

Hasanova R.Sh.¹, Asgarov F.Sh.²

^{1,2}Institute of Information Technology of ANAS, Baku, Azerbaijan

¹rahasanova@gmail.com, ²firudinasgarov@gmail.com

CONCEPTION OF CREATING THE NATIONAL ANTI-PLAGIARISM SERVICE IN AZERBAIJAN

Received: 01.12.2020

Revised: 28.12.2020

Accepted: 05.01.2021

Currently, fighting against plagiarism is one of the most actual problems in the scientific community and society. The fast development of information and communication technologies, the openness and accessibility of scientific information in the Internet, create a favorable environment for plagiarism. Every author may encounter such situations unintentionally or for some other reason. The article highlights conception of creating the national anti-plagiarism service in Azerbaijan. It provides brief information about plagiarism and its types. Information about characteristics of anti-antiplagiarism systems is also provided. In the example of Azerbaijan, the stages of rendering national anti-plagiarism service are listed. The proposed national anti-plagiarism service can be used by teachers, students and researchers of all universities and research organizations. Thus, with the submitted reference for each scientific work, it is possible to decide whether it is original or not.

Keywords: *plagiarism, piracy, plagiarism checker systems, anti-antiplagiarism systems.*

Introduction

Plagiarism is representing thoughts and ideas of another author as one's own. This event is closely related to theft and piracy [1]. Plagiarism is considered a type of a cybercrime such as computer virus, spam, phishing and etc. Plagiarism constitutes of directly coping or making a few changes to a text taken from the Internet, book, journal etc.

According to currently used anti-plagiarism systems, such as Turnitin, Plagiarism and Research Resources, plagiarism is presenting others' work as one's own; expressing others' ideas and thoughts as one's own without referencing them; not marking other's idea as a citation; providing wrong information about citation source; copying the main structure by changing the words without referencing source; copying majority of words and ideas of a material from other sources regardless of reference [2-4]. In modern period, following two classes of plagiarism are mainly used:

- Text plagiarism. This type of plagiarism is usually committed by students and researches in scientific institutions. In this case, plagiarized text is exactly similar to the text of original article, thesis, and scientific work.
- Source code plagiarism. This type of plagiarism is also committed by students, i.e. they copy an entire text or a part of it by changing the code without indicating the source. This type of plagiarism is difficult to determine.

Text and source code plagiarisms can be detected using 2 methods: physical and automatic detection [5-8].

- During the physical detection, expert reads the material and decides whether it's plagiarized or not. For example, teacher personally checks assigned term papers and evaluates its originality. But this method is not effective for checking many documents, as it requires the teacher a lot of time and effort.
- Automatic detection is performed by computer, thus, there are many software that detect plagiarism: PlagAware, PlagScan, Check for Plagiarism, iThenticate, PlagiarismDetection, Academic Plagiarism, The Plagiarism Checker, Urkund, Docoloc.

Types of Plagiarism

In general, plagiarism is differently classified by experts [9-13]. Comprehensively, plagiarism can be divided into following types: I type plagiarism: The Direct Copy. An author copies another's phrases, paragraphs or entire page word-by-word and uses it on behalf of himself/herself. II type plagiarism: The Minimalist. Several words or phrases are changed from source material, but main context, structure remains the same. III type plagiarism: The Combo. A certain part of information is cited, hoping it will not be detected in the future, but other parts are copied without citation. IV type plagiarism: The Quilt. Some authors only cite one source despite using several sources. V type plagiarism. The parrot. A mistake made by the majority of researchers: they copy the main structure and idea of the cited source. VI type plagiarism: The Insecure. This case is also known as "Little mermaid complex", so that, the research rejects his own idea, and only uses others' idea; i.e., the author writes an article completely consisting of ideas cited in sources of other authors, instead of his own. VII type plagiarism: The Lost-at-Sea. We can encounter this when there's a requirement for the number of sources in reference section of the article. VIII type plagiarism: The Self-Obsessed. Authors think that, owning an idea allows them to use these ideas at different times and different journals. Presenting written article to the academic society in a new form is also considered self-plagiarism.

National Anti-Plagiarism Service

In addition to anti-plagiarism software, there are legal, public reproach and educational methods of fighting plagiarism [14-16].

Nowadays, legislative framework for plagiarism has been established in many countries. Plagiarism in these countries is considered a criminal offense such as "fraud", "theft" etc. and a criminal has to pay for the damage caused by plagiarism.

International experience shows that punishing plagiarist using social reproach is a useful mechanism. These individuals are deprived of their duties and academic titles without being liable to legal remedies.

One of the ways to combat plagiarism is education, as this process involves the training of specialists in combating plagiarism and teaching this knowledge in higher education institutions.

Detecting plagiarism prior to publishing the scientific work is considered more secure for the author. For this reason, it is very important to apply anti-plagiarism systems. Each institution must plan pre-checking every work that will be submitted to journal or conference and running it through anti-plagiarism software [17, 18]. At the Institute of Information Technology of ANAS anti-plagiarism service is a process consisting of three stages (Figure 1):

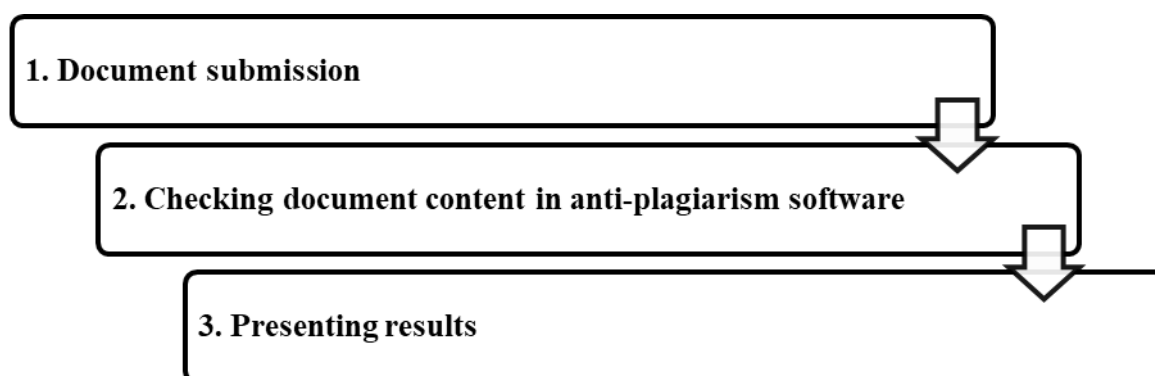


Figure 1. Stages of anti-plagiarism service at the Institute of Information Technology of ANAS

At the "Document submission" stage, electronic version of scientific works – thesis, article, term paper, graduation paper, dissertation etc. is submitted to the relevant department by the author. Checking document content in anti-plagiarism software includes entering text and object

into anti-plagiarism software part-by-part and evaluating it. Results obtained during the last stage are presented to the author one way or another.

Document submission process in the first stage is schematically described in Figure 2.

During document submission process, initially submitted scientific works are classified; after documents' classification as book, article, dissertation etc., the process of clearing the documents from objects starts.

While clearing the document from objects, figures, tables, diagrams etc. referenced in the text are determined and collected in a new folder. Then it is intended to check these objects individually in the anti-plagiarism system. While dividing the whole text, it is important to divide the whole text into sections such as introduction, problem solution, result, chapter 1, chapter 2 etc. This allows detecting which sections of the text contain more facts of plagiarism while presenting the results. Separating the reference section determines the authenticity of cited sources and to which degree the authors uses these resources.

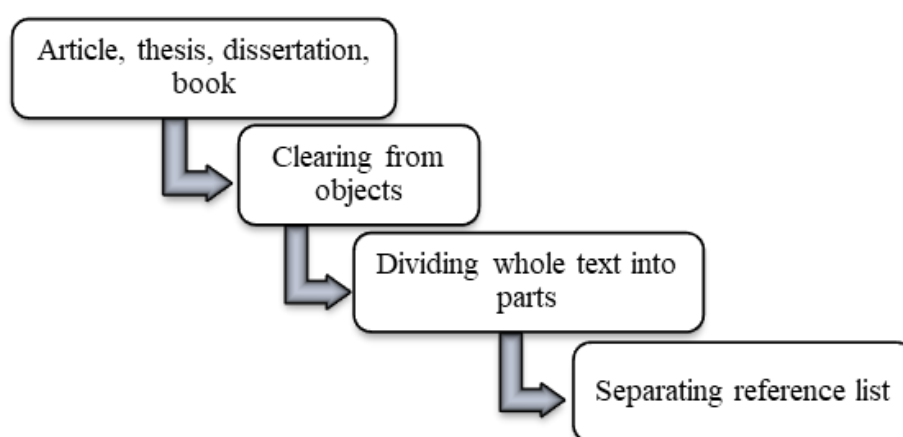


Figure 2. Document submission process

Results can be presented to the authors in two forms: simple and justified. Simple form contains originality percentages obtained from entering the whole text into anti-plagiarism software without analysis of the results by the expert. Justified results presentation lists several points (Figure 3).

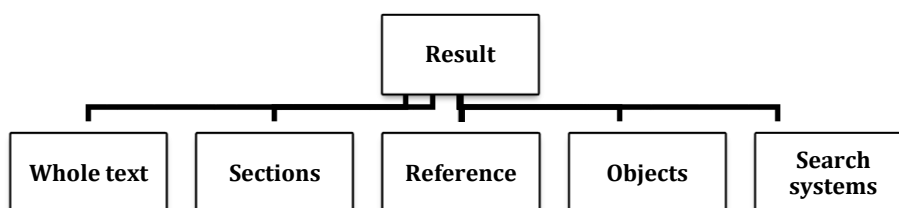


Figure 3. Presentation of the results

Justified results are presented to the author along with percentage indicators covered in simple results. Here, originality and plagiarism percentages obtained after checking separate parts of the text to anti-plagiarism software are presented. Anti-plagiarism software determines how the authors used the sources indicated in reference list (i.e., copying idea, text etc. as is) and the relevant indicators are provided. Each object separated from text is compared to similar figures in search systems through anti-plagiarism system and encountered identities are provided in results. Each anti-plagiarism system searches plagiarism both online from the Internet and from the documents in its local base. Therefore, data base representing electronic content of all scientific works published in the country (i.e., National Reference Index) must be accessible for plagiarisms systems functioning in the country. From this standpoint, creating national electronic content is the most important point for developing a single anti-plagiarism service in the country. Currently,

creation of national content in Azerbaijani is one of the topical issues. A very small number of over 200 scientific journals have website online versions, which do not meet international standards. Thus, the content of all published articles must be accessible and existing anti-plagiarism system must be able to carry out the revision using this base. In fact, the availability of scientific works in the Internet can lead to increased plagiarism, but at the same time they can be easily detected by anti-plagiarism programs.

Overall operation principle of anti-plagiarism software is described in following figure (Figure 4):

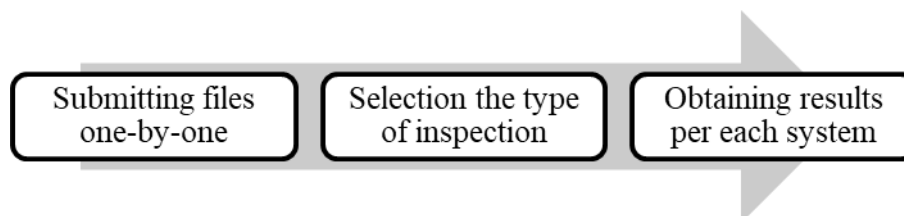


Figure 4. Operation principle of anti-plagiarism system

Anti-plagiarism Systems

Currently, along with anti-plagiarism systems, anti-plagiarism systems that deceive these systems function in internet environment. This service is rendered on a paid basis or free of charge. Basically said, operation principle of an anti-plagiarism system is based on changes made in the text in order to avoid detection by anti-plagiarism software and implemented in several ways.

- All sentences in the given text are processed one by one. Plagiarists use a method called “legitimate” to change the text (rewrite) using synonyms, as long as the meaning is kept. “Synonymizer” software is used for this purpose.
- The letters “a” of the Latin alphabet in the given text are replaced with a similar letter of Cyrillic, Greek or the least commonly used alphabet in the Internet. Although traditional anti-plagiarism systems cannot detect this change, this method can now be detected by many anti-plagiarism systems.
- Many anti-plagiarism systems deceive anti-plagiarism system using translation effect. Translation effect is translation of a text from one language to another and then translating it back using i.e., Google Translate. As a result, originality of the text increases from 0% to 79%.
- It is possible to increase originality of a text from 0.05% to 68.75% by changing scale and font size. Here, a successful result is obtained by entering the same text as completely different text into the anti-plagiarism system using the text editor's automatic splitting function etc.

Conclusion

Despite following multilingualism principles in available anti-plagiarism systems, so far there are no universal anti-plagiarism systems that are not dependent on language characteristics. For this reason, anti-plagiarism systems must consider characteristics of different languages. Following characteristics can be listed for Azerbaijani language:

Anti-plagiarism systems can compare texts not only using symbols, but also by taking screenshots and comparing texts, or automatic auto-abstract of texts can be made and checked.

Relevant possible variations of Latin and Cyrillic alphabets can be checked, i.e., checking by inter-changing Latin “o” and Cyrillic “o” letters etc. Text can be checked using ASCII and Unicode reverse coding.

Plagiarism detected in anti-plagiarism system must be analyzed based on date. Sometimes in internet environment detected identical text is considered plagiarism, but submitted text is actually the original, i.e., written at an earlier date than the other text.

It is possible to find and compare the texts of the sources provided in reference section, not only comparing the titles of the sources in reference list.

The establishment of the National Anti-Plagiarism Service in Azerbaijan will provide more reliable information sent to scientific journals and conferences from all scientific organizations and universities. Simultaneously, it will allow detecting plagiarism facts before the discussion of course papers, graduate works, abstracts and dissertations submitted for defense. Thus, a decision will be made on the authenticity of each scientific work based on the submitted reference on the results of the plagiarism test.

References

1. Əliquliyev R.M., Ağayev N.B., Alıquliyev R.M. Plagiatlıqla mübarizə texnologiyaları, Bakı: İnformasiya Texnologiyaları nəşriyyatı, 2015, 165 s.
2. Education with Integrity, <http://www.turnitin.com>
3. What is Plagiarism? 2017, <http://www.plagiarism.org>
4. Asim M., Hussam M., Snasel V. Overview and Comparison of Plagiarism Detection Tools / Proceedings of the DATESO 2011: Annual International Workshop on DATABASES, TEXTS, SPECIFICATIONS and OBJECTS, Pisek, Czech Republic, April 20, 2011, pp.161-172.
5. T. Mallon, Stolen words: Forays into the origins and ravages of Plagiarism, Ticknor and Fields Publications, 1989, 300 p.
6. Guardian news reports, November 24 1986, March 10 1997, July 11 1997, November 22, 2005, <http://www.theguardian.com>
7. Encyclopedia Britannica, <http://www.britannica.com/EBchecked/topic/462640/plagiarism>.
8. How Do I Know If I'm Plagiarizing? http://www.scribendi.com/advice/how_do_i_know_if_im_plagiarizing.en.html
9. Reliably plagiarism detector, <http://www.plagiarismdetecting.org>
10. Alzahrani S., Salim N., Abraham A. Understanding Plagiarism Linguistic Patterns, Textual Features and Detection Methods // IEEE Transactions on Systems, Man, and Cybernetics – Part C: Applications and Reviews, 2012, vol.42, no.2, pp.133-143.
11. Hoard T.C., Zobel J. Methods for Identifying Versioned and Plagiarized Documents // Journal of the American Society for Information Science and Technology, 2003, vol.54, no.3, pp.203-215.
12. Jadalla A., Elnagar A. PDE4Java: Plagiarism Detection Engine for Java source code: a clustering approach // International Journal of Business Intelligence and Data Mining, 2008, vol.3, no.2, pp.121-135.
13. Zechner M., Muhr M., Kern R., Michael G. External and intrinsic plagiarism detection using vector space models / Proceedings of the SEPLN'09 Workshop on Uncovering Plagiarism, Authorship and Social Software Misuse, 2009, pp.47-55.
14. Jooste C. The Plagiarist on Trial – a legal perspective on plagiarism, 2018, <http://www.blogs.sun.ac.za>
15. Education minister loses doctorate for plagiarism, <http://www.thelocal.de/20130205>.
16. Əliquliyev R.M., Alıquliyev R.M., Mahmudov R.Ş. Azərbaycanca plagiatlıqla mübarizə problemləri və həll yolları // İnformasiya cəmiyyəti problemləri, 2019, №1, s.34-43.
17. Həsənova R.Ş., Əsgərov F.Ş. Plagiatlığın növləri və ondan qorunmağın üsulları // “İnformasiya təhlükəsizliyinin aktual multidissiplinar elmi-praktiki problemləri” IV respublika konfransı, 14 dekabr 2018-ci il, s.221-223.
18. Həsənova R.Ş., Əsgərov F.Ş. Plagiatlıq və onunla mübarizə vasitələri: antiplagiat sistemləri // İnformasiya cəmiyyəti problemləri, 2019, №2, s.94-101.

UOT 004.04

Həsənova R.Ş.¹, Əsgərov F.Ş.²

^{1,2}AMEA İnformasiya Texnologiyaları İnstitutu, Bakı, Azərbaycan

¹rahasanova@gmail.com, ²firudinasgarov@gmail.com

Azərbaycanda milli antiplagiat xidmətinin yaradılması konsepsiyası

Hazırda elmi cəmiyyətdə plagiatlıqla mübarizə aktual məsələlərdən biridir. İKT-nin sürətli inkişafı, elmi informasiyanın İnternet mühitində açıq və əlyətərli olması plagiatlıq kimi neqativ hallar üçün əlverişli mühit yaradır. Hər bir müəllif istəmədən və ya digər səbəbdən bu kimi hallarla qarşılaşa bilər. Məqalə Azərbaycanda milli antiplagiat xidmətinin yaradılması konsepsiyasına həsr edilmişdir. Burada plagiatlıq və onun növləri haqqında qısa məlumat verilmişdir. Eləcə də, anti-plagiat sistemlərinin xüsusiyyətləri haqqında məlumat verilmişdir. Azərbaycan nümunəsində milli antiplagiat xidmətinin göstərilməsi mərhələləri sadalanmışdır. Təklif edilən milli antiplagiat xidmətindən bütün elmi təşkilatların, ali təhsil müəssisələrinin müəllim-tələbə heyəti və elmi tədqiqatçıları yararlanı bilər. Belə ki, nəticədə hər bir elmi iş üçün təqdim olunan arayışla onun orijinal olub-olmaması haqqında qərar qəbul etmək mümkündür.

Açar sözlər: plagiatlıq, piratçılıq, antiplagiat sistemləri, anti-antiplagiat sistemləri.

УДК 004.04

Гасанова Рахилия Ш.¹, Аскеров Фирудин Ш.²

^{1,2}Институт Информационных Технологий НАНА, Баку, Азербайджан

¹rahasanova@gmail.com, ²firudinasgarov@gmail.com

Концепция создания национальной службы антиплагиата в Азербайджане

В настоящее время борьба с плагиатом – одна из самых актуальных проблем в научном сообществе. Стремительное развитие информационных и коммуникационных технологий, открытость и доступность научной информации в Интернете создают благоприятную среду для неприятных ситуаций, таких, как плагиат. Любой автор может столкнуться с подобными ситуациями невольно или по любой другой причине. Статья посвящена концепции создания в Азербайджане национальной службы антиплагиата. Здесь приведена краткая информация о плагиате и его видах. В статье также предоставляется информация об особенностях систем антиантиплагиата. На примере Азербайджана перечислены этапы оказания национальных услуг по борьбе с плагиатом. Предлагаемый национальный сервис антиплагиата может использоваться преподавателями, студентами и научными исследователями университетов и научных организаций. Таким образом, по представленной справке на каждую научную работу можно решить, является она оригиналом или нет.

Ключевые слова: плагиат, пиратство, система антиплагиата, система антиантиплагиата.