SciencePG
Science Publishing Group

Survey and Comparison between Plagiarism Detection Tools

Mahmoud Nadim Nahas

Faculty of Engineering, King Abdulaziz University, Jeddah, Saudi Arabia

Email address:
mnahas@kau.edu.sa

Abstract: With the spread of using Internet in teaching and learning, and in literature reviews, plagiarism has spread widely, where researchers take readymade researches and claim them to themselves. The action of "copy and paste" has become extensive in all fields, and the scientific research domain is no less than other spheres. This has prompted many international organizations to produce detection tools to spot plagiarism in scientific articles and research reports. Many computer programs exist to identify plagiarism in scientific papers and essays in order to establish scientific honesty and to upgrade the level of university researches. Some of these programs are free, while some are commercially available. This paper presents comparison between the most famous plagiarism detection programs. The comparison helps individuals as well as organizations to select the plagiarism detection program that is most relevant for their objectives.

Keywords: Plagiarism, Research Misconduct, Research Ethics, Internet Cheating, Plagiarism Detection Programs, Plagiarism Detection Tools

1. Introduction

Plagiarism is to take the work of others and claim it to be yours [1]. This is fraud, forgery, piracy and copyright infringement. In the internet world, it can be similar to other "cybercrimes". Taking the work of others can be in different forms. The most important one is not mentioning the source. It can be direct "copy and paste" or taking the idea but modifying the wording. Plagiarism as such includes any improper use of any work of others without giving them a credit.

Since the emergence of the internet, plagiarism has become an easy job through the copy and paste procedure, which can take few minutes, while the original work may have taken months of hard work.

The importance of detecting plagiarism is clear, although the consequences differ from one case to another. For instance, there is no comparison between the case of a secondary school student copying homework from a classmate, and a graduate research student taking the results of other researchers and claiming them as obtained from lab work. This is not a call to tolerate the former case, because it leads to weakness in the knowledge of the secondary school student. However, in the latter case, the graduate student will obtain an unmerited qualifications or promotion.

Since the internet facilitated plagiarism, the cure came also from the internet. All the programs that detect plagiarism are online. Some of these programs are free, while the rest are paid for.

This paper surveys the different available programs for identifying plagiarism and gives brief descriptions of each. Some of these programs are good for schoolteachers and university professors to detect plagiarism in the students' essays and homework, while the more sophisticated ones are good for scientific journals and research funding organizations.

2. Methodology

It is mentioned above that there are several online plagiarism detection tools. The author has selected the tools that are currently most popular in the field of plagiarism detection. Brief descriptions of their main features are given and the personal experience is also mentioned.

3. The Open Programs for Plagiarism Detection

These programs are available free of charge and can be used easily online. Here are the most famous ones:

3.1. Google Search Engine

Although google site [2] is not intended for plagiarism detection, it can be used for this purpose. It is suitable for short texts and gives quick results showing the sites that have published similar texts (in all forms including htm, docs and pdf). Long texts must be divided into several short texts to be inspected separately. This is why it is of limited features.

3.2. Dustball

The University of Maryland in the USA provides this program [3] for academic staff to check for plagiarism in the students' assignments. Although it is free, but when I tried to use it, I got a message saying: "Free Limit Reached. The free service has already reached its daily limit of free uses (1,000 per day, worldwide!). This service costs money to maintain. Please help pay to keep the service running. The cost is $8 a month and you can cancel at any time".

According to the instructions given, it can be used either by copying the text and pasting it in the provided window, or by loading a Word file. In both cases, it will take you to a google page to show you the probable sources for the text you are inspecting for plagiarism. Although it is designed for academic staff to scrutinize the students' assignments, it cannot be used to compare two texts or two files to discover copying an assignment from each other.

3.3. Dupli Checker

This program [4] can also be used by either pasting the text in the provided window, or by loading a Docs or Text file. It is limited by 1000 words per search. After pressing the "Search" button, it will show you the texts that may have been taken from other sources. Each text will have a button, which, when pressed, show you the probable source. Alternatively, you may have this message "No Plagiarism Detected". If you try to use it again you will get a message saying: "Your Maximum Limit To Search Exceeded. Please Signup Free Or Login If Already Registered".

3.4. Plagiarisma

This program [5] can be used in three different ways. You can use it directly online without registration, online with registration, or by downloading it on your computer to use it from thereof. Online without registration works by pasting the text in a special window, then pressing the search button. With registration, one can load a complete document for inspection. The document can be a Word, Text, Acrobat or htm file. There is also option in this method for comparing two texts against each other, giving the possibility to check any similar assignments or essays.

The result of the search comes in a list of possible plagiarized texts together with the probable sources. I have tried the program by copying and pasting an abstract of a paper from a scientific journal. The program showed 14 probable sources. The first one was the source that I have taken the abstract from, while the rest of the sources were the papers that cited my source.

In the third method of the downloaded program, the program indicates to the text that is not taken from any source, by the word "Unique".

3.5. Academic Plagiarism

Academic Plagiarism [6] is free online program that requires registration to be used. Both methods of copy and paste (of maximum 700 words), and uploading files are available. The result is given in a report that can be seen by pressing the "Open Report" button. However, in the free version of the program, the result shows the probable plagiarized texts without giving the possible sources. The use of this free version is also limited by three times a day. The full version of the program has all the features but it is not free.

3.6. The Plagiarism Checker

The Plagiarism Checker [7] is a free easy-to-use online program. It uses the method of copy and paste only. No file can be loaded. It can be used in any language because it depends on google search engine.

3.7. Plagiserve

This is a simplified program [8] that requires registration to enable you to use it. However, the result is not immediate. It is sent by e-mail within 12 hours.

4. The Commercial Programs for Plagiarism Detection

These programs are available online through payment or subscription. Here are the most famous ones:

4.1. EVE2

This program [9] is available at reasonable price. It is suitable to students and teachers alike. However, it may not be that good for scientific journals and research funding agencies. It searches the internet as well as the databases on the net. It gives a good report about the texts that are similar to the examined text as well as their sites.

4.2. Plag Aware

This program [10] searches the internet to detect plagiarism. One of its important features is the possibility to compare several texts together giving the opportunity for teachers to check plagiarism of students from each other. The program was originally introduced to detect plagiarism in German language. However, it can now be used for English and Japanese.

4.3. Plag Scan

PlagScan [11] detects plagiarism through the internet as well as in its own database. Some scientific publishers use this program because it can be used for all languages that use UTF-8 code as well as Latin and Arabic.

4.4. Check for Plagiarism

This program [12] detects plagiarism by searching the internet and its own database. It can compare few texts against each other. It is available for students and teachers at reasonable price. It is used for few European languages as well as Arabic, Korean and Chinese.

4.5. Plagiarism Detection

This program [13] checks the internet, the databases and the electronic libraries that are available on the internet. It is used by university professors as well as students, because it is available at reasonable price list depending on the number of words in the document to be inspected (ranging from 10,000 to 500,000 words).

4.6. Write Check

This program [14] is the simplest of three programs for plagiarism detection presented by Turnitin Company. It is intended for students, as it provides other services for them while writing their researches. It correct the language (spelling, grammar and style). It also correct the citation style. At the same time, it examines the text to make sure it is free of plagiarized texts. The company claims that this program uses the same technology used by its more advanced program, Turnitin. This is to say that "writecheck" is presented at a special reduced price for students. It can be used once for $8.00 without signing any contract with the company. Some discount is also available for three times or five times uses.

4.7. Turnitin

As mentioned above, this program [15] is presented by Turnitin Company. It is the second one in its three programs. It is intended for the academic establishments. More than one million professors and more than twenty million students use this program. It detects for plagiarism in the internet as well as the company databases. It gives the result in what is called "originality report" which shows, in different colors, the probable plagiarized texts. According to Bull *et al.* [16], Turnitin is better than other programs, but it is slower and more expensive.

4.8. Ithenticate

This is the last program in the series of Turnitin programs [17]. It is intended for scientific publishers, research organizations and research funding agencies. The program checks the internet and the company databases for spotting plagiarism. It can also compare few texts together. iThenticate can be used with many European languages as well as Arabic, Turkish and some eastern languages. This is why it is used by about 80% of the journals that are in the Master Journal List of ISI (Web of Knowledge). The program is also available for individuals. However, it is the most expensive program for them.

5. Best Practices

Below I have identified four salient best practices for using plagiarism detection programs.

(1) According to the situation, the user should first decide whether to use the freely available plagiarism detection tools, or to go to the commercial tools, which involve cost.

(2) In the case of checking cheating in students' assignments, the free online software will do the job perfectly. The teacher may try to use them all in the beginning in order to choose the most appropriate one for her/his need.

(3) Master and doctorate students can also try the free online programs to make sure that their reports and theses are free of plagiarism. However, they may find that better tools would be preferred. They can select the software that are based on one-time payment. Better still is to use the ones that check the writing language (spelling, grammar and style), such as WriteCheck. This program costs few dollars for one time without signing any contract with the company. Some discount is also available for three times or five times uses.

(4) Organizations need to go direct to the paid tools. Research funding bodies, universities and publishers of books and journals require the most rigorous plagiarism detection tools. Turnitin and iThenticate are the best programs for research and education establishments. However, the latter of the two programs is the most expensive tool, but can be used with many European languages as well as Arabic, Turkish and some eastern languages.

6. Research Agenda

Challenges to the ethical conduct of research will continue to evolve as the Internet makes it easy for everyone to "copy and paste". Effort should continue to evaluate research articles and essays so as to seek solutions to the protection of intellectual property of people and to discover cheating and plagiarism in all work, starting from students' assignments to research work. Collaborations between research bodies' management and software developers should continue to overcome any difficulties and repair any bug in the plagiarism detection tools. In future research, attention should be paid to inform all students as well as researchers that their work will be subjected to checking, and they should be trained to use the available detection tools to check their work before submission. This can constitute a basic level of ethics among them all, and would further help to minimize the effort of checking.

7. Educational Implications

Educating of students about research ethics is necessary to ensure that they are sensitized to the "copy and paste" from other people work. They should learn to respect the intellectual property of others and make this as top priority in their research. Universities should consider introducing courses and/or seminars about research ethics and teaching their students how to avoid plagiarism.

8. Conclusions

When I started this work, I thought it would be an easy job. However, I discovered that it is not as I thought. Registration to use some of the programs is not a problem, but some programs require subscriptions. There are more programs other than the ones mentioned above, but apparently, their developers have stopped supporting them due to the emergence of more advanced programs. My conclusion is that the best comprehensive plagiarism detection program is "iThenticate" as per today. However, it is also the most expensive for individual users. Only universities, research centers and organizations, research funding agencies and scientific publishers of books and journals can afford the high cost of the program.

The free plagiarism detection programs differ from each other in the features, and it is advisable that one can use them alternatively to reach to the best suitable one according to the personal needs. There used to be a site for comparison of the plagiarism programs [18] but it is not updated since April 2011. This is why it is of no use.

References

[1] Ali, A. M. E. T., Abdulla, H. M. D. & Snasel, V. (2011), Overview and Comparison of Plagiarism Detection Tools. In V. Snasel, J. Pokorny & K. Richta (Eds.), *CEUR Workshop Proceedings* (pp. 161–172), Písek, Czech Republic: VŠB-Technical University of Ostrava.

[2] Google (2016). Retrieved from http://www.google.com.

[3] Klug, B. (2014). Retrieved from http://www.dustball.com/cs/plagiarism.checker.

[4] Dupli Checker (2012). Retrieved from http://www.duplichecker.com.

[5] Plagiarisma (2015). Retrieved from http://plagiarisma.net.

[6] Academic Plagiarism (2014). Retrieved from https://academicplagiarism.com.

[7] Plagiarism Checker (2016). Retrieved from http://www.plagiarismchecker.com.

[8] PlagiServe (2015). Retrieved from http://www.plagiserve.com.

[9] EVE Plagiarism Detection System (2000). Retrieved from http://www.canexus.com/eve/index.shtml.

[10] Plag Aware (2012). Retrieved from http://www.plagaware.com.

[11] PlagScan (2016). Retrieved from http://www.plagscan.com.

[12] Academic Paradigms (2016). Retrieved from http://www.checkforplagiarism.net.

[13] Academic Plagiarism (2016). Retrieved from http://plagiarismdetection.org.

[14] Turnitin (2015a). Retrieved from http://en.writecheck.com.

[15] Turnitin (2015b). Retrieved from http://www.turnitin.com.

[16] Bull, J., Collins, C., Coughlin, E. and Sharp, D. (2001). *Technical Review of Plagiarism Detection Software Report*, Luton, UK: University of Luton and Computer-assisted Assessment Centre.

[17] Turnitin (2015c). Retrieved from http://www.ithenticate.com.

[18] Top Ten Reviews (2011). Retrieved from http://plagiarism-checker-review.toptenreviews.com.