

Ткаченко В.А.,
молодший науковий співробітник відділу мережних технологій і баз даних
Інституту інформаційних технологій і засобів навчання НАПН України

ПРОБЛЕМИ ЗБЕРЕЖЕННЯ УНІКАЛЬНОСТІ НАУКОВИХ ПУБЛІКАЦІЙ В ЕЛЕКТРОННИХ БІБЛІОТЕКАХ ТА ІНШИХ ЕЛЕКТРОННИХ РЕСУРСАХ

Повнотекстові електронні бібліотеки становлять важливу частину української та світової наукової інфраструктури. При експлуатації електронних бібліотек перед власниками виникають проблеми збереження унікальності електронного ресурсу і гарантії дотримання авторських прав. При трудомісткості створення якісного контенту і низької вартості відтворення цифрових копій зазначені проблеми є найбільш актуальними. Якщо зловмисник отримає необмежений доступ до вмісту електронної бібліотеки, він може скопіювати електронні документи та відтворити ресурс під іншим ім'ям для його використання з метою отримання прибутку в обхід інтересів правовласників. Таким чином, проблеми збереження унікальності електронного ресурсу і гарантування дотримання авторських прав безпосередньо пов'язані із завданням захисту від повного несанкціонованого копіювання цифрових документів електронної бібліотеки.

Законодавства розвинутих країн світу в тій чи іншій мірі захищають бази даних цифрових документів електронних бібліотек від повного несанкціонованого копіювання. Згідно з чинним законодавством України, база даних охороняється авторським правом, якщо вона є результатом творчої діяльності з добору та/або розташування включених в неї матеріалів.

Одночасно з цим база даних може охороняється суміжним правом, яке визнається за її виробником незалежно від наявності авторських прав на цю базу даних. При цьому виключне право поширюється на ті бази даних, створення яких вимагає істотних фінансових, матеріальних, організаційних чи інших витрат. У Європейському союзі існує спеціальний правовий режим *suī generis*, згідно з яким виробник бази даних має право забороняти витяг або повторне використання сукупності або істотної частини змісту бази даних. В США обмеження на копіювання значущої частини бази даних часто прописуються в контракті і захищаються в рамках контрактного права.

Можна відзначити факти судового переслідування організацій і приватних осіб, які в порушення користувальницького договору завантажують велику кількість документів електронної бібліотеки. Однак крім юридичного захисту, велике значення має використання технічних засобів захисту від повного несанкціонованого копіювання. Один з найбільш популярних методів – це обмеження завантаження користувачем кількості документів в певний період часу. Такі обмеження, як правило, прописуються також в ліцензійних договорах і правилах доступу до електронних бібліотек.

Інший спосіб – захист від автоматичного скачування електронних документів за допомогою різних реалізацій каптча. Каптча (CAPTCHA) – це англійська абревіатура, що означає «Повністю автоматизований публічний тест Тюрінга для розрізнення комп'ютерів і людей» та полягає в розпізнаванні символів, мови або рішення найпростіших математичних або логічних завдань. Такі тести може легко пройти людина, але для комп'ютера це викликає труднощі. Каптча застосовується зазвичай для того, щоб запобігти проникненню програм-ботів на інтернет-сервіси, але не гарантує захисту, тому що паралельно з розвитком засобів автоматичного розрізнення програм, не менш успішно розвиваються і технології їх автоматизованого подолання.

Зазначені методи захисту від повного несанкціонованого копіювання при відносній простоті реалізації мають істотні недоліки. Розробка та впровадження ефективних механізмів захисту від повного несанкціонованого копіювання дозволить зберегти

унікальність загальнодоступних електронних бібліотек і гарантувати захист авторських прав. Це дасть можливість одним електронним бібліотекам відкрити свій доступ більш широкому колу читачів, а іншим – більше ефективно стежити за дотриманням договорів з авторами і видавництвами.

З метою формування науково-методологічних основ утвердження академічної доброчесності та ефективних механізмів запобігання та протидії плагіату в наукових і навчальних роботах із використанням потенціалу судового та позасудового захисту авторських прав, методів і засобів судової експертизи в Україні, а також програмних продуктів, які дозволяють виявляти плагіат у наукових роботах, 30 вересня 2016 року у Міністерстві юстиції України відбувся круглий стіл на тему «Формування академічної доброчесності та механізмів запобігання та протидії плагіату у наукових текстах в Україні».

На цьому заході було обговорено такі питання:

- Шляхи і методи формування академічної доброчесності в сфері науки і освіти в Україні;
- Академічний плагіат: поняття, критерії, методи і суб'єкти виявлення;
- Практика судового розгляду спорів щодо недобросовісного запозичення у академічних творах фрагментів робіт інших авторів;
- Потенціал державних науково-дослідних установ судової експертизи Міністерства юстиції України і інших суб'єктів експертної діяльності щодо запобігання плагіату в наукових творах;
- Використання спеціальних знань при встановленні фактів запозичення праць інших авторів при створення наукових творів;
- Проблеми і перспективи розробки і акредитації програмних продуктів, які дозволяють виявляти плагіат у наукових роботах.

Надання в останній час більшої свободи академічним спільнотам передбачає збільшення їх відповідальності, у тому числі й щодо виявлення випадків плагіату – запозичення ідей, результатів наукових досліджень. У цьому контексті можна виокремити три аспекти проблеми: організаційно-технічний (створення репозитарію текстів, електронної системи для їх порівнянь), правовий, який вимагає уніфікованого підходу до тлумачення тих чи інших норм, аналіз змістового наповнення тексту (висновок про наявність чи відсутність плагіату обов'язково повинен бути верифікований експертом).

Термінологічний апарат, який сьогодні використовується під час досудового розслідування та судового розгляду для встановлення фактів відтворення праць інших авторів у наукових творах, трактує *плагіат* як оприлюднення результатів інших авторів під чужим іменем. У царині наукового плагіату чітко простежуються такі тенденції: заміна назв текстів і прізвищ авторів із абсолютно запозиченим текстом, збільшення кількості справ про порушення авторських прав, створення робіт на основі копії, заміна прізвищ російських авторів, сфер дослідження, назв галузей українськими, недобросовісне відтворення праць інших авторів. З огляду на те, що встановлення факту плагіату потребує багато трудовитрат і часу, лєвова частка якого витрачається на створення таблиць зіставлення текстів, вкрай важливо створити програмні засоби для полегшення проведення експертизи, які повинні пройти відповідну сертифікацію.

Однак, необхідно розмежовувати плагіат ідеї і плагіат форми. Розвиток ідеї може відбуватися різними шляхами, тоді форма вираження твору буде різною. Методи боротьби з плагіатом існують, однак необхідно усвідомлювати, що це довготривалий процес. У цьому контексті вагомим значення набуває широка соціальна ініціатива, оскільки не у всіх сферах (медицина, математика) юрист в змозі самотужки виявити плагіат. Аналізуючи методи виявлення плагіату, необхідно усвідомлювати, що факт його встановлення є виключно прерогативою суду. Важливо вибудувати механізм виявлення плагіату, який би чітко дозволив визначити у який спосіб відбулося використання одного

твору у іншому, якщо такий факт мав місце. Для оцінки та висновку про наявність чи відсутність плагіату чітко визначеного граничного значення не існує.

Плагіат все частіше зустрічається в журнальних публікаціях, в чужому тексті електронних видань або в запозиченні фрагментів авторського тексту без посилання на джерело запозичення.

В освіті та науці плагіат поділяється на *три типи* [1]:

- Текстовий плагіат – повне або часткове запозичення фрагментів тексту (не видозмінених або модифікованих), що присутній у статтях, тезах, звітах, монографіях, рукописах кваліфікаційних робіт, тощо.

- Плагіат програмних кодів – повне або часткове копіювання програмного коду, написаного іншою особою (наприклад, студентами) та поданого як власна розробка.

- Плагіат в нетекстових джерелах – копіювання даних у електронних таблицях, діаграмах, даних наукових експериментів, фото- та медіаресурсах.

Існує два способи виявлення наявності плагіату:

1. *Ручний пошук*, що здійснюється безпосередньо викладачами, науковцями, редакторами, читачами журналів. Наявність в редакційній колегії наукових часописів добросовісних, чесних, свідомих рецензентів, що є експертами в своїй галузі та ґрунтовно аналізують рукописи, може значно мінімізувати проблему, однак не усунути її цілком.

2. *Автоматичний пошук* за допомогою комп'ютерної техніки та програмних засобів.

Технічні засоби дійсно допомагають у виявленні плагіату

Антиплагіат – система автоматичної перевірки текстів на наявність запозичень із загальнодоступних джерел.

Серед науковців поширений плагіат, що можна поділити на такі *типи*:

1. *Запозичення* (плагіат) чужий наукової роботи або декількох робіт і опублікування результату під своїм ім'ям.

2. *Компілювання* – створення суміші власного і запозиченого тексту без належного цитування джерел.

3. *Рерайт* (перефразування чужої праці без згадування оригіналу автора). Цей тип плагіату, на сьогоднішній день, досить важко виявити технічними засобами пошуку плагіату. Оригінальність тексту – це поняття, протилежне плагіату. Чим більше плагіату в тексті, тим менше його оригінальність і навпаки – чим менше плагіату, тим вище рівень антиплагіату.

Для простоти можна вивести **формулу антиплагіату**:

Оригінальність% = 100% - % плагіат = Рівень антиплагіату%

Наприклад, якщо програма перевірки рівня плагіату (антиплагіат) (antiplagiat.ru, etxt.ru, advego.ru) показує, що оригінальність певного тексту становить 70%, це означає, що 30% цього тексту – не оригінальні, запозичені. Тобто з точки зору програми перевірки 30% в даному випадку – плагіат.

Можна з упевненістю відзначити, що з точки зору не машини, а людини це твердження не коректне. Наприклад, якщо процитовано фрагмент тексту і на нього виставлені коректні посилання на джерело – це не плагіат, а запозичення. Але сучасні програми з виявлення рівня плагіату в тексті не мають інтелект, і на даний момент не можуть відокремити цитування від плагіату, а отже, вони вважають коректно процитоване запозичення плагіатом. Навіть більше, найчастіше програми антиплагіат вважають за плагіат будь-які три слова йдуть один за одним у тому ж порядку. Саме тому будь-який навіть самий «авторський» текст навряд чи може мати оригінальність 100%.

Використання в Україні досвіду антиплагіатної ініціативи Дисергейт показав, що із 57 робіт, які пройшли перевірку, лише у 7 не виявлено плагіату (текстових запозичень без належних посилань). Озвучення цього факту є аспектом академічної доброчесності. Науковий плагіат в Україні – це ганебне явище, подолання якого уможливить процес розбудови академічної доброчесності. Право інтелектуальної власності потрібно поважати, як і будь-яке інше.

Науковцями Інституту інформаційних технологій і засобів навчання НАПН України виконується дослідження в межах науково-дослідної роботи «Система інформаційно-аналітичної підтримки педагогічних досліджень на основі електронних систем відкритого доступу» (2015-2017 рр., державний реєстраційний № 0115U002234). В результаті цього дослідження здійснено узагальнення зарубіжного і вітчизняного досвіду щодо електронних систем відкритого доступу для інформаційно-аналітичної підтримки проведення наукових досліджень, що дало змогу визначити основні з них, зокрема: наукові електронні бібліотеки та їх інформаційно-аналітичні сервіси, відкриті журнальні системи, системи антиплагіату тощо [2].

Проаналізувавши наукову літературу та джерела Інтернету, визначено, що нині у вітчизняній науковій літературі недостатньо публікацій, у яких розглянуто специфіку добору програмних засобів пошуку плагіату, потребує ґрунтовнішого дослідження аналіз існуючих та проектування і розробка нових програмних засобів. Також, важливими є підготовка рекомендації щодо використання автоматизованих програм для конкретних цілей: для редакцій наукових видань (перевірка рукописів статей), для спеціалізованих вчених рад (перевірка рукописів дисертацій), для наукових установ (перевірка рукописів

Остаточне рішення про наявність факту плагіату залишається на розсуд користувача, тому потреба в ручній перевірці та людському оцінюванні (рецензуванні) матеріалу все одно існує. Оскільки кожен такий ресурс має свої переваги і недоліки згідно програмно-технічних характеристик, не можна очікувати від машинної перевірки стовідсотково точних результатів. Основна ж мета використання даного класу програмних засобів – максимально загальмувати тенденцію до стрімкого поширення плагіату і вивести вітчизняну науку на новий якісний рівень.

Наразі, у мережі Інтернет існує багато різних спеціалізованих програм для перевірки унікальності текстів. В результаті дослідження було виділено низку популярних он-лайн ресурсів та програмних засобів, що можуть бути корисними як для окремих дослідників, так і для редакцій наукових видань.

Це **програмне забезпечення**: Anti-Plagiarism, eTXT Антиплагиат, Advego Plagiat, Double Content Finder, Praide Unique Content Analyser, Viper, Плагиата НЕТ, HPIDOL.

Он-лайн ресурси: DupliChecker, PaperRater, Plagiarisma.net, PlagiarismChecker, Plagium, PlagTracker, SeeSources, PlagScan, Plagiarism Detector, Защита уникальности контента, FindCopy (Miratools), Grammarly, Docoloc, Text.ru, Антиплагиат.

eTXT Антиплагиат – (<http://www.etxt.ru/antiplagiat/>) – сервіс перевірки текстів на унікальність, що доступний як програмне забезпечення для установки на персональний комп'ютер, так і у форматі он-лайн ресурсу. Можливо здійснювати перевірку текстових фрагментів, окремих файлів та пакетів файлів, а також сторінок сайтів. Передбачено дві функції перевірки тексту: на плагіат (наявність дослівних збігів) за допомогою опції «Метод виявлення копій», та на наявність рерайта (сміслових збігів) – опція «Метод виявлення рерайта». За результатами тестування цю програму рекомендовано для застосування в наукових установах НАПН України та редакціях фахових видань.

Серед інструментаріїв для пошуку текстового плагіату можна виділити *он-лайн ресурси FindCopy (MiraTools), Защита уникальности контента та Text.ru*, які теж за результатами експерименту рекомендовано для застосування в наукових установах НАПН України.

FindCopy (<http://www.miratools.ru/>) – проект холдингу Mirafox (Росія), що являє собою безкоштовний он-лайн ресурс для перевірки унікальності текстів.

Защита уникальности контента (<http://www.content-watch.ru/text>) – безкоштовний он-лайн сервіс, що дозволяє здійснювати перевірку унікальності текстів та контенту сайтів.

Text.ru (<http://text.ru/>) – безкоштовний онлайн-сервіс перевірки тексту на унікальність, що дозволяє розпізнавати плагіат, навіть якщо в тексті було здійснено

перестановку слів і фраз, змінено відмінки, часи та інші граматичні категорії слова та додані нові слова.

Висновок. Якщо Україна прагне відходити від постолігархічної системи, то вона мусить запроваджувати академічну доброчесність, як ознаку цивілізованої держави, для якої важливою є її репутація. Сьогодні простежується тенденція закладення нових цінностей у старі існуючі інститути. Публічний осуд – значно дієвіша й сильніша інституційна форма боротьби з плагіатом, аніж будь-які каральні форми чи суди. Кадрове очищення не може відбутися руками держави, а лише на рівні академічної громади навчального закладу чи установи. Особливої гостроти набувають проблеми формування академічної етики та культури. Боротися із проявами академічної недоброчесності можна моральними заходами і частково юридичними запобіжниками. Революційний шлях тут неможливий. Інколи доводиться вдаватися до крайніх форм боротьби з плагіатом, таких як замовлення на проведення судових експертиз задля встановлення фактів використання запозичених текстів, ідей, думок.

Починати формувати доброчесність необхідно з дитинства, школи, сім'ї. У програмах навчальних дисциплін потрібно ліквідувати репродуктивні завдання, які спонукають до проявів академічної недоброчесності. Одним із напрямів подолання академічної недоброчесності є «відміна ганебної практики присвоєння фаховості журналам».

В освітній-науковій сфері ми маємо наразі систему поглядів, запозичених без дозволу. Відповідальність – це сьогодні найголовніше. Цінність інтелектуальної роботи в тому, що у своїй праці науковець має бути першим.

Плагіат не може бути виправданим у жодному разі. Треба докласти максимум зусиль, щоб подолати плагіат. Він негативно впливає як на внутрішній розвиток системи освіти і науки в Україні, так і на її репутацію за кордоном. Є труднощі з отриманням бази для порівнянь через відсутність бази даних дисертацій і наукових статей. Саме тому в Україні виникла ідея створення «*Національного репозитарію українських текстів*», до якого б увійшли публікації науковців, а згодом й студентів. Безумовно він може спростити пошук схожих текстів, які підлягатимуть подальшій експертній оцінці. Репозитарій – необхідна, але не достатня умова боротьби з плагіатом, як однією з форм прояву академічної недоброчесності. Важливими також є зміна академічної культури, вдосконалення правової бази та наявність політичної волі.

Для використання у наукових установах і редакціях наукових видань (суспільні та гуманітарні науки) можливо використовувати такі автоматизовані програми виявлення унікальності тексту: eTXT Антиплагіат, FindCopy (MiraTools), Захист унікальності контенту, Text.ru (безкоштовні); Anti-Plagiarism, Docol©с (комерційні); Advego Plagiatu, Плагиата.НЕТ (довготривала перевірка). Ці системи перевірки наукових робіт на плагіат є важливим інструментом для інформаційно-аналітичної підтримки педагогічних досліджень Підсумовуючи викладене вище, можна додати, що використання програмних засобів пошуку текстового плагіату є актуальним і затребуваним для проведення результатів наукових досліджень, а тому потребує подальших ґрунтовних досліджень.

Список використаних джерел

1. Автоматизована система перевірки текстів на плагіат / В. Б. Мокін, В. В. Войтко, С. В. Бевз [та ін.] // Вісник Вінницького політехнічного інституту, 2010. – № 5. – С. 12-17.
2. Лупаренко Л. А. Інструментарій виявлення плагіату в наукових роботах: аналіз програмних рішень [Електронний ресурс] / Л. А. Лупаренко // Інформаційні технології і засоби навчання, 2014. – № 2 (40). – С. 151-169. – Режим доступу: <http://journal.iitta.gov.ua/index.php/itlt/article/view/1050/775>.
3. Михайловський Ю. Б. Система Anti-Plagiarism як інструмент запобігання плагіату в навчальній та науковій діяльності / Ю. Б. Михайловський, Н. А. Длугунович // Вісник Хмельницького національного університету. Технічні науки, 2013. – № 3. – С. 162-168.

4. Використання електронних систем відкритого доступу для інформаційно-аналітичної підтримки педагогічних досліджень [Електронний ресурс] / О. М. Спірін, А. В. Яцишин, С. М. Іванова та ін. // Інформаційні технології і засоби навчання, 2016. – №5 (55). – С. 136-174. – Режим доступу: <http://journal.iitta.gov.ua/index.php/itlt/article/view/1501/10>.