

МАГІСТЕРСЬКА КВАЛІФІКАЦІЙНА РОБОТА НА ТЕМУ:

# «Інформаційна веб-технологія для створення робочих зв'язків»

*Виконав: магістрант групи ІКН-17м*

*Білий В.А.*

*Керівник: к.т.н., доц. Арсенюк І.Р.*

# МЕТА ТА ЗАВДАННЯ ДОСЛІДЖЕННЯ

Актуальність Необхідність створення і розвитку високоефективної мережі робочих зв'язків з можливістю цілодобового доступу до даних.

Метою дослідження магістерської кваліфікаційної роботи є підвищення швидкодії та функціональності веб-орієнтованих засобів створення робочих зв'язків шляхом використання інтелектуальних технологій.

Для досягнення поставленої мети необхідно розв'язати такі завдання:

- провести аналіз проблеми створення робочих зв'язків в мережі Інтернет;
- розглянути існуючі методи вирішення задачі створення робочих зв'язків в мережі Інтернет та обрати й обґрунтувати вибір методу, який задовольняє мету даної магістерської кваліфікаційної роботи;
- розробити метод кластеризації даних користувачів при створенні робочих зв'язків в мережі Інтернет;
- сформулювати стадії інформаційної технології, розробити структуру та алгоритм роботи програмного засобу;
- виконати програмну реалізацію запропонованої інформаційної веб-технології;
- провести тестування програмного продукту та виконати аналіз отриманих результатів.

# ОБ'ЄКТ, ПРЕДМЕТ ТА МЕТОДИ ДОСЛІДЖЕННЯ

Об'єкт дослідження – процес організації веб-сервісу для створення робочих зв'язків.

Предмет дослідження – інформаційна веб-технологія та програмні засоби створення робочих зв'язків з використанням інтелектуальних технологій, їх швидкодія та функціональність.

## Методи дослідження

У роботі використані наступні методи наукових досліджень:

- системного аналізу,
- теорії інтелектуального аналізу даних для реалізації інформаційної технології створення робочих зв'язків,
- методи математичної статистики,
- об'єктно-орієнтованого програмування.

# НАУКОВА НОВИЗНА ОДЕРЖАНИХ РЕЗУЛЬТАТІВ

4

полягає в наступному:

1. Набула подальшого розвитку інформаційна технологія створення робочих зв'язків, яка відрізняється використанням інтелектуального аналізу даних, що дозволило підвищити швидкодію та функціональність веб-орієнтованих засобів створення робочих зв'язків.
2. Удосконалено метод к-середніх для кластеризації даних, який відрізняється способом початкового вибору центрів кластерів, що пришвидшило сходження алгоритму і збільшило швидкодію пошуку потрібних записів.

## ПРАКТИЧНЕ ЗНАЧЕННЯ ОДЕРЖАНИХ РЕЗУЛЬТАТІВ

- розроблено структуру сайту створення робочих зв'язків, структуру бази даних та діаграму сутність-зв'язок, діаграму класів;
- розроблено алгоритм роботи програмного забезпечення створення робочих зв'язків на основі інтелектуального аналізу даних;
- розроблено програмні засоби для створення робочих зв'язків на основі інтелектуального аналізу даних.

# АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ СТВОРЕННЯ РОБОЧИХ ЗВ'ЯЗКІВ У МЕРЕЖІ ІНТЕРНЕТ

5

У розроблюваній інформаційній веб-технології для створення робочих зв'язків необхідно виконувати таку функцію як пошук колег або потенційних працівників у великій базі даних за певними критеріями їх професійної діяльності. Для швидкого та ефективного виконання цієї функції потрібно застосувати методи інтелектуального аналізу даних (ІАД).

Задачі ІАД: класифікація,  
**кластеризація**,  
прогнозування,  
асоціація,  
візуалізація,  
аналіз виявлення відхилень,  
оцінювання,  
підведення підсумків

Було обрано кластеризацію

Методи кластеризації:

## 1 Графові методи кластеризації

- Алгоритм виділення зв'язних компонент
- Алгоритм ФОРЕЛ
- Функціонали якості кластеризації

## 2 Ієрархічна кластеризація (таксономія)

- Агломеративна ієрархічна кластеризація
- Дендрограма і властивість монотонності
- Властивість стиснення, розтягування і редукованості

## 3 Статистичні методи кластеризації

- EM-алгоритм
- Метод **k-середніх**

Було обрано метод k-середніх



# ВИБІР І ОБҐРУНТУВАННЯ АНАЛОГУ

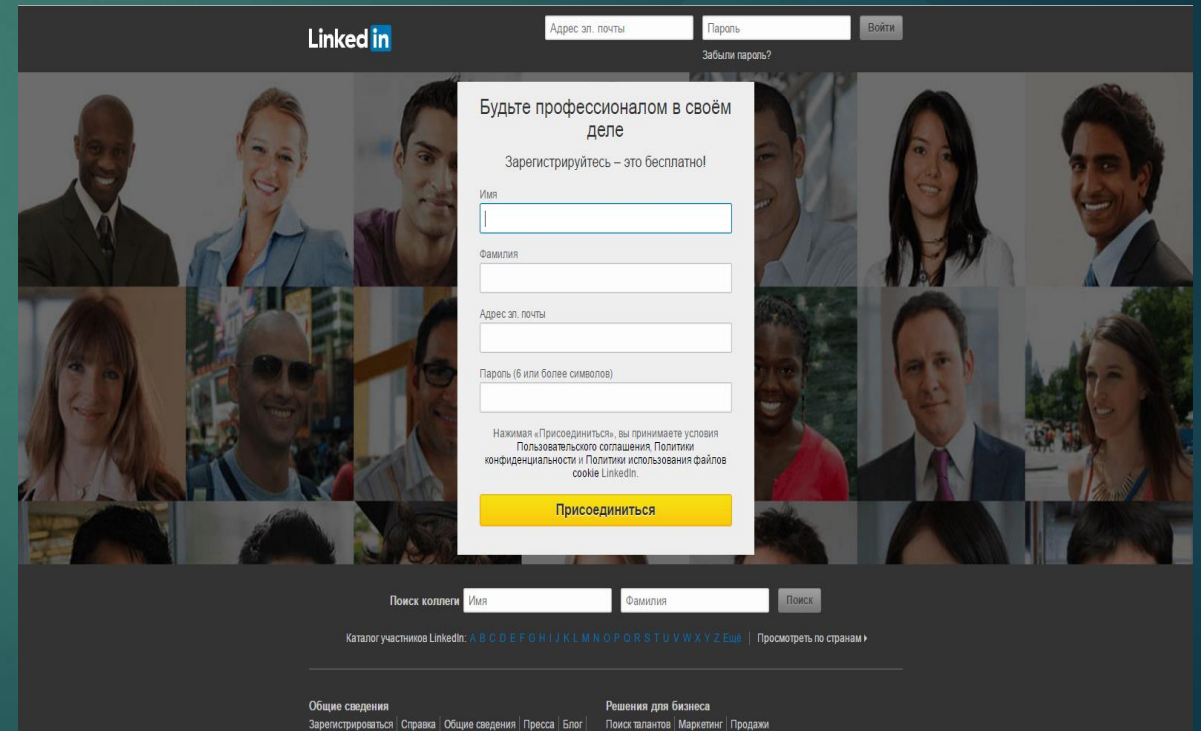
6

Таблиця 1.1 – Порівняння розроблювального сайту з аналогами

	«LinkedIn»	«XING»	«BranchOut»	Розроблювальний сайт
Інтерактивна взаємодія користувачів	1	1	1	1
Форум	0	1	0	1
Налаштування оповіщень	1	0	1	1
Блог	1	0	0	1
Використання ІАД	0	0	0	1
Сумарний коефіцієнт	3	2	2	5

Примітка: 0 – відсутність, 1 – наявність.

Аналог – веб-сервіс  
LinkedIn



# Розробка методу кластерного аналізу даних користувачів

7

У нас є дані про відвідувачів сайту для створення робочих зв'язків. Ці дані вони внесли під час реєстрації. До цих даних відносяться: професійні вміння та навички, досвід, якими технологіями програмування володіє і так далі. Всі ці параметри кодуються певним чином і отримується вектор, що представляє відвідувача. Приклад фрагменту закодованого запису користувача представлено у табл. :

Таблиця 2.1 – Фрагмент закодованого запису користувача

Прізвище		Стаж роботи, років	PHP	Javascr ipt	MySQL L	Git	Linux		CSS
Іванюк	...	5	1	1	1	0	1	...	1

Модифікований метод к-середніх Мак-Кіна (відрізняється тим, що центри кластерів перераховуються кожного разу після зарахування точки до найближчого кластера та способом вибору центрів кластерів):

1. Задається число  $K$  – кількість кластерів.
2. Серед усієї множини об'єктів обирається  $K$  точок таким способом:  $K$  типових об'єктів. Кожна з цих  $K$  точок вважається центром окремого кластера.
- 3 Далі кожен об'єкт зараховуємо до кластера, центр якого найближчий.

Тобто для кожної точки  $X_i, i = \overline{1, n}$  обчислюємо відстані

$$d(X_i, M_j), j = \overline{1, K} \quad (1)$$

та обираємо номер того кластера, де буде досягтися мінімум.

- 4 Після того, як об'єкт розподілено, перераховуються центри кластерів

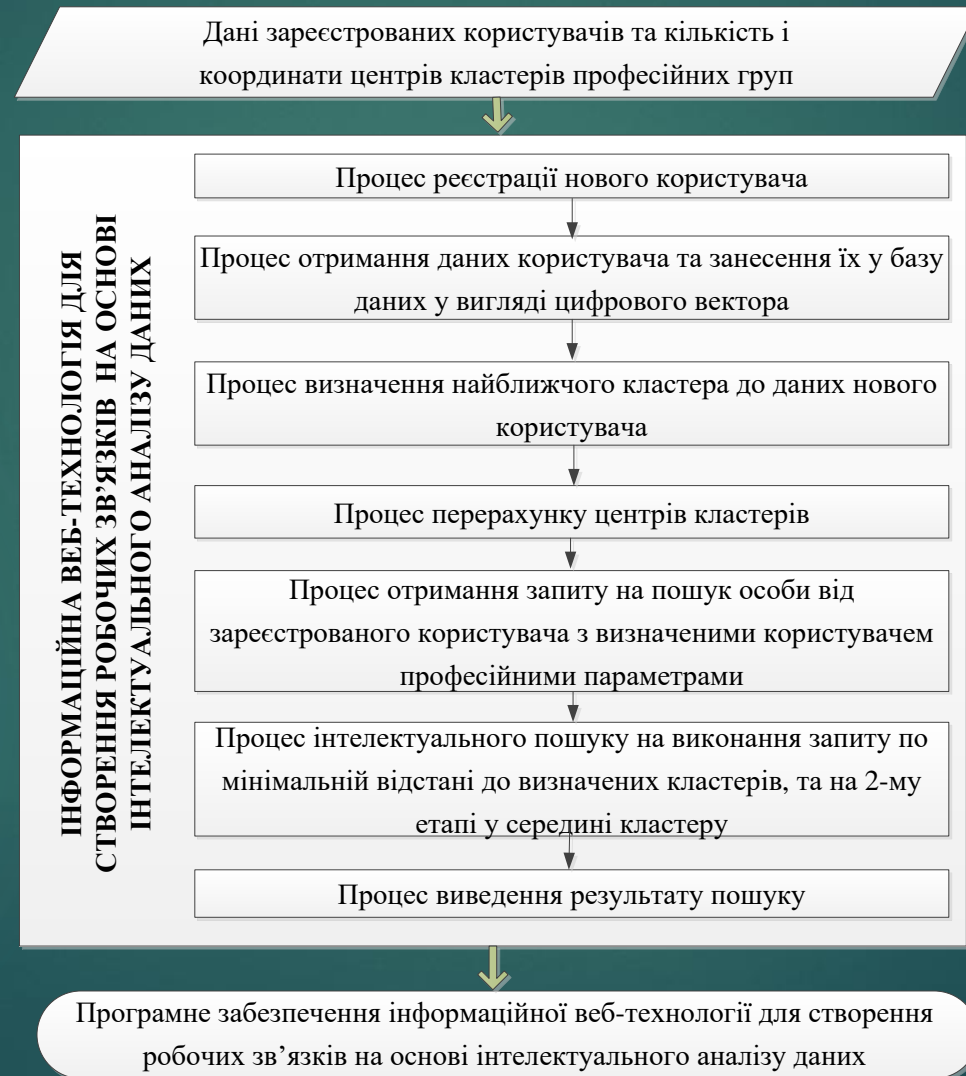
$$M_i = (\bar{x}_{i1}, \bar{x}_{i2}, \dots, \bar{x}_{ip}), i = \overline{1, K}, \quad (2)$$

$$\bar{x}_{ij} = \frac{1}{n_i} \sum_{k=1}^{n_i} x_{kj}, j = \overline{1, p}.$$

де

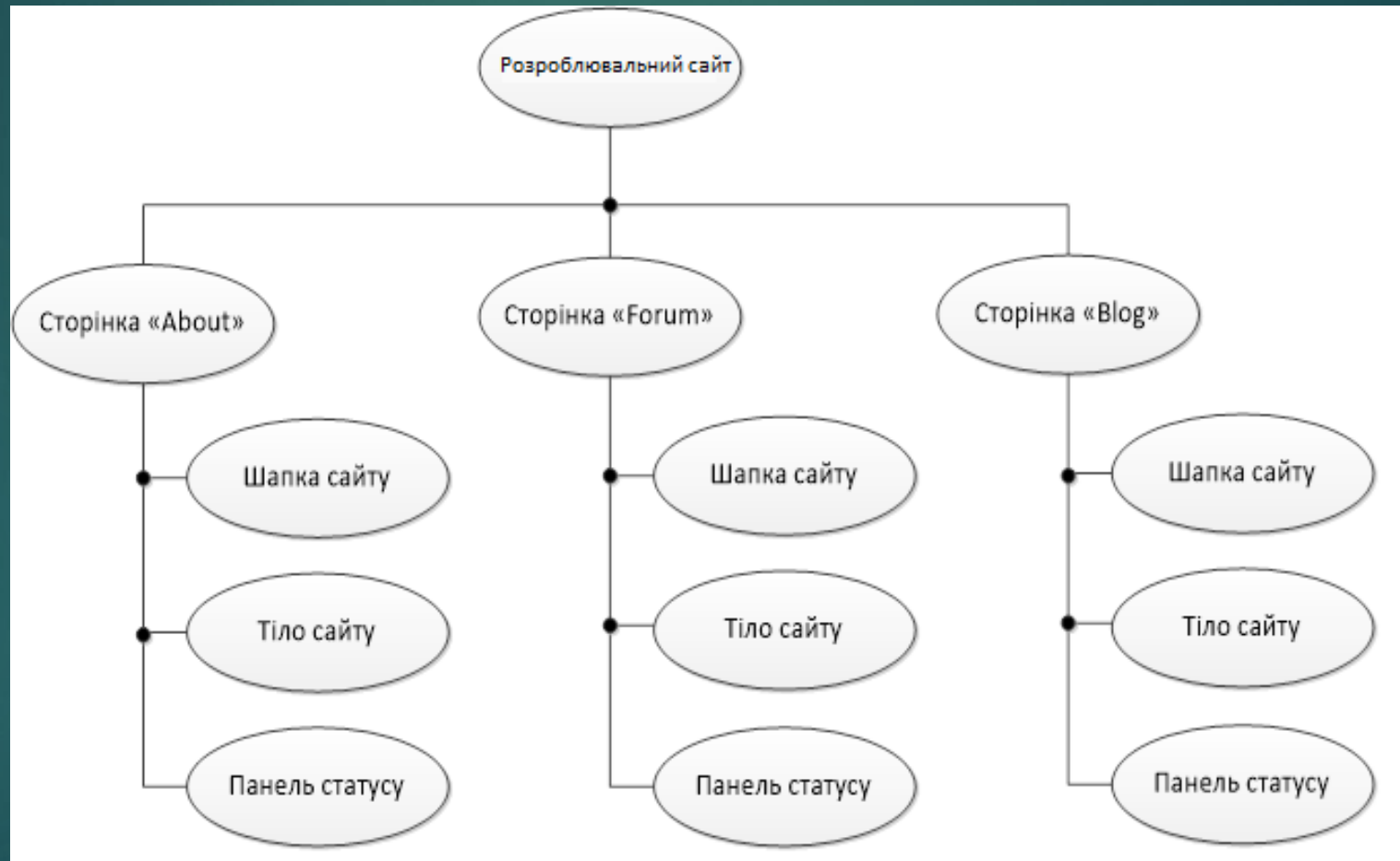
- 5 Повторюємо кроки 3–4, доки центри не стабілізуються або кількість ітерацій не перевищить задану.

# Структура інформаційної веб-технології для створення робочих зв'язків на основі інтелектуального аналізу даних





# Структурна схема сайту



# Розробка алгоритмів

Схема алгоритму роботи веб-сервісу для створення робочих зв'язків

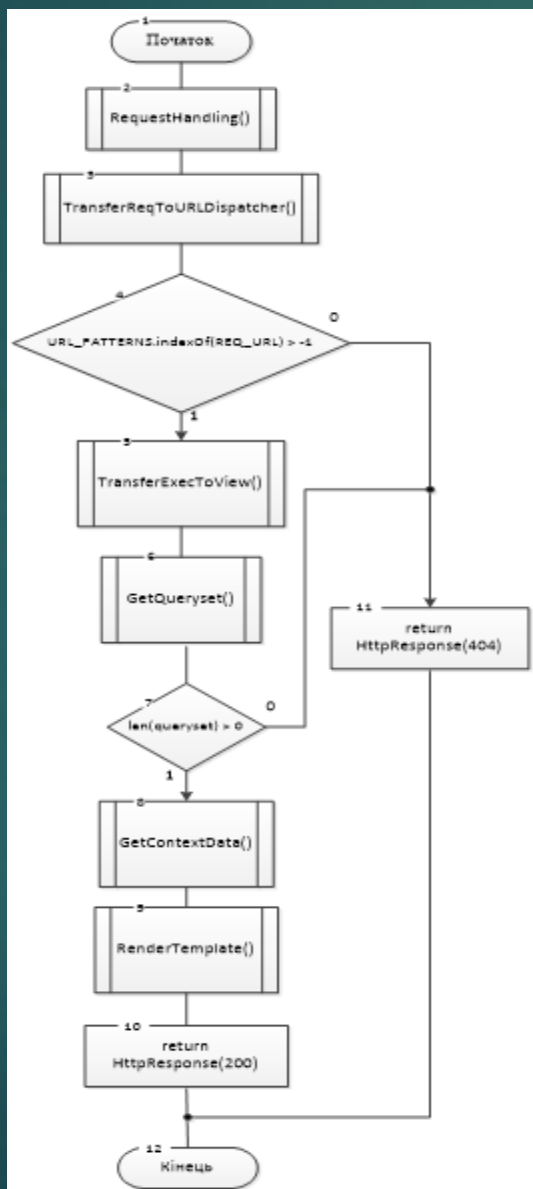
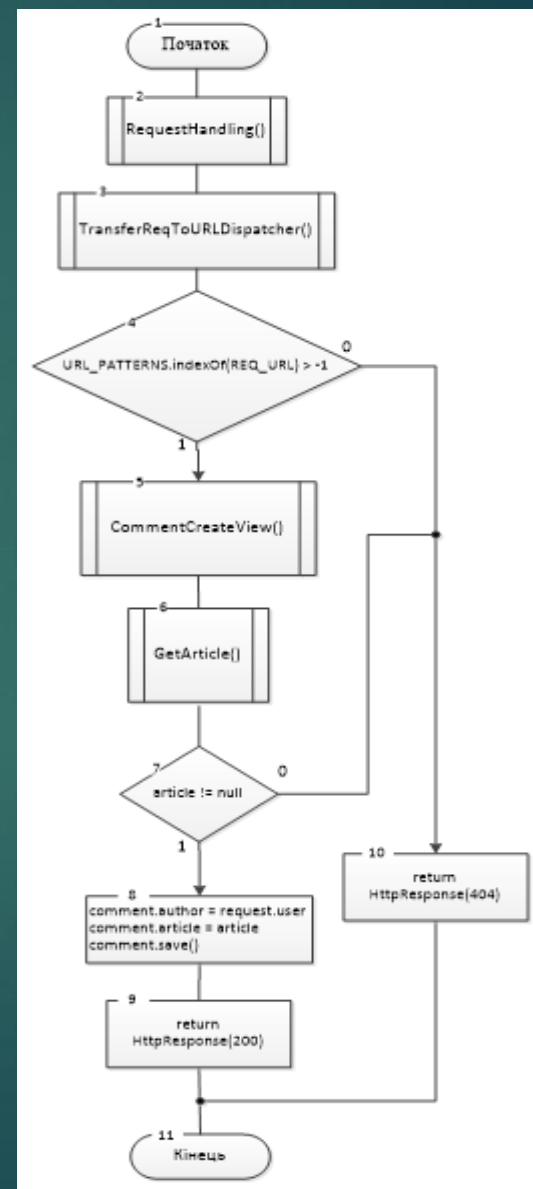


Схема алгоритму створення коментаря



# Діаграма класів та схема бази даних

11

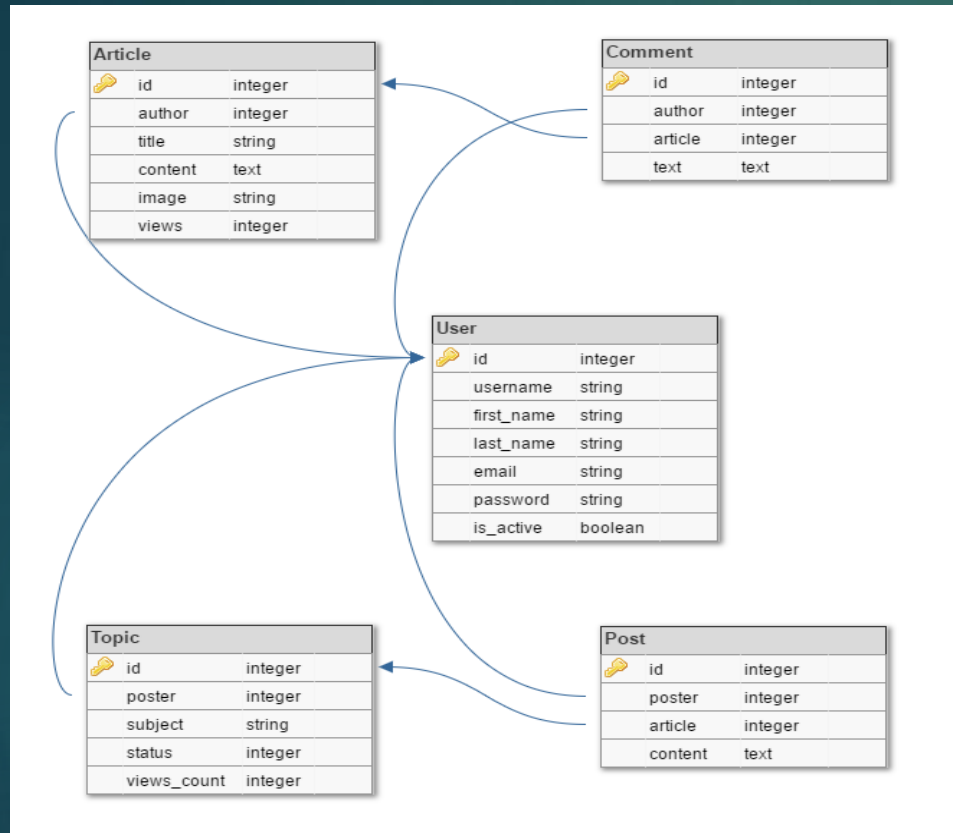
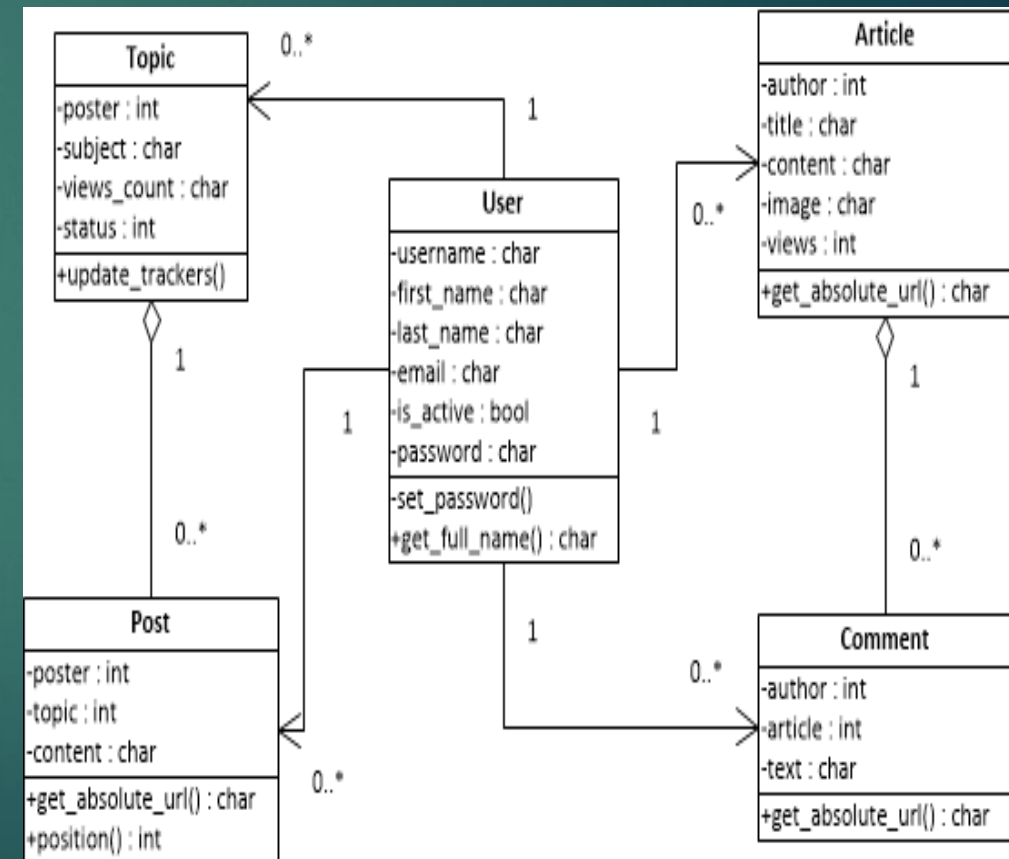


Схема бази даних веб-сервісу для створення робочих зв'язків

Діаграма класів веб-сервісу для створення робочих зв'язків



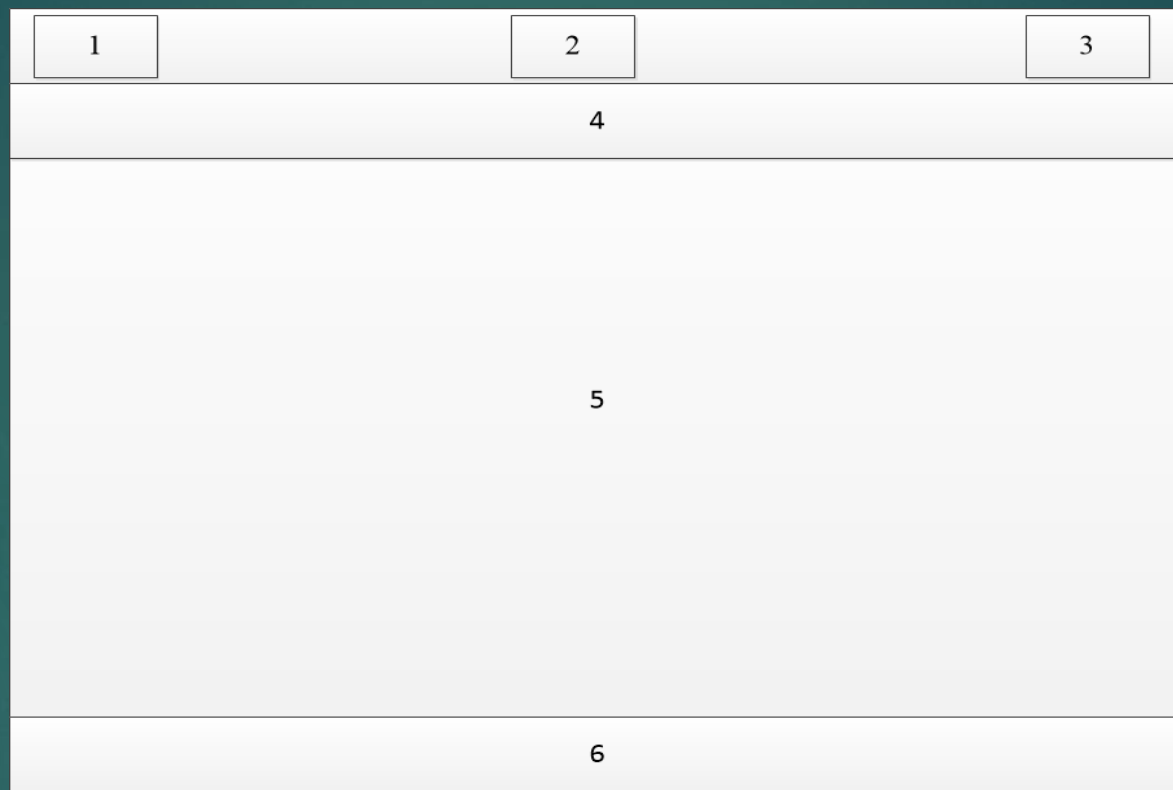
# Засоби розробки

	Python	C#	Java
ООП	1	1	1
Інтерпретатор командного рядка	1	1	0
Динамічна типізація	1	0	0
Множинне успадкування	1	1	1
Сумарний коефіцієнт	4	3	2

	PyCharm	Eric	Eclipse
Ліцензія	Пропріетарна, LGPL	LGPL	LGPL
Windows	1	0	1
Linux	1	1	0
MacOS	1	0	1
Відладчик	1	1	1
Розробка GUI	1	0	1
Браузер класів	1	0	1
Автодоповнення	1	1	0
Сумарний коефіцієнт	7	3	5

Програмне забезпечення веб-сервісу створене на мові програмування Python у середовищі розробки PyCharm з використанням Django Framework

# Схема інтерфейсу



1) логотип сайту;

4) панель вкладок;

2) пошукова стрічка;

5) робоча область;

3) панель керування особистим кабінетом;

6) панель статусу.



# Тестування сайту

Тест	Дія	Очікуваний результат	Отриманий результат	Результат тесту
Реєстрація нового користувача	Відкрити форму реєстрації, заповнити поля, натиснути кнопку «Submit».	Додано користувача	Додано користувача	ОК
Редагування профіля	Відкрити особистий кабінет, редагувати одне з полів, зберегти зміни.	Профіль Редаговано	Профіль Редаговано	ОК

JOIN THE NETWORK TODAY AND EXPAND YOUR PRESENCE

1 step — 2 step — 3 step

**Step 2. Individual registration:**  
Setup a username (email address) and password

E-mail

E-mail address

Password (6 Or More Characters)

Password (6 or more characters)

Verify Password

Verify password

[Previous](#) [NEXT](#)

JOBSEEKER

**Elin Musk** [EDIT](#)

LOCATION [EDIT](#)  
Minsk, Minsk City, Belarus

MY COUNTRY OF ORIGIN [CANCEL](#) [SAVE](#)

CHOOSE YOUR COUNTRY OF ORIGIN

Minj

Dominica

Dominican Republic

United States Minor Outlying Islands

[EDIT](#)

Click edit button to select secondary expertise

Followers 0 Following 0

SETTINGS [EDIT](#)

CERTIFICATIONS [ADD CERTIFICATIONS](#) LANGUAGES SPOKEN [ADD LANGUAGE](#)  
Click here to describe your certificate Click here to describe your language spoken

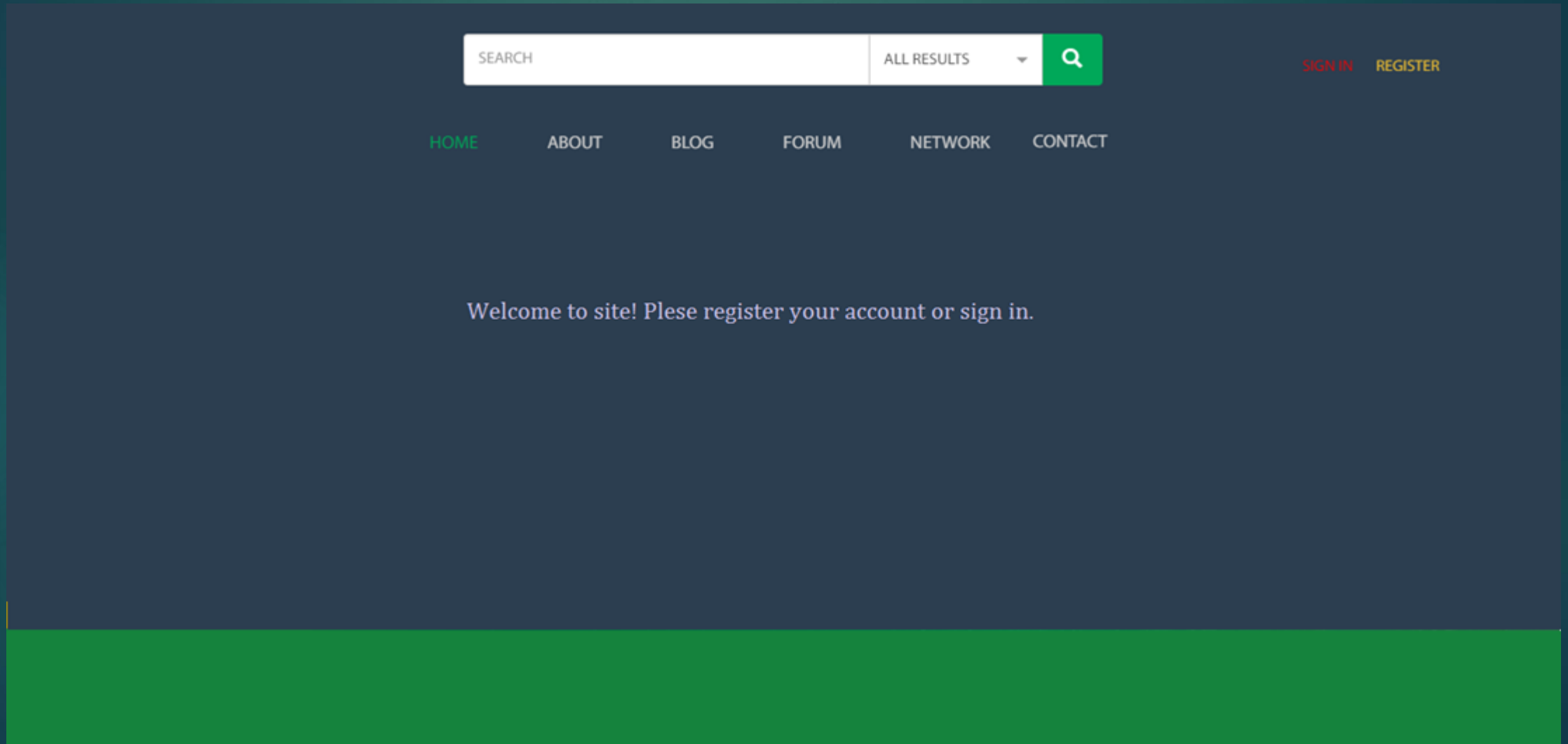
ACADEMIC BACKGROUND/DEGREES [EDIT](#)  
Add your degrees here separated by a comma

SKILLS [EDIT](#)  
ttt

ABOUT ME [EDIT](#) INTERESTS [EDIT](#)  
1 Click here to describe your interests

# ГОЛОВНА СТОРІНКА

15



# Аналіз результатів роботи веб-сервісу для створення робочих зв'язків

16

Щоб показати, що мета роботи досягнута, потрібно довести, що розроблений веб-сервіс має більшу швидкодію виконання запитів пошуку та більше функціональних можливостей.

Що стосується функціональних можливостей, то із табл. 1.1 видно, що розроблений веб-сервіс має 5 функцій, а сервіс-аналог -3, тобто кількість функцій збільшилась, а значить функціональні можливості розширились.

Таблиця 1.1 – Порівняння розроблювального сайту з аналогами

	«LinkedIn»	«XING»	«BranchOut»	Розроблювальний сайт
Інтерактивна взаємодія користувачів	1	1	1	1
Форум	0	1	0	1
Налаштування оповіщень	1	0	1	1
Блог	1	0	0	1
Використання ІАД	0	0	0	1
Сумарний коефіцієнт	3	2	2	5

Примітка: 0 – відсутність, 1 – наявність.

Що стосується швидкодії, то для доведення її збільшення потрібно провести експериментальні дослідження з розробленим сервісом та з сервісом-аналогом при однакових умовах (однакова кількість записів у базі даних та розмірність кожного запису).

Технічно виконати це дуже складно, тому доведемо підвищення швидкодії на теоретичному рівні. Звичайній базі даних, яка має  $N$  записів, знадобиться у середньому  $N/2$  тактів. А у нашому випадку при використанні кластеризації на  $K$  кластерів, для пошуку знадобиться  $(K + [N/K])/2$  тактів, де вираз  $[N/K]$  означає округлення до найближчого більшого цілого. Тобто виграш у швидкодії складе  $(N/2) / ((K + [N/K])/2) = N / (K + [N/K])$  разів.

Наприклад, при  $N=10000$  і  $K=100$  виграш складе  $\sim 16$  раз.

Було проведено економічне обґрунтування доцільності розробки програмного забезпечення веб-сайту для створення робочих зв'язків. Нова розробка має рівень комерційного потенціалу вище середнього. Загальна сума витрат на виконання робіт склала "131543,15" грн. Загальні витрати на виконання та впровадження результатів виконаної наукової роботи— 154756,65 грн. Абсолютна ефективність вкладених інвестицій становить 1651512,19 грн, і це свідчить про те, що вкладання коштів на виконання та впровадження результатів НДДКР є доцільним. Відносна (щорічна) ефективність вкладених в наукову розробку інвестицій – 127 %, отже інвестор буде зацікавлений у фінансуванні даної наукової розробки. Термін окупності - 0,79 року. Виходячи з отриманих результатів, можна зробити висновок, що впровадження розробки програмного забезпечення веб-сайту для створення робочих зв'язків є прибутковим і вкладення інвестицій буде доцільним.

# АПРОБАЦІЯ РЕЗУЛЬТАТІВ РОБОТИ ТА ПУБЛІКАЦІЇ

18

## Апробація результатів роботи.

Результати досліджень апробовані на конференції «Молодь в науці: дослідження, проблеми, перспективи-2019», Вінниця, 2019

<https://conferences.vntu.edu.ua/index.php/mn/index/pages/view/zbirn2019>

## Публікації.

За результатами магістерської кваліфікаційної роботи опубліковано 1 тези доповідей на конференції та свідоцтво про реєстрацію авторського права на твір (Комп'ютерну програму).



# ВИСНОВОК

В результаті виконання роботи було розроблено інформаційну веб-технологію для створення робочих зв'язків на основі інтелектуального аналізу даних. Програмну реалізацію технології здійснено мовою програмування Python у середовищі розробки PyCharm з використанням Django Framework. Розроблене програмне забезпечення характеризується збільшеною швидкістю пошуку запитів зареєстрованих користувачів (в  $N/(K+[N/K])$  разів) та розширеною функціональністю, а значить мета роботи досягнута.

Дякую за увагу!