

УДОСКОНАЛЕННЯ ІНФОРМАЦІЙНОЇ СИСТЕМИ ІНТЕЛЕКТУАЛЬНОГО СОРТУВАННЯ ДАНИХ ДЛЯ ІНТЕГРАЦІЇ В «GOOGLE DRIVE API» НА ОСНОВІ КОНТЕНТУ

D.A. Boiko

IMPROVEMENT OF THE INFORMATION SYSTEM OF INTELLIGENT DATA SORTING FOR INTEGRATION INTO “GOOGLE DRIVE API” BASED ON CONTENT

Інформаційні системи інтелектуального сортування даних «на основі контенту» передбачають аналіз змісту файлів та їхніх атрибутів (текстового наповнення, метаданих, семантики, типу та структури), а не лише за допомогою стандартних правил і назв. Такий підхід дає змогу автоматично визначати призначення або категорію документа, розпізнавати взаємозв'язки між файлами та оптимізувати процес їх класифікації. У контексті інтеграції з Google Drive API, використання аналізу контенту дозволяє підвищити точність сортування, пришвидшити доступ до релевантних даних та покращити користувацьку взаємодію, що відповідає актуальним задачам сучасних технологій щодо автоматизації обробки інформації та впровадження інтелектуальних систем управління даними. [1].

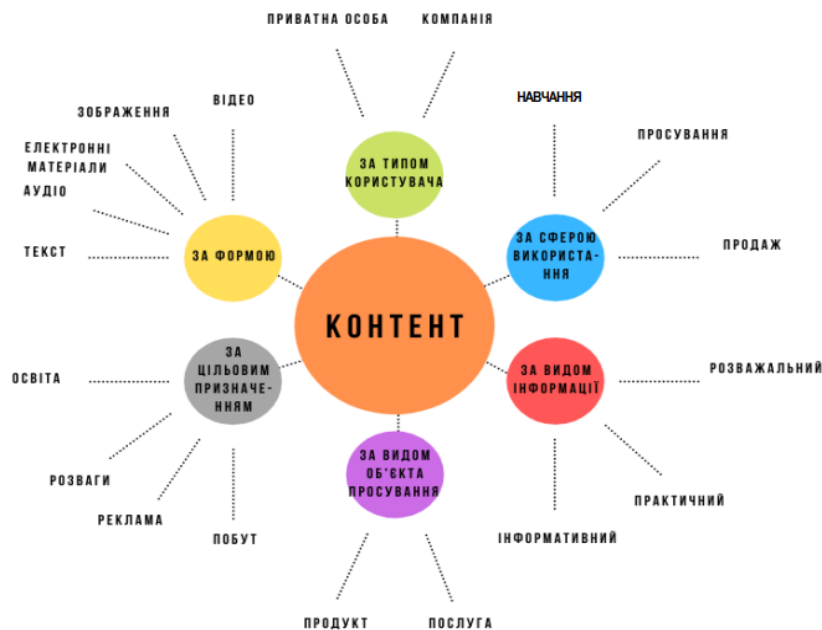


Рисунок 1. Структура класифікації контенту за різними ознаками

Однією з ключових актуальних технологій у процесі цифрової трансформації є використання прикладних програмних інтерфейсів (API), що забезпечують взаємодію між системами та доступ до хмарних сервісів. У контексті інтеграції інформаційної системи інтелектуального сортування даних із Google Drive API, API дозволяє автоматично отримувати метадані файлів, здійснювати доступ до вмісту, виконувати операції класифікації та переміщення без участі користувача. Завдяки підтримці сучасних стандартів безпеки (OAuth 2.0), API дає змогу реалізувати персоналізовану обробку даних, працювати з великими обсягами інформації у реальному часі та масштабувати алгоритми машинного навчання на хмарній інфраструктурі. Такий підхід сприяє підвищенню ефективності цифрових рішень, що відповідає тенденціям розвитку інтелектуальних веб-технологій і систем управління даними [2].

Для підвищення точності контент-аналізу в системі інтелектуального сортування даних можуть використовуватися гібридні моделі машинного навчання, що поєднують статистичні та математичні методи [3]. На етапі первинної класифікації файлів доцільним є застосування алгоритмів векторного представлення та порівняння, зокрема k-means та методу головних компонент (PCA) для виявлення структурної подібності на основі метаданих (розмір, формат, дата створення, частота появи ключових параметрів). Для оптимізації рішень щодо належності до певної категорії може бути використана логістична регресія або SVM, які дозволяють визначити ймовірність віднесення об'єкта до класу на підставі обраних ознак.

Реалізацію моделей можна виконати як у середовищі Python з використанням бібліотек машинного навчання, так і через інші технології (PHP, JavaScript) у разі інтеграції з відповідними сервісами. Такий підхід забезпечує адаптивність системи, можливість подальшого донавчання та масштабування моделі разом зі зростанням обсягів контенту [4].

У порівнянні з підходами, орієнтованими виключно на пошук за ключовими словами або метаданими, запропонована модель забезпечує багатовимірний аналіз, включаючи статистичні, структурні та змістові ознаки. Крім того, інтелектуальне сортування дозволяє автоматично оновлювати категорії на основі нового користувацького контенту, тоді як класичні системи потребують ручного коригування. Це значно підвищує швидкість обробки даних, знижує ризик помилки та робить систему масштабованою в умовах зростання обсягів інформації [5].

Подальші дослідження можуть зосередитися на інтеграції додаткових методів аналізу контенту, таких як глибинне навчання для розпізнавання семантичних зв'язків та контекстуальних залежностей між файлами, а також на впровадженні адаптивних механізмів рекомендацій для користувачів. Ці напрямки відкривають можливості для підвищення ефективності класифікації, забезпечення більш інтелектуальної взаємодії із даними та розвитку масштабованих систем управління інформацією у хмарних середовищах. Крім того, варто звернути увагу на оптимізацію продуктивності систем при обробці великих обсягів даних, безпеку при взаємодії з хмарними сервісами, управління версіями та історією змін файлів, а також на можливість інтеграції з різноманітними джерелами даних.

Література

1. Структури даних та алгоритми URL: <https://studfile.net/preview/10025806/page:21/>. Доступ до ресурсу: 30.11.2025
2. Brenda Jin. Designing Web APIs / Brenda Jin, Saurabh Sahni, Amir Shevat.
3. Григорова Т.А. Дослідження методів машинного навчання для пошуку інформації / Т.А. Григорова, В.П. Ляшенко, О.О. Москаленко (Статті. Кременчуцький національний університет ім. М. Остроградського)
4. Карабін Р.М., Литвиненко Я.В. Огляд бібліотек машинного навчання для мови python.
5. Google Drive API Overview URL: <https://developers.google.com/workspace/drive/api/guides/about-sdk> (дата звернення: 30.11.2025)