: questionable indicators of research quality. Allergy. 52: 1050-
1056.

Benson Honig& Akanksha Bedi (2012): The fox in the hen house: A critical examination of plagiarism among members of
the academy of management. Academy of Management Learning and Education, 11(1), 101-123.