**ПРОБЛЕМИ ПЕРЕВІРКИ ДОСТОВІРНОСТІ ДАНИХ**

**DATA VERIFICATION PROBLEMS**

Науковий керівник: к.т.н., доцент кафедра інформаційних технологій

Рудніченко Микола Дмитрович

Здобувач бакалаврату Предеін Нестор Дмитрович

Supervisor: (Ph.D), assistant professor Department of Information Technology

Rudnichenko Mykola Dmytrovych

Bachelor's candidate Predein Nestor Dmitrovich

**Анотація:** Мета роботи полягає в оцінці достовірності даних текстового контенту на базі автоматизації процесу перевірки його унікальності із застосуванням пошукових систем. Призначення програмного забезпечення полягає у перевірці текстової інформації на унікальність з метою запобігання плагіату текстового контенту. Мовою розробки є Java, для написання коду і компіляції застосування було використано середовище розробки NetBeans. Інтерфейс користувача створений за допомогою використання технології Swing.

**Annotation:** The aim of the work is to assess the reliability of text content data based on the automation of the process of verifying its uniqueness using search engines. The purpose of the software is to test textual information for uniqueness in order to prevent plagiarism of textual content. The development language is Java, and the NetBeans development environment was used to write the code and compile the application. The user interface is created using Swing technology.

**Ключові слова:** інформаційна унікальність даних, достовірність даних, системи перевірки плагіату.

**Key words:** information uniqueness of data, data reliability, plagiarism verification systems.

Сьогодні, як ніколи, достовірна інформація – це актуальна проблема кожної людини. В епоху диджиталізації та засилля простих рішень, що пропонуються інтернет-лобістами, створюються гіпер-умови, в яких користувачам прищеплюється стереотип, що авторитетним особам, які написали низку книг, можна вірити [1].

Інформаційна революція стала одним із ключових факторів утворення глобального інформаційного суспільства. Нові знання, вміння та технології відкривають численні можливості для забезпечення різноманітних потреб індивідів, а також їхнього захисту від внутрішніх та зовнішніх загроз. Проте перехід до нової моделі соціальних відносин характеризується і низкою негативних явищ. Наприклад, комп'ютерні технології та Інтернет надали фактично безмежні можливості для вільного створення, розповсюдження та зберігання інформації. Одночасно це призвело до вдосконалення форм та методів маніпулювання даними, використання інформації з метою неправомірного впливу на свідомість та поведінку не лише окремих громадян або їх груп, а й соціуму загалом [2].

Незважаючи на очевидні переваги Інтернету як джерела оперативно значущої інформації, необхідно звернути увагу на проблемні питання, що виникають внаслідок його використання. Серед них слід виділити: відсутність систематизації даних (взаємозв'язок між ними який завжди очевидна внаслідок несистемного характеру інформації); швидкість зміни відомостей – для видалення інформації потрібно дуже мало часу – від блокування контенту до його повного видалення; достовірність розміщеної інформації не завжди є абсолютною і вимагає додаткової перевірки та зіставлення [3].

Останнім часом значно поширилося таке явище, як фейкова інформація, яка розуміється як новини, сторінки в соціальних мережах, блогах, на форумах, що імітують рейтингові сайти, за допомогою яких певні групи людей або окремі особи привертають увагу суспільства до недостовірних подій некоректним. чи неприпустимим шляхом. Подібна інформація (зокрема, «фейкові» новини) розповсюджується з великою швидкістю, поповнюючись новими подробицями, які є реакцією користувачів соціальних мереж [4].

Слід зазначити, що достовірність інформації визначається трьома факторами: повнотою; цілісністю; істинністю.

Таким чином, при навчанні методики організації процесу пошуку оперативно значущої інформації слід приділяти увагу всім трьома компонентами, що визначають якість отриманих даних. Питання повноти інформації можна вирішити, зокрема, з допомогою грамотного побудови пошукового запиту. Цілісність інформації в електронному середовищі розглядається з позицій науки інформатики та має на увазі її стійкість щодо спотворюючих впливів різного характеру [5]. У разі наявності в оперативного співробітника сумніву в цілісності інформації, розміщеної на певному ресурсі, доцільно звернутися за допомогою спеціаліста у цій сфері знань для вироблення подальших дій. Однак зазначена проблема виходить за рамки предмета даного дослідження і потребує подальшої, більш детальної теоретичної розробки.

Найбільш складним і важливим є вміння визначати, чи отримана інформація є істинною. Від цього зрештою залежатиме достовірність отриманої інформації. Визначення обґрунтованості та точності інформації є досить складним завданням для користувача [6].

Контент-аналіз достовірності даиих складається з трьох основних етапів:

* визначаються одиниці аналізу та їх машиночитабельний вид;
* проводиться підрахунок частоти згадок досліджуваних змінних, виявляються взаємозв'язки між ними, розраховуються відносні індекси, наприклад індекс інформаційної присутності або прес-індекс, індекс емоційної спрямованості, індекс активності та інші;
* інтерпретація отриманих результатів, опис позитивних та негативних тенденцій, прогнозування.

При веденні звітності у програмах, як придбаних, так і створених або доопрацьованих самостійно, ступінь достовірності даних може знижуватися через:

* несанкціонованого доступу до даних та їх зміни;
* втрати цілісності інформації;
* відмовитися від обліку у програмному продукті останніх змін законодавства чи своєчасних оновлень;
* помилок користувачів.

Для контролю ступеня точності інформації та проведення аудиту достовірності даних на практиці необхідно:

* перевірити наявність правил розмежування доступу користувачів до файлів системи;
* ознайомитись із політиками безпеки;
* самостійно отримати доступ до необхідної програми;
* залучити технічного спеціаліста.

Висновки. Проведений аналіз досліджуваної тематики дозволив встановити актуальність та затребуваність на сучасному ринку систем перевірки достовірності та унікальності текстового контенту, через що подальшим розвитком роботи може стати розробка повноцінного програмного застосування для автоматизацій таких процесів.

**Література:**

1. 10 способів перевірити текст на оригінальність [Електроній ресурс]. – Режим доступу: https://raskrutka.com.ua/blog/10-sposobov-proverit-tekst-na-unikalnost/. - Дата доступу: 16.04.2022.

2. Що таке унікальність тексту? [Електронний ресурс]. - Режим доступу: https://wiki.rookee.ru/unikalnost/. - Дата доступу: 16.04.2022.

3. Ушакін С. Плагіат? Про етику в науці / С. Ушакін. - М.: Суспільні науки і сучасність, 2001. - 191 с.

4. Рудниченко Н. Д. Специфика разработки системы проверки уникальности текста / Н. Д. Рудниченко, В. Н. Овчинников // Інформатика, інформаційні системи та технології: тези доповідей п’ятнадцятої всеукраїнської конференції студентів і молодих науковців. Одеса, 27 квітня 2018р. - Одеса, 2018. – С. 112.

5. Рудниченко Н. Д. Концепция микросервисной модели поиска и извлечения знаний из медийного контента веб-ресурсов / Н. Д. Рудниченко, Н. О. Шибаева, Д. С. Шибаев //
21-й міжнародний молодіжний форум «Радіоелектроніка та молодь у ХХI столітті». Зб. Матеріалів форуму. – Харьків: Хнуре. – 2018. – Т.6. – С.61-62.

6. Щербина А. Д. Порівняльний аналіз існуючих напрямів у інтелектуальному аналізі даних / А. Д. Щербина, Д. С. Шибаєв, М. Д. Рудніченко, Н. О. Шибаєва // Project, Program, Portfolio Management The Third International Scientific-practical Conference, Odesa, ONPU.
07–08 Dec 2018. – С.88-90