Applied Medical Informatics
Original Research **Vol. 44, No. 2 /2022, pp: 73-86**

Comparison of Plagiarism Detection Performance between some Commercial and Free Software

Ilie-Andrei CONDURACHE* and Sorana D. BOLBOACĂ

Department of Biostatistics and Medical Bioinformatics, "Iuliu Haţieganu" University of Medicine and Pharmacy Cluj-Napoca, Louis Pasteur Str., No. 6, 400349 Cluj-Napoca, Romania
E-mails: conduandrei@yahoo.com; sbolboaca@umfcluj.ro

* Author to whom correspondence should be addressed;

Received: June 20, 2022/Accepted: June 30, 2022/ Published online: June 30, 2022

Abstract
Introduction: The act of plagiarism is represented by using someone else's information or research without the author's consent and/or without the author's full acknowledgement. Detection of plagiarism can be easily made using computer software that identifies fragments of texts as not original. *Aim:* This study aimed to highlight and compare the performance in detecting a specific type of plagiarism (copy-paste) in different types of medical documents between free and commercial software. *Material and Method:* A document of 808 words was created using eight fragments of texts from eight different sources. Two other versions of the document were then created: one with approximately 43% of the text similar and another with the entire text paraphrased. Seven software programs (Turnitin – international commercial software used for plagiarism detection, PlagScan, Smallseotools, Prepostseo, Plagiarismdetector, Plagiarism Checker X and Sistemantiplagiat) were used for the similarity analysis of each of the three texts (4 commercial and 3 free, each software was representative for their category (free or commercial) and the selection was based on this criterion). *Results:* When all software is taken into consideration, commercial software had a worse performance than free software. The original document showed differences in detection performances (97% observed similarity by Turnitin and 93% observed similarity by Plagiarism detector, both of them had the highest performance). In the document with 43% similarity, the performance was affected across all programs (smaller percentage of identified sources and bigger differences between identified and exact similarities in comparison to the first version), but Plagiarism detector had the best performance (43% observed similarity). None of the evaluated software could detect the original sources in the entire paraphrased document. *Conclusion:* Among the tested software, **Turnitin** proved to be the best commercial software and **Plagiarism detector** the best free software for testing academic documents similarity, differences between them being minimal. Overall, in this sample of analyzed software, commercial software had a worse performance than free software. Differences between identification of open access/closed access sources were not relevant.

Keywords: Plagiarism; Similarity; Performance; Software

Introduction

According to Oxford University, plagiarism is "*the act of presenting another person's work or ideas as their own, with or without the consent of the author, by incorporating them into their own data without full acknowledgment of the author. All published and unpublished materials, whether handwritten, printed or electronic, are covered by this definition. Plagiarism can be intentional, reckless or unintentional.*" [1].

The principles of ethics in academic research state that *"a research paper, which belongs to an author, is a contract between the author and the readers of that paper*" and that "*the author assumes that he is the sole creator of the paper and that any information that does not belong to him is clearly marked in his paper where appropriate* " according to the Office of Research Integrity (ORI) [2]. Office of Research Integrity recommends that in the act of conducting research, when it is necessary to take over information whose property we do not own, it is recommended to summarize the information taken using our own language and syntax, concluding the paragraph taken by quoting the original source, to avoid intentional plagiarism like copy-paste or other types [2].

To facilitate easy detection of intentional / unintentional plagiarism acts, the authors and the competent authorities use specialized software to detect plagiarism. Software to detect plagiarism can be online or client type, commercial (subscription, contract with an academic institution, or purchasable tokens equivalent to a number of pages/words verified), or free (a limited number of pages/words which are available to verify without costs) and have different performances. Their function is to compare the tested text with a database available for the software and to report what percentage of the text is similar to the sources observed by the software using a similarity coefficient [3].

No international guidelines define an accepted percentage for the similarity coefficient by which a work is considered original or plagiarized. In Romania, the authority that verifies the originality of scientific papers is the National Council for Attestation of University Degrees, Diplomas and Certificates (CNATDCU). In the MENCS order no. 3485/2016 published by the Ministry of National Education , the software programs recognized and used by CNATDCU to validate scientific papers and thesis are listed: **iThenticate**, **Turnitin**, **Plagiarism detector + PDAS**, **Safe Assign**, **SEMPLAG** and **Sistemantiplagiat** [4].

Existence of plagiarism is dependent on many factors. Debnath considers that the high accessibility of online information, the existence of publish or perish academic movement, the lack of academic education or academic morality among publishers or the time-pressured writings are the main factors that determine plagiarism [5].

Plagiarism detection tools have been developed along with the appearance of academic plagiarism. Turnitin and iThenticate – were software developed by iParadigms LLC in 1997 for verifying and testing similarities of academic documents, but students also use them for checking documents for any missing citations or for assignments.

Turnitin is a highly used product, licensed by the company Turnitin LLC, for 30 million students and 15000 academic institutions worldwide. At the time of acquirement by Turnitin LLC (subsidiary of Advance Publications) the Turnitin company (along with the product) was evaluated at $1.75 billion [6]. Turnitin LLC shares two main products as anti-plagiarism detection tools, with the same database, according to the University of Waterloo: Turnitin (which is orientated mainly for students assignments and coursework) and iThenticate (which is mainly used as a plagiarism detection software by researchers and academic institutions) [7].

Meo & Talha emphasized on the idea that although software can be used for detecting similar text, an analysis of an originality report is tricky and must be made with thorough inspection. The software reports are mainly used for suspicions of possible plagiarism and a similarity percentage is not always a plagiarism percentage [8].

Among university students, plagiarism can exist mainly because of poor academic writing and lack of paraphrasing skills. A study made by Gallant et al. on 135 students from the University of California and 255 laboratory reports analyzed using Turnitin, showed an average of 29/28 matches, depending on the analyzed laboratory report (the research team considered the majority of the matches as insignificant because the matches were directed to the laboratory manual, clearing them out resulting in an average of 1.6/2.2 matches). More than half of reports (53%) exhibited at least one significant match named as "source material incorporation problem" and 8% of the reports exhibited more than four significant matches. 87% of the significant matches were identified as patchwriting (as defined by authors: "a clear (but unsuccessful) attempt to restructure the original sentence" by having a vaguely modified source text "through word substitutions or deletions") (59%) or technical parroting (as defined by authors: "the repetition of methods, processes, or procedures from the laboratory manual, with little or no change from the original" (28%) [9].

Usage and acceptance of plagiarism detection tools is different among academic ranks. Arabyat et al. demonstrated in their study that full professors were more likely to use anti-plagiarism detection tools (APT) than assistant professors, and also academic staff who were very experienced with the software usage were more likely to use APTs than staff who were poorly experienced. Their data showed that the most frequently used softwares were Turnitin, iThenticate and PlagScan, mostly for analyzing papers before submission or for analyzing theses/dissertations [10].

Differences in the methods of reporting similarity coefficients between the various programs and in the setting of thresholds beyond which there could be suspicions of plagiarism in a manuscript exist. For example, Sistemantiplagiat considers as a threshold a similarity coefficient 1 (CS1) of 50% and a similarity coefficient 2 (CS2) of 5% (the differences between the two coefficients being the minimum length of words identified as similar) [11]. Instead, Turnitin considers a similarity coefficient of 25% as the threshold value. [12].

This study aimed to highlight and compare the performance in the detection of a particular type of plagiarism (copy-paste) in different kinds of medical documents (websites/abstracts/full-text articles/books) with different access (open or closed access) between several free and commercial software.

Material and Method

Text Identification and Compilation

A document of 808 words was created consisting of eight fragments taken (using copy-paste) from eight different sources to detect differences in performance between the used software. Each fragment was meant to have between 90-120 words and between 4-5 segments (delimited as sentences / phrases). Four kinds of sources were used: websites, abstracts, full-text articles, and books. Four sources were available online for free or were visible using the Google search engine (Open Access) and four sources were not visible using the Google search engine but only based on accessing the original research through a paid subscription or other methods (Closed Access). The characteristics of the compiled document are presented in Table 1.

Table 1. Original document used for analysis - structure

Fragment number (Ref)	Words no.	Segments no.	Source Type	Access Type
F1 [13]	96	5	Website	Open
F2 [14]	92	5	Website	Closed
F3 [15]	96	5	Abstract	Open
F4 [16]	106	4	Abstract	Closed
F5 [17]	96	4	Full-Text	Open
F6 [18]	97	5	Full-Text	Closed
F7 [19]	118	5	Book	Open
F8 [20]	107	5	Book	Closed

Ref = Reference of the used source

The limit between each segment was defined as the last word in the sentence and was marked by bold and italic formatting (Table 2). Two additional versions of the document were created starting from the original version of the document: ~43% similarity and 0% similarity (Table 2).

Table 2. All three versions of the document used for analysis

	Original	~ 43% similarity	~ 0% similarity
F1 [13] (web)	"The most consistent forms of relief from disabling dystonia are baclofen, trihexyphenidyl, and	Baclofen, trihexyphenidyl, and clonazepam are by far the most consistent therapies for disabling	Baclofen, trihexyphenidyl, and clonazepam are by far the most consistent therapies for disabling

	Original	~ 43% similarity	~ 0% similarity
	clonazepam. These medications can be taken ***orally***. Later in disease, a baclofen pump can be used to administer regular doses automatically into the central nervous ***system***. Intramuscular botulinum toxin may also help treat specific regions where dystonia is problematic. Levodopa/carbidopa does not generally appear to help patients with PKAN, although there may be ***exceptions***. These treatments may have a role in the treatment of other causes of NBIA; however, their overall effectiveness is unknown and the responsiveness in individual cases is ***unpredictable***."	***dystonia***. "*These medications can be taken* ***orally***. " Later in the illness, a baclofen pump can be administered to continuously deliver regular volumes further into the central nervous ***system***. "*Intramuscular botulinum toxin may also help treat specific regions where dystonia is problematic. Levodopa/carbidopa does not generally appear to help patients with PKAN, although there may be* ***exceptions***. " These therapies may be useful in treating various causes of NBIA; however, their overall efficacy is uncertain, and their reactivity in individual patients is ***inconsistent***.	***dystonia***. For these drugs an oral administration is ***advised***. Later in the illness, a baclofen pump can be administered to continuously deliver regular volumes further into the central nervous ***system***. Intramuscular botox injections may also aid in the treatment of dystonia in certain areas. While there can be exceptional cases, levodopa or carbidopa don't seem to assist people with ***PKAN***. These therapies may be useful in treating various causes of NBIA; however, their overall efficacy is uncertain, and their reactivity in individual patients is ***inconsistent***.
F2 [14] (web)	"In addition, patients are surviving longer with improved quality of life compared with pretransplantation ***status***. However, this prolonged longevity has brought about new concerns, such as the long-term effects of immunosuppression, as they relate to effects on the cardiovascular system, infections, and propensity for ***malignancy***. Thus, the search for newer immunosuppressive strategies to minimize these adverse effects continues ***today***. Excessive alcohol consumption negatively impacts long-term survival after liver transplant, regardless of the primary ***indication***. Mortality is due largely to the recurrence of liver disease and non-hepatic cancer, along with cardiovascular ***disease***."	Furthermore, compared to pretransplantation state, patients are living much longer and have a higher standard of ***living***. "*However, this prolonged longevity has brought about new concerns, such as the long-term effects of immunosuppression, as they relate to effects on the cardiovascular system, infections, and propensity for* ***malignancy***. " As a result, the hunt for improved immunosuppressive techniques to reduce these side effects continues to this day. Heavy alcohol use, regardless of the underlying reason, has a deleterious influence on long-term survival following liver ***transplant***. "*Mortality is due largely to the recurrence of liver disease and non-hepatic cancer, along with cardiovascular* ***disease.*** "	Furthermore, compared to pretransplantation state, patients are living much longer and have a higher standard of ***living***. However, this increased lifespan has raised additional issues, including the long-term impact of immunodeficiency mostly on the cardiovascular system, rate of possible infections, and predisposition for ***neoplasia***. As a result, the hunt for improved immunosuppressive techniques to reduce these side effects continues to this ***day***. Heavy alcohol use, regardless of the underlying reason, has a deleterious influence on long-term survival following liver ***transplant***. The reappearance of hepatic conditions and non-hepatic cancer, as well as cardiovascular events, are the leading causes of ***death***.
F3 [15] (abst.)	"Membrane and protein traffic in the secretory and endocytic pathways is mediated by vesicular ***transport***. Recent studies of certain key regulators of vesicular transport, the Rab GTPases, have linked Rab dysfunction to human ***disease***. Mutations in Rab27a result in Griscelli syndrome, caused by defects in melanosome transport in melanocytes and loss of cytotoxic killing activity in ***Tcells***. Other genetic diseases are caused by partial dysfunction of multiple Rab proteins resulting from mutations in general regulators of Rab ***activity***; Rab escort protein-1 (choroideremia), Rab geranylgeranyl transferase (Hermansky-Pudlak syndrome) and Rab GDP dissociation inhibitor-alpha (X-linked mental ***retardation***)."	Vesicular transport facilitates membrane and protein transport in the secretion and endocytic ***processes***. "*Recent studies of certain key regulators of vesicular transport, the Rab GTPases, have linked Rab dysfunction to human* ***disease***. " Rab27a mutations cause Griscelli syndrome, which is characterized by abnormalities in melanosome transportation in melanocytes and a lack of cytotoxic killing capacity in ***Tcells***. "*Other genetic diseases are caused by partial dysfunction of multiple Rab proteins resulting from mutations in general regulators of Rab* ***activity***; " in the case of choroideremia the Rab escort protein-1 is affected, in the case of the Hermansky-Pudlak syndrome the Rab geranylgeranyl transferase is affected and in the case of X-linked mental retardation a protein which acts as a Rab GDP	Vesicular transport facilitates membrane and protein transport in the secretion and endocytic ***processes***. Current findings on Rab GTPases, which are essential mediators of vesicular transport, suggests a correlation between Rab malfunction and a pathological ***status***. Rab27a mutations cause Griscelli syndrome, which is characterized by abnormalities in melanosome transportation in melanocytes and a lack of cytotoxic killing capacity in ***Tcells***. Several inherited illnesses are caused by incomplete malfunction of numerous Rab proteins induced by alterations in Rab activity ***regulation***; in the case of choroideremia the Rab escort protein-1 is affected, in the case of the Hermansky-Pudlak syndrome the Rab geranylgeranyl transferase is affected and in the case of X-linked mental retardation a protein which

	Original	~ 43% similarity	~ 0% similarity
		alpha-inhibitor for dissociation is ***affected***.	acts as a Rab GDP alpha-inhibitor for dissociation is ***affected***.
F4 [16] (abst.)	"The radio-activity of potassium salts was investigated by the method in which one vessel is placed within, and insulated from, a second and the electrical charge which the insulated body or vessel more or less rapidly acquires is then ***observed***. Preliminary experiments were made on the charging action of polonium, and that of the secondary rays excited in aluminum by the R-rays from ***radium***; the charges acquired by uranium salts at low pressures were also ***observed***. In the case of potassium salts the results show that in high vacua both potassium nitrate and potassium sulphate emit an excess of charged particles of the ***β-type***. "	"*The radio-activity of potassium salts was investigated by the method in which one vessel is placed within, and insulated from, a second and the electrical charge which the insulated body or vessel more or less rapidly acquires is then* ***observed***. " The charging activity of polonium was studied, as well as the secondary rays stimulated in aluminum by R-rays from ***radium***; the charges gained by uranium complexes at reduced pressure were also ***researched***. "*In the case of potassium salts the results show that in high vacua both potassium nitrate and potassium sulphate emit an excess of charged particles of the* ***β-type***. "	Radioactive potassium salts were studied using a procedure during which one container is put into and isolated from another, and the electric potential which the separated body or container rapidly accumulates is then ***examined***. The charging activity of polonium was studied, as well as the secondary rays stimulated in aluminum by R-rays from ***radium***; the charges gained by uranium complexes at reduced pressure were also ***researched***. For potassium salts, the findings suggest that in high vacuum, potassium compounds (sulphate and nitrate) produce an excessive amount of electrically charged ***β-particles***.
F5 [17] (fullt)	"This promotion of the β-oxidation of fatty acids reduced the availability of fatty acids for very-low density lipoprotein (VLDL) synthesis and ***secretion***. Fenofibrate also increased the expression of the gene for lipoprotein lipase and decreases ApoC-III expression in the ***liver***. Thus, fenofibrate lowered the concentration of TG both by reducing the rate of synthesis and increasing the rate of hydrolysis of triglyceride-rich ***lipoproteins***. Moreover, fenofibrate treatment reduced the proportion of small, dense LDL, with the formation of larger, less dense LDL particles with a higher affinity for the LDL receptor and thus catabolized more ***rapidly***."	"*This promotion of the β-oxidation of fatty acids reduced the availability of fatty acids for very-low density lipoprotein (VLDL) synthesis and* ***secretion***. *Fenofibrate also increased the expression of the gene for lipoprotein lipase and decreases ApoC-III expression in the* ***liver***." Therefore, fenofibrate reduced trygliceride concentrations by slowing production and increasing the speed for the hydrolysis of triglyceride-rich ***lipoproteins***. Furthermore, fenofibrate therapy lowered the fraction of tiny, dense LDL particles, resulting in the production of bigger, less dense LDL molecules with a higher potency for the LDL receptor and therefore catabolized more ***quickly***.	The increased disponibility of fatty acids for VLDL production and excretion was lowered by the stimulation of fatty acids ***β-oxidation***.Fenofibrate also increases lipoprotein lipase gene expression while decreasing the hepatic expression of ***ApoC-III***. Therefore, fenofibrate reduced trygliceride concentrations by slowing production and increasing the speed for the hydrolysis of triglyceride-rich ***lipoproteins***. Also, fenofibrate treatment reduced the percentage of small, compact lipoproteins, resulting in the creation of larger, less compact lipoproteins molecules with increased potency for the lipoprotein receptor and thus metabolised ***faster***.
F6 [18] (fullt)	"Oppenheim described a diffuse weakness which was usually most marked in the legs but spared the extraocular, tongue, pharyngeal, and diaphragmatic ***musculature***. Electrical stimulation demonstrated a marked reduction of excitability in the muscles of some of these children, as well as a possible reaction of degeneration in the more severely ***affected***. There was no impairment of intelligence or ***sensation***. Since a similar state was not found in older children, Oppenheim thought that recovery must ***occur***. However, in this report there was no information as to the future course of these children nor were pathological studies ***included***."	Oppenheim identified a generalized impairment that mostly affected the limbs yet left the extraocular, oral, pharyngeal, and diaphragmatic regions ***unaffected***. "*Electrical stimulation demonstrated a marked reduction of excitability in the muscles of some of these children, as well as a possible reaction of degeneration in the more severely* ***affected***. *There was no impairment of intelligence or* ***sensation***." Oppenheim reasoned that because a comparable situation was not observed in older children, recovery had to occur. However, there was no evidence in this study about the subsequent course of these children, nor were pathology findings ***provided***.	Oppenheim identified a generalized impairment that mostly affected the limbs yet left the extraocular, oral, pharyngeal, and diaphragmatic regions ***unaffected***. Using artificial electrical stimuli showed a significant drop in muscle responsiveness in some cases, and often a potential consequence of reactional deterioration in the more seriously ***affected***. No negative effects were reported in cases of intellect or ***sensibility***. Oppenheim reasoned that because a comparable situation was not observed in older children, recovery had to occur. However, there was no evidence in this study about the subsequent course of these children, nor were pathology findings ***provided***.
F7 [19]	"As with all new therapeutic agents, the clinical evaluation of anti-HIV	"*As with all new therapeutic agents, the clinical evaluation of anti-HIV drugs is*	The clinical assessment of anti-HIV medications, like that of all novel

	Original	~ 43% similarity	~ 0% similarity
(book)	drugs is divided into a series of (more or less) sequential ***phases***. Phase I studies are typically performed in healthy volunteers following extensive safety, toxicity, genotoxicity, and pharmacologic studies done both in cell culture in vitro and in ***animals***. Phase I evaluations are short-term and are designed to assess the pharmacology and toxicity of the test compound in ***humans***. If the data and observations are acceptable, phase II studies are performed in patient populations for whom the drug is ***intended***. These studies are also limited in size and are intended primarily for determining dosages, assessing tolerability, and, importantly, assessing the in vivo activity of the ***compound***."	*divided into a series of (more or less) sequential* ***phases***." Following rigorous safety, toxicology, genotoxicity, and pharmacologic experiments in cell lines in vitro and in animals, phase I trials are often conducted in healthy ***individuals***. Phase I assessments are brief and are aimed to investigate the test molecule pharmacology and toxicity in ***individuals***. If the results and findings are satisfactory, phase II investigations in clinical groups for whom the medicine is designed are ***conducted***. "*These studies are also limited in size and are intended primarily for determining dosages, assessing tolerability, and, importantly, assessing the in vivo activity of the* ***compound***."	therapeutic treatments, is split into a number of successive ***phases***. Following rigorous safety, toxicology, genotoxicity, and pharmacologic experiments in cell lines in vitro and in animals, phase I trials are often conducted in healthy ***individuals***. Phase I assessments are brief and are aimed to investigate the test molecule pharmacology and toxicity in ***individuals***. If the results and findings are satisfactory, phase II investigations in clinical groups for whom the medicine is designed are ***conducted***. These trials are also narrow in scope and are meant primarily for calculating doses, testing tolerance, and, most critically, determining the molecule's in vivo ***efficacy***.
F8 [20] (book)	"Interference in the epithelial barrier present in the gut may authorize an unfettered entry in the lamina propria by the intestinal microbiota, where the cells of the defence system is ***located***. Cells of the immune system reside in systematized arrangements in the intestine, jointly known as gut-associated lymphoid tissues (***GALT***). GALT is exceedingly flexible and is colonized by bacteria. Immune system is highly functional within the ***intestine***. Large amounts of macrophages and lymphocytes are spread all over the lamina propria and present upto basal ***epithelium***. Macrophages that inhabit the intestine are mostly insensitive to bacteria and their constituents, as there is absence of lipopolysaccharide (LPS) ***co-receptor***."	Disturbances in the mucosal layer at the intestinal level may allow the gastrointestinal microbiome unhindered access into the lamina propria, in which the lymphocytes are ***found***. "*Cells of the immune system reside in systematized arrangements in the intestine, jointly known as gut-associated lymphoid tissues (****GALT****). GALT is exceedingly flexible and is colonized by bacteria. Immune system is highly functional within the* ***intestine***." A large number of monocytes and lymphocytes are found throughout the lamina propria and up to the basal ***epithelium***. In cases which lack co-receptors for lipopolysaccharides (LPS), intestine-dwelling monocytes are generally unresponsive to microorganisms and their ***components***.	Disturbances in the mucosal layer at the intestinal level may allow the gastrointestinal microbiome unhindered access into the lamina propria, in which the lymphocytes are ***found***. Immune system cells exist in structured formations in the gut, which are referred as the gut-associated lymphoid tissues (***GALT***). GALT is extremely adaptable and is inhabited by microorganisms. Within the gut, the immune system is very ***active***. A large number of monocytes and lymphocytes are found throughout the lamina propria and up to the basal ***epithelium***. In cases which lack co-receptors for lipopolysaccharides (LPS), intestine-dwelling monocytes are generally unresponsive to microorganisms and their ***components***.

Closed Access - F2, F4, F6, F8

Three segments from the fragments with five segments, respectively two segments from the fragments with four segments were paraphrased in the version of ~ 43% similarity. Random Integer Set Generator software from Random.org (https://www.random.org/integer-sets/) was used to decide the fragments to be paraphrased (Table 3).

The same researcher did the paraphrasing using the QuillBot software (https://quillbot.com/). In the process of paraphrasing, the syntax and vocabulary were changed, without changing the original idea (thus masking the act of plagiarism). In this document version, 353 of the 821 words are from the original document, assuring an exact similarity of 43%. The 0% similarity version was obtained by paraphrasing the entire original text.

The assumed similarity percentage is ~ 43% and respectively 0% if the software appropriately recognizes all sources used. If a higher percentage exists, it would mean that the software would recognize the paraphrased text as similar (by means of ideas and concepts, which would be correct under the correct identification of the source).

Table 3. Segments selected at random to be paraphrased from each fragment

Fragment no. (Ref)	Randomly selected segments
F1 [13] (web)	1, 3, 5
F2 [14] (web)	1, 3, 4
F3 [15] (abst.)	1, 3, 5
F4 [16] (abst.)	2, 3
F5 [17] (fullt)	3, 4
F6 [18] (fullt)	1, 4, 5
F7 [19] (book)	2, 3, 4
F8 [20] (book)	1, 4, 5

Ref = Reference of the used source

Programs for Plagiarism Detection

The verification of an uploaded document against a database is made by plagiarism detection software. Seven software programs were used in this study (Table 4). Turnitin and Sistemantiplagiat are approved by CNATDCU for verification of academic papers and theses [4]. All evaluated programs (excepting Plagiarism Checker X which is client-based) are web-based programs.

The selection of the software was made based on the following criteria:

- In both commercial and free software groups there should be an approximately equal number of the software used in each group;
- Each software selected could be used optimally by Romanian users (whether it is approved by CNADTCU or it is frequently used mainly because of a Romanian interface);
- Turnitin and Sistemantiplagiat were selected as commercial software because of CNADTCU validation. Along with them, PlagScan and Plagiarism Checker X were selected because they were commercial softwares;
- Smallseotools and Prepostseo were selected because they were free software with Romanian interface. Along with them Plagiarismdetector was selected because it was a free software with similar functions as the previous two (similarity check, grammar check, online upload via Google Drive/Dropbox).

Performance Assessment

The performance of the plagiarism detection used software was judged based on the criteria presented in Table 5.

Table 5. Plagiarism performance metrics

Abb	Description	Interpretation
%Frag	The percentage of correct identification of all sources	the highest the better
%Key	The percentage of correct identification of sources with closed access	the highest the better
%Sim	The value of the reported similarity percentage	the highest the better
DifSim	The difference between the percentage of similarity considered real and the identified one	the smaller the better

The similarity reports were retrieved in 11 April 2022.

Table 4. Plagiarism software detection: main characteristics of the used programs

Software	Type (Owner)	Language/ Services	Account	Similarity index
Turnitin	Commercial (Turnitin LLC)	International (>15,000 institutions; universities and high schools), similarity check	Linked with an institution	≥25% - Possible plagiarism
PlagScan	Free (limited credit) and Commercial (Markus Goldbach and Johannes Knabe)	International (in 2018 by over 1500 institutions and more than 1.5 million users.), similarity check	Registration as a single user or as an organization	≥5% - Possible plagiarism
Smallseotools	Free (1000 words) (Tausif Akram)	Romanian interface, similarity check, grammar correction or paraphrasing	Not needed Documents uploaded locally or online (Google Drive/Dropbox)	N/S
Prepostseo	Free (1000 words) (Ahmad Sattar & AR AS)	Romanian interface, similarity check	Not needed Documents uploaded locally or online (Google Drive/Dropbox)	N/S
Plagiarismdetector	Free (1000 words)	English interface, similarity check, grammar correction or paraphrasing	Not needed Documents uploaded locally or online (Google Drive/Dropbox)	N/S
Plagiarism Checker X	Free (up to 120 words) Commercial	International, similarity check	Not needed Local documents only	≥20% - Possible plagiarism
Sistemantiplagiat	Commercial (Poland)	Romanian interface (international users), similarity check, check legal documents database, check paraphrased text (SmartMarks)	Individual (Token – PayPal or SMS) Institutional	*CS1* (% of similar text with phrases that contain ≥5 similar words) ≥50% - Possible plagiarism *CS2* (% of similar text with phrases that contain ≥25 similar words) ≥5% - Possible plagiarism *BDL* (% of similar text with phrases that contain ≥8 similar words from legal documents database) *CIT* (% of quoted text)

CS1 = Similarity coefficient 1, CS2 = Similarity coefficient 2, BDL = Legal documents database, CIT = Quoted text coefficient

Results

The best performances in identification of plagiarism was obtained by the Turnitin software, with a similarity difference of 3%, identifying all sources correctly (Table 6). Plagiarism detector software proved the best free software, with a similarity difference of 7%, identifying almost all sources correctly (Table 6). When all software is taken into consideration, overall, commercial software had a worse performance than free software (in terms of correct identification of all sources and differences between the identified and exact similarity of the document). Identification of open or closed access fragments was mixed in results (some software could not identify open access sources but could identify closed access sources, some software would have the opposite behavior and some software could identify neither open access nor closed access sources). As such, differences according to this criterion are not relevant to this sample of analyzed software.

Table 6. Original document verification results (actual similarity of 100%): performances of plagiarism checker software

Software / Criteria	Turnitin	PlagScan	Small seotools	Pre postseo	Plagiarism detector	Plagiarism Checker X	Sistem antiplagiat
F1 (web)	✓	×	✓	✓	✓	×	✓
F2 (web)	✓	×	✓	✓	✓	✓	✓
F3 (abst.)	✓	✓	✓	✓	✓	✓	✓
F4 (abst.)	✓	✓	✓	×	✓	×	✓
F5 (fullt.)	✓	✓	✓	✓	✓	✓	✓
F6 (fullt.)	✓	×	×	×	×	×	×
F7 (book)	✓	✓	✓	✓	✓	×	×
F8 (book)	✓	×	×	✓	✓	×	×
%Frag	100	50	75	75	87.5	37.5	62.5
%Key	100	25	50	50	75	25	50
%Sim	97	45.6	56	66	93	36	55.68 (CS1) 47.22 (CS2)

%Frag = % of correct identification of all sources; %Key = % of correct identification of sources with closed access; %Sim = % of the reported similarity; ✓ = identified; × = not identified; CS = coefficient of similarity; F = fragment number as presented in Table 2

The best performance for the documents with a real similarity of 43% was obtained by Plagiarism detector (Table 7) with a similarity difference of 0%, identifying 87.5% of all sources correctly. The next best performing program is the Turnitin (Table 7). When all software is taken into consideration, overall, commercial software had a worse performance than free software (in terms of correct identification of all sources and differences between the identified and exact similarity of the document), similar as in the original version. Most software had a worse performance in the second version than the original version (smaller percentage of identified sources and bigger differences between identified and exact similarities in comparison to the first version).

Identification of open or closed access fragments was mixed in results (some software could not identify open access sources but could identify closed access sources, some software would have the opposite behavior and some software could not identify neither open access nor closed access sources). As such, differences according to this criterion are not relevant on this sample of analyzed software.

No differences exist between the similarity programs when the document with a similarity of approximately 0% is evaluated (Table 8).

The overall performances of the investigated similarity checker software in the three scenarios is presented in Figure 1.

Table 7. Almost 43% similarity in the document (actual similarity considered 43%): performances of plagiarism checker software

Software / Criteria	Turnitin	PlagScan	Small seotools	Pre postseo	Plagiarism detector	Plagiarism Checker X	Sistem antiplagiat
F1 (web)	×	×	✓	✓	✓	×	✓
F2 (web)	×	×	✓	✓	✓	✓	×
F3 (abst.)	✓	×	✓	✓	✓	✓	✓
F4 (abst.)	×	×	✓	×	✓	×	✓
F5 (fullt.)	✓	×	✓	✓	✓	✓	✓
F6 (fullt.)	✓	×	✓	×	×	×	×
F7 (book)	✓	×	✓	✓	✓	×	×
F8 (book)	✓	×	×	×	✓	×	×
%Frag	62.5	0	87.5	62.5	87.5	37.5	50
%Key	50	0	75	25	75	25	25
%Sim	26	0	31	27	43	18	30.08 (CS1) 25.45 (CS2)

%Frag = % of correct identification of all sources; %Key = % of correct identification of sources with closed access; %Sim = % of the reported similarity; ✓ = identified; × = not identified; CS = coefficient of similarity; F = fragment number as presented in Table 2

Table 8. Totally rephrased document (actual similarity considered 0%): performances of plagiarism checker software

Software / Criteria	Turnitin	PlagScan	Small seotools	Pre postseo	Plagiarism detector	Plagiarism Checker X	Sistem antiplagiat
F1 (web)	×	×	×	×	×	×	×
F2 (web)	×	×	×	×	×	×	×
F3 (abst.)	×	×	×	×	×	×	×
F4 (abst.)	×	×	×	×	×	×	×
F5 (fullt.)	×	×	×	×	×	×	×
F6 (fullt.)	×	×	×	×	×	×	×
F7 (book)	×	×	×	×	×	×	×
F8 (book)	×	×	×	×	×	×	×
%Frag	0	0	0	0	0	0	0
%Key	0	0	0	0	0	0	0
%Sim	0	0	0	0	0	8	1.46 (CS1) 0.00 (CS2)

%Frag = % of correct identification of all sources; %Key = % of correct identification of sources with closed access; %Sim = % of the reported similarity; ✓ = identified; × = not identified; CS = coefficient of similarity; F = fragment number as presented in Table 2

Discussion

Our results show that Turnitin (software used by CNADTCU) was the most accurate commercial software (with a difference of 3% from the actual similarity in the first version and a difference of 17% in the second version) and Plagiarism detector was the most accurate free software (with a difference of 7% from the actual similarity in the first version and a difference of 0% in the second version). Sistemantiplagiat (another software highly used by CNADTCU) performed poorly (with a difference of 44.32% from the actual similarity in the first version and 12.92% from the actual similarity in the second version), unable to recognize 3 out of 8 sources in the first version and 4 out of 8 sources in the second version, according to Tables 6 and 7.

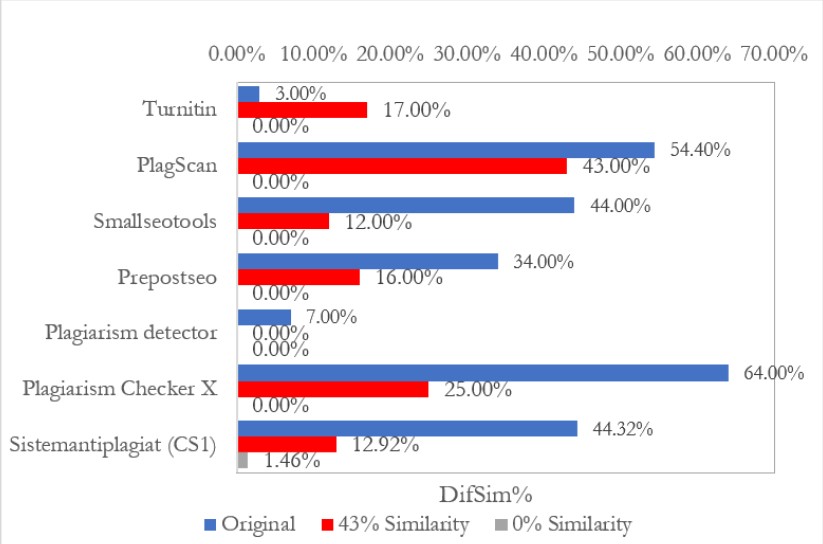

Figure 1. Overall performance of the similarity check software expressed as DifSim (%) = difference between the percentage of similarity considered real and the identified one

In the entire paraphrased document, none of the utilized software could detect any of the eight sources, especially Sistemantiplagiat, which had a key feature in this direction (SmartMarks), according to Table 8.

A main reason behind the differences of the software performances in terms of accuracy would be the size of the databases used (especially for closed access documents). Unfortunately, information of this type is scarce, software companies refuse to disclose this type of information (in case of free software) or give some brief data for the databases used (in case of commercial software). For instance:

- Sistemantiplagiat uses an Internal Database, an External Database, the RefBooks database (with about 24 million publications) and a Legal Database for legal documents [21];
- PlagScan has a database with over 10800 journals (such as BMJ, Springer, Taylor&Francis, Wiley Blackwell, Gale) and 14 million articles included [22];
- Turnitin uses a database with 89.4 million articles, 56000 journals, 13000 open access repositories. They also mention that 95% of the top 10000 journals world-wide are included in their database. [23]

All of the software was tested for the same function, to verify the similarity of a document against the software's database and output the result in an originality report. Some software has additional functions such as correcting grammar or paraphrasing the original text (e.g., Smallseotools, Plagiarism detector), which could be beneficial for students or other academic members in writing papers. Most of the software could be easily accessible using the Internet (excepting Plagiarism Checker X that is client-based), some of them could facilitate the document transfer by Google Drive/Dropbox file upload (e.g., Smallseotools, Prepostseo, Plagiarism detector), although the main flaw of these three programs was the lack of guideline or threshold for the Similarity index for detecting possible Plagiarism (such as Turnitin - ≥25%, Plagscan - ≥5%, Plagiarism Checker X - ≥20% or Sistemantiplagiat (CS1 - ≥50%, CS2 - ≥5%)).

Although this study's results show a high performance mainly for **Turnitin**, because it was tested only on a three-versions document, the small number of tested samples is not enough to prove this fact, conferring a low research quality. Ideal future studies should implement multilingual documents (not only in English) with different types of plagiarism (not only copy-paste) in a large sample number (e.g., 100-200 documents) that could be tested in other software too (e.g. iThenticate, Safe Assign, SEMPLAG) along with statistical comparisons of similarity performances among softwares (inter-group comparisons or intra-group comparisons for repeated tests – paraphrased/non-paraphrased versions). This study lacks the concepts mentioned above, therefore, it has it's limits.

According to other studies, Turnitin appears to be a highly used software. Arabyat et al. showed that Turnitin (43%) and iThenticate (32.8%) were the most frequently used software among faculty members enrolled in their study [10]. Another example is The Medical Journal of Armed Forces from India (MJAFI), which regularly use iThenticate to verify the submitted manuscripts, according to Debnath [5].

Literature data is in favor of Turnitin's high performance. A study made by Turnitin LLC on 55 million documents demonstrated that throughout five years, academic institutions which had used Turnitin had their members unoriginal writing reduced by 39.1% (with a median reduction of 44%). The highest reduction was found in colleges with a 2-year education period and a student population of about 3000-5000 students (77.9%). Their data also showed that from 2004 to 2013, the number of electronic submissions increased from 500,000 papers to more than 45 million, increasing the transparency of the submission process and the demand for electronic verification [24].

Baker et al. evaluated two groups of US university student graduates, one with and the other without access to the plagiarism software and demonstrated that usage of the Turnitin program reduced the similarity index significantly with an average of 2.71% (95% C.I.: 0.67-4.76) [25].

Although many software programs exist to detect plagiarism, literature data show that comparison of detection performances does not exist. Garner specifies that although the software for plagiarism offers many functions (reporting results, identifying similarity using a coefficient, providing grammar correction or paraphrasing functions), no comparative analyzes between the types of software related to their function and performance were found [26].

As expected, our results showed the existence of the differences between tested software considering a relatively small document. It would be beneficial to evaluate the software performances for larger documents, including tables and figures. None of the softwares that were tested were able to identify the paraphrased text. A question that needs to be answered is "paraphrasing a text conceals the original ideas copied - does this mask or not an act of plagiarism?" By using massive paraphrasing of a document, plagiarism (especially copy-paste type) can be undetectable, but the question is: does paraphrasing give a false impression of originality to a scientific paper?

Plagiarism software checking should be supplemented with human supervision, with domain knowledge. Automated plagiarism detection can be seen as a screening tool, not as a final decision system. Human reasoning should be applied to judge each case. This might prevent high percentage of paraphrasing to be missed by the software. Nevertheless, no fixed percentages or rules, nor human reasoning will be perfect indicators of plagiarism, since any threshold is arbitrary. This will remain a subject of debate for a long time. The reported results are meant to be an opportunity for future studies that could show the actual comparative performance of current software and what algorithms could be implemented to increase the performances of future software.

Conclusions

Among the tested software, **Turnitin** proved to be the best commercial software and **Plagiarism detector** the best free software for testing academic documents similarity, differences between them being minimal. Overall, in this sample of analyzed software, commercial software had a worse performance than free software. Differences between identification of open access/closed access sources were not relevant. The evaluated softwares are not able to identify the paraphrased text.

List of abbreviations

Abb: Abbreviation
APTs: Anti-plagiarism detection tools
DifSim: Difference between calculated and reported similarity percentages
%Frag: Percentage of correct identification of sources
%Key: Percentage of correct identification of sources with Closed Access
MENCS: Ministry of National Education and Scientific Research

N/S: Not specified.
PDAS: Plagiarism Detector Accumulator Server
Ref: Reference of the used source.
%Sim: Reported similarity percentage

Conflict of Interest

The authors declare no conflict of interest.

References

1. University of Oxford. Plagiarism [Internet]. 2022 [cited 2022 Apr 11]. Available from: https://www.ox.ac.uk/students/academic/guidance/skills/plagiarism
2. Miguel R. Avoiding plagiarism, self-plagiarism, and other questionable writing practices: A guide to ethical writing [Internet]. Office of Research Integrity; 2015 [cited 2022 Apr 12]. Available from: https://ori.hhs.gov/sites/default/files/plagiarism.pdf
3. Luksanapruksa P, Millhouse PW. Guidelines on What Constitutes Plagiarism and Electronic Tools to Detect it. Clin Spine Surg. 2016 Apr;29(3):119-20.
4. Ministerul Educaţiei Naţionale. Ordinul MENCS nr. 3485/2016 - lista programelor recunoscute de CNATDCU şi utilizate la nivelul instituţiilor de învăţământ superior organizatoare de studii universitare de doctorat şi al Academiei Române, în vederea stabilirii gradului de similitudine pentru lucrările ştiinţifice [Internet]. 2016 [cited 2022 Apr 12]. Available from: http://www.cnatdcu.ro/wp-content/uploads/2011/04/ordin-programe-similitudini1.pdf
5. Debnath J. Plagiarism: A silent epidemic in scientific writing - Reasons, recognition and remedies. Med J Armed Forces India. 2016 Apr;72(2):164-7.
6. Sydney Johnson. Turnitin to Be Acquired by Advance Publications for $1.75B [Internet]. EdSurge. 2019 [cited 2022 Jun 15]. Available from: https://www.edsurge.com/news/2019-03-06-turnitin-to-be-acquired-by-advance-publications-for-1-75b
7. University of Waterloo. Turnitin and iThenticate [Internet]. [cited 2022 Jun 15]. Available from: https://uwaterloo.ca/academic-integrity/integrity-instructors-and-tas/turnitin-and-ithenticate
8. Meo SA, Talha M. Turnitin: Is it a text matching or plagiarism detection tool? Saudi J Anaesth. 2019 Apr;13(Suppl 1):S48-51.
9. Bertram Gallant T, Picciotto M, Bozinovic G, Tour E. Plagiarism or not? investigation of Turnitin®-detected similarity hits in biology laboratory reports. Biochem Mol Biol Educ. 2019 Jul;47(4):370-9.
10. Arabyat RM, Qawasmeh BR, Al-Azzam SI, Nusair MB, Alzoubi KH. Faculty Members' Perceptions and Attitudes Towards Anti-Plagiarism Detection Tools: Applying the Theory of Planned Behavior. J Empir Res Hum Res Ethics. 2022 Jul;17(3):275-83.
11. Sistemantiplagiat. FAQ [Internet]. 2022 [cited 2022 Apr 12]. Available from: https://panel.sistemantiplagiat.ro/#/info
12. Turnitin. Interpreting the Similarity Report [Internet]. Interpreting the Similarity Report. 2022 [cited 2022 Apr 12]. Available from: https://help.turnitin.com/feedback-studio/turnitin-website/student/the-similarity-report/interpreting-the-similarity-report.htm
13. National Organization for Rare Disorders. Pantothenate Kinase-Associated Neurodegeneration [Internet]. [cited 2022 Apr 11]. Available from: https://rarediseases.org/rare-diseases/pantothenate-kinase-associated-neurodegeneration/
14. Cosme M. Liver Transplantation [Internet]. Medscape. 2021 [cited 2022 Apr 11]. Available from: https://emedicine.medscape.com/article/431783-overview#a4
15. Seabra MC, Mules EH, Hume AN. Rab GTPases, intracellular traffic and disease. Trends Mol Med. 2002 Jan;8(1):23–30.

16. McLeanan JC. Radio-activity of potassium salts. Journal of the Franklin Institute [Internet]. 1910 Jul [cited 2022 Apr 11];170(1):45. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0016003210909794
17. Farnier M. Update on the clinical utility of fenofibrate in mixed dyslipidemias: mechanisms of action and rational prescribing. VHRM. 2008 Oct;4:991-1000. A
18. Byers RK, Banker BQ. Infantile Muscular Atrophy. Archives of Neurology [Internet]. 1961 Aug 1;5(2):140-64. 1
19. Coffin JM, Hughes SH, Varmus H, editors. Retroviruses. Plainview, NY: Cold Spring Harbor Laboratory Press; 1997. 843 p.
20. Sayyed RZ, Khan M, editors. Microbiome-Gut-Brain Axis: Implications on Health. 1st ed. 2022. Singapore: Springer Singapore; 2022. 443 p.
21. Sistemantiplagiat. Sistemantiplagiat - How it works [Internet]. [cited 2022 Jun 29]. Available from: https://sistemantiplagiat.ro/en/about-the-system/how-it-works/
22. PlagScan. Algorithm and Sources [Internet]. PlagScan; [cited 2022 Jun 29]. Available from: https://www.plagscan.com/en/algorithm-and-sources
23. Turnitin. The Turnitin Difference: the largest and fastest growing database [Internet]. Turnitin LLC; [cited 2022 Jun 29]. Available from: https://marketing-tii-statamic-assets-us-west-2.s3-us-west-2.amazonaws.com/marketing/our-content-databases_brochure_us_0322.pdf
24. Turnitin. The Effectiveness of Turnitin in Higher Education [Internet]. Turnitin LLC; [cited 2022 Jun 15]. Available from: https://go.turnitin.com/paper/effectiveness-turnitin-higher-education
25. Baker RK, Thornton B, Adams M. An Evaluation Of The Effectiveness Of Turnitin.Com As A Tool For Reducing Plagiarism In Graduate Student Term Papers. CTMS. 2008 Sep 1;4(9):1-4.
26. Garner HR. Combating unethical publications with plagiarism detection services. Urol Oncol. 2011 Feb;29(1):95-9.