

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
“КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ
ІМЕНІ ІГОРЯ СІКОРСЬКОГО”

Факультет прикладної математики
Кафедра програмного забезпечення комп'ютерних систем

КУРСОВИЙ ПРОЕКТ
ТЕХНІЧНЕ ЗАВДАННЯ
з дисципліни “Бази даних”

спеціальність 121 – Програмна інженерія

на тему: **Моніторингова система новинних повідомлень в Інтернеті**

Студент

групи КП-73

Защик Михайл Олександрович

(підпис)

Викладач

к.т.н, доцент кафедри

СПіСКС

Петрашенко А.В.

(підпис)

Київ – 2020

ЗМІСТ

1. НАЙМЕНУВАННЯ ТА ГАЛУЗЬ ЗАСТОСУВАННЯ РОЗРОБКИ	3
2. ДАТА ПОЧАТКУ ТА ЗАКІНЧЕННЯ ПРОЕКТУ	3
3. МЕТА РОЗРОБКИ	3
4. ВИМОГИ ДО ПРОГРАМНОГО ЗАБЕЗПЕЧЕННЯ	3
5. ОБҐРУНТУВАННЯ ВИБОРУ СУБД	4
6. ВИМОГИ ДО ІНТЕРФЕЙСУ КОРИСТУВАЧА	4
7. ВИБІР ЗАСОБІВ РОЗРОБКИ	5
8. ЕТАПИ РОЗРОБКИ	5

1. НАЙМЕНУВАННЯ ТА ГАЛУЗЬ ЗАСТОСУВАННЯ РОЗРОБКИ

Найменування: Моніторингова система новинних повідомлень в Інтернеті.

Галузь застосування: покращення таргетингу новин; аналіз даних з відповідним висновком відносно ефективності розподілення новин.

2. ДАТА ПОЧАТКУ ТА ЗАКІНЧЕННЯ ПРОЕКТУ

Дата початку проекту: 10.03.2020

Дата закінчення проекту: 21.05.2020

3. МЕТА РОЗРОБКИ

Метою розробки є створення програмного забезпечення, яке буде взаємодіяти з постреляційною базою даних, проводити заміри досліджуваного об'єкта і подальший аналіз, оцінку, порівняння отриманих результатів для виявлення певних закономірностей, тенденцій, змінних і їх динаміки. А також здобуття навичок використання СУБД та інструментальними засобами підтримки розробки додатків для подібних баз даних, оформлення відповідного текстового, програмного та ілюстративного матеріалу у формі проектної документації.

4. ВИМОГИ ДО ПРОГРАМНОГО ЗАБЕЗПЕЧЕННЯ

- Вимоги до генерації даних:

Має бути реалізовано ПЗ для збору новин в Україні, що розміщені на відповідних веб-сайтах: [/www.ukr.net/ua](http://www.ukr.net/ua), ua.korrespondent.net, rivne1.tv.

- Вимоги до фільтрації та валідації даних:

Додаткове програмне забезпечення з функціоналом фільтрації та перевірки даних на коректність, відсіювання дублікацій.

- Вимоги до засобів реплікації та масштабування:

Розроблений програмний комплекс має використовувати СУБД MongoDB, тому засіб реплікації використовуватимемо Replica Sets.

- Вимоги до аналізу даних:

Отримані дані повинні бути проаналізовані наступним чином:

1. Структурувати усі отримані дані з різних веб-ресурсів, зокрема джерела, теми та назви новин.
2. Здійснити валідацію даних.
3. Провести кореляцію по підготовлених даних.
4. Здійснити аналіз залежності популярності новини від категорії та ресурсу публікації та ключових слів (заголовки, джерела, текст статті).

- Вимоги до резервування та відновлення даних:

Резервування даних здійснювати шляхом використання реплікації. У разі втрачання зв'язку з певним вузлом, при підключенні вже нового вузла дані мають бути відновлені в ньому.

5. ОБҐРУНТУВАННЯ ВИБОРУ СУБД

Для поставленої задачі було обрано СУБД MongoDB — документо-орієнтована система керування базами даних з відкритим вихідним кодом, яка не потребує опису схеми таблиць. MongoDB підтримує зберігання документів в JSON-подібному форматі, має досить гнучку мову для формування запитів, може створювати індекси для різних збережених атрибутів, ефективно забезпечує зберігання великих бінарних об'єктів, підтримує журналювання операцій зі зміни і додавання даних в БД, підтримує реплікацію. У MongoDB є вбудовані засоби із забезпечення шардінгу, комбінуючи який з реплікацією даних можна побудувати горизонтально масштабований кластер зберігання, в якому відсутня єдина точка відмови, підтримується автоматичне відновлення після збою і перенесення навантаження з вузла, який вийшов з ладу. Розширення кластера або

перетворення одного сервера на кластер проводиться без зупинки роботи БД простим додаванням нових машин.

Також причиною вибору саме цієї бази даних є те, що відповідна БД використовувалась в минулому і тому є доволі знайомою та зрозумілою.

6. ВИМОГИ ДО ІНТЕРФЕЙСУ КОРИСТУВАЧА

Для спрощення роботи з проектом, інтерфейс користувача було виконано в консольному варіанті з виводом необхідної інформації в тому випадку, якщо у користувача виникла потреба у її перегляді. Також інтерфейс буде використано для запуску на виконання ПЗ, його налаштування та передачі параметрів для збору та аналізу даних.

7. ВИБІР ЗАСОБІВ РОЗРОБКИ

У якості засобів розробки було обрано мову C# та середовище розробки Visual Studio 2019. Для роботи з СУБД використовується відповідний драйвер MongoDB.Driver.

8. ЕТАПИ РОЗРОБКИ

№	Назва етапів розроблення	Термін виконання
1	Затвердження теми курсової роботи. Опрацювання відповідної літератури. Розроблення та узгодження технічного завдання.	12.03.2020
2	Аналіз постановки задачі	20.03.2020
3	Розробка засобів генерації даних.	01.04.2020
4	Додавання засобів фільтрації та валідації даних.	15.04.2020
5	Реалізація зберігання, реплікації та масштабування інформації розробленої моніторингової системи.	25.04.2020

6	Додавання засобів аналізу даних.	30.05.2020
7	Додавання засобів резервування та відновлення даних (з урахуванням необхідності підключення додаткового комп'ютера як елемента горизонтального масштабування).	10.05.2020
8	Тестування програми	15.05.2020
9	Аналіз результатів. Підготовка матеріалів курсового проекту та оформлення пояснювальної записки	19.05.2020
10	Захист курсової роботи	21.05.2020