

Гайтан О.М.

Полтавський національний технічний університет імені Юрія Кондратюка

Горошко А.І.

Полтавський національний технічний університет імені Юрія Кондратюка

ПОРІВНЯЛЬНИЙ АНАЛІЗ СИСТЕМ АВТОМАТИЧНОЇ ПЕРЕВІРКИ ОРИГІНАЛЬНОСТІ ТЕКСТІВ

У статті розглядається питання неправомірного використання інтелектуальної власності, а також досліджуються наявні системи контролю для перевірки автентичності змісту тексту. У роботі здійснено перевірку вибраних текстів із метою виявлення рівня їх оригінальності та визначення переваг чи недоліків представлених на ринку програмного забезпечення програмних комплексів автоматичного аналізу текстів.

Ключові слова: оригінальність тексту, онлайн-сервіс, програмний комплекс, антиплагіат.

Постановка проблеми. На сучасному етапі розвитку суспільства в людей з'явився широкий доступ до різних інформаційних джерел у зв'язку з поширенням мережі Інтернет, а високий розвиток інформаційно-комунікаційних технологій полегшив завдання пошуку та зберігання необхідної інформації. Сьогодні достатньо ввести назву теми в пошукову систему для подальшого копіювання необхідної інформації. Унаслідок цього почав поширюватися метод написання робіт (зокрема студентами), який полягає в простому копіюванні інформації з одного чи декількох джерел із мінімальним редагуванням. Одним із варіантів вирішення проблеми є використання систем автоматичної перевірки оригінальності тексту з метою вдосконалення організації та контролю навчального процесу в університеті, забезпечення самостійності виконання випускних кваліфікаційних і курсових робіт, збереження прав інтелектуальної власності. Отже, дуже важливим є питання про порівняння інформаційних систем подібного роду.

Аналіз останніх досліджень і публікацій. Проблемі аналізу систем антиплагіату та використанню їх для перевірки наукових і академічних текстів протягом останнього десятиліття приділяється значна увага. Публікації із цього питання можна поділити на декілька категорій:

– технології й особливості використання таких систем у навчальному процесі для перевірки студентських робіт на оригінальність і захисту інтелектуальної власності авторів навчальних курсів у рамках питання забезпечення академічної чесності, яке є сьогодні є дуже актуальним для України. Зараз цьому питанню приділяється багато

уваги як на рівні держави, так і на рівні окремих навчальних закладів. Американські ради з освіти за участі Міністерства освіти й науки України та за підтримки Посольства Сполучених Штатів Америки започаткували проект «Сприяння академічній доброчесності в Україні» (SAIUP project), який нині охоплює 10 українських вузів [1];

– технології й особливості використання цих систем у науковому середовищі для перевірки наукових робіт на оригінальність, зокрема й наукових статей і дисертацій. Наразі більшість статей, які подаються до наукових журналів, і всі дисертації, представлені до захисту в спеціалізованих вчених рада, перевіряються системами антиплагіату. Однак ця перевірка є поверховою та залежить від можливостей системи антиплагіату, а також бази джерел, що використовується в цій системі [2; 3];

– аналіз наявних і розроблення нових алгоритмів і методів антиплагіату. У якості прикладу можна навести модифікацію алгоритму шинглів для пошуку нечітких дублікатів для складених документів [4; 5];

– аналіз наявних і розроблення нових систем антиплагіату; технології, що використовуються, і їх програмна реалізація [6–14];

– класифікація програмних засобів цієї категорії за різними типами показників: за колекцією джерел, що використовується, за характеристиками ресурсу, за складністю метрик, що використовуються, тощо [15; 16].

Однак, незважаючи на велику кількість досліджень, присвячених цьому питанню, залишається питання ефективності конкретних систем і вибору оптимальної з орієнтацією на українські роботи, написані українською чи російською мовами, з

урахуванням використання простого перекладу між цими мовами.

Постановка завдання. Основна мета статті полягає в дослідженні особливостей функціонування й результативності різних інформаційно-програмних систем автоматичної перевірки текстів для вдосконалення аналізу студентських наукових робіт на оригінальність.

Об'єкт дослідження – інформаційно-програмні системи автоматичної перевірки текстів.

Виклад основного матеріалу дослідження. У процесі дослідження систем автоматичної перевірки текстів за критеріями оригінальності використовуються різні підходи. Одні з них базуються на використанні власної оффлайн-бази текстів, а інші можуть коректно працювати лише за умови підключення до Інтернету. Але ми вважаємо, що найбільшою популярністю користуються такі системи:

ETXT-антиплагіат (<https://www.etxt.ru/antiplagiat/>) [19] – система, розповсюджена в Україні, принцип роботи якої полягає в порівнянні тексту із запитом в пошукових системах (Google, Яндекс, Bing тощо).

Переваги системи:

- 1) підтримка української мови;
- 2) можливість виявлення російсько-українського та зворотного перекладу;
- 3) ефективність пошуку дублікатів;
- 4) додаткові можливості перевірки на рерайт, комплексної перевірки всіх сторінок сайту, перевірки зображення та SEO-перевірки.

Недоліки системи:

- 1) обмеження обсягу тексту, що перевіряється, у безкоштовній онлайн-версії (до 3000 символів – без реєстрації, до 5000 символів – після реєстрації). Однак програма має безкоштовну десктопну версію, яку можна завантажити із сайту програми;
- 2) тривалість перевірки великих документів до 1 години та необхідність багаторазового введення капчі для порівняння з пошуковими запитом Google (наприклад, перевірка цієї статті вимагає введення капчі 17 разів).

Адвего-Плагиатус (<http://advego.ru/plagiat/>) [20] – система, що досить розповсюджена в Україні. Принципи роботи системи: текст розбивається на маленькі частини, які аналізуються за декількома алгоритмами: алгоритм шинглів для перевірки точних збігів фраз і сторінок-першоджерел, алгоритм лексичних збігів для перевірки схожості набору лем, термінів і значущих слів, алгоритм псевдоунікалізації для перевірки сторонніх символів. Програма має онлайн і десктопну версію,

яку можна завантажити із сайту програми.

До переваг системи можна віднести підтримку української мови, відсутність обмежень на кількість символів у документі.

Недоліки: менша ефективність перевірки порівняно із системою ETXT-антиплагіат. Крім того, проблематичною є перевірка великих файлів, незважаючи на формальну відсутність обмежень на кількість символів у документі, і необхідність багаторазового введення капчі для порівняння з пошуковими запитом Google.

Наступною за популярністю серед користувачів є така система, як text.ru-Антиплагіат (<http://text.ru/antiplagiat/>). Система використовує вдосконалену методику перевірки текстів, відмінну від інших систем.

Переваги системи:

- 1) підтримка української мови;
- 2) можливість виявлення російсько-українського перекладу та незначного рерайту (переставлення слів, фраз і речень місцями, зміна граматичних форм слів, додавання чи видалення декількох слів);

3) додаткові можливості: SEO-аналіз тексту та перевірка орфографії.

Основний недолік системи полягає в тому, що система не підтримує перевірку великих текстів, які необхідно розбивати на шматки, що змінює оригінальність тексту.

Наступною для дослідження була обрана польська інформаційна система «Антиплагіат» (<https://www.plagiat.pl/webplagiat/main.action?menu=home&lang=en&lang2=en>), оскільки саме за цією системою багато українських вузів перевіряють роботи.

Як зазначає компанія-розробник, однією з переваг цієї системи є акумуляція баз даних усіх вищих навчальних закладів, з якими співпрацює компанія. Компанія постійно розширює свою присутність на ринку України. 4–5 червня 2018 р. у Львівському державному університеті внутрішніх справ пройшов круглий стіл, присвячений системі StrikePlagiarism.com (основний продукт системи «Антиплагіат»), під час якого представники компанії представили можливості системи та новий інтерфейс програми, а представники українських вузів (Львівського державного університету внутрішніх справ, Київського національного економічного університету імені Вадима Гетьмана, Харківського національного університету ім. В.Н. Каразіна та Прикарпатського національного університету імені Василя Стефаника) поділилися позитивними враженнями від

упровадження системи та досвідом її використання. Принципи роботи цієї системи схожі з ЕТХТ: можливість перевірки текстів українською та російською мовами, використання модуля перекладу для виявлення російсько-українського та українсько-російського перекладу, можливості виявлення незначної модифікації тексту, зокрема перекладів-синонімів, інфлексії тощо. Перевірка здійснюється онлайн: після завантаження документа, його аналізу та перевірки система видає детальний звіт.

Таким чином, перевагами системи є підтримка української мови, використання внутрішньої бази університетів і досить потужний алгоритм пошуку дублікатів.

До недоліку системи можна віднести платність ресурсу. Тарифи для університетів обговорюються окремо залежно від їх можливостей. Плата для звичайних користувачів становить 2,44 євро за 20 000 символів.

Антиплагиат.ру [18] – російська система, яка впроваджена з 2005 р. і працює лише з російською мовою.

Завантажені тексти перевіряються за такими базами:

- модуль пошуку в мережі Інтернет, який містить 200 млн джерел російською, англійською й іншими мовами;
- колекція дисертацій російської державної бібліотеки;
- модуль пошуку «Elibrary»;
- модулі «Lexpro» і «Гарант», які містять нормативно-правову документацію;
- модуль «Медицина», який містить наукову та навчальну літературу з медицини та фармацевтики;
- модуль «Патенти».

Принципи роботи системи:

- 1) розбивання файлу на невеликі фрагменти;
- 2) пошук дублікатів за аналогією з роботою системи «Яндекс».

До переваг системи можна віднести ефективність роботи алгоритму перевірки.

Недоліки системи:

- 1) система не підтримує українську мову, тому в Україні майже не використовується;
- 2) вартість виведення повних звітів про перевірку, яка залежить від кількості перевірок і періоду доступу до системи. Мінімальна вартість – 150 руб. за 10 перевірок. У безкоштовному тарифі виконується пошук лише в модулі Інтернету.

Система Антиплагиат.ру має розширення Антиплагиат-вуз, яке, крім вищевказаних моду-

лів, містить базу робіт російських вузів. Але це розширення також не використовується в Україні через відсутність підтримки української мови.

Після представленого вище опису систем аналізу текстів, що покладені в основу нашого дослідження, доцільно перейти до розгляду принципів їх функціонування й проведення порівняння результатів роботи.

У системі ЕТХТ-антиплагиат (<https://www.etxt.ru/antiplagiat/>) [19] робота з перевірки тексту на оригінальність складається з таких етапів:

1) збір інформації з різних джерел: сайтів Інтернету, бази наукових статей і рефератів тощо. Завантажені документи проходять процедуру фільтрації, засновану на унікальній технології очищення тексту, у результаті якої відкидається марна з погляду потенційного цитування інформація;

2) кожен з отриманих текстів певним чином форматується й заноситься в системну базу даних. Таким чином, на вході система приймає файл, а на виході видає всі фрагменти тексту, щодо яких система знайшла схожі аналоги в Інтернеті (ці фрагменти відповідно виділяються), видається посилання, де вони могли бути скопійовані. Також підраховується ймовірність запозичення фрагмента: 100% – фрагмент запозичений у початковому вигляді, якщо ймовірність менше, то у фрагменті можуть бути переставлені слова, змінені словоформи, розбиті речення, слова замінені синонімами, зроблені інші маніпуляції з текстом. Користувачі, які підключили платний аккаунт, отримують можливість перевірки за додатковими базами, перегляду повного звіту з посиланнями на джерела отриманої інформації.

Завантажені тексти перевіряються за такими базами:

- 1) модуль пошуку в мережі Інтернет, що включає колекції рефератів, наукові публікації, статті Вікіпедії;
- 2) колекції дипломних робіт деяких ВНЗ;
- 3) колекції дисертацій державної бібліотеки;
- 4) колекції юридичних документів LEXPRO;
- 5) колекції наукової електронної бібліотеки elibrary.ru.

Відзначимо, що пункти 3–5 доступні користувачам із платними аккаунтами.

Інший онлайн-сервіс (Findcopy.ru (<http://findcopy.ru/>)) перевіряє текст на унікальність, спираючись здебільшого на сайти мережі Інтернет. Для роботи достатньо безкоштовно зареєструватися на сайті й пройти авторизацію. У середньому перевірка тексту здійснюється про-

тягом 45 секунд. Текст, підготовлений для перевірки, необхідно розмістити в спеціальному полі на сторінці ресурсу. Альтернативним способом є перевірка оригінальності тексту веб-сторінки, для чого необхідно вказати її URL-адресу [21].

Згідно з отриманою нами інформацією [10; 22], система «Автор.NET» здійснює перевірку як за джерелами, доступними в Інтернеті, так і за власними джерелами (бази статей, курсових і контрольних робіт, дипломних проектів тощо). За результатами перевірки формується звіт з висвітленням знайдених запозичень і можливістю перегляду знайдених джерел.

Як зазначається в огляді «Автор.NET» [22], система містить два модулі, що функціонують незалежно один від одного:

1) перший модуль здійснює перевірку за внутрішньою базою джерел;

2) другий модуль проводить перевірку за джерелами мережі Інтернет.

Для цієї мети текст документа розбивається на інформативні фрагменти, кількість яких залежить від розміру документа. Проводиться пошук джерел, що містять ці фрагменти, із використанням модуля пошукової системи Яндекс, онлайн-пошуку систем Google.ru, Rambler.ru, Aport.ru, Поиск.Mail.ru, Nigma.ru тощо. Отримані джерела перевіряються на відповідність вхідному документу. Для цього визначається формат джерела (html, txt, doc, rtf або pdf-файл). Із гіпертекстових документів видаляються теги розмітки. Далі джерела проходять попередню обробку, а потім проводиться оцінка їх подібності з вихідним документом.

Сервіс Copyscape (www.copyscape.com) [23] використовується для перевірки оригінальності контенту російськомовних і англо-мовних текстів, опублікованих у мережі Інтернет. Для перевірки необхідно вказати URL-адресу ресурсу, який цікавить користувача. У цьому сервісі кількість запитів з однієї IP-адреси обмежена, також обмеження поширюються на перегляд результатів пошуку збігів: у безкоштовному аккаунті надаються тільки первинні джерела.

На основі робіт [9; 10; 24] і власного аналізу розглянутих сер-

вісів можна сформулювати головні принципи функціонування систем перевірки тексту.

По-перше, вихідний текст проходить попередню обробку, яка містить такі дії:

1) розділення тексту на частини;

2) перетворення регістра;

3) перетворення латинських літер в українських словах на аналогічні літери українського алфавіту для текстів українською мовою;

4) видалення стоп-слів (прийменники, прислівники тощо) і знаків пунктуації;

5) фільтрація тексту: видалення неінформативних, найбільш розповсюджених і слів, що рідко трапляються;

6) відкидання закінчень слів.

По-друге, для оригінального чи обробленого тексту визначається індекс його оригінальності. При цьому різні способи індексування переслідують різні цілі – прискорення пошуку, скорочення розміру пошукової бази, стійкість до помилок чи пропусків, стійкість до перетворень тексту тощо.

Для порівняння роботи сервісів під час перевірки оригінальності текстів нами був проведений експеримент. Було підготовлено кілька текстів.

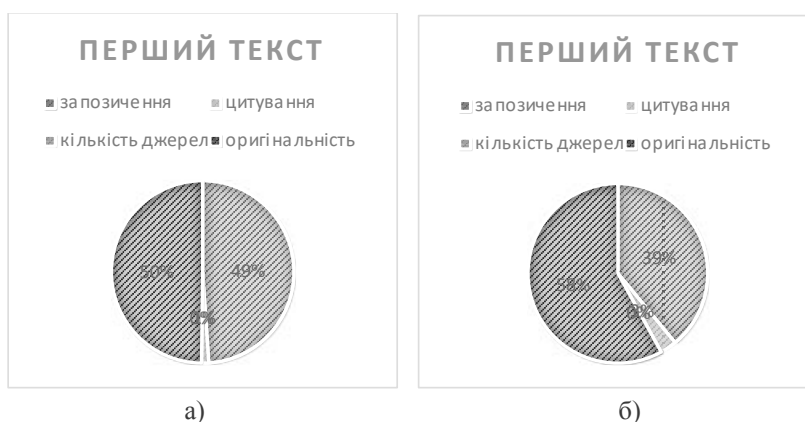


Рис. 1. Результати перевірки першого тексту в ETXT-антиплагіат (а) і Findcopy.ru (б)



Рис. 2. Результати перевірки другого тексту в ETXT-антиплагіат (а) і Findcopy.ru (б)

1. Текст, створений на основі лекції з дисципліни «Моделювання систем» зі спеціальною термінологією.

2. Текст, отриманий на основі статті з електронного журналу із заміною деяких слів синонімами. У тексті також використовуються спеціальні терміни.

3. Текст без використання термінів, але з наявністю складних оборотів, складносурядних і складнопідрядних речень. Текст є повністю оригінальним, оскільки був складений автором статті.

Ці тексти були перевірені на системах Findcopy.ru та ETXT-антиплагіат.

За першим підготовленим текстом у безкоштовному аккаунті системи ETXT-антиплагіат були отримані такі результати: оригінальність – 50,44%, запозичення – 49,56%, цитування – 1%.

Результати перевірки цього самого тексту на ресурсі Findcopy.ru показали оригінальність тексту 58% і 3 джерела.

Другий текст був узятий зі статті електронного журналу із заміною деяких слів синонімами. У цьому тексті, так само, як і в першому, використовується складна термінологія, кількість символів – 217.

В обох версіях системи ETXT-антиплагіат отримані такі дані: оригінальність – 100%, запозичення – 0%, цитування – 0%.

Такий результат може означати те, що в базі системи перевірки тексту на оригінальність немає статей із низки вельми відомих електронних журналів. Питання про внесення їх до бази поки залишається відкритим.

Перевірка другого тексту на Findcopy.ru виявила оригінальність тексту 16,2%; наведено одне джерело – lasola.ru (83,8%).

Третій текст був написаний автором статті у 2017 році для участі у Всеукраїнському конкурсі журналістики і є фрагментом літературної прози, у ньому немає запозичень і спеціальної термінології, кількість символів – 309.

У безкоштовній версії системи ETXT-антиплагіат (<https://www.etxt.ru/antiplagiat/>) отримані такі дані: оригінальність – 44,56%, запозичення – 55,44%, цитування – 0%.

Перевірка третього тексту на сервісі Findcopy.ru показала стовідсоткову оригінальність тексту.

Результати порівняння показали, що система ETXT-антиплагіат здійснює пошук за кількома напрямками, представленими вище. Проте система має низку недоліків.

По-перше, система не здійснює пошук серед усіх документів, доступних у мережі Інтернет. Особливо це стосується тематичних сайтів і новинних порталів: багато запозичень здійснюється саме з таких джерел. Доказом цього є приклад – перевірка другого тексту, повністю взятого з електронного видання. Ця система, перевіривши його, показала унікальність тексту 100%.

По-друге, у безкоштовній версії є обмеження розміру тексту – 3000 або 5000 символів (доступно після реєстрації).

По-третє, обмежений перегляд документів, що частково відповідають тексту. Крім того, система обмежує можливість перевірки в базі наявних робіт.

До переваг системи Findcopy.ru можна віднести можливість пошуку не за окремими колекціями, а за всіма документами, доступними в мережі Інтернет. Прикладом є перевірка другого тексту, де система показала 16,2% унікальності, а ETXT-антиплагіат – 100%. Також є обмеження довжини тексту до 10 000 знаків. Основним недоліком Findcopy.ru є те, що система не описує детально джерела, а тільки вказує посилання на них і кількість відсотків, узятих із кожного джерела. У результаті порівняння можна зробити висновок, що відмінності в отриманих результатах зумовлені насамперед відмінностями баз систем, де перевіряється завантажений текст. Деякі системи перевіряють текст тільки в Інтернеті, на певних сайтах, однак багатьох є й свої бази.

Висновки. 1. Сучасність диктує необхідність створення та використання систем контролю для перевірки автентичності змісту тексту.

2. Інформаційно-програмні системи автоматичної перевірки текстів дозволяють виховати в студентів оригінальність науко-

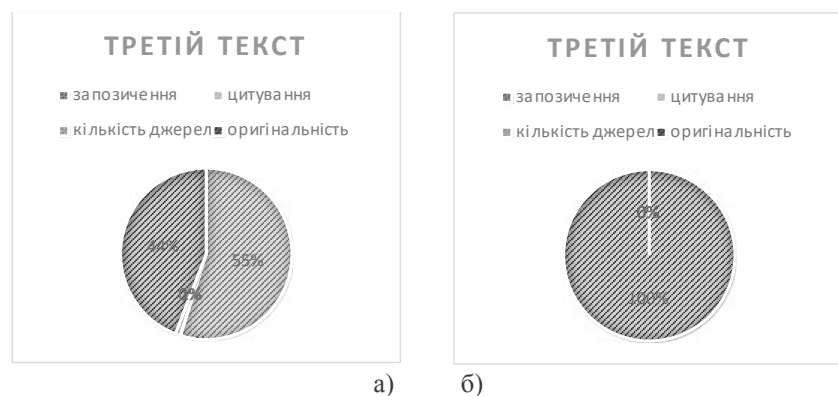


Рис. 3. Результати перевірки третього тексту в ETXT-антиплагіат (а) і Findcopy.ru (б)

вого мислення та забезпечити збереження інтелектуальної власності.

3. Результати перевірки деяких систем, отримані під час дослідження, виявили низку недоліків, пов'язаних із їх роботою.

4. Для отримання найбільш оптимального й об'єктивного результату під час перевірки текстів варто створювати гібридну версію або ж використовувати декілька інформаційних систем.

Список літератури:

1. Яворська Т.М. Академічна доброчесність – запорука якісної освіти. URL: http://www.library.vn.ua/Konf2017/text/2017_26.doc (дата звернення: 20.09.2018).
2. Арсентьев Д.А. Использование программы антиплагиат в условиях современного вуза. Вестник Московского государственного университета печати им. Ивана Федорова. 2015. № 6. С. 9–10.
3. Хачецуков З.М. Проверка на оригинальность научных текстов: питання теорії і практики. Гуманітарій Півдня Росії. 2014. № 1. С. 166–179.
4. Квашина Ю.А. Методы поиска дубликатов скомпонованных текстов научной стилистики. Технологический аудит и резервы производства. 2013. № 3 (3). С. 16–20.
5. Чижова А.А. Алгоритмы пошуку плагіату. Східно-Європейський журнал передових технологій. 2010. № 2 (46), С. 13–16.
6. Чиркин Е.С. Системы автоматизованої перевірки на неправомірні запозичення. Вісник Тамбовського університету. Серія: Гуманітарні науки. 2013. № 12 (128). С. 164–174.
7. Shenoy M., Shet K.C., Acharya U.D. Automatic Plagiarism Detection Using Similarity Analysis. Advanced Computing: An International Journal. 2012. № 3 (3). Pp. 59–62.
8. Власова М.С., Суханов О.В. Технічні засоби для боротьби з плагіатом з метою підвищення якості навчального процесу. Вісник цивільних інженерів. 2012. № 6 (35). С. 187–192.
9. Шарапова Е.В. Исследование возможностей системы «антиплагиат.» для обнаружения заимствований. Перспективы науки и образования. 2013. № 3. С. 215–219.
10. Sharapova E.V., Sharapov R.V. Universal System of text checking for plagiarism «Avtor.net». Informatika i ejo primenenija. 2012. № 3 (6). Pp. 52–58.
11. Tschuggnall M., Specht G. Detecting Plagiarism in Text Documents through Grammar-Analysis of Authors. Database Systems for Business, Technology and Web: proceedings of the 15th GI-Symposium. (11th March–15th March, 2013). 2013. Pp. 241–259. URL: <http://cs.emis.de/LNI/Proceedings/Proceedings214/241.pdf> (дата звернення: 20.09.2018).
12. Singh R., Dutta C. Duplicity Detection System for Digital Documents. International Journal of Soft Computing and Engineering. 2012. № 5 (2). Pp. 24–28.
13. Meyer zu Eissen S., Stein B. Intrinsic Plagiarism Detection. Advances in Information Retrieval: Proceedings of the 28th European Conference on IR Research, ECIR. 2006. Pp. 565–569.
14. Ерьсь А.В. Анализ современных систем обхода заимствований текста. Постулат. 2018. № 1. URL: <http://e-postulat.ru/index.php/Postulat/article/viewFile/1097/1124> (дата звернення: 20.09.2018).
15. Лупаренко Л.А. Инструментарій виявлення плагіату в наукових роботах: аналіз програмних рішень. Інформаційні технології і засоби навчання. 2014, Т. 40. № 2. С. 151–169.
16. Lancaster T., Culwin F. Classifications of Plagiarism Detection Engines. Innovation in Teaching and Learning in Information and Computer Sciences. 2005. №2 (4). 16 pp.
17. Мустафина А.К., Кальпеева Ж.Б., Маженов А.К., Сабыржан А.А. О создании университетской системы проверки текстов на плагиат. Издательство, нэтижелер: Исследования, результаты. 2017. № 3. С. 407–413.
18. URL: <https://www.antiplagiat.ru/> (дата звернення: 20.09.2018).
19. ЕТХТ-антиплагиат. URL: <https://www.etxt.ru/antiplagiat/> (дата звернення: 20.09.2018).
20. Программа для проверки уникальности текста Advego Plagiatus. URL: <https://advego.com/plagiatus/> (дата звернення: 20.09.2018).
21. Findcopy. URL: http://mastersloga.ru/news/novyj_servis_proverki_unikalnosti_find_copy_ru/ (дата звернення: 20.09.2018).
22. Огляд Автор.NET. URL: <http://ceur-ws.org/> (дата звернення: 20.09.2018).
23. Огляд copyscape.com. URL: <http://profcopywriting.ru/copyscape.html> (дата звернення: 20.09.2018).
24. Шарапов Р.В., Шарапова Е.В. Система проверки текстов на заимствование из других источников. Электронные библиотеки: Перспективные методы и технологии, электронные коллекции: труды 13-й всероссийской научной конференции. Воронеж. 2011. С. 121–126.

**СРАВНИТЕЛЬНЫЙ АНАЛИЗ СИСТЕМ
АВТОМАТИЧЕСКОЙ ПРОВЕРКИ ОРИГИНАЛЬНОСТИ ТЕКСТОВ**

В статье рассматривается вопрос неправомерного использования интеллектуальной собственности, а также исследуются существующие системы контроля для проверки подлинности содержания текста. В работе проведена проверка выбранных текстов с целью выявления уровня их оригинальности и определения преимуществ и недостатков представленных на рынке программного обеспечения программных комплексов автоматического анализа текстов.

Ключевые слова: оригинальность текста, онлайн-сервис, программный комплекс, антиплагиат.

**COMPARATIVE ANALYSIS OF SYSTEMS
FOR TEXTS ORIGINALITY AUTOMATED CHECKING**

The issue of intellectual property misuse is considered in the article. The existing control systems for verifying of the text content authenticity are also examined. In this paper, we checked the selected texts in order to identify the level of their originality and determine the advantages and disadvantages of the automated text analysis software packages on the software market.

Key words: text originality, online service, software package, anti-plagiarism.