

# РОЗРОБКА КОМП'ЮТЕРНОЇ СИСТЕМИ ПЕРЕВІДКИ ТЕКСТУ НА ПЛАГІАТ

Вінницький національний технічний університет

## **Анотація**

*Сучасна система перевірки і порівняння тексту на плагіат представляє собою середовище Інтернет-сервісу в якому зібрані дані по яких перевіряється на унікальність текст. Системи буде реалізована за декількома методами, ідентифікація співпадання речень та підрахунок співпаливаючих слів. Після чого користувачеві буде показано факти наявності плагіату.*

**Ключові слова:** Інтернет-сервіс, плагіат,

## **Abstract**

*The modern system of checking and comparing text to plagiarism is an environment of the Internet service in which the collected data through which is checked for the uniqueness of the text. The systems will be implemented by several methods, the identification of the matching of sentences and the counting of the matching words. After that, the user will be shown the facts of plagiarism*

**Keywords:** Internet-service, plagiarism,

## **Вступ**

На сьогоднішній день в світовій павутині можна знайти безліч різноманітної інформації тому зараз є досить актуальною тема плагіату. Плагіат - це присвоєння чужого відкриття, винаходу та використання у своїх працях чужого твору без посилання на автора. Він є порушенням патентного законодавства та може призвести до юридичної відповідальності. Найрозповсюдженішою формою плагіату є повне або часткове копіювання тексту із одного або декількох джерел.

## **Результати дослідження**

Система виявлення плагіату може бути розроблена для: текстових даних таких як нариси, статі, дослідження і так далі. Коли користувач вводить текстовий документ, щоб перевірити його на оригінальність. Перевірку оригінальності виконує модуль перевірки, або через сервер куди буде імпортовано документ для перевірки, після чого буде показаний звіт користувачеві. Для обробки мов потрібне забезпечення надійного та конкретного механізму який здатний виявити плагіат у будь-якому документі. У цій роботі я намагаюсь розробити систему основу на принципах вилучення ключових слів та речень, яке виявляє плагіат.

Однак є аналоги досліджуваної системи, прикладом якої може служити сервіси: UNICHECK, ANTIPLAGIAT, CONTENT-WATCH.

UNPLAG — це сервіс з перевірки робіт на ознаки плагіату, який вчасно надає підтримку своїм користувачам та може адаптувати систему відповідно до потреб навчальних закладів.

Створений українськими розробниками у 2014, він може використовуватись онлайн або інтегруватись з навчальними системами ВНЗ (так звані learning management systems). Серед них – Moodle, Canvas, Blackboard, Schoology, Google Classroom та інші. Наразі сервісом користуються більш ніж 50 українських університетів. Навчальні заклади США, Іспанії, Бельгії та інших країн світу також перевіряють роботи з Unicheck.

ANTIPLAGIAT — ідея сервісу полягала спочатку в перевірці на плагіат курсових та дипломних робіт. Він дійсно корисний для студентів і аспірантів, але дуже слабенький при перевірці звичайних текстів. У загальному і цілому, можете користуватися для будь-якої початкової перевірки на унікальність, але до результатів варто ставитися з підозрою. Краще підкріпити їх результатами в інших системах. Без реєстрації можна перевіряти за раз до 5000 символів, що цілком прийнятно і цього навіть майже завжди вистачить, але можна і зареєструватися.

CONTENT-WATCH — це один онлайн сервіс. При перевірці на унікальність він використовує власний алгоритм пошуку в Інтернеті сайти можуть містити матеріали повні або часткові копії заданого тексту. Реєструватися для перевірки на сайті не потрібно. На основі запропонованих варіантів підраховується загальна унікальність тексту у відсотках, а також унікальність щодо кожної знайденої сторінки з збігами. Є можливість подивитися, які частини тексту були знайдені на кожній з проаналізованих сторінок. При цьому існує відмінна можливість ігнорувати певний сайт. До обмежень відноситься довжина тексту до 3000 символів (розширюється до 10.000 символів після реєстрації); до 5 запитів в день на одного користувача (20 після реєстрації).

Існують дуже багато систем аналогів подібного типу однак вони відрізняються кількістю перевірки слів та кількістю запитів на добу також безплатними і платними послугами того чи іншого сервісу.

### Висновки

Отже система перевірки тексту на плагіат – це взаємопов'язана сукупність методів і алгоритмів для обробки інформації. Основною задачею якої є порівняння оригінальності тексту. Також у роботі буде розглянуто поняття плагіату, його поширення та класифікації. Виконано огляд існуючих систем виявлення плагіату, виділено їх функціональні характеристики, визначені формати вхідних і вихідних даних.

### СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Поповський О. І. Огляд програм порівняльного аналізу на збіг.
2. Шарапова Е.В., Шарапов Р.В. Универсальная система проверки текстов на плагиат
3. <https://www.sciencedirect.com/science/article/pii/S1877042812052986>
4. <http://library.kubg.edu.ua/poslygy-ta-servis/dostup-do-internet.html>
5. ANTIPLAGIAT <http://www.antiplagiat.ru/>
6. CONTENT-WATCH <http://www.content-watch.ru/text/>

**Бондар Роман Володимирович** — студент групи КІВ-16б, факультет комп'ютерних систем і автоматики, Вінницький національний технічний університет, Вінниця, e-mail: [bondarroman1999@gmail.com](mailto:bondarroman1999@gmail.com)

**Науковий керівник: Ковалюк Олег Олександрович** — доцент кафедри КСУ, Вінницький національний технічний університет, м. Вінниця, e-mail: [ksu.kovalyuk@gmail.com](mailto:ksu.kovalyuk@gmail.com)

**Roman Bondar** - student of the KIV-16b group, Faculty of Computer Systems and Automatics, Vinnytsia National Technical University, Vinnytsia, e-mail: [bondarroman1999@gmail.com](mailto:bondarroman1999@gmail.com)

**Scientific supervisor: Kovaluk Oleg Oleksandrovich** - Associate Professor of the Department of KSU, Vinnytsia National Technical University, Vinnytsya, e-mail: [ksu.kovalyuk@gmail.com](mailto:ksu.kovalyuk@gmail.com)