EDITORIAL

The boundary between plagiarism and research is surprisingly faint. After all, advanced research is only possible "standing on other's shoulders", as is often said [1]. Contrary to popular belief, plagiarism is not a problem reserved for students, but a phenomenon that seems to be increasing, as it is also found in professionals, scientists, and researchers. For example, a recent study on medical research found that more than a quarter of retracted publications in the last 20 years has been due to plagiarism and duplication issues [2].

There are different behaviors that lead to describe plagiarism as [3, 4]:

• slightly changing someone else's work and submitting it as your own.
• copying someone's words or ideas without giving credit.
• giving incorrect information about the source of a quotation.
• changing words but copying the sentence structure of a source without giving credit
• copying most of your text words or ideas from any source, giving credit or not

Figure 1 shows different methods of plagiarism in common practice. They are classified according to frequency and severity [1, 2, 5].
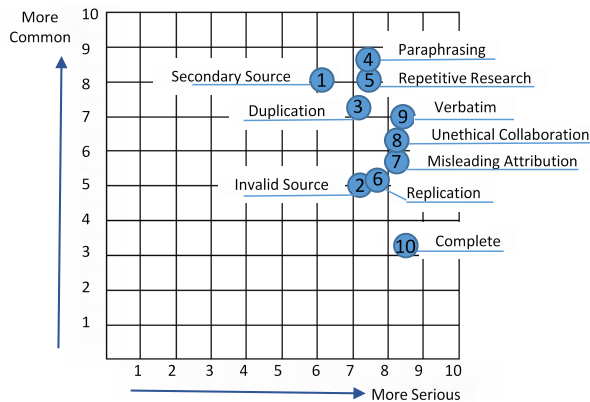

Figure 1 Types of plagiarism in research [5]

1. Secondary source: When a secondary source is used but only the primary source is cited. Commonness: 6.9/10, Seriousness: 6.4/10.
2. Invalid source: When incorrect or non-existent sources are cited. Commonness: 3.9/10, Seriousness: 7.3/10
3. Duplication: When a researcher uses material from their previous studies without attribution. Commonness: 6.3/10, Seriousness: 7.5/10
4. Paraphrasing: Rewriting someone else's words, but making the idea or even the research look like an original work. Commonness: 7.5/10, Seriousness: 7.6/10
5. Repetitive research: taking data or text from a similar research without attribution. Commonness: 7.1/10, Seriousness: 7.6/10
6. Replication: Submitting an article to multiple publications so it may be published more than once. Commonness: 4.2/10, Seriousness: 7.7/10

7. Misleading attribution: Providing inaccurate or insufficient information about the author in an article. Commonness: 4.8/10, Seriousness: 8.2/10
8. Unethical collaboration: Working together with other researchers in an unethical manner, such as using a written text or ideas resulting from a collaboration without citing the collaborative nature of the study nor participants involved. Commonness: 5.3/10, Seriousness: 8.2/10
9. Verbatim: Copying and pasting words without attribution. Commonness: 2.3/10, Seriousness: 8.4/10
10. Complete: Submitting an entire article from another researcher as one's own. Commonness: 2.3/10, Seriousness: 8.8/10

There are different methods to detect plagiarism, which may be classified into three main categories: i) by comparing documents word by word in order to detect partial or full similarity, ii) taking a paragraph from the document and running a search with a good search engine, and iii) analysing document style or comparing writing style to texts previously written by the same author (stylometry) [1].

Currently, there are several tools and web applications to help detect plagiarism in documents. Some of the most common are:

Turnitin: It is a product from iParadigms [1] which is based on a web service for processing and remote sensing. User uploads their document to the system database, which creates a fingerprint of the document, stores it and uses algorithms to search on 4.5 billion web pages, books and journals on ProQuestTM database, besides 10 million documents already contained in Turnitin database.

SafeAssign: This is a service provided by Mydropbox. SafeAssign is based on an unique matching-text algorithm able to detect exact and inexact match between a document and its source. Search is performed on an index containing 8 billion web documents, more than 1,100 book titles, and about 2.6 million articles uploaded since 1990, which are contained in ProQuest ABI/Inform database, FindArticlesTM, LookSmartTM and other school data bases, besides 300,000 documents offered by Paper Mills [1, 6].

Docol©c: It is a web service offered by Institut Angewandte für Lerntechnologien (IFALT)9. It uses search and classification from Google API. User uploads their document to check on a server. The software provides a console to configure size, date restrictions, among others. The analysis report provides a header with the title, author name, file name, and date of plagiarism verification. It then shows marked sentences and percentage of sentences found in other documents [1, 7].

Urkund: It was created by the company Prioinfo, located in Sweden. No software installation or registration is required. E-mail is used both for sending documents to be analyzed and for returning plagiarism results from Urkund. This tool is supported in different languages, such as English, German, French, Turkish, Polish, Spanish, and Swedish. Documents to be analyzed are compared against 10 billion web pages, material published in e-books, encyclopedias, and journals. It also has information providers such as DIVA (Digital Scientific Archive), Iustus, Nationalencyklopedin, and ProQuest [8].

Copycatch: Users must register for personal licenses for Google API. Copycatch identifies documents with similar words, highlighting the text that needs further analysis to identify similarity. It uses two algorithms to detect document similarity [1, 9, 10].

WCopyfind: It is an open source tool for detecting word and phrase similarities between documents. It is available under the GNU Public License, which means you can do whatever you want with the tool, but selling it to someone else. WCopyfind is a single executable file with no installation required [11].

Eve2 (Essay Verification Engine): It determines whether the analysed material has been plagiarized from the World Wide Web. It accepts plain text files, Microsoft Word or Corel Word Perfect. When evidence of similarity is found, Eve2 records the URL and generates a full report on each plagiarized document, containing plagiarism percentage and a copy of the document with plagiarized parts highlighted in red [1, 12].

GPSP-Glatt Plagiarism Screening Program: this tool exploits the uniqueness of each individual's style and language patterns. Its process assumes that the writing style of each person is as unique as their fingerprints. This tool is a sensitive measure for discriminating plagiarists from non-plagiarists, and it is especially useful in situations where you can not locate the original source materials [13].

MOSS-a Measure of Software Similarity: It is an automatic system for determining software similarity. Its algorithm is an important improvement over other algorithms for plagiarism detection. It is not a completely automatic plagiarism detection system: it automatically detects software similarity, but it does not identify which codes are similar, so a human must revise the parts of the code highlighted by Moss and make the decision on whether or not plagiarism exists [14].

JPlag: It finds similarities between multiple sets of source code files, allowing to detect software plagiarism. JPlag is aware of the programming language syntax and program structure, enabling to detect plagiarism [15].

References

1. H. Maurer, F. Kappe and B. Zaka, "Plagiarism-A Survey", *Journal of Universal Computer Science,* vol. 12, no. 8, pp. 1050-1084, 2006.
2. PlagiarismToday, *The most common and serious types of research plagiarism*, 2013. [Online]. Available: https://www.plagiarismtoday.com/2013/10/08/ithenticate-survey-identifies-common-serious-types-research-plagiarism/. Accessed on: Jul. 28, 2016.
3. iParadigms, LLC., *What is plagiarism?*, 2014. [Online]. Available: http://www.plagiarism.org/plagiarism-101/what-is-plagiarism/. Accessed on: Jul. 22, 2016.
4. University College London, *Plagiarism,* 2013. [Online]. Available: http://www.ucl.ac.uk/current-students/guidelines/plagiarism. Accessed on: Jul. 26, 2016.
5. iParadigms, *Types of plagiarism in research infographic,* 2013. [Online]. Available: http://www.ithenticate.com/resources/infographics/types-of-plagiarism-research. Accessed on: Jul. 21, 2016.
6. Blackboard Inc., *Use SafeAssign in Assignments.* [Online]. Available: https://en-us.help.blackboard.com/Learn/Instructor/Assignments/060_Use_SafeAssign. Accessed on: Aug. 2, 2016.
7. Docoloc, *The Docol©c-Report.* [Online]. Available: https://www.docoloc.de/report2012.hhtml. Accessed on: Jul. 28, 2016.
8. Universidad Politécnica de Madrid, *Urkund.* [Online]. Available: https://plagiostop.wordpress.com/de-pago/urkund/. Accessed on: Jul. 27, 2016.
9. Pearltrees, *Plagiarism Checker.* [Online]. Available: http://www.pearltrees.com/u/1072048-plagiarism-checker. Accessed on: Jul. 27, 2016.
10. A. Beutel, W. Xu, V. Guruswami, C. Palow and C. Faloutsos, "CopyCatch: Stopping Group Attacks by Spotting Lockstep Behavior in Social Networks", *in International World Wide Web Conference Committee*, Rio de Janeiro, Brazil, 2013.
11. L. Bloomfield, *WCopyfind.* [Online]. Available: http://plagiarism.bloomfieldmedia.com/wordpress/software/wcopyfind/. Accessed on: Jul. 26, 2016.
12. Essay Verification Engine (EVE2), *What is EVE2?* [Online]. Available: http://www.canexus.com/eve/abouteve.shtml. Accessed on: Jul. 28, 2016.
13. Glatt Plariarism Services, *Glatt Plagiarism Screening Program/GPSP.* [Online]. Available: http://www.plagiarism.com/glatt-plagiarism-screening-program.html. Accessed on: Jul. 27, 2016.
14. Stanford University, *Moss: A System for Detecting Software Plagiarism.* [Online]. Available: https://theory.stanford.edu/~aiken/moss/. Accessed on: Jul. 27, 2016.
15. Karlsruhe Institute of Technology, JPlag - *Detecting Software Plagiarism.* [Online]. Available: https://jplag.ipd.kit.edu/. Accessed on: Aug. 1, 2016.

Maryory Astrid Gómez Botero
Editor-in-Chief
Revista Facultad de Ingeniería
https://orcid.org/0000-0001-9685-3080