

**НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ БІОРЕСУРСІВ
І ПРИРОДОКОРИСТУВАННЯ УКРАЇНИ
Факультет інформаційних технологій**

ПОГОДЖЕНО

Декан факультету (Директор ННІ)

інформаційних технологій

(назва факультету (ННІ))

(підпис)

Ігор БОЛБОТ

(ім'я ПРИЗВИЩЕ)

“ ___ ” _____ 2025 р.

ДОПУСКАЄТЬСЯ ДО ЗАХИСТУ

Завідувач кафедри

комп'ютерних наук

(назва кафедри)

(підпис)

Белла ГОЛУБ

(ім'я ПРИЗВИЩЕ)

“ ___ ” _____ 2025 р.

МАГІСТЕРСЬКА КВАЛІФІКАЦІЙНА РОБОТА

на тему Експертна система оцінки наукової діяльності педагогічного складу за інформацією з міжнародних джерел

Спеціальність 122 Комп'ютерні науки

(код і найменування)

Освітня програма Інформаційні управляючі системи і технології

(назва)

Орієнтація освітньої програми освітньо-професійна

(освітньо-професійна або освітньо-наукова)

Гарант освітньої програми

к.т.н., доцент

(науковий ступінь та вчене звання)

(підпис)

Белла ГОЛУБ

(ім'я ПРИЗВИЩЕ)

Керівник магістерської кваліфікаційної роботи

к.т.н., доцент

(науковий ступінь та вчене звання)

(підпис)

Іван ПАРХОМЕНКО

(ім'я ПРИЗВИЩЕ)

Виконав

(підпис)

Микола МАРУЩАК

(ім'я ПРИЗВИЩЕ здобувача)

КИЇВ – 2025

НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ БІОРЕСУРСІВ
І ПРИРОДОКОРИСТУВАННЯ УКРАЇНИ

Факультет інформаційних технологій

ЗАТВЕРДЖУЮ
Завідувач кафедри

к.т.н., доцент _____ Белла ГОЛУБ
(науковий ступінь, вчене звання) (підпис) (ім'я ПРІЗВИЩЕ)
“ 01 ” листопада 2024 року

З А В Д А Н Н Я

ДО ВИКОНАННЯ МАГІСТЕРСЬКОЇ КВАЛІФІКАЦІЙНОЇ РОБОТИ ЗДОБУВАЧУ

Марушаку Миколі Миколайовичу

(прізвище, ім'я, по батькові)

Спеціальність 122 «Комп'ютерні науки»

Освітня програма Інформаційні управляючі системи і технології

Орієнтація освітньої програми освітньо-професійна

Тема магістерської кваліфікаційної роботи Експертна система оцінки наукової діяльності педагогічного складу за інформацією з міжнародних джерел

затверджена наказом ректора НУБіП України від “01” листопада 2024 р. №1964 «С»

Термін подання завершеної роботи на кафедру 01.12.2025

Вихідні дані до магістерської кваліфікаційної роботи дані про викладачів, показники оцінювання наукової діяльності, інформація з міжнародних джерел, база даних для збереження, аналітичні результати та рекомендації, програмне забезпечення для розробки.

Перелік питань, що підлягають дослідженню:

1. Аналіз предметної області оцінки наукової діяльності та визначення ключових наукометричних показників, що характеризують наукову діяльність педагогічного складу за даними міжнародних джерел.

2. Дослідження та обґрунтування структури критеріїв оцінки, вагових коефіцієнтів і правил прийняття рішень в експертній системі.

3. Оцінка отриманих результатів та практичної придатності розробленої експертної системи для підтримки прийняття рішень.

Дата видачі завдання “01” листопада 2024 р.

Керівник магістерської кваліфікаційної роботи _____

(підпис)

Іван ПАРХОМЕНКО
(ім'я ПРІЗВИЩЕ)

Завдання прийняв до виконання _____

(підпис)

Микола МАРУЩАК
(ім'я ПРІЗВИЩЕ здобувача)

Зміст

ВСТУП	5
РОЗДІЛ 1 СИСТЕМНИЙ АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ.....	9
1.1. Аналіз сутності процесу оцінювання наукової діяльності	9
1.2. Огляд і аналіз існуючих підходів та програмних рішень	11
1.3. Постановка завдання.....	15
РОЗДІЛ 2 МОДЕЛЮВАННЯ СИСТЕМИ	17
2.1. Об'єктне та функціональне моделювання	17
2.2. Діаграма прецедентів	18
2.3. Діаграма послідовності	22
2.4. Діаграма класів.....	25
РОЗДІЛ 3 РОЗРОБКА СИСТЕМИ	29
3.1. Архітектура системи.....	29
3.2. Модуль збору інформації з міжнародних джерел.....	30
3.3. Модуль обробки інформації та формування показників оцінювання.....	32
3.4. Модуль зберігання даних.....	34
3.4.1. Створення бази даних.....	34
3.4.2. Створення сховища даних.....	38
3.4.3. Заповнення сховища даних	45
РОЗДІЛ 4 РЕЗУЛЬТАТИ ДОСЛІДЖЕННЯ	48
4.1. Вимоги до апаратного та програмного забезпечення	48
4.2. Хід виконання дослідження	49
4.3. Аналіз результатів роботи	51
4.3.1. Оцінка публікаційної активності та цитування	52
4.3.2. Результати кластерного аналізу та ключових показників ефективності	58
ВИСНОВКИ	65
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ.....	68
ДОДАТКИ.....	71

СПИСОК УМОВНИХ ПОЗНАЧЕНЬ

ЕС - експертна система;

БД - База Даних;

SSMS - SQL Server Management Studio (Середовище управління SQL Server);

SSAS - SQL Server Analysis Services(Служби аналізу SQL Server);

SSIS - SQL Server Integration Services (Служби інтеграції даних SQL Server);

OLAP - Online Analytical Processing (Аналітична обробка у реальному часі);

KPI - Key Performance Indicators (Ключові показники активності);

DSV - Data source view (Уявлення джерела даних);

DWH - Data Warehouse (Сховище даних).

ВСТУП

Актуальність теми дослідження. У сучасному світі цифровізації освіти та науки особливої ваги набуває проблема об'єктивного оцінювання результативності наукової діяльності педагогів. Показники публікаційної активності та цитованості дедалі частіше розглядаються як ключові критерії професійного зростання викладачів у відповідних сферах.

Сьогодні міжнародні наукометричні бази даних, такі як Scopus, Web of Science, Google Scholar та інші, відіграють провідну роль у моніторингу наукової продуктивності. Вони надають доступ до великої кількості даних, що характеризують наукові публікації, цитування та індекси впливу авторів. Проте, попри розвиненість цих ресурсів, процес збору, узагальнення та інтерпретації інформації залишається складним та трудомістким і часто потребує ручного втручання. Це створює ризики виникнення неточностей і суб'єктивних оцінок під час аналізу результатів діяльності педагогічних працівників.

Традиційні методи оцінювання наукової активності зазвичай спираються на часткові показники або обмежені джерела, що не дозволяє отримати цілісну картину дослідницької діяльності викладача. Крім того, відсутність єдиної системи збору та обробки інформації ускладнює формування прозорих рейтингів і прийняття управлінських рішень у науково-освітньому середовищі.

У зв'язку з цим розроблення експертної системи, здатної автоматично збирати, уніфікувати, аналізувати та візуалізувати дані про наукову діяльність педагогічного складу, є актуальним напрямом дослідження. Така система має забезпечити автоматизовану інтеграцію інформації з міжнародних наукометричних баз, усунення дублювань, а також створення узагальнених показників продуктивності наукової праці.

Впровадження подібного підходу сприятиме підвищенню об'єктивності оцінювання, прозорості академічних рейтингів і формуванню аналітичної бази для прийняття управлінських рішень у сфері освіти. Крім того, використання інтелектуальних методів аналізу даних створює умови для підвищення мотивації викладачів до активнішої участі у наукових дослідженнях, міжнародних проектах і публікаціях.

Об'єктом дослідження є процес оцінювання наукової діяльності педагогічних працівників на основі даних, отриманих із міжнародних наукометричних баз. Цей процес охоплює етапи збору, обробки, інтеграції та аналізу інформації про публікаційну активність, цитування та інших показників результативності.

На практиці цей процес часто характеризується відсутністю інструментів для роботи з даними, що зберігаються у різних джерелах. У багатьох закладах вищої освіти оцінювання здійснюється вручну або із застосуванням неповних даних, що ускладнює порівняльний аналіз і може знижувати об'єктивність результатів. Автоматизація цього процесу дозволяє не лише зменшити кількість помилок, але й сформувати цілісну картину наукової активності кожного педагога в контексті міжнародних стандартів.

Предметом дослідження є експертна система оцінювання наукової діяльності педагогічного складу, яка реалізує алгоритми збору, аналізу та візуалізації інформації з міжнародних джерел. Предмет включає методи обробки наукометричних даних, формування узагальнених показників продуктивності, інтеграцію з міжнародними джерелами та застосування інтелектуальних алгоритмів для підвищення прозорості та об'єктивності оцінки.

Метою дослідження роботи є підвищення об'єктивності та прискорення процесу оцінювання наукової діяльності педагогічного складу шляхом використання даних міжнародних наукометричних баз і технологій автоматизованого аналізу та завантаження інформації.

Для досягнення поставленої мети у магістерській роботі необхідно розв'язати такі **завдання**:

1. Провести аналіз сучасних підходів та методів оцінювання наукової діяльності педагогічних працівників, що використовуються для цієї мети.
2. Розробити архітектуру програмного додатку, що забезпечує збір, уніфікацію та обробку даних з міжнародних наукометричних баз.
3. Реалізувати модулі аналізу та формування узагальнених показників продуктивності.
4. Використати методи аналізу для відображення результатів оцінювання, рейтингів та візуалізації даних у вигляді таблиць і графіків.
5. Провести тестування програмного додатку, оцінити коректність визначених розрахунків, роботу аналітичних модулів та практичну придатність розробленої системи.

Новизна магістерської роботи полягає у використанні методів автоматизованого аналізу наукометричних даних для підвищення прозорості та об'єктивності оцінювання наукової діяльності педагогічного складу.

Розроблено та реалізовано підхід до збору та уніфікації даних із різних міжнародних наукометричних баз, який дозволяє забезпечити повноту та точність інформації про публікаційну активність та цитованість.

Розроблена архітектура програмного додатку забезпечує наявність модулів обробки даних та візуалізації результатів з базою даних, що підвищує швидкість доступу до інформації та зменшує ймовірність дублювання даних.

Отримані результати становлять практичну та наукову цінність, оскільки демонструють доцільність поєднання методів автоматизованого аналізу наукометричних даних із експертними правилами для оптимізації процесу оцінювання наукової діяльності педагогів і формування прозорих рейтингів.

Пояснювальна записка складається з вступу, чотирьох розділів, висновків, списком використаних літературних джерел, який складається з 22 найменувань та додатків. Основний обсяг роботи викладено на 71 сторінці.

У вступі описано актуальність теми, визначено об'єкт, предмет і мету дослідження, представлено новизну роботи та практичне значення отриманих результатів.

Перший розділ присвячено системному аналізу предметної області. У ньому розглянуто сутність процесу оцінювання наукової діяльності, проаналізовано існуючі підходи та програмні рішення, що використовуються на практиці. Також сформульовано постановку завдання та визначено основні вимоги до експертної системи, що розробляється.

Другий розділ містить етап моделювання системи. У ньому наведено об'єктне й функціональне моделювання, представлені діаграми прецедентів, послідовності та класів. Отримані моделі дозволяють формалізувати логіку роботи системи, визначити ролі користувачів та взаємодію між її елементами.

У третьому розділі розглянуто процес розробки експертної системи. Описано архітектуру ЕС та роботу основних модулів, а саме, збору інформації з міжнародних джерел, обробки даних і формування показників оцінювання, а також модулів зберігання інформації. Окрему увагу приділено створенню бази даних, організації сховища даних та його наповненню.

Четвертий розділ містить результати проведеного дослідження. У ньому визначено вимоги до апаратного і програмного забезпечення, наведено опис реалізації системи та проведення експериментів. Також здійснено аналіз отриманих результатів.

У висновках підсумовано виконану роботу, наведено основні досягнення, результати яких вдалось отримати в ході дослідження та визначено перспективи подальшого застосування системи.

РОЗДІЛ 1 СИСТЕМНИЙ АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ

1.1. Аналіз сутності процесу оцінювання наукової діяльності

У сучасних умовах розвитку вищої освіти наукова діяльність педагогічних працівників є одним із ключових чинників формування конкурентоспроможності закладу освіти, його репутації та позицій у міжнародних рейтингах. Рівень наукової активності викладачів безпосередньо впливає на якість освітнього процесу, оскільки забезпечує впровадження нових знань, методик і результатів досліджень у навчальний процес.

Оцінювання наукової діяльності викладачів покликане відображати їхню результативність у проведенні досліджень, публікаційній активності та наукових проектах. Таке оцінювання виступає основою для ухвалення управлінських рішень, а саме, розподілу фінансування, підвищення кваліфікації, присвоєння наукового ступеню і стимулювання професійного зростання.

Наукова активність традиційно характеризується кількісними та якісними показниками. До основних кількісних показників належать кількість опублікованих праць, цитованість та індекс Гірша. Якісні показники охоплюють рівень впливу публікацій та внесок автора у розвиток певного наукового напрямку.

Значну роль у процесі оцінювання відіграють міжнародні наукометричні бази даних - Scopus, Web of Science, Google Scholar тощо. Вони забезпечують централізований доступ до великого масиву бібліографічної інформації, показників цитованості та аналітичних звітів про діяльність науковців. Інформація з цих баз використовується для формування інституційних рейтингів, визначення рівня наукової продуктивності викладачів, а також для моніторингу тенденцій розвитку науки у певних галузях.

Попри високий потенціал таких ресурсів, процес збору та аналізу даних залишається складним і потребує значних часових затрат. Часто дані з різних баз не є узгодженими, інколи існують дублікати записів. Через це виникають труднощі у формуванні єдиних, достовірних і об'єктивних показників.

Крім того, у багатьох закладах вищої освіти оцінювання наукової діяльності все ще має ручний або напівформалізований характер. Викладачі самостійно подають звіти про власні публікації, що створює ризик суб'єктивності, помилок або неповноти даних. Такий підхід ускладнює порівняння показників між різними підрозділами чи кафедрами та не забезпечує прозорості управлінських рішень.

Водночас із розвитком інформаційних технологій зростає потреба у створенні інтелектуальних систем, здатних автоматично збирати, обробляти та аналізувати дані з кількох джерел, забезпечуючи об'єктивність і повноту оцінки. Саме такі системи повинні забезпечувати інтеграцію наукометричних показників, їх уніфікацію, формування аналітичних звітів і динамічних рейтингів наукової активності.

Отже, процес оцінювання наукової діяльності педагогічних працівників є багатокomпонентним і потребує використання сучасних засобів обробки даних.

Його ефективність визначається здатністю системи поєднувати інформацію з різних міжнародних джерел, забезпечувати достовірність і точність аналітики, а також мінімізувати людський фактор у процесі прийняття рішень. Саме тому автоматизація цього процесу на основі експертних систем є перспективним напрямом наукових досліджень та практичного впровадження у сфері управління науковою діяльністю закладів вищої освіти.

1.2. Огляд і аналіз існуючих підходів та програмних рішень

У сфері аналізу та оцінювання наукової діяльності сьогодні використовується низка інформаційних систем, що забезпечують збір, обробку й відображення показників результативності науковців. Більшість таких рішень функціонують як наукометричні бази даних, орієнтовані на збереження бібліографічної інформації, індексацію цитувань і формування статистичних метрик. Попри значну кількість подібних платформ, жодна з них не забезпечує повного автоматизованого оцінювання продуктивності педагогічного складу закладів освіти з урахуванням даних із кількох міжнародних джерел одночасно.

Однією з найвідоміших систем є Scopus, розроблена компанією Elsevier. Вона надає доступ до великої кількості публікацій у рецензованих журналах, матеріалах конференцій та книгах. Основною перевагою Scopus є висока якість даних та розвинені аналітичні можливості - система дозволяє відстежувати цитованість, розраховувати індекс Гірша, порівнювати авторів, установи та наукові напрями. Для університетів Scopus пропонує аналітичні модулі Scopus Author Preview і Scopus Affiliation Analyzer, які можуть використовуватись для формування внутрішніх рейтингів. Водночас платформа має низку недоліків, а саме, обмежений доступ (переважно платний для комерційного використання), відсутність інтеграції з іншими базами без API-ключів та складність у формуванні локальних аналітичних звітів.

Іншою провідною базою даних є Web of Science, створена компанією Clarivate Analytics. Вона містить менше джерел порівняно зі Scopus, однак відзначається більшою строгістю відбору журналів. Web of Science підтримує низку власних метрик - Impact Factor, Journal Citation Reports, Essential Science Indicators - що робить її одним із головних інструментів для оцінювання наукової продуктивності у міжнародному масштабі. До переваг системи належить висока достовірність інформації, можливість побудови динамічних звітів та доступ до

пов'язаних сервісів (наприклад, Publons). Основний недолік - комерційна модель доступу і складність інтеграції з університетськими інформаційними системами.

Більш відкритим і доступним рішенням є Google Scholar, яке забезпечує автоматичну індексацію публікацій з відкритих джерел. Система охоплює наукові статті, тези конференцій, дисертації, книги та навіть неофіційні препринти. Google Scholar дозволяє формувати авторські профілі, автоматично обчислює індекс Гірша та загальну кількість цитувань. Його головною перевагою є безкоштовність і широта охоплення, однак відсутність ручної модерації призводить до появи помилок, дублікатів і неточностей у відображенні публікацій. Крім того, система не забезпечує стандартизації даних, що ускладнює інтеграцію результатів у комплексні експертні системи.

Значного поширення набув також науковий соціальний сервіс ResearchGate, який виконує функції мережевої платформи для дослідників. У системі можна розміщувати власні публікації, обмінюватися відгуками, вести дискусії та спостерігати за показниками переглядів і цитувань. ResearchGate формує власний індекс RG Score, який базується на активності користувача, кількості публікацій, відгуків та взаємодій з іншими науковцями. Проте цей показник не є стандартизованим і не може бути використаний як офіційна метрика при оцінюванні наукової діяльності. Крім того, система не має прямої інтеграції з базами Scopus чи Web of Science.

Сервіс Publons, нині інтегрований до Web of Science, призначений для обліку рецензентської активності. Він дозволяє фіксувати кількість проведених рецензій, перевіряти участь у редагуванні журналів і формувати профілі науковців із деталізацією їх внеску у рецензування. Хоча Publons і забезпечує підвищення прозорості оцінювання рецензентської роботи, він не враховує інші аспекти наукової діяльності, тому має допоміжний характер.

Окреме місце посідає система ORCID, що надає дослідникам унікальні цифрові ідентифікатори, які використовуються для уніфікації даних між різними

наукометричними платформами. Вона спрощує ідентифікацію авторів, зменшує кількість помилок у відображенні публікацій, однак не виконує функцій аналітики чи оцінювання, тому найчастіше виступає як допоміжний інструмент інтеграції.

Крім міжнародних рішень, окремі університети та наукові установи впроваджують власні внутрішні системи моніторингу наукової активності, побудовані на локальних базах даних. Такі системи дають змогу вести облік публікацій, конференцій, грантів і проектів, але зазвичай не мають автоматизованого імпорту інформації з міжнародних джерел. Це обмежує їх актуальність і призводить до дублювання даних під час оновлення.

Для узагальнення проведеного аналізу наведено таблицю порівняння основних існуючих систем (табл. 1.1).

Таблиця 1.1

Порівняльна характеристика існуючих інформаційних систем оцінювання наукової діяльності

Система	Основне призначення	Аналітичні можливості	Доступність	Переваги	Недоліки
Scopus	Індексація публікацій, аналітика цитувань	Індекс Гірша, аналітичні звіти, порівняння авторів	Варіюється	Висока якість даних, глибока аналітика, авторитетність	Обмежений доступ, складна інтеграція
Web of Science	Аналітика цитованості та рейтингів	Impact Factor, JCR, ESI, звіти	Частково платна	Надійність, перевірені джерела, глибока статистика	Висока вартість, обмежене охоплення
Google Scholar	Відкрита індексація наукових робіт	h-index, i10-index, цитованість	Безкоштовна	Доступність, широке охоплення	Дублікати

Продовження таблиці 1.1

Система	Основне призначення	Аналітичні можливості	Доступність	Переваги	Недоліки
Research Gate	Соціальна мережа для науковців	RG Score, перегляди, статистика активності	Безкоштовна	Можливість співпраці, швидкий обмін інформацією	Неофіційні метрики, суб'єктивність
Publons	Облік рецензентської активності	Звіти про рецензії	Безкоштовна	Підвищення прозорості рецензування	Обмежена функціональність
ORCID	Ідентифікація авторів і об'єднання профілів	Відсутні	Безкоштовна	Уніфікація авторських профілів, підтримка інтеграцій	Не виконує аналітику
Внутрішні університетські системи	Облік публікацій і проектів	Обмежені звіти, рейтинги	Варіюється	Гнучкість налаштування, зручність у локальному використанні	Відсутність автоматизації, дублювання даних

Проведений аналіз показує, що сучасні інформаційні системи оцінювання наукової діяльності педагогічних працівників мають широкий функціонал, проте орієнтовані переважно на загальносвітовий або міждисциплінарний рівень. Такі системи, як Scopus та Web of Science, забезпечують високу точність і глибину аналітики, але є платними й обмеженими щодо інтеграції з локальними базами даних навчальних закладів. Безкоштовні рішення, зокрема Google Scholar або ResearchGate, натомість характеризуються простотою використання, проте не гарантують достовірності та повноти зібраної інформації.

Аналіз показав, що жодна з розглянутих платформ не забезпечує комплексного автоматизованого підходу до оцінювання наукової активності педагогічного складу з урахуванням даних із кількох міжнародних джерел одночасно. Відсутність механізмів перевірки достовірності даних, уніфікації показників та можливості формування гнучких звітів суттєво обмежує використання таких систем у внутрішніх процесах освітніх установ.

Отже, виникає потреба у створенні експертної системи нового типу, яка б поєднувала можливості автоматизованого збору, обробки та порівняння даних з міжнародних наукометричних джерел, а також надавала інструменти для комплексної оцінки наукової діяльності педагогічних працівників. Це завдання визначає подальший напрямок магістерського дослідження.

1.3. Постановка завдання

На основі проведеного аналізу предметної області та існуючих інформаційних систем виявлено, що жодне з доступних рішень не забезпечує повного автоматизованого процесу оцінювання наукової діяльності педагогічного складу з урахуванням даних із кількох міжнародних джерел. Існуючі системи або не враховують специфіку освітніх установ, або мають обмежену функціональність і не підтримують інтеграцію з міжнародними базами.

У зв'язку з цим постає завдання розроблення експертної системи, яка забезпечуватиме:

- автоматизований збір інформації про публікаційну активність педагогічних працівників із таких джерел, як Scopus, Web of Science, Google Scholar;
- аналіз отриманих даних із використанням аналітичних і експертних методів для визначення рівня наукової продуктивності;

- уніфікацію показників цитованості, публікацій та інших метрик для формування комплексного рейтингу;
- візуалізацію результатів оцінювання у зручній формі звітів і діаграм.

Основна увага у дослідженні приділяється моделюванню процесів оцінювання, визначенню критеріїв експертної оцінки та розробленню архітектури програмного забезпечення, що забезпечує високу точність, об'єктивність і швидкість обробки даних.

Таким чином, у рамках магістерської роботи передбачається створення програмного прототипу експертної системи, який дозволить перевірити ефективність запропонованого підходу та підтвердити доцільність його впровадження у практику оцінювання наукової діяльності педагогічних працівників.

РОЗДІЛ 2 МОДЕЛЮВАННЯ СИСТЕМИ

2.1. Об'єктне та функціональне моделювання

Побудова ефективної експертної системи вимагає застосування методів моделювання, які дозволяють описати як функціональні особливості системи, так і її внутрішню структуру. На практиці для моделювання програмних систем найчастіше використовують два основних підходи - функціональний та об'єктно-орієнтований.

Функціональний підхід орієнтується на опис системи через функції, які вона виконує. Він дозволяє відобразити логічні процеси, що відбуваються під час роботи системи, та взаємодію користувачів із цими процесами. Основними засобами такого моделювання є діаграми прецедентів, що відображають сценарії використання та послідовність дій.

Натомість об'єктно-орієнтований підхід розглядає систему як сукупність взаємопов'язаних об'єктів, кожен із яких має власні атрибути, методи та взаємодії з іншими компонентами. Такий підхід є більш гнучким, масштабованим і придатним для систем, які потребують розширення функціоналу та подальшої інтеграції з іншими інформаційними ресурсами. Основними інструментами об'єктно-орієнтованого моделювання є діаграми класів, послідовності та взаємодії.

Зважаючи на специфіку теми магістерського дослідження - розроблення експертної системи оцінки наукової діяльності педагогічного складу за інформацією з міжнародних джерел - доцільним є поєднання двох підходів. Функціональний підхід буде використано для опису загальних сценаріїв роботи системи та взаємодії користувачів із нею, тоді як об'єктно-орієнтований підхід забезпечить глибоке моделювання структури даних і програмних компонентів.

2.2. Діаграма прецедентів

Для визначення основних функціональних можливостей експертної системи та взаємодії користувачів із нею використовується діаграма прецедентів. Цей тип діаграм належить до функціонального підходу моделювання й дозволяє відобразити зовнішню поведінку системи - тобто, які дії може виконувати користувач і які функції при цьому виконує система.

Діаграма прецедентів забезпечує наочне представлення взаємозв'язків між користувачами (акторами) та процесами, які реалізуються системою. Вона є основою для подальшого створення сценаріїв роботи користувачів і переходу до об'єктно-орієнтованого моделювання, зокрема побудови діаграм послідовності та класів.

Для побудови діаграми прецедентів розроблюваної системи перш за все необхідно визначити акторів, які взаємодіють з нею. На Рис. 2.2.1., показано перелік акторів, які взаємодіють з системою.

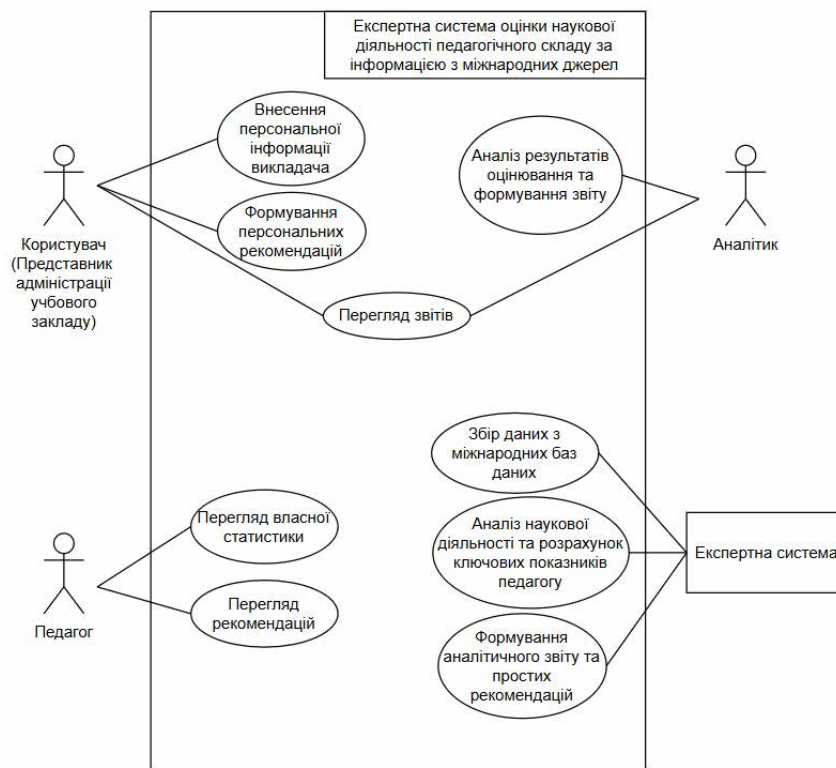


Рис. 2.2.1. Діаграма прецедентів

Основні актори системи:

- **Користувач** (Представник адміністрації навчального закладу) - виконує введення та редагування персональної інформації викладачів, переглядає звіти й формує рекомендації на основі результатів оцінювання.
- **Аналітик** - аналізує результати роботи системи, формує узагальнені звіти, виконує порівняльний аналіз оцінювання ЕС.
- **Педагог** - отримує доступ до власних результатів, переглядає статистику наукової діяльності, ознайомлюється з рекомендаціями згенерованими системою.
- **Експертна система** - автоматично збирає дані з міжнародних баз, виконує аналітичну обробку, розраховує показники та формує рекомендації та звіти.

Нижче наведені основні прецеденти системи:

Прецедент «Внесення персональної інформації викладача»

1.1 Передумови:

- Користувач має роль «Адміністратор».
- Користувач авторизований у системі.
- Дані про викладача підготовлені до внесення.

1.2 Тригер:

- Користувач ініціює дію створення або редагування профілю викладача через інтерфейс системи.

1.3 Головний потік:

1. Система відображає форму для введення або редагування даних викладача;
2. Користувач заповнює інформацію про користувача;
3. Система перевіряє формат та наявність обов'язкових полів (A1);
4. Дані зберігаються у базі даних;

5. Система повідомляє користувача про успішне створення або оновлення запису.

1.4 Альтернативні потоки:

- A1: Якщо деякі поля не заповнено або введено некоректно, система відображає повідомлення про помилку й пропонує користувачу виправити дані.

Прецедент «Аналіз наукової діяльності викладача»

1.1 Передумови:

- Система має збережені профілі викладачів.
- Експертна система налаштована на доступ до міжнародних баз даних.
- Наявне підключення до мережі для збору актуальних даних.

1.2 Тригер:

- Користувач ініціює процес аналізу наукової діяльності.

1.3 Головний потік:

1. Система ідентифікує профіль викладача за унікальним ідентифікатором;
2. Виконується автоматизований збір даних із відкритих наукометричних джерел;
3. Дані проходять попередню обробку – перевірка на наявність дублікатів та нормалізація назв(A1);
4. Система розраховує ключові показники продуктивності - кількість публікацій, цитованість, h-index та i10-index;
5. Результати зберігаються та стають доступними для подальшого аналізу;
6. Система повідомляє про успішне завершення процесу.

1.4 Альтернативні потоки:

- A1: Якщо виявлений дублікат, система не зберігає ідентичний запис.

Прецедент «Перегляд рекомендацій»

1.1 Передумови:

- Для викладача вже виконано аналіз наукової діяльності.
- Система має сформовані рекомендації на основі результатів оцінювання.
- Користувач має роль «Педагог» і авторизований у системі.

1.2 Тригер:

- Викладач входить до свого облікового запису та переходить до розділу «Рекомендації».

1.3 Головний потік:

1. Система ідентифікує користувача за обліковими даними;
2. Вибирає з бази даних індивідуальні результати аналізу;
3. На основі ключових показників система формує персональні поради щодо підвищення наукової активності (A1);
4. Користувач переглядає рекомендації на екрані або завантажує їх у звітному форматі.

1.4 Альтернативні потоки:

- A1: Якщо дані ще не проаналізовано або результатів немає, система повідомляє користувача про відсутність рекомендацій і пропонує повторити запит після оновлення бази.

Діаграма прецедентів дозволяє візуально відобразити взаємодію користувачів із системою, визначити ключові функції та сценарії використання програмного продукту.

Дана діаграма забезпечує чітке розуміння функціональних вимог до системи, допомагає визначити межі відповідальності акторів, а також слугує основою для побудови подальших моделей.

2.3. Діаграма послідовності

Діаграма послідовності є одним із ключових інструментів моделювання, що дозволяє відобразити логіку взаємодії між об'єктами системи. Вона описує порядок обміну повідомленнями між користувачами, модулями та компонентами програмного забезпечення. Такий підхід дає змогу чітко простежити, у якій послідовності виконуються дії в системі, які об'єкти взаємодіють між собою та в якому порядку відбувається обробка даних.

Для експертної системи оцінки наукової діяльності педагогічного складу було побудовано три основні діаграми послідовності, а саме, «Збір даних із міжнародних джерел», «Аналіз наукової діяльності викладача» та «Формування звіту і рекомендацій».

Процес починається з того, що користувач ініціює додавання інформації про наукову діяльність викладачів через експертну систему програми. ЕС оброблює запит збору даних, який встановлює з'єднання із зовнішніми наукометричними базами. Система надсилає запити до кожного джерела, використовуючи параметри пошуку. Отримані результати надходять до модуля збору, де перевіряються на повноту та унікальність. Далі відбувається обробка даних, яка нормалізує структуру записів та усуває дублікати.

Після завершення процесу система повідомляє користувача про успішне внесення даних. Така послідовність дій забезпечує достовірність і повноту зібраної інформації, а також створює основу для подальшого аналізу. Рис.2.3.1.

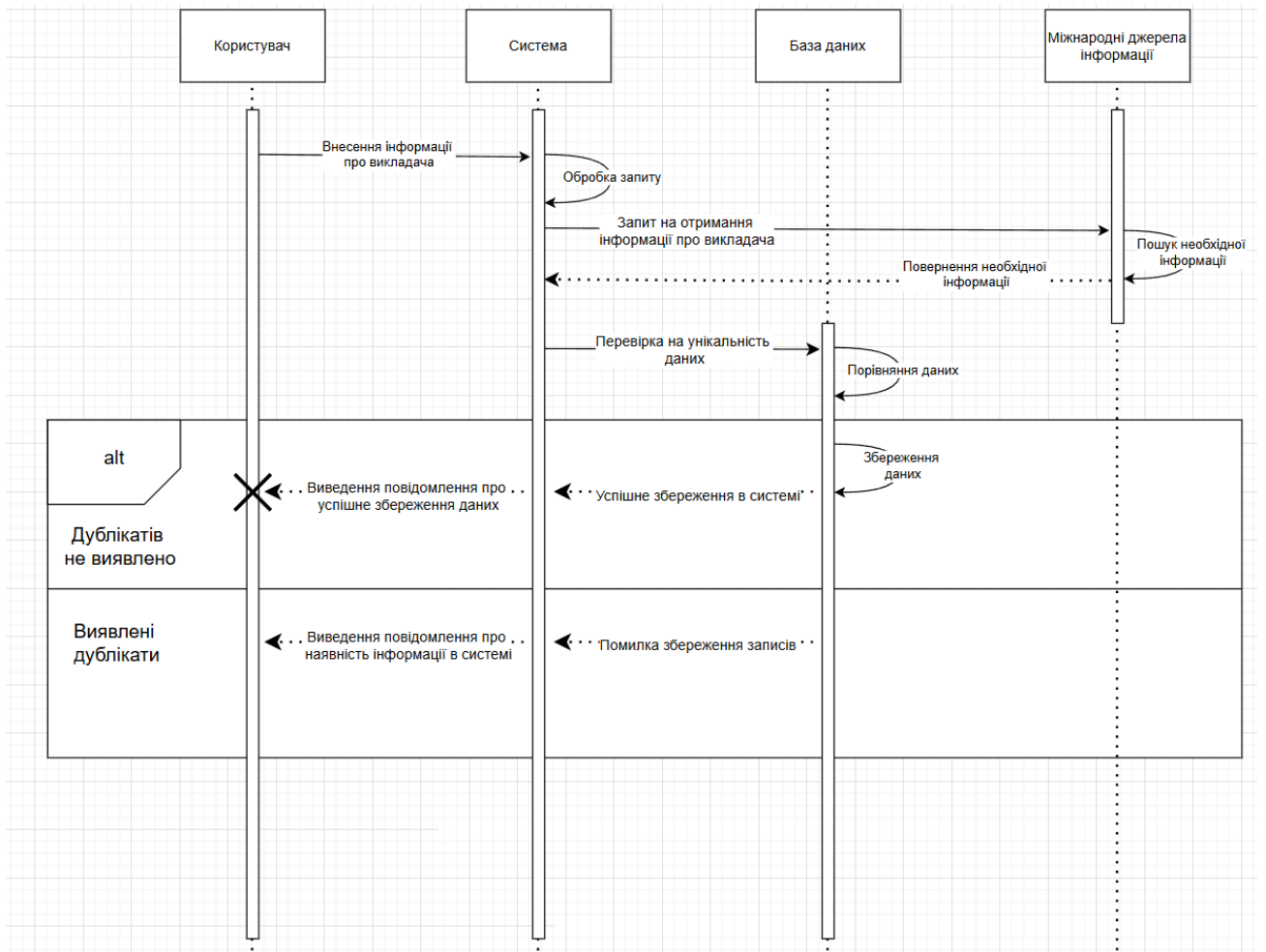


Рис.2.3.1. Діаграма послідовності «Збір даних із міжнародних джерел»

Для аналізу наукової діяльності викладача, процес починається з того, що користувач обирає конкретного викладача у системі для проведення аналізу. Після отримання інформації у системі, ЕС обчислює рейтинг по ключовим метрикам. Після визначення рейтингу по метрикам, визначається загальний рейтинг з урахуванням вагових коефіцієнтів.

У результаті користувач отримує аналітичний звіт із ключовими показниками наукової діяльності викладача, який може бути використаний для порівняльного аналізу чи внутрішнього рейтингування. Рис.2.3.2.

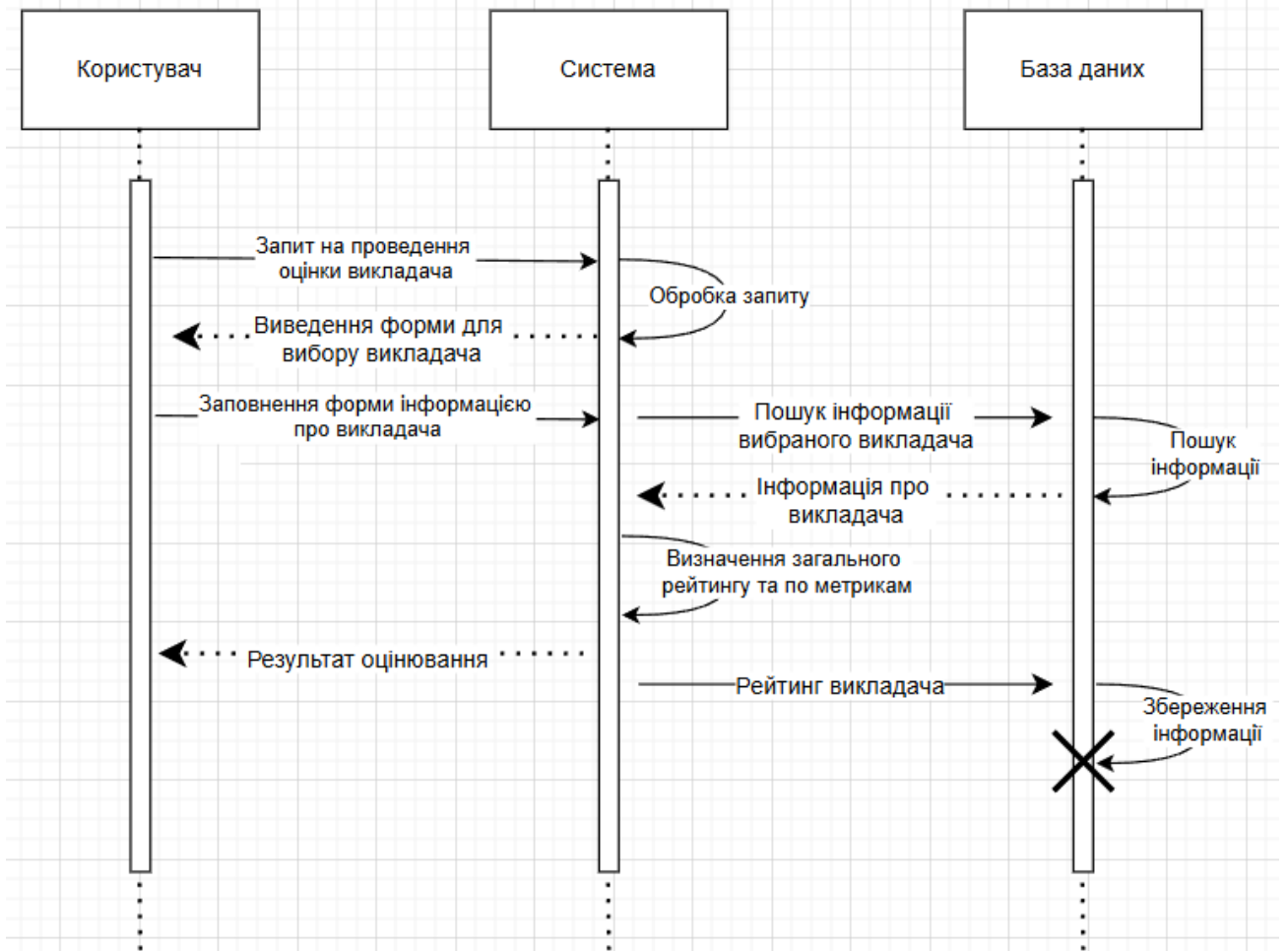


Рис.2.3.2 Діаграма послідовності «Аналіз наукової діяльності викладача»

В наступній діаграмі, користувач ініціює створення звіту, обираючи відповідний пункт у меню системи. Система надсилає запит до бази даних для отримання результатів попереднього аналізу. Після чого дані по викладачу структуруються у табличному або графічному вигляді. Далі використовуючи правила експертної системи формуються персоналізовані поради для викладачів. Наприклад, система може запропонувати підвищення активності у написанні публікацій.

Після завершення формування результатів звіт зберігається у базі даних і стає доступним для перегляду. Користувач отримує повідомлення про успішне створення документа. Рис.2.3.3.

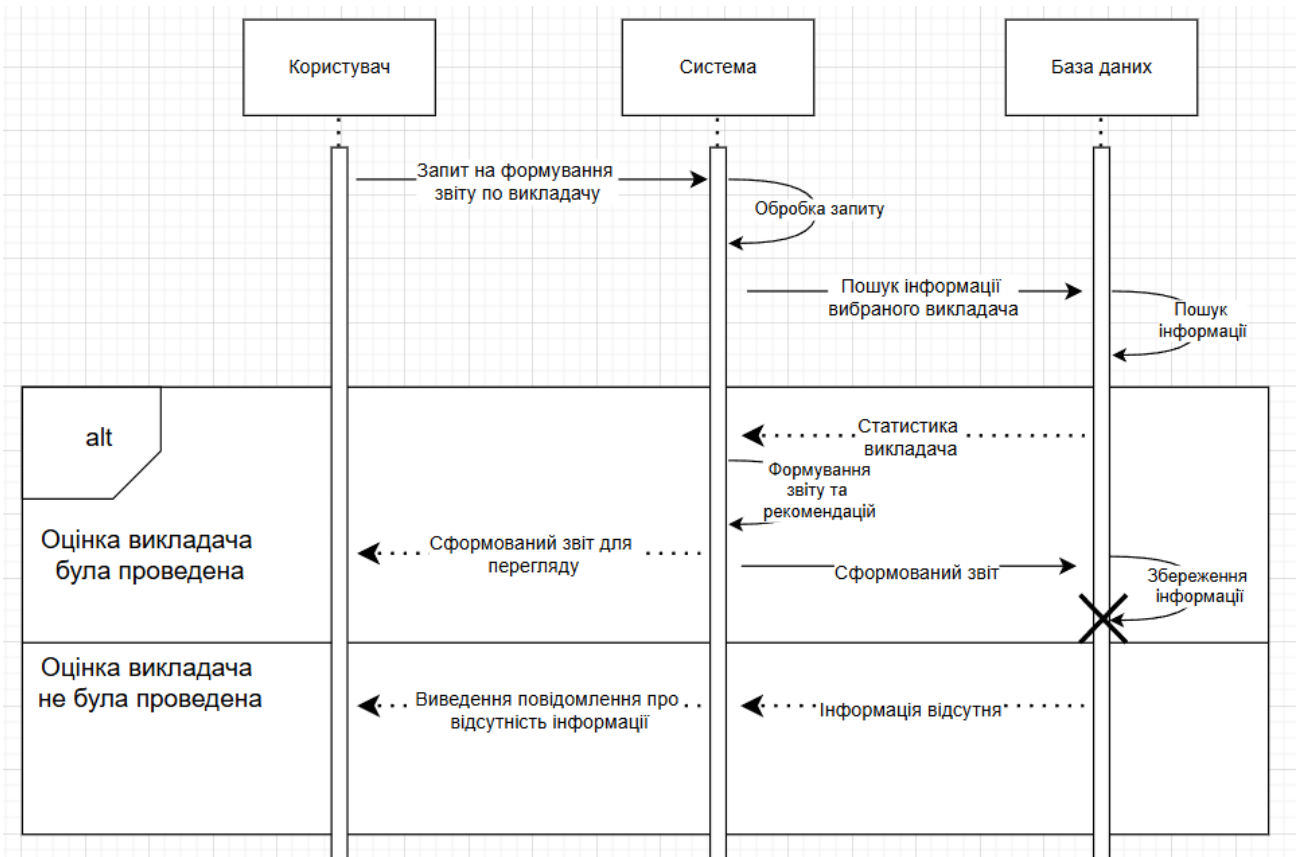


Рис.2.3.3 Діаграма послідовності «Формування звіту і рекомендацій»

Побудовані діаграми послідовностей дають змогу відобразити повний життєвий цикл функціонування експертної системи - від збору даних із зовнішніх джерел до формування рекомендацій для педагогічного складу. Вони чітко демонструють логіку взаємодії користувачів із системою, послідовність обробки інформації та взаємозв'язок між основними модулями. На основі цих моделей надалі здійснюється формування діаграм класів.

2.4. Діаграма класів

Діаграма класів є основним елементом об'єктно-орієнтованого моделювання, що описує статичну структуру програмної системи. Вона відображає основні класи, їх атрибути, методи та взаємозв'язки між ними. На основі діаграми класів формується архітектура системи, визначається логіка зберігання та обробки даних, а також реалізуються механізми взаємодії між компонентами програмного забезпечення.

Для ЕС оцінки наукової діяльності педагогічного складу за інформацією з міжнародних джерел побудована діаграма класів, що відображає зв'язки між сутностями, які відповідають основним таблицям бази даних. Рис.2.4.1.

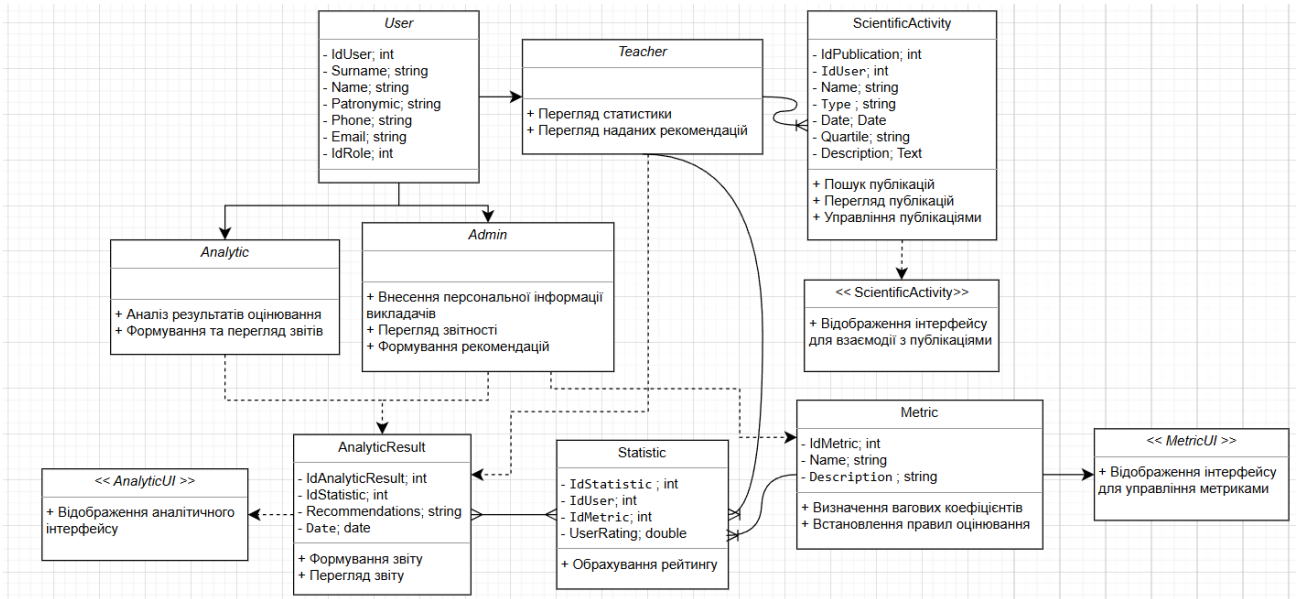


Рис.2.4.1 Діаграма класів

На даній діаграмі представлені наступні класи:

User - базовий клас користувача системи. Атрибути: *IdUser*, *Surname*, *Name*, *Patronymic*, *Phone*, *Email*, *IdRole*. Методи: авторизація користувача, оновлення особистих даних. Взаємозв'язки: виступає батьківським класом для *Admin*, *Analytic* та *Teacher*, які наслідують його основні властивості та мають власні спеціалізовані функції.

Admin - клас користувача, який виконує адміністративні функції в системі. Методи: внесення персональної інформації викладачів, перегляд звітності, формування рекомендацій. Взаємозв'язки: взаємодіє з класами *ScientificActivity*, *Statistic* та *AnalyticResult* для перегляду, аналізу та створення рекомендацій щодо наукової діяльності педагогічного складу.

Analytic - клас користувача, який здійснює аналітичну обробку даних. Методи: аналіз результатів оцінювання, формування та перегляд звітів. Взаємозв'язки: взаємодіє з класом *Statistic* для отримання даних та з класом

`AnalyticResult` для створення записів результатів аналізу. Має асоціацію з інтерфейсом `AnalyticUI`, який відображає аналітичну інформацію.

`Teacher` - клас користувача, що представляє викладача. Методи: перегляд власної статистики та рекомендацій, наданих системою. Взаємозв'язки: пов'язаний із класом `ScientificActivity` для перегляду наукових публікацій і з `AnalyticResult` для перегляду персональних рекомендацій та аналітичних висновків.

`ScientificActivity` - клас, що містить інформацію про наукові публікації викладачів. Атрибути: `IdPublication`, `IdUser`, `Name`, `Type`, `Date`, `Quartile`, `Description`. Методи: пошук, перегляд і управління публікаціями. Взаємозв'язки: пов'язаний із класом `User` через атрибут `IdUser`; дані публікацій використовуються в `Statistic` для обрахунку наукової активності. Має інтерфейс `ScientificActivityUI` для відображення даних користувачу.

`Metric` - клас, який описує показники, за якими проводиться оцінка наукової діяльності. Атрибути: `IdMetric`, `Name`, `Description`. Методи: визначення вагових коефіцієнтів, встановлення правил оцінювання. Взаємозв'язки: пов'язаний із `Statistic`, оскільки метрики є основою для формування рейтингових оцінок. Для керування метриками використовується інтерфейс `MetricUI`.

`Statistic` - клас, що зберігає результати оцінювання діяльності викладачів. Атрибути: `IdStatistic`, `IdUser`, `IdMetric`, `UserRating`. Методи: обрахунок рейтингу. Взаємозв'язки: асоційований із `User` (викладачем, який оцінюється), `Metric` (показником оцінювання) та `AnalyticResult` (результатом аналізу).

`AnalyticResult` - клас, який зберігає результати аналітичної обробки та рекомендації. Атрибути: `IdAnalyticResult`, `IdStatistic`, `Recommendations`, `Date`. Методи: формування звітів і перегляд результатів аналізу. Взаємозв'язки: пов'язаний зі `Statistic` для отримання базових даних та з `Analytic` і `Admin` для формування рекомендацій. Результати відображаються користувачам через `AnalyticUI`.

ScientificActivityUI, MetricUI, AnalyticUI - допоміжні інтерфейсні класи, що реалізують відображення відповідних даних для користувачів системи. Взаємозв'язки: кожен інтерфейс асоційований із відповідним класом(ScientificActivity, Metric, AnalyticResult), забезпечуючи взаємодію користувача з інформацією через графічний інтерфейс.

Діаграма класів, відображає об'єктно-орієнтовану структуру системи та взаємозв'язки між її складовими. Вона показує, як організовано зберігання даних у базі, які атрибути та методи реалізує кожен клас, а також яким чином відбувається спадкування та асоціація між ними. На її основі можливо перейти до етапу проектування компонентів та розроблення архітектури програмного забезпечення, що забезпечить логічну узгодженість і розширюваність системи.

РОЗДІЛ 3 РОЗРОБКА СИСТЕМИ

3.1. Архітектура системи

Розроблена експертна система для оцінювання наукової діяльності педагогічного складу побудована на основі багаторівневої архітектури, яка забезпечує логічне розділення основних процесів - від збору даних до їх аналізу та представлення результатів користувачу. Такий підхід сприяє підвищенню масштабованості та гнучкості програмного забезпечення, а також спрощує його подальшу модернізацію.

Загальна схема роботи системи

Робота системи розпочинається з етапу збору даних про наукову діяльність педагогів із різних міжнародних наукометричних баз. Зібрана інформація включає дані про кількість публікацій, цитованість, індекс Гірша та інші показники продуктивності. Після отримання даних система виконує етап уніфікації та очищення.

На наступному етапі аналітичний модуль проводить обробку даних, а саме, розраховує узагальнені наукометричні показники, формує статистику та порівнює результати між викладачами. Отримані результати зберігаються у базі даних і використовує їх для формування аналітичних звітів і рекомендацій.

Користувачі системи взаємодіють із нею через інтерфейс користувача, який дозволяє переглядати статистику, аналітичні звіти, персональні рекомендації або вносити нові дані.

Розроблена система складається з таких основних підсистем:

- Підсистема збору та оновлення даних - здійснює завантаження та оновлення інформації про наукові публікації, цитування та показники з міжнародних баз;

- Підсистема аналітичної обробки - реалізує алгоритми обчислення метрик, рейтингів, аналізу динаміки наукової активності та формування рекомендацій;
- Підсистема візуалізації - забезпечує подання результатів у вигляді таблиць, діаграм і звітів для зручного сприйняття інформації користувачами;
- Підсистема збереження даних - відповідає за структуроване зберігання даних та доступ до них.

Таким чином, архітектура системи реалізує чітку логіку обміну інформацією між підсистемами, де кожен компонент виконує свою функцію - від збору даних до аналітики та представлення результатів у зручному вигляді.

3.2. Модуль збору інформації з міжнародних джерел

Підсистема збору даних є одним з найголовніших компонентом експертної системи, оскільки саме від її ефективної роботи залежить повнота, точність та об'єктивність оцінки наукової діяльності педагогів. Основним завданням цього модуля є отримання та уніфікація інформації про наукові досягнення викладачів з різноманітних джерел. Система пропонує різні режими роботи - повністю автоматичний збір даних та напівавтоматичний режим, який передбачає участь користувачів у процесі введення або редагування інформації.

Автоматичний збір даних реалізується через інтеграцію з провідними міжнародними наукометричними базами даних, наприклад, Scopus або Google Scholar. Використовуючи сучасні API та передові механізми веб-парсингу, система здійснює збір структурованої інформації про публікаційну активність, показники цитувань та інші. Після отримання первинних даних система проводить перевірку унікальності записів та ідентифікацію типу наукової роботи.

Для збору даних з джерел, які не надають публічного API, система використовує спеціалізований парсер на основі Python, що імітує поведінку реального користувача. Цей модуль застосовує бібліотеки Selenium та

BeautifulSoup для автоматизації взаємодії з веб-ресурсами, що дозволяє обходити обмеження на автоматичний доступ. Парсер виконує послідовні дії, характерні для людини, а саме перехід по посиланнях, очікування завантаження контенту, прокручування сторінок та клік по елементах інтерфейсу, забезпечуючи тим самим надійний збір необхідної інформації.

У випадках, коли автоматичний збір даних виявляється неможливим або отримана інформація потребує додаткового уточнення, система передбачає можливість ручного введення даних через спеціально розроблений інтерфейс користувача. Адміністратори системи або самі викладачі мають змогу додавати нові записи про наукові публікації, а також вносити корективи в існуючі дані. Для оптимізації цього процесу система надає інтуїтивно зрозумілі форми введення з обов'язковими полями для всіх необхідних атрибутів.

Після завершення процесу збору та попередньої обробки відбувається завантаження інформації до центральної бази даних, де вона стає доступною для подальшого аналітичного опрацювання. Система реалізує розширені механізми перевірки інформації - коли одна й та сама публікація фіксується в кількох джерелах, система автоматично ідентифікує потенційні дублікати та об'єднує їх у єдиний коректний запис, забезпечуючи цілісність даних.

Таким чином, підсистема збору даних забезпечує комплексне формування повної, структурованої та достовірної інформаційної бази, що є критично важливою передумовою для формування об'єктивних аналітичних показників та науково обґрунтованих рекомендацій. Цей модуль грає ключову роль у підтримці безперервного оновлення наукових даних та забезпеченні актуальності оцінювання педагогічного складу, що в кінцевому підсумку сприяє підвищенню якості науково-дослідної діяльності у закладі вищої освіти.

3.3. Модуль обробки інформації та формування показників оцінювання

Для забезпечення об'єктивного оцінювання наукової діяльності педагогічних працівників система реалізує комплексну модель, що базується на взаємопов'язаних метриках, а саме, публікаційна активність, цитування, індекс Гірша (h-index), індекс i10.

Кожна з метрик нормалізується, зважується відповідно до її значущості, а потім сумується у єдину фінальну оцінку викладача.

Метрика «Публікації» є показником наукової продуктивності викладача. Вона визначає не лише кількість робіт, але й якість публікацій, рівень видань, у яких вони опубліковані, тип наукового матеріалу (монографія, стаття, тези тощо) та актуальність у часі. Вона допомагає оцінити внесок викладача у розвиток науки через опубліковані результати, стимулюючи публікацію якісних матеріалів у міжнародних джерелах (Ваговий коефіцієнт - 0.25).

$$M_1 = \frac{1}{N} \sum_{i=1}^N (QF_i * TF_i * DF_i), \quad (3.1)$$

- Де N - загальна кількість публікацій;
- QF_i - коефіцієнт якості журналу, який завантажується з Scimago Journal & Country Rank;
- TF_i - коефіцієнт типу публікації (Наукова стаття – 1, Стаття – 0.8, Монографія – 0.9, Тези – 0.6);
- DF_i - коефіцієнт давності, що враховує актуальність публікації, $DF_i = e^{-0,1(Y_{curr}-Y_i)}$, де e - стале значення 2,71, Y_{curr} - це поточний рік, Y_i - рік публікації.

Метрика «Цитування» враховує не просто загальну кількість цитувань, а також середню цитованість публікацій і розподіл цитувань протягом певного проміжку часу, що дає можливість визначити, наскільки стійкий науковий інтерес до робіт викладача (Ваговий коефіцієнт - 0.35).

$$M_2 = 0,6 * \frac{\log(1+C_{total})}{\log(1+C_{max})} + 0,4 * \frac{C_{avg}}{C_{avg,max}}, \quad (3.2)$$

- Де C_{total} - загальна кількість цитувань викладача;
- C_{max} - максимальна кількість цитувань серед усіх викладачів;
- C_{avg} - середня кількість цитувань на одну публікацію;
- $C_{avg,max}$ - максимальне середнє цитування серед усіх користувачів.

Метрика «Індекс Гірша» враховується як стабілізуючий показник, що дозволяє оцінити збалансованість між кількістю публікацій і якістю їх цитування.

Індекс Гірша показує, скільки публікацій автора мають не менше ніж h цитувань кожна. Тобто, якщо у дослідника є 5 статей, і кожна з них процитована щонайменше 5 разів, тоді його $h = 5$. Цей показник береться напряду через міжнародні джерела даних (Ваговий коефіцієнт - 0.25).

Метрика «Індекс i10» - це бібліометричний показник, який був запроваджений науковою соціальною мережею Google Scholar як простий, але ефективний інструмент для оцінки високоцитованих публікацій дослідника. Формальне визначення звучить так - це кількість публікацій дослідника, які отримали щонайменше 10 цитувань кожна (Ваговий коефіцієнт - 0.15).

На відміну від більш складних метрик (наприклад, h -індексу, який поєднує продуктивність і вплив у єдине число), $i10$ -index прагне дати дуже конкретну і легку для розуміння відповідь на питання - «Скільки значних, помітних наукою робіт зробив цей учений?».

Формула зведення рейтингу викладача обчислюється як зважена комбінація усіх метрик.

$$R = \sum_{i=1}^4 (w_i * M_i), \quad (3.3)$$

- Де w_i – вагові коефіцієнти метрик;
- M_i – обчислений рейтинг кожної метрики.

Отримане значення нормалізується в інтервалі [0;1] та інтерпретується за такими рівнями, після чого конвертується в 10-бальну систему і зберігається в БД:

- 0.00 - 0.39 - початковий рівень наукової активності;
- 0.40 - 0.69 - середній рівень;
- 0.70 - 0.89 - високий рівень;
- 0.90 - 1.00 - відмінний рівень.

3.4. Модуль зберігання даних

3.4.1. Створення бази даних

База даних є ключовим елементом експертної системи оцінки наукової діяльності педагогічного складу. Вона забезпечує централізоване зберігання, обробку, систематизацію та подальший аналіз інформації, отриманої з міжнародних наукометричних джерел.

Основна мета побудови бази даних - створити структуру, яка дозволяє об'єктивно відображати наукову активність кожного викладача, забезпечуючи при цьому гнучкість у розрахунках та цілісність інформації.

База даних виступає головним компонентом, що зв'язує між собою всі підсистеми експертної системи. Вона зберігає персональні дані педагогів, їхні

публікації, результати аналітичних розрахунків та рекомендації, сформовані системою. Рис.3.4.1.1.

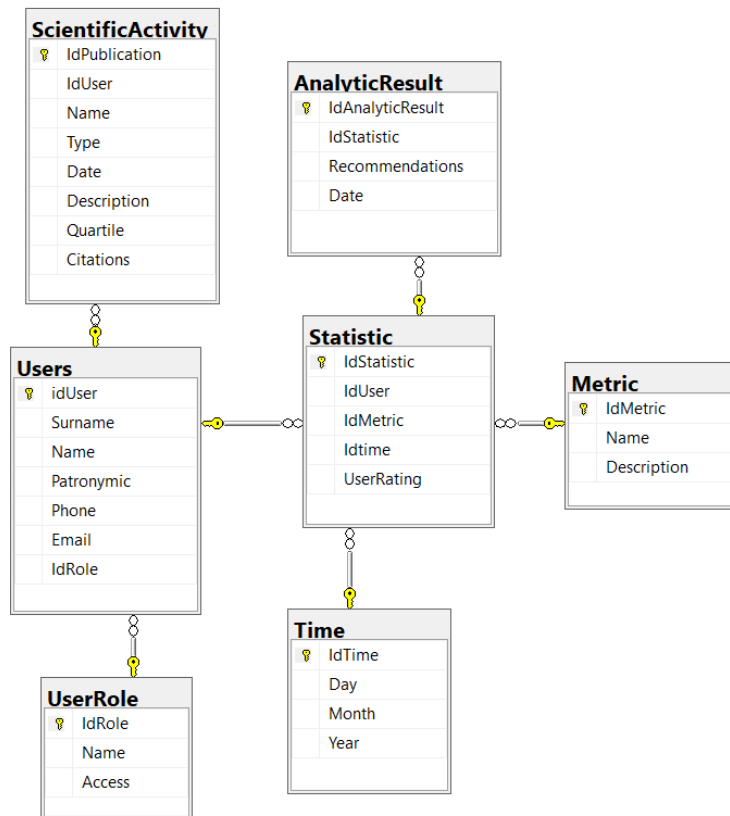


Рис. 3.4.1.1 Схема бази даних

Таблиця UserRole

Ця таблиця містить інформацію про ролі користувачів у системі. Вона необхідна для реалізації механізму авторизації та розмежування прав доступу до функціональних можливостей.

1. IdRole - унікальний ідентифікатор ролі;
2. Name - назва ролі (Адміністратор, Аналітик, Педагог);
3. Access - рівень доступу, який визначає, які дії може виконувати користувач.

Таблиця є базовою для побудови системи прав доступу, що гарантує безпечну роботу з персональними даними.

Таблиця Users

Таблиця Users зберігає інформацію про всіх користувачів системи, включаючи педагогів, аналітиків та адміністраторів. Вона є однією із найголовніших таблиць в базі даних, оскільки більшість зв'язків реалізовано саме через неї.

1. IdUser - унікальний ідентифікатор користувача;
2. Surname, Name, Patronymic - персональні дані користувача;
3. Phone, Email - контактна інформація;
4. IdRole - зовнішній ключ, який вказує на роль користувача у таблиці UserRole.

Таблиця забезпечує можливість фіксування індивідуальних наукових досягнень кожного викладача та їх подальшої оцінки системою.

Таблиця ScientificActivity

Таблиця зберігає відомості про наукову діяльність педагогів - публікації, тези, статті, монографії тощо. Вона є ключовим джерелом даних для розрахунку метрик, пов'язаних з публікаційною активністю.

1. IdPublication - унікальний ідентифікатор запису;
2. IdUser - ідентифікатор автора (зовнішній ключ до Users);
3. Name - назва публікації або наукової роботи;
4. Type - тип наукової активності (стаття, тези, книга, доповідь тощо);
5. Date - дата публікації або представлення роботи;
6. Quartile - кuartиль видання за класифікацією Scopus (Q1–Q4);
7. Description - короткий опис, посилання на джерело, DOI або інше;
8. Citations – кількість цитувань на публікацію.

Ця таблиця дозволяє відстежувати динаміку наукової активності кожного викладача та враховувати якісні характеристики публікацій (зокрема, престижність журналу через кuartиль).

Таблиця Metric

Таблиця містить перелік усіх показників, за якими здійснюється оцінка наукової діяльності. Вона виконує роль керування аналітичними критеріями, що можуть бути розширені або оновлені для зміни структури оцінювання всієї системи.

1. IdMetric - унікальний ідентифікатор метрики;
2. Name - назва метрики (Публікації, Цитування, Індекс Гірша, i10-index);
3. Description - короткий опис метрики.

Таблиця забезпечує можливість зміни критеріїв оцінювання.

Таблиця Statistic

Таблиця Statistic містить розраховані кількісні значення для кожного користувача за кожною метрикою.

1. IdStatistic - унікальний ідентифікатор запису;
2. IdUser - посилання на користувача, для якого розраховано показник;
3. IdMetric - посилання на метрику, яку розраховано;
4. UserRating - рейтинг, отриманий у результаті аналізу.

Ця таблиця дозволяє здійснювати порівняльний аналіз педагогів між собою, а також формувати загальний рейтинг наукової активності.

Таблиця AnalyticResult

Таблиця зберігає кінцеві результати роботи системи - висновки та рекомендації щодо підвищення продуктивності наукової діяльності викладачів.

1. IdAnalyticResult - унікальний ідентифікатор запису;
2. IdStatistic - посилання на запис із таблиці Statistic;
3. Recommendations - автоматично сформовані поради або написані рекомендації аналітика;
4. Date - дата створення рекомендацій.

Ця таблиця є також дуже важливою, адже саме на її основі користувач отримує висновки по його діяльності.

Тому можна сказати, що спроектована база даних виступає одним з найголовніших компонентів експертної системи оцінки наукової діяльності, забезпечуючи цілісне зберігання та ефективну обробку інформації.

3.4.2. Створення сховища даних

Сховище даних (Data Warehouse) - це спеціалізована інформаційна система, призначена для зберігання та аналітичної обробки великих обсягів даних. На відміну від звичайних баз даних, які орієнтовані на оперативну діяльність (запис, зміну чи видалення окремих записів), сховище даних призначене для довгострокового збереження історичних даних, виконання складних аналітичних запитів, побудови звітів, а також підтримки процесів прийняття рішень.

Основна концепція DWH полягає в об'єднанні різнорідних даних у єдину структуру, яка дозволяє виконувати аналітичні операції без навантаження на оперативні системи. У контексті експертної системи оцінки наукової діяльності педагогів, воно забезпечує формування об'єктивних показників продуктивності наукової роботи викладачів на основі даних з різних міжнародних джерел.

Служби аналізу SQL Server (SSAS) - це технологія Microsoft, призначена для створення аналітичних рішень та обробки великих обсягів даних у форматі OLAP (Online Analytical Processing). Основна мета SSAS полягає у забезпеченні швидкого доступу до даних і можливості багатовимірного аналізу, що є особливо корисним для прийняття управлінських рішень та побудови систем бізнес-аналітики (BI).

SSAS дозволяє створювати багатовимірні моделі даних, у яких інформація представлена через виміри і показники. Це дає змогу користувачам аналізувати різні дані, наприклад, за роками, викладачами, напрямками досліджень чи кількістю публікацій. Також SSAS підтримує табличний режим роботи, який

базується на технології колонкових баз даних і забезпечує ще швидшу обробку запитів.

Під час розробки куба використовувалося середовище Visual Studio із розширенням для SSAS. На початковому етапі визначається джерело даних, після чого виконується створення вимірів. Рис. 3.4.2.1–3.4.2.2.

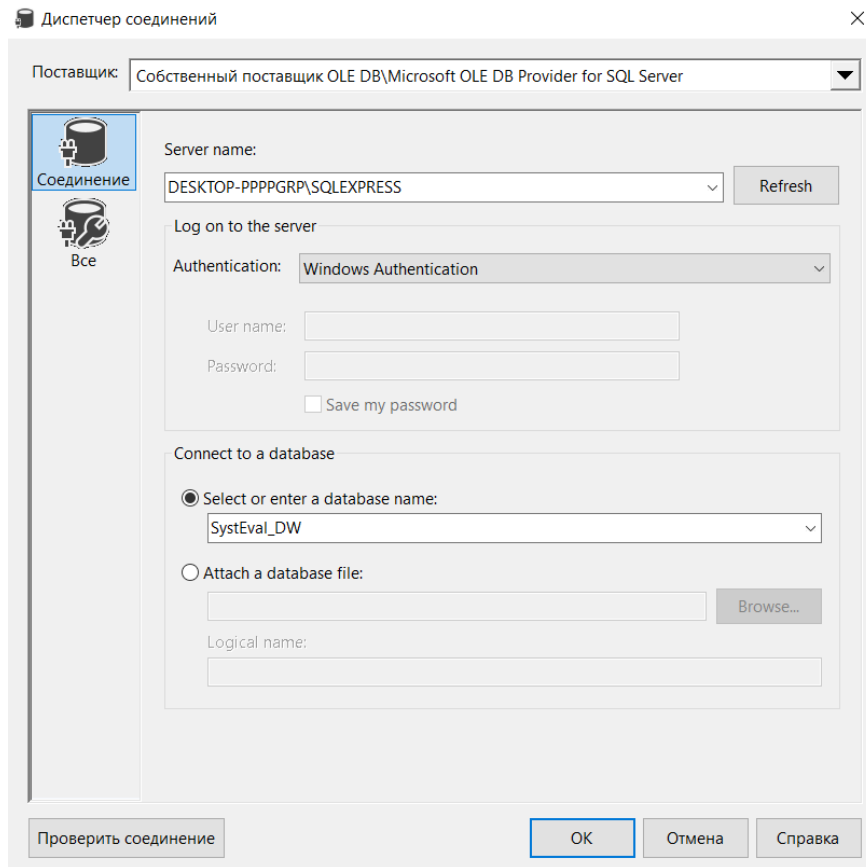


Рис.3.4.2.1 Процес налаштування підключення до джерела даних

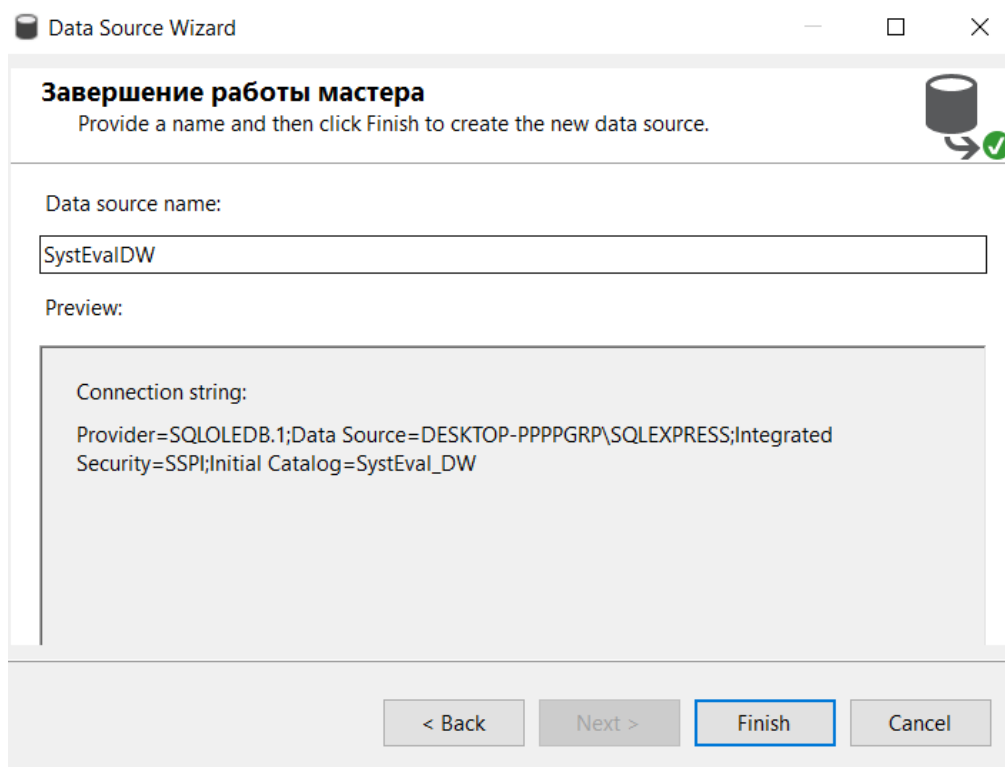


Рис.3.4.2.2 Підтвердження вибору налаштованого підключення

Наступним етапом побудови куба є створення уявлення джерела даних (Data Source View). DSV забезпечує зручну роботу з даними, виступаючи як віртуальне представлення структури джерела. Воно дозволяє організувати таблиці та підготувати дані для побудови кубів і вимірів у багатовимірному проекті.

Основне призначення DSV полягає у спрощенні роботи з даними без необхідності змінювати вихідну базу. Завдяки цьому можна перейменовувати атрибути, створювати обчислювані стовпці, об'єднувати таблиці або застосовувати фільтри, не впливаючи на початкове джерело.

У межах одного проекту можна створювати декілька уявлень джерел даних, що базуються на різних джерелах або їх комбінаціях. Це дозволяє адаптувати структуру даних під конкретні аналітичні завдання. Рис.3.4.2.3 - 3.4.2.4.

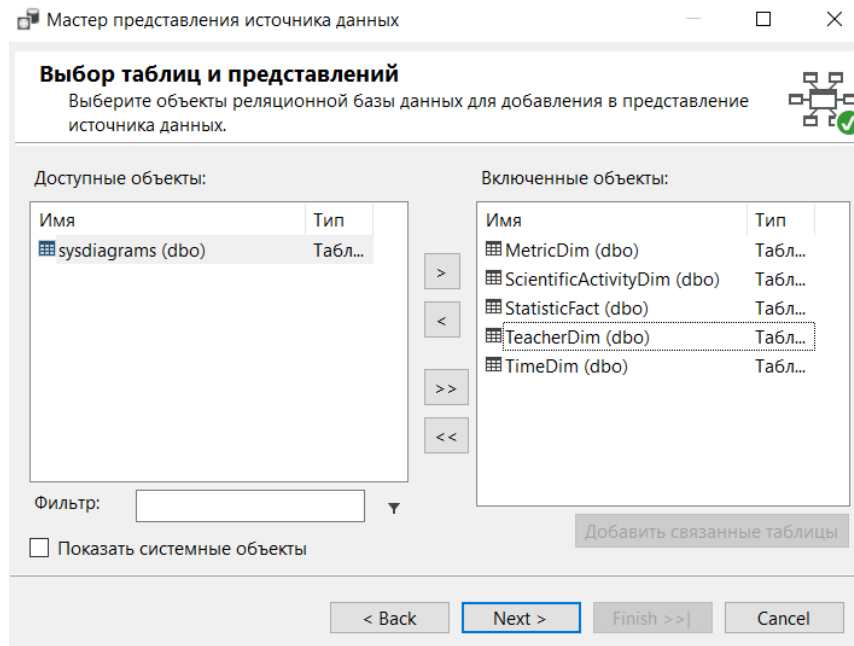


Рис.3.4.2.3 Вибір потрібних таблиц на основі джерела SysEvalDW

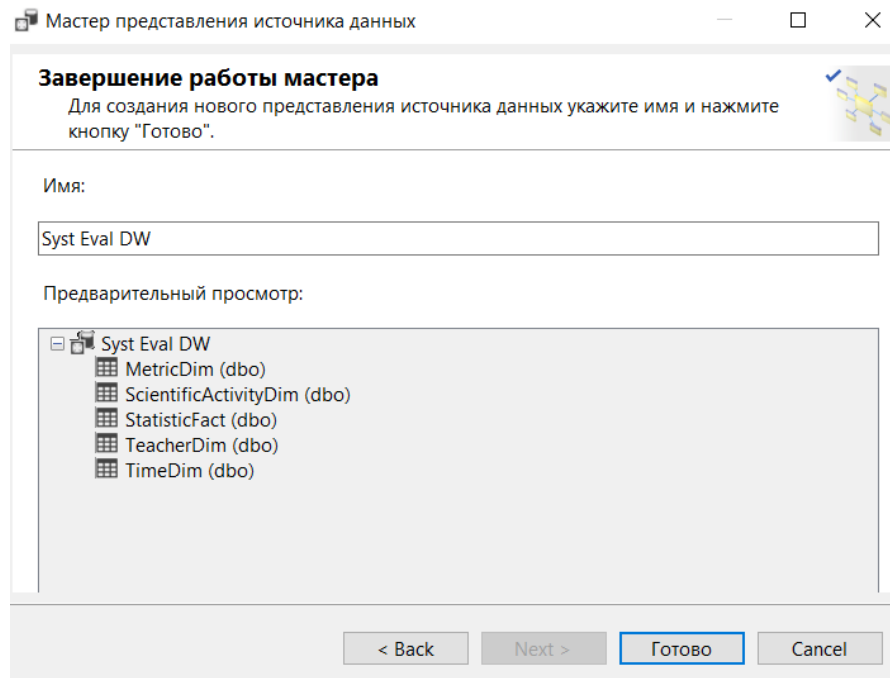


Рис.3.4.2.4 Підтвердження вибору таблиц

Обработка куба - це один із ключових етапів роботи з аналітичною моделлю. Вона полягає у завантаженні, оновленні та агрегуванні даних із джерела у створений куб, що забезпечує можливість швидкого виконання аналітичних

запитів. Без обробки куб містить лише структуру - виміри, показники та зв'язки між ними, але не фактичні дані.

Після обробки у нас повністю готовий до використання аналітичний куб, який містить актуальні дані зі сховища та доступний для аналізу користувачами через інструменти бізнес-аналітики. Рис.3.4.2.5.

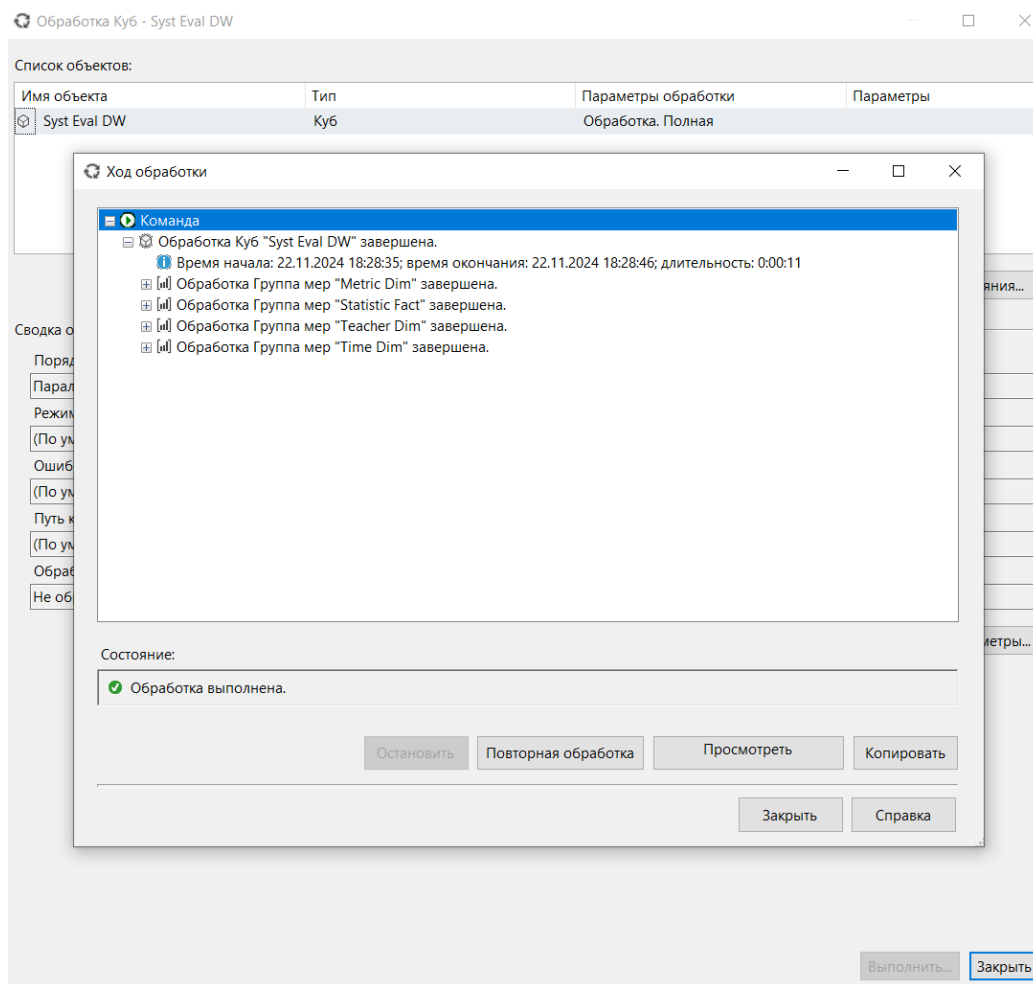


Рис.3.4.2.5 Обробка кубу

Сховище реалізовано за принципом зіркової схеми, де центральною таблицею є фактова таблиця, що містить показники оцінювання, а навколо неї розташовані вимірні таблиці, які описують контекст цих показників. Такий підхід дозволяє побудувати OLAP-куб, що дає змогу аналізувати наукову діяльність педагогів у різних розрізах - за часом, видами публікацій, рівнями цитування, та іншими метриками. Рис.3.4.2.6.

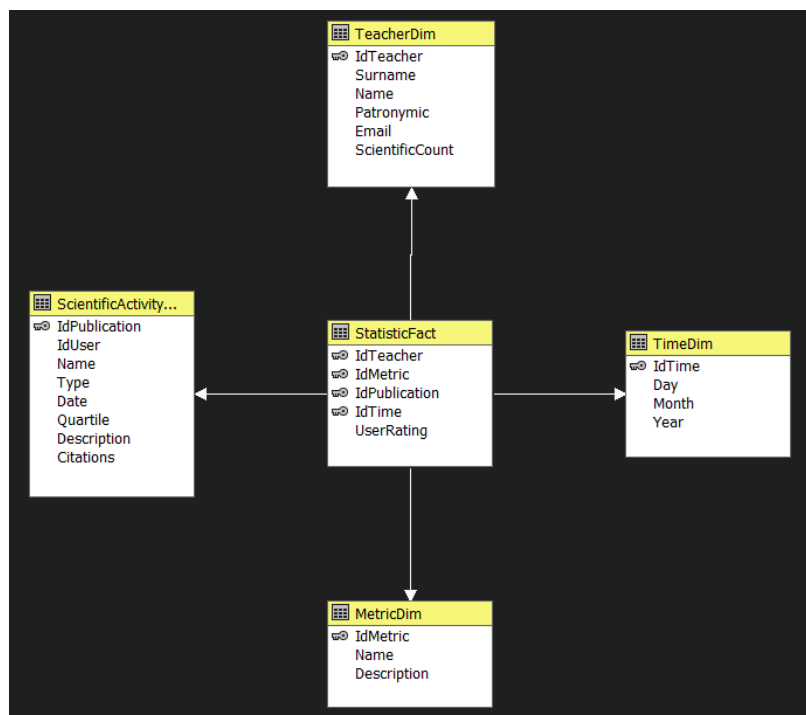


Рис.3.4.2.6 Схема сховища даних

Таблиця **TimeDim** використовується як вимір часу і зберігає детальну інформацію про часові параметри, такі як день, місяць і рік. Це дозволяє здійснювати аналітичні запити з різною глибиною часової деталізації - наприклад, аналізувати публікаційну активність за певний рік, визначати динаміку зростання індексу цитування за місяцями або оцінювати кількість наукових робіт у певний період. Таким чином, таблиця часу відіграє ключову роль у побудові тенденцій і динаміки наукових показників.

Таблиця **TeacherDim** містить дані про викладачів, наукова діяльність яких підлягає оцінюванню. До основних атрибутів цієї таблиці належать прізвище, ім'я, по батькові, електронна адреса та кількість наукових праць. Ці дані дозволяють однозначно ідентифікувати кожного педагога у системі та забезпечують можливість формування індивідуальних і колективних звітів, побудови рейтингів наукової активності, а також порівняльного аналізу між викладачами.

Таблиця **MetricDim** описує метрики, за якими здійснюється оцінювання. У ній зберігаються назви та описи показників, таких як кількість публікацій, цитування, індекс Гірша та i10-index. Ця таблиця забезпечує гнучкість системи, оскільки дозволяє легко розширювати набір аналітичних показників без необхідності змінювати загальну структуру сховища даних.

Таблиця **ScientificActivityDim** відображає вимір наукової активності викладачів. Вона містить детальну інформацію про публікації, включаючи їхню назву, тип (стаття, монографія тощо), дату публікації, квартали видання (Q1–Q4), а також можливий опис до роботи. Додатково зберігається ідентифікатор користувача, що дозволяє пов'язати кожну наукову роботу з конкретним викладачем. Ця таблиця є основним джерелом даних для аналізу структури та обсягів наукової продуктивності викладачів.

Центральне місце у сховищі даних займає фактова таблиця **StatisticFact**, у якій зберігаються показники оцінювання, отримані в результаті аналітичної обробки даних. У таблиці зберігаються зв'язки з усіма вимірними таблицями - викладачами, метриками, видами наукової активності та часом - а також обчислений показник **UserRating**, який відображає значення метрик для певного викладача в заданий період часу. Завдяки цьому структура сховища дозволяє виконувати багатовимірний аналіз, наприклад - оцінити динаміку наукових публікацій за останні три роки, визначити середній рівень цитування чи порівнювати загальні показники викладачів.

Таким чином, розроблене сховище даних забезпечує не лише централізоване зберігання аналітичної інформації, але й створює основу для побудови OLAP-кубу, який надає можливість глибокого багатовимірного аналізу. Його використання підвищує достовірність і оперативність оцінювання наукової діяльності педагогів, що, у свою чергу, сприяє прийняттю обґрунтованих рішень щодо розвитку наукового потенціалу закладу вищої освіти.

3.4.3. Заповнення сховища даних

Для управління процесами внесення даних було застосовано SQL Server Integration Services (SSIS). Він надає комплексний інструментарій для створення процесів внесення даних, забезпечуючи надійне транспортування та трансформацію інформації між різними джерелами.

Ключовим компонентом реалізації виступив Data Flow, де було налаштовано спеціалізовані потоки даних. В рамках цих потоків відбувалося не лише перенесення, але й активне перетворення даних для подальшого завантаження в цільові таблиці вимірів і фактів. Візуальне представлення усіх потоків обробки даних зображено на рисунках 3.4.3.1 - 3.4.3.2.



Рис. 3.4.3.1 Загальна структура потоків

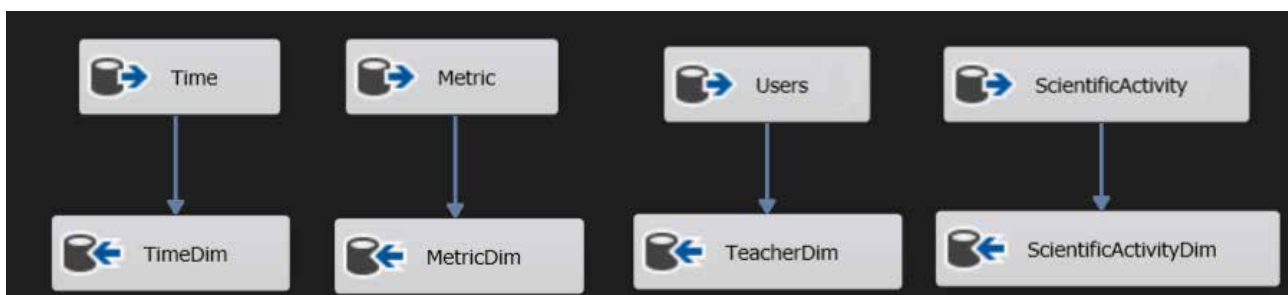


Рис. 3.4.3.2 Структура потоків наповнення вимірів

Для забезпечення точності даних далі співставлено стовпців для кожної таблиці. Цей процес є критично важливим етапом, оскільки саме він визначає, як дані трансформуються та інтегруються в цілісну структуру вимірів. Рис.3.4.3.3.

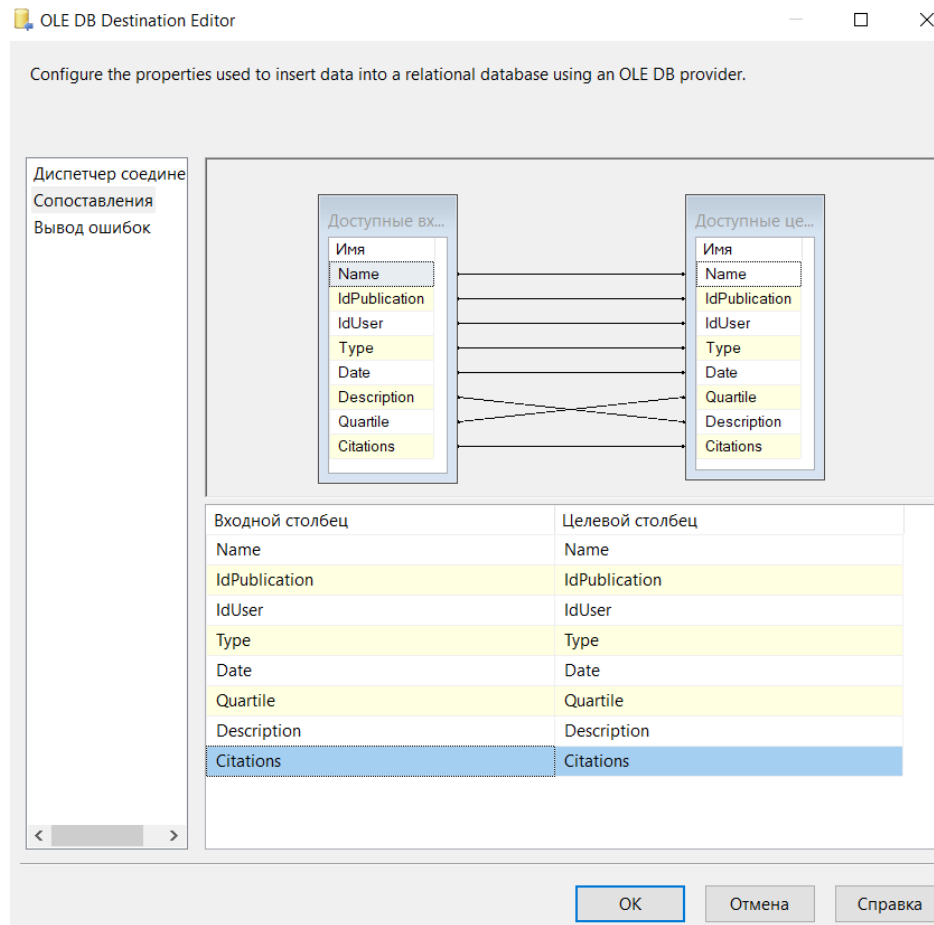


Рис. 3.4.3.3 Встановлення зв'язків між відповідними таблицями

Наступним етапом стало наповнення таблиць фактів, яке, хоча і має схожість з заповненням вимірів, має ключові відмінності в механізмі отримання даних. Якщо для вимірів дані надходили з відповідних джерел, то для фактів основними джерелами виступили вже сформовані таблиці вимірів. Таким чином, заповнення фактів є фінальним етапом побудови сховища даних, де всі елементи - виміри та факти - починають функціонувати як єдина аналітична система.

Рис.3.4.3.4 - 3.4.3.6.

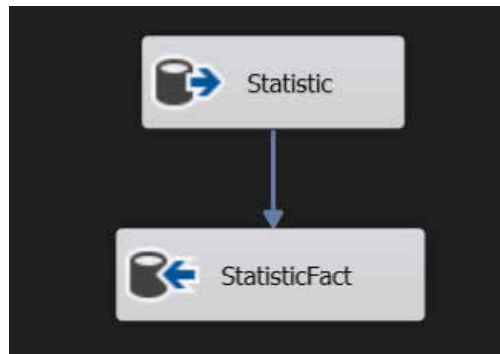


Рис.3.4.3.4 Вигляд потоку заповнення фактів

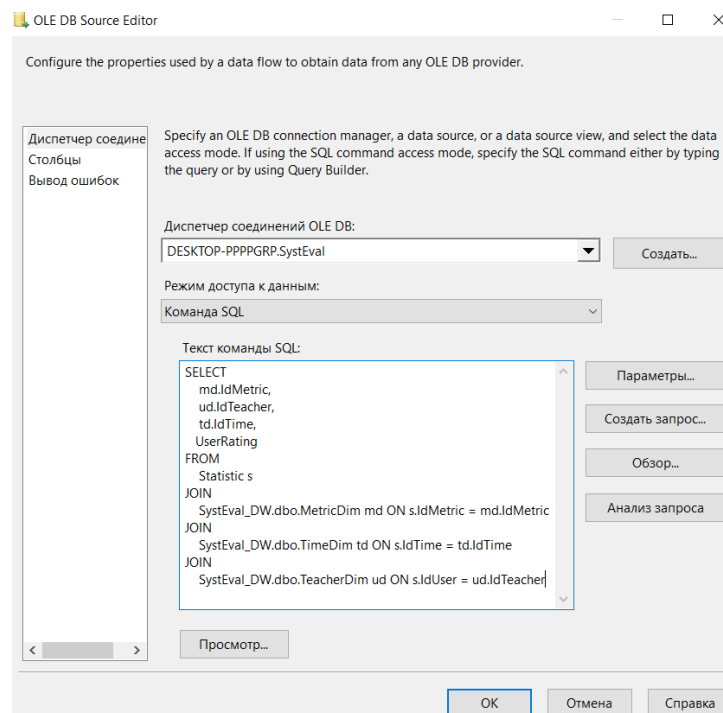


Рис.3.4.3.5 Вибірка для заповнення таблиці фактів даними з вимірів

Після виконання всіх попередніх етапів, дані завантажуються з бази до сховища даних. На цьому етапі інформація з різних таблиць об'єднується та структурується. Це забезпечує цілісність і узгодженість даних, що дозволяє в подальшому проводити якісні аналітичні аналізи. Рис.3.4.3.6.



Рис.3.4.3.6 Успішний процес занесення даних в СД

РОЗДІЛ 4 РЕЗУЛЬТАТИ ДОСЛІДЖЕННЯ

4.1. Вимоги до апаратного та програмного забезпечення

Для коректної роботи розробленої експертної системи оцінювання наукової діяльності педагогічного складу необхідно визначити вимоги до апаратного та програмного забезпечення, що забезпечать стабільність та продуктивність обробки даних. Оскільки система використовує алгоритми кластеризації, аналізу наукометричних показників, обробку структурованих даних та формування звітів, апаратна частина має бути здатною виконувати ці операції без затримок, а програмна частина забезпечувати відповідне середовище виконання.

Апаратні вимоги

Апаратне забезпечення повинно мати достатню обчислювальну здатність для виконання експертної оцінки, обробки масивів даних та візуалізації результатів.

Мінімальні рекомендовані характеристики для робочої станції

- Багатоядерний CPU з частотою не менше 2.0 ГГц (рекомендується Intel Core i5 / AMD Ryzen 5 або аналог). Це забезпечує достатній ресурс для виконання обчислень під час кластеризації та статистичного аналізу;
- Для стабільної роботи, потрібна оперативна пам'ять, мінімум 8 ГБ, рекомендовано 16 ГБ. Достатній обсяг ОЗП необхідний для роботи з таблицями даних та запобігання уповільнення під час інтенсивних обчислень;
- Для швидкодії системи рекомендується SSD накопичувач обсягом від 256 ГБ. Твердотільний накопичувач забезпечує високу швидкість доступу до даних та зменшує час виконання операцій читання/запису;
- Щодо графічного адаптеру, інтегрованої графіки достатньо, оскільки система не використовує GPU-прискорення. Однак наявність дискретного GPU може прискорити генерацію графіків при роботі з великими вибірками;

- Мережеве підключення також є обов'язковим, стабільний доступ до Інтернету зі швидкістю від 10 Мбіт/с, якщо передбачається отримання наукометричних даних з відкритих міжнародних джерел (Google Scholar, Scopus, Web of Science).

Програмні вимоги

Для розробки, підтримки та тестування аналітичних рішень застосовуються середовища керування SQL Server, зокрема SQL Server Management Studio (SSMS). SSMS забезпечує адміністрування бази даних та виконання SQL-запитів.

Окремим програмним компонентом є Python 3.10+, який застосовується для додаткової обробки даних, реалізації методів експертної системи та створення аналітики. Для роботи в Python використовуються бібліотеки обробки даних (pandas, numpy), аналітики та експертної оцінки, а також засоби візуалізації (matplotlib, seaborn).

Для зручності розробки та супроводу системи використовується професійне середовище програмування, таке як Visual Studio Code, яке забезпечує можливість швидкої обробки коду, інтеграції з Git та підключення до серверів баз даних.

У комплексі ці апаратні та програмні ресурси дозволяють забезпечити стабільну роботу системи, швидку генерацію аналітичних звітів, автоматизацію процесів збору та збереження даних та масштабування платформи відповідно до потреб навчального закладу.

4.2. Хід виконання дослідження

Процес проведення дослідження складався з кількох основних етапів, кожен з яких мав своє цільове призначення та був спрямований на поступове створення, перевірку та вдосконалення експертної системи для оцінювання наукової діяльності педагогічного складу.

На першому етапі було проведено аналіз предметної області та виявлено основні показники, що характеризують результативність наукової діяльності викладачів. Було визначено набір ключових метрик, які можуть об'єктивно відображати рівень наукової активності, а саме кількість публікацій, кількість цитувань, індекс Гірша та i10-index. Для кожної метрики розроблено окремі методи збору та нормалізації.

На другому етапі здійснювалося проектування структури бази даних та сховища даних для забезпечення зручного зберігання, аналітики та подальшої побудови OLAP-куба. Було розроблено модель бази даних, що включає таблиці користувачів, наукової активності, метрик, статистичних показників та аналітичних результатів. Для підтримки багатовимірної аналізи створено окреме сховище даних із вимірами часу, викладачів, наукових активностей та метрик. Центральною частиною сховища стала фактова таблиця, у якій накопичуються результати оцінювання викладачів за всіма визначеними критеріями.

Далі було реалізовано підсистему збору даних. Для цього використовувалися відкриті API міжнародних наукометричних баз даних (Google Scholar, Scopus), а також методи веб-парсингу на Python (бібліотеки Selenium та BeautifulSoup), що дозволили автоматизувати процес отримання інформації про публікації, цитування та інші показники. Отримані дані проходили етапи очищення та уніфікації перед завантаженням до бази даних.

Після реалізації збору даних, було розроблено модуль аналітики, який реалізує алгоритми обчислення рейтингових показників педагогів. Кожна метрика оцінювалася окремо з урахуванням її ваги, після чого формувався показник наукової результативності. Для цього використовувалася формула з нормалізацією даних та обчисленням середніх значень. Додатково проводилося зведення результатів у OLAP-куб, що дозволило виконувати порівняльний аналіз діяльності окремих викладачів у різних часових зрізах.

Передостаннім етапом було створення підсистеми рекомендацій. На основі отриманих аналітичних результатів система формувала автоматичні поради для користувачів, наприклад, рекомендації щодо підвищення цитованості або публікацій у журналах з вищим квантилем. Цей модуль реалізовано за принципом експертних правил, які визначають зв'язок між поточними показниками викладача та можливими шляхами їх покращення.

Завершальним етапом став етап тестування та оцінки ефективності системи. Проводилося тестування працездатності всіх модулів, від перевірки правильності імпорту даних до достовірності сформованих рекомендацій. Для цього використовувалися як реальні дані викладачів, так і тестові набори, змодельовані на основі типових сценаріїв наукової активності. У результаті було підтверджено, що система здатна коректно збирати, обробляти та аналізувати інформацію, формуючи об'єктивну оцінку наукової діяльності педагогів.

Таким чином, проведене дослідження охоплює повноцінну реалізацію експертної системи - від аналізу предметної області та збору даних до побудови аналітичних моделей, формування рекомендацій та тестування функціональності. Розроблена система дозволяє автоматизувати процес оцінювання наукової діяльності викладачів, забезпечуючи комплексний підхід до аналізу публікаційної активності, цитованості та інших показників. Вона є допоміжним інструментом підтримки прийняття рішень для адміністрації закладів освіти, спрямованим на підвищення якості та ефективності науково-дослідної роботи педагогічного складу.

4.3. Аналіз результатів роботи

Після розроблення експертної системи оцінки наукової діяльності педагогічного складу було проведено тестування її функціональності та перевірку коректності реалізованих алгоритмів розрахунку показників. Основною метою даного етапу стало оцінювання здатності системи об'єктивно

аналізувати наукову продуктивність викладачів на основі даних, отриманих із міжнародних наукометричних джерел.

Під час дослідження було сформовано вибірку з кількох викладачів, для яких зібрано показники публікаційної активності, цитованості, індексу Гірша та і10-індексу для їх подальшого аналізу. Особлива увага приділялася перевірці відповідності вагових коефіцієнтів кожної метрики.

Отримані результати дозволили провести всебічний аналіз наукової активності викладачів, виявити закономірності між кількістю публікацій, рівнем цитованості та індексами впливу.

4.3.1. Оцінка публікаційної активності та цитування

Для перевірки роботи експертної системи, а саме, відповідності присвоєного системою рейтингу та фактичної наукової діяльності викладачів, було обрано 5 випадкових спеціалістів з подальшим аналізом їхньої діяльності.

Таблиця 4.1

Таблиця базових показників наукової діяльності

Викладач	Публікації	Цитування	Середня цитованість	Визначений рейтинг
Ткаченко В.П.	162	785	4,84	6.3
Ковпак В.А.	89	115	1,29	3.7
Зварич В.О.	113	721	6,38	6.2
Степаненко А.В.	39	1303	33,41	6.4
Кропивко М.Ф.	62	1289	20,79	7.1

Як видно з таблиці, між кількістю публікацій і цитувань існує певна кореляція, однак не завжди більша кількість публікацій означає вищу якість або впливовість досліджень. Наприклад, Степаненко А.В. має всього 39 публікацій, але дуже високий показник середньої цитованості - 33,41, що свідчить про значний вплив кожної його роботи на наукову спільноту. Аналогічно, Кропивко

М.Ф. демонструє високі значення як за кількістю цитувань, так і за середньою цитованістю, що закономірно впливає на отримання високого рейтингу - 7,1.

Водночас Ковпак В.А. має порівняно велику кількість публікацій (89), але низький рівень цитування (115), що вказує на потребу підвищення якості або міжнародної видимості досліджень.

Також в таблиці видно що у Зварича В.О та Степаненка А.В. приблизно однаковий рейтинг, але є значна різниця між кількістю робіт та їх якістю, в свою чергу це свідчить про те, що система опирається не лише на одиничні показники але розглядає роботу викладача в цілому за багатьма факторами.

Оцінювання наукової діяльності викладачів у системі проводиться на основі багатокритеріального підходу, який дозволяє врахувати не лише кількість наукових робіт, а й якість, вплив та актуальність досліджень. Для цього використовується модель, що складається з чотирьох основних компонентів, а саме, публікаційна активність, цитованість, індекс Гірша та показник i10. Кожен із цих параметрів має власну вагу, яка відображає його відносний вплив на загальну оцінку. Зокрема, публікації мають вагу 0.25, цитування - 0.35, індекс Гірша – 0.25, а i10-index - 0.15.

Оцінювання публікаційної активності (M1)

Показник публікаційної активності відображає кількість і якість наукових робіт викладача. Для його обчислення враховується тип публікації (наукова стаття, монографія, тези тощо), якість журналу, у якому розміщено публікацію, а також її давність.

Публікації з високим квантилем журналу (Q1 - Q4 за Scimago) мають більшу вагу в оцінці, ніж роботи, опубліковані в менш рейтингових виданнях. Також враховується актуальність, більш свіжі публікації мають вищу оцінку, ніж ті, що були опубліковані багато років тому. Це дозволяє стимулювати постійну наукову активність викладачів, а не лише накопичення історичних публікацій.

Оцінювання цитованості (M2)

Другим важливим показником є рівень цитованості. Він характеризує науковий вплив робіт викладача в міжнародному науковому середовищі. Для цього порівнюється загальна кількість цитувань викладача з максимальним показником серед усіх досліджуваних осіб, а також враховується середня кількість цитувань на одну публікацію. Таким чином, оцінка враховує як загальну впливовість науковця, так і якість кожної окремої роботи.

Більш висока кількість цитувань свідчить про значну зацікавленість наукової спільноти у працях викладача, що підвищує його загальну оцінку. Тому, для викладачів із невеликою кількістю публікацій, але високим середнім показником цитованості, система також забезпечує відповідне підвищення рейтингу, що гарантує збалансованість оцінювання.

Оцінювання за індексом Гірша (M3)

Індекс Гірша використовується для відображення стабільності та довготривалості наукової продуктивності. Він показує, скільки публікацій автора мають принаймні таку ж кількість цитувань. Чим вищий індекс, тим більш системною є наукова діяльність викладача.

Для оцінювання показники h-індексу нормалізуються відносно максимального значення серед усіх викладачів. Ця метрика дозволяє об'єктивно порівняти науковців з різними масштабами діяльності та стажем роботи.

Оцінювання за показником i10-index (M4)

Показник i10-index характеризує кількість публікацій, які отримали понад 10 цитувань. Цей показник є додатковим індикатором сталого впливу науковця у своїй галузі. Чим більше таких публікацій, тим більший внесок у розвиток наукової спільноти. Як і у випадку з індексом Гірша, значення i10-index нормалізується відносно максимального серед усіх викладачів.

Нормалізація та підсумковий рейтинг

Після розрахунку всіх чотирьох метрик їх результати зважуються відповідно до заданих ваг і підсумовуються для отримання фінальної оцінки наукової діяльності. Отриманий результат нормалізується у межах від 0 до 1, після чого конвертується у 10-бальну шкалу для зручності сприйняття та подальшого порівняння.

Для демонстрації роботи системи оцінювання було використано реальні дані, отримані з міжнародних наукометричних джерел. Нижче наведено приклад розрахунку фінальної оцінки для п'яти викладачів.

Таблиця 4.2

Таблиця всіх ключових показників для обрахунку рейтингу викладачів

Викладач	Публікації	Цитування	h-індекс	i10-індекс	Підсумкова оцінка	Рік з найбільшою продуктивністю
Ткаченко В.П.	162	785	16	24	6.3	2017р.
Ковпак В.А.	89	115	5	1	3.7	2014р.
Зварич В.О.	113	721	15	24	6.2	2025р.
Степаненко А.В.	39	1303	16	17	6.4	2005р.
Кропивко М.Ф.	62	1289	18	22	7.1	2010р.

Отримані результати демонструють, що найвищу оцінку (7.1 балів) отримав викладач Кропивко М.Ф., який має значний рівень цитованості та найвищий індекс Гірша серед усіх. Його наукова діяльність відзначається як кількістю, так і якістю публікацій, що підтверджується стабільними показниками h- та i10-індексу.

Степаненко А.В. та Зварич В.О. мають схожі підсумкові бали (6.4 і 6.2 відповідно). Загалом, Степаненко характеризується вищою цитованістю, що відповідно компенсує меншу кількість публікацій, це мало би свідчити про вищий рейтинг чим він є насправді. Але беручи до уваги що основна наукова діяльність Степаненко А.В. не є надто сучасною, відповідно система знизилася ваговий вплив цих робіт. Через це фінальний результат викладача не є надто високим, хоча інші показники відображають гарні результати. Це свідчить про значну наукову цінність його робіт в минулому.

Ткаченко В.П. отримав оцінку 6.3, що обумовлено великою кількістю публікацій та стабільною цитованістю. Високі показники індексу Гірша та i10 також підтверджують продуктивності його наукової діяльності.

Ковпак В.А. має найнижчу оцінку (3.7), що пояснюється низькими показниками цитованості та обмеженим впливом її наукових публікацій у міжнародному просторі. Проте навіть у цьому випадку система дозволяє побачити загальну тенденцію до розвитку наукової активності.

Для аналізу продуктивності наукової діяльності викладача Зварича В.О. було побудовано графік динаміки зміни основних наукометричних показників за останнє десятиріччя. Графік показує розвиток ключових метрик по кожному році.

Рис.4.3.1.1.

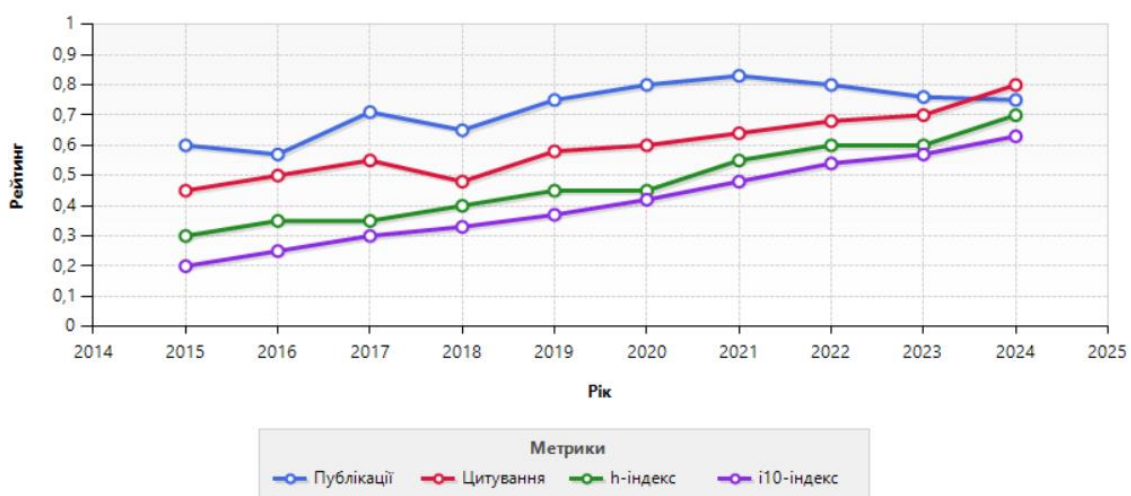


Рис.4.3.1.1 Тенденція зміни наукової діяльності викладача протягом часу

Згідно з графіком, показник публікацій протягом аналізованого періоду демонструє стабільне зростання. Починаючи з 2015 року значення становило 0.6, і до 2024 року підвищилось до 0.75. Це свідчить про поступове збільшення публікаційної активності викладача, а також про стабільну наукову продуктивність.

Показник цитувань також має позитивну динаміку. Його значення зросло з 0.45 у 2015 році до 0.8 у 2024 році. Така тенденція вказує на підвищення якості наукових робіт та збільшення інтересу до них у науковому середовищі.

Щодо h-індексу, його розвиток доволі плавний, від 0.3 у 2015 році до 0.7 у 2024 році. Цей індекс характеризує поєднання кількості та впливовості публікацій, тому його зростання свідчить про збільшення як кількісних, так і якісних показників наукової роботи. Аналогічно, i10-індекс, який відображає кількість публікацій із щонайменше десятьма цитуваннями, також зростає протягом періоду з 0.2 до 0.63, демонструючи поступове підвищення впливовості наукових результатів.

Отже, побудований графік чітко демонструє позитивну динаміку розвитку наукової активності викладача Зварича В.О. за останнє десятиріччя. Спостерігається зростання всіх ключових показників, що свідчить про стабільний розвиток наукової кар'єри та підвищення рівня визнання результатів досліджень у міжнародному науковому просторі. Ця візуалізація підтверджує ефективність використаної моделі оцінювання наукової діяльності та може бути застосована для подальшого порівняльного аналізу серед педагогічного складу.

Для оцінки ефективності експертної системи було проведено порівняльний аналіз її роботи з ручним підходом, який виконується відповідальним працівником. Порівняння охоплює ключові параметри, що впливають на якість, швидкість та об'єктивність процесу оцінювання наукової діяльності педагогічних працівників. У таблиці наведено усереднені показники, отримані на основі тестових сценаріїв та реалістичних умов роботи.

Таблиця 4.3

Порівняльна таблиця затрачених ресурсів

Параметр порівняння	Експертна система	Робота працівника
Час на збір інформації про викладача	1 – 2 хвилини	35 – 50 хвилин
Швидкість аналітичної обробки (KPI, рейтинг, динаміка)	5 – 10 секунд	1 – 2 години
Прозорість методів оцінки	100% (алгоритм фіксований)	50 – 90% (залежить від людського фактору)
Необхідність спеціальної фахової підготовки	базові навички роботи з системою	висока кваліфікація для аналітики та пошуку даних
Середній час формування рейтингу викладачів	~ 28 секунд	2 – 3 години

Згідно з результатами, система значно швидше збирає дані, виконує аналітичні розрахунки та формує рейтинги викладачів. Вона забезпечує повну прозорість методів оцінювання й не потребує глибокої фахової підготовки оператора. У той час ручний метод є більш трудомістким, залежить від кваліфікації працівника та характеризується нижчою швидкістю й більшою схильністю до впливу людського фактора.

4.3.2. Результати кластерного аналізу та ключових показників ефективності

Кластерний аналіз є одним із найпоширеніших методів аналітичної обробки даних, який дозволяє виявити природні групи об'єктів за схожістю їх характеристик. У контексті оцінювання наукової діяльності педагогічного складу кластеризація дає змогу згрупувати викладачів відповідно до рівня їхньої наукової продуктивності. Це дозволяє не лише формально присвоювати рейтингові оцінки, а й проводити глибший аналіз - визначати закономірності,

виявляти групи з потенціалом розвитку або навпаки, ті, що потребують підвищення наукової активності або їх якості.

Перший графік, отриманий за допомогою методу ліктя, демонструє залежність показника «inertia» (суми квадратів відстаней точок до центрів кластерів) від кількості кластерів k .

На початку крива різко спадає, що означає значне покращення якості розбиття. Після певного моменту спад стає більш плавним, утворюючи «лікоть» - це точка, у якій додавання нового кластера вже не дає суттєвого покращення.

У нашому випадку «лікоть» спостерігається в районі 4 кластерів, що й було обрано як оптимальну кількість. Такий вибір дозволяє отримати збалансований розподіл викладачів за рівнями наукової активності: від початкового до відмінного. Рис.4.3.2.1.

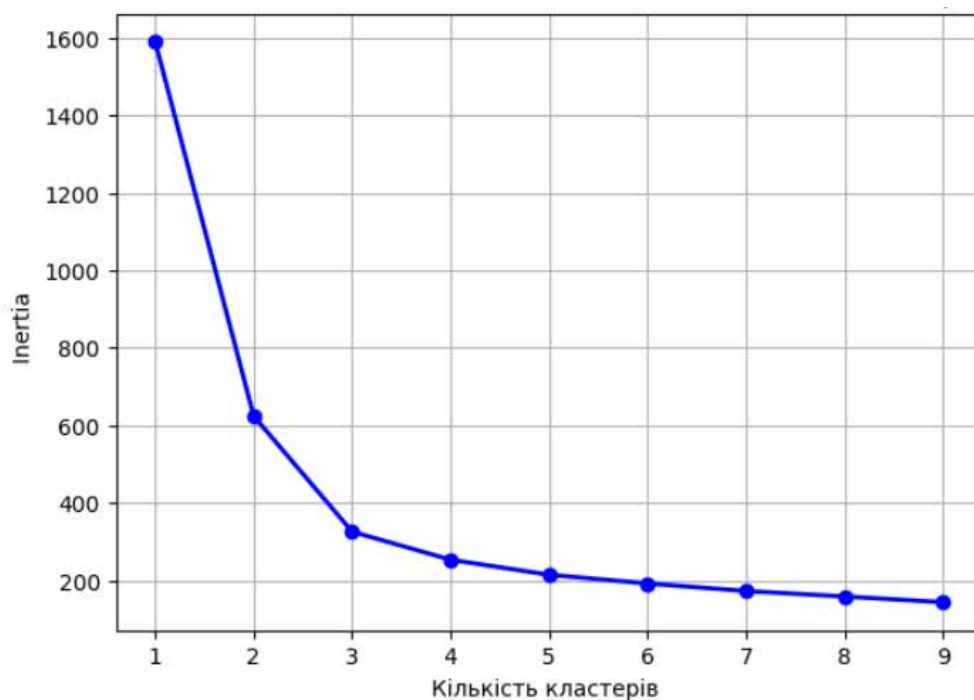


Рис.4.3.2.1 Результат визначення оптимальної кількості кластерів

Другий графік демонструє розподіл викладачів у двовимірному просторі, де вісь X відображає кількість публікацій, а вісь Y рейтинг викладача. Рис.4.3.2.2.

Таблиця 4.4.

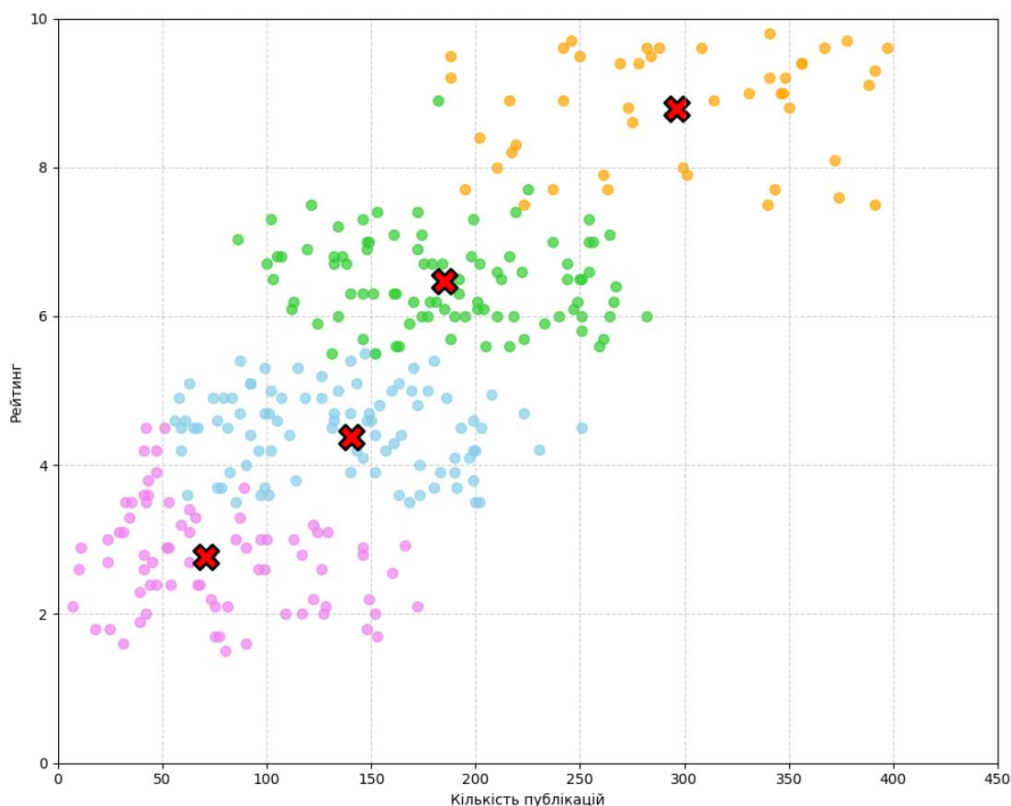


Рис.4.3.2.2 Результат кластеризації

Таблиця 4.4

Таблиця груп викладачів в кластеризації та їх середні показники

Кластер	Публікації	Цитування	Індекс Гірша	H10-index	Рейтинг
0	77	154	5.9	4.2	2.6
1	150	368	9.4	7.6	4.5
2	200	941	14.3	19.3	6.5
3	290	1549	18.2	34.8	8.8

В результаті кластерного аналізу, зважаючи на їх рейтинг та кількість робіт, викладачі поділились на 4 основні групи.

Початковий рівень (Рожевий колір). Ця група характеризується відносно невеликою кількістю публікацій (~77), низькими показниками цитованості та рейтингом до 3,5 балів. Це переважно молоді викладачі або ті, хто лише починає активно займатися науковою діяльністю. Згідно статистики для підвищення рейтингу, цим викладачам потрібно підвищувати якість своїх робіт.

Середній рівень (Помаранчевий колір). Викладачі цього кластеру мають приблизно 150 публікацій, помірну цитованість і рейтинг у межах 4,5 балів. Вони ведуть регулярну публікаційну діяльність, але здебільшого у виданнях нижчих кватилів.

Високий рівень (Зелений колір). До цього кластеру належать викладачі з вищою кількістю публікацій (~200), більш активними цитуваннями і рейтингом близько 6,5, а також вони часто мають h-індекс понад 10.

Відмінний рівень (Фіолетовий колір). Це найменша, але найбільш значуща група, до якої входять найкращі викладачі з понад 200 публікаціями, h-індексом понад 16 та рейтингом вище 8 балів. Що свідчить про їх високий вплив на розвиток відповідних наукових галузей.

В свою чергу, КРІ - це ключовий показник ефективності, що використовується для вимірювання результативності працівників, процесів або організацій загалом. Його основна мета - не просто відобразити поточний стан діяльності, а й надати кількісні орієнтири, за допомогою яких можна оцінювати динаміку розвитку, виявляти проблемні зони та формувати управлінські рішення. У контексті оцінювання наукової активності педагогів КРІ дозволяють уніфікувати підхід, забезпечити прозорість оцінки та побачити, наскільки викладачі досягають визначених стандартів наукової продуктивності.

У межах дослідження було сформовано спеціальний КРІ, який дозволяє порівнювати середні результати педагогів за два звітні періоди - поточний та попередній рік. Це дає змогу не лише визначити абсолютний рівень наукової активності, а й оцінити динамічний прогрес або регрес викладачів.

КРІ було поділено на чотири рівні, які відповідають групам викладачів із різними показниками наукової діяльності, а саме - низький, середній, високий та відмінний. Для кожної категорії розраховано два параметри, а саме, фактичне середнє значення рейтингу за цей та за минулий рік. Рис.4.3.2.3.

Отобразить структуру	Значение	Цель	Состояние	Тренд
 Середня оцінка викладачів з високими показниками	6.4	6.1		↑
 Середня оцінка викладачів з відмінними показниками	8.8	8.9		↘
 Середня оцінка викладачів з низькими показниками	2.9	2.7		↑
 Середня оцінка викладачів з середніми показниками	4.3	4.5		↘

Рис.4.3.2.3 КРІ для перегляду тенденції зміни рейтингу викладачів

КРІ викладачів з низькими показниками. До цієї групи віднесено викладачів, у яких рівень наукової активності суттєво нижчий ніж в інших. Фактичне середнє значення рейтингу становить 2.9, тоді як рейтинг за минулий рік дорівнює 2.7. Це означає, що хоч і не сильно суттєво, але наукова продуктивність цієї групи підвищилась.

Викладачі з середніми показниками мають фактичний бал 4.3, а значення минулого року дорівнювало 4.5. Тобто, у цій групі фіксується незначне погіршення, адже фактичний показник нижчий за попередній. Таке відхилення може свідчити про недостатню участь у наукових проєктах, зменшення кількості або якості публікацій.

КРІ викладачів із високими показниками – загалом вони демонструють найкращу тенденцію результатів, фактичний показник становить 6.4, тоді як минулорічне значення було 6.1. Різниця вказує на стійке та рівномірне зростання наукової продуктивності. Викладачі цієї категорії підтримують високий темп роботи, регулярно публікуються та покращують якість своїх робіт, що позитивно впливає на їхній персональний рейтинг.

КРІ викладачів із відмінними показниками. У категорії педагогів із відмінним рівнем результативності середнє фактичне значення становить 8.8, тоді як минулорічне було 8.9. Незначне зниження показника свідчить про те, що викладачі залишаються на дуже високому рівні, але за темпом зростання дещо поступаються торішнім результатам. Водночас їхні показники все ще значно перевищують середні. Для таких викладачів важливо підтримувати уже досягнуту високу планку та розширювати наукові колаборації, щоб уникати подальшого зниження.

Також був побудований ще один КРІ який дозволяє комплексно відстежувати динаміку розвитку наукової діяльності щодо публікаційної активності, цитованості та результативності викладачів у цілому. КРІ побудований на середніх значеннях по всій вибірці, що забезпечує об'єктивний аналіз змін у загальній продуктивності. Рис.4.3.2.4.



Отобразить структуру	Значение	Цель	Состояние	Тренд
 Середній рейтинг всіх викладачів	5.6	5.55		
 Середня кількість публікацій	6.4	7.1		
 Середня кількість цитувань	71.6	63.2		

Рис.4.3.2.4 КРІ для відстеження загальної динаміки зміни наукової діяльності

Першим показником КРІ є середній рейтинг викладачів, який враховує загальну оцінку наукової активності. У поточному році він становить 5.6, що дещо перевищує показник минулого року - 5.55. Зростання є незначним, проте воно свідчить про загальну позитивну динаміку в підвищенні наукової ефективності. Навіть мінімальне підвищення цього рейтингу, характеризує поступове посилення наукової продуктивності. Це означає, що частина викладачів поліпшила індивідуальні результати, що загалом підняло середній показник.

Другим компонентом КРІ є середня кількість публікацій на одного викладача. Тут спостерігається зворотна тенденція, а саме, фактичне значення поточного року 6.4, тоді як минулорічне становило 7.1. Зменшення середнього показника вказує на певне уповільнення публікаційної активності. Цей тренд може бути пов'язаний із різними чинниками, а саме, підвищенням складності досліджень, зосередженням викладачів на підготовці якісніших робіт або участю в довготривалих проектах.

Третім показником КРІ виступає середня кількість цитувань, яка демонструє суттєво позитивну динаміку. У поточному році середнє значення становить 71.6 цитувань, що значно перевищує показник минулого року - 63.2 цитування. Зростання на понад 8 цитувань у середньому свідчить про підвищення

міжнародної видимості досліджень викладачів, їхньої актуальності та відповідності запитам наукової спільноти. Така динаміка компенсує зменшення середньої кількості публікацій і вказує, що хоча робіт менше, але вони стають більш значущими та мають більший вплив.

Зведений аналіз цих трьох компонентів дозволяє дійти важливого висновку - загальна наукова продуктивність викладачів демонструє помірне покращення, навіть попри зниження кількості публікацій. Ріст цитованості вказує на те, що підвищується якість наукових досліджень, а збільшення середнього рейтингу підтверджує позитивний результат комплексної оцінки. Таким чином, побудований KPI дає змогу повноцінно оцінити зміни у науковій активності, поєднуючи кількісні та якісні характеристики роботи педагогів.

ВИСНОВКИ

У магістерській роботі була розроблена експертна система оцінювання наукової діяльності педагогічного складу, яка забезпечує підвищення об'єктивності та швидкості оцінювання, а також, прозорості аналізу продуктивності викладачів за рахунок використання міжнародних наукометричних даних, OLAP-технологій та методів інтелектуального аналізу даних. Виконане дослідження дозволило комплексно автоматизувати процес збору та обробки показників наукової активності педагогів та сформувати єдиний інструмент для підтримки управлінських рішень у закладах вищої освіти.

У рамках виконання роботи було проаналізовано сучасні методики оцінювання наукової діяльності, досліджено та визначено ключові показники, що характеризують рівень наукової діяльності викладачів. На основі цих показників було реалізовано сховище даних та створено багатовимірний аналітичний куб у SSAS, що надає можливість проводити різносторонній аналіз та порівнювати результати викладачів за періодами.

Важливим етапом дослідження стало застосування алгоритму кластеризації k-середніх, який дозволив виділити групи викладачів з різним рівнем наукової діяльності. Це створило можливість виявлення закономірностей, визначення сильних і слабких груп, а також формування рекомендацій щодо підвищення продуктивності. Аналіз кластерів підтвердив істотні відмінності між групами викладачів та надав можливість визначити показники наукової діяльності для кожної групи відповідно.

Окрему увагу було приділено розробці системи ключових показників ефективності. Побудована модель дозволила порівняти середні значення рейтингу, кількості публікацій та цитувань за поточні і попередні роки. Результати показали позитивну динаміку за більшістю критеріїв, наприклад, середній рейтинг викладачів підвищився та кількість цитувань зросла, навіть за

наявності зниження середньої кількості публікацій покращення цитованості вказує на підвищення якості та впливовості робіт. Таким чином, КРІ-система дала можливість не просто фіксувати показники, а й оцінювати тенденції розвитку наукової діяльності.

Розроблена експертна система дозволила досягти низки важливих практичних результатів, що суттєво впливають на процес оцінювання наукової діяльності педагогічного складу. Насамперед було автоматизовано процес збору та обробки інформації з міжнародних наукометричних баз, що значно скоротило час підготовки звітів та майже повністю усунуло ризик помилок, пов'язаних із ручним введенням.

Також, варто зазначити, що система базується на даних, які отримуються безпосередньо з незалежних міжнародних джерел, що дозволяє уникнути суб'єктивності та забезпечити коректність і актуальність усіх показників. Завдяки цьому, рейтинги педагогів стали більш обґрунтованими, а результати оцінювання більш прозорими. Це дозволяє мінімізувати вплив людського фактору та зробити процес оцінки структурованим, однаковим для всіх учасників та незалежним від зовнішніх факторів.

Завдяки прозорості, регулярності та об'єктивності оцінювання педагогічні працівники отримують чітке розуміння того, за якими критеріями оцінюється їхня діяльність та що саме впливає на їхній рейтинг. Можливість порівняти себе з іншими та перегляд власної динаміки розвитку стимулюють викладачів активніше брати участь у наукових роботах, підвищувати рівень цитування та збільшувати кількість якісних публікацій.

Водночас, слід підкреслити, що експертна система не може повністю замінити комплексне оцінювання діяльності науково-педагогічних працівників. Наукові публікації та цитування - лише один із напрямів професійної активності викладача. Значну роль відіграють педагогічна майстерність, методична робота, участь у виховному процесі, проведення практичних і лабораторних занять та

інше. Ці аспекти потребують додаткових, здебільшого якісних підходів до оцінювання, які не можуть бути повністю автоматизовані.

Отже, розроблену експертну систему варто розглядати як сучасний інструмент підтримки прийняття рішень, який доповнює, але не замінює традиційні механізми оцінювання. Вона забезпечує об'єктивну наукометричну основу та значно підсилює аналітичні можливості закладу вищої освіти, створюючи умови для більш збалансованої, прозорої та мотивуючої системи оцінювання наукової діяльності викладачів.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Денисова Анастасія. Використання методу експертних оцінок у системі моніторингу якості підвищення кваліфікації фахівців галузі професійної освіти, 2019р. [Електронний ресурс]. URL: [https://doi.org/10.33272/2522-9729-2019-5\(188\)-39-43](https://doi.org/10.33272/2522-9729-2019-5(188)-39-43). (дата звернення: 20.08.2025).
2. Ольга Іваницька. Алла Панченко. Експертне оцінювання як інструмент управління загальноосвітнім навчальним закладом, 2015р. [Електронний ресурс]. URL: <https://doi.org/10.28925/2226-3012.2015.4.6570>. (дата звернення: 22.08.2025).
3. Шевчук С.С. Система педагогічного контролю та оцінювання як фактор забезпечення якості професійної освіти, 85с., 2023. [Електронний ресурс]. URL: <https://surl.li/dmifey>. (дата звернення: 24.08.2025).
4. Server Integration Services, 2024. [Електронний ресурс]. URL: <https://learn.microsoft.com/en-us/sql/integration-services/sql-server-integration-services?view=sql-server-ver17>. (дата звернення: 12.04.2025).
5. SQL Server Analysis Services overview, 2024. [Електронний ресурс]. URL: <https://learn.microsoft.com/en-us/sql/integration-services/sql-server-integration-services?view=sql-server-ver17>. (дата звернення: 24.04.2025).
6. Баклан І.В. Експертні системи курс лекцій, 132с., 2012р. [Електронний ресурс]. URL: https://baklaniv.at.ua/MSAI/ekspertni_sistemi-kurs_lekcij.2012.pdf. (дата звернення: 28.07.2025).
7. Visual Studio Code documentation, 2025. [Електронний ресурс]. URL: <https://code.visualstudio.com/docs>. (дата звернення: 10.05.2025).
8. SQL Server technical documentation, 2025. [Електронний ресурс]. URL: <https://learn.microsoft.com/en-us/sql/sql-server/?view=sql-server-ver17>. (дата звернення: 11.04.2025).

9. Visual Studio documentation, 2025. [Электронный ресурс]. URL: <https://learn.microsoft.com/en-us/visualstudio/windows/?view=visualstudio>. (дата звернения: 09.05.2025).
10. Python 3.14.0 documentation, 2025. [Электронный ресурс]. URL: <https://docs.python.org/3/>.(дата звернения: 25.04.2025).
11. Online analytical processing, 2025. [Электронный ресурс]. URL: <https://learn.microsoft.com/en-us/azure/architecture/data-guide/relational-data/online-analytical-processing>. (дата звернения: 29.04.2025).
12. Data Mining Concepts, 2023. [Электронный ресурс]. URL:<https://learn.microsoft.com/en-us/analysis-services/data-mining/data-mining-concepts?view=asallproducts-allversions>. (дата звернения: 23.04.2025).
13. Pandas documentation, 2025. [Электронный ресурс]. URL: <https://pandas.pydata.org/docs/>. (дата звернения: 11.06.2025).
14. Introduction to Unified Modeling Language, 2023. [Электронный ресурс]. URL: <https://www.omg.org/uml/>. (дата звернения: 15.07.2025).
15. Use Case Diagram - Unified Modeling Language (UML), 2025. [Электронный ресурс]. URL: <https://www.geeksforgeeks.org/system-design/use-case-diagram/>.(дата звернения: 16.07.2025).
16. Explore the UML sequence diagram, 2023. [Электронный ресурс]. URL: <https://developer.ibm.com/articles/the-sequence-diagram/>.(дата звернения: 17.07.2025).
17. UML Class Diagram, 2025. [Электронный ресурс]. URL: <https://www.geeksforgeeks.org/system-design/unified-modeling-language-uml-class-diagrams/>.(дата звернения: 20.07.2025).
18. Peter J.F. Lucas & Linda C. van der Gaag. Principles of Expert Systems, 426с., 2014 р. [Электронный ресурс]. URL: <https://www.cs.ru.nl/P.Lucas/proe.pdf>. (дата звернения: 29.08.2025).

19. Python Tutorial, 2025. [Електронний ресурс]. URL: <https://www.geeksforgeeks.org/python/python-programming-language-tutorial/>. (дата звернення: 30.04.2025).

20. Ляшенко О. І., Лукіна Т. О., Булах І. Є., Мруга М. Р. Методика і технології оцінювання діяльності загальноосвітнього навчального закладу, 160с., 2012р. [Електронний ресурс]. URL: <https://files.core.ac.uk/download/pdf/20054342.pdf>. (дата звернення: 05.06.2025).

21. Т.М.Канівець. Основи педагогічного оцінювання, 102с., 2012р. [Електронний ресурс]. URL: https://moodle.ndu.edu.ua/file.php/1/Konivets_Osnovy_Ped_Jtsinuvannia.pdf. (дата звернення: 08.06.2025).

22. What is SQL Server Management Studio. [Електронний ресурс]. URL: <https://learn.microsoft.com/en-us/ssms/sql-server-management-studio-ssms>. (дата звернення: 26.05.2025).

ДОДАТКИ

Код створення таблиць бази даних

```

create table UserRole(
  IdRole int Primary key,
  Name varchar(25) not null,
  Access varchar(30) not null
)

create table Users(
  idUser int Primary key,
  Surname varchar(30) not null,
  Name varchar(30) not null,
  Patronymic varchar(35) not null,
  Phone varchar(13) not null,
  Email varchar(50) not null,
  IdRole int not null,
  Foreign Key(IdRole) references UserRole(IdRole)
)

create table ScientificActivity(
  IdPublication int Primary key,
  IdUser int not null,
  Name varchar(40) not null,
  Type varchar(40) not null,
  Date Date not null
  Quartile varchar(10),
  Citations int,
  Description text,
  Foreign Key(IdUser) references Users(IdUser)
)

create table Metric(
  IdMetric int Primary key,
  Name varchar(20) not null,
  Description varchar(200) not null
)

create table Statistic(
  IdStatistic int Primary key,
  IdUser int not null,
  IdMetric int not null,
  UserRating float not null,
  Foreign Key(IdUser) references Users(IdUser),
  Foreign Key(IdMetric) references Metric(IdMetric)
)

create table AnalyticResult(
  IdAnalyticResult int Primary key,
  IdStatistic int not null,
  Recommendations varchar(100) not null,
  Date date not null,
  Foreign Key(IdStatistic) references Statistic(IdStatistic)
)

```

Код створення таблиць сховища даних

```

create table TimeDim(
  IdTime int Primary key,
  Day INT not null,
  Month INT not null,
  Year INT not null
)

create table TeacherDim(
  IdTeacher int Primary key,
  Surname varchar(30) not null,
  Name varchar(30) not null,
  Patronymic varchar(35) not null,
  Email varchar(50) not null,
  ScientificCount int not null
)

create table MetricDim(
  IdMetric int primary key,
  Name varchar(20) not null,
  Description varchar(200) not null
)

create table ScientificActivityDim(
  IdPublication int Primary key,
  IdUser int not null,
  Name varchar(40) not null,
  Type varchar(40) not null,
  Date Date not null,
  Quartile varchar(10),
  Citations int ,
  Description text
);

create table StatisticFact(
  IdTeacher int not null,
  IdMetric int not null,
  IdPublication int not null,
  IdTime int not null,
  UserRating float,
  PRIMARY KEY (IdTeacher, IdMetric, IdPublication, IdTime),
  Foreign key (IdTeacher) references TeacherDim(IdTeacher),
  Foreign key (IdMetric) references MetricDim(IdMetric),
  Foreign key (IdPublication) references ScientificActivityDim(IdPublication),
  Foreign key (IdTime) references TimeDim(IdTime));

```

Частина основного функціоналу додатку

```

class DataCollector:

    def __init__(self, db_manager: DatabaseManager):
        self.db = db_manager
        self.google_scholar_base_url = "https://scholar.google.com"

    def collect_teacher_data(self, teacher_id: int) -> Dict:
        teacher_data = {}
        teacher_info = self.get_teacher_info(teacher_id)
        if teacher_info.empty:
            return teacher_data

        teacher_data['basic_info'] = teacher_info.iloc[0].to_dict()
        teacher_data['scopus'] = self.collect_scopus_data(teacher_info.iloc[0])
        teacher_data['google_scholar'] = self.collect_google_scholar_data(teacher_info.iloc[0])

        return teacher_data

    def get_teacher_info(self, teacher_id: int) -> pd.DataFrame:
        query = """
        SELECT u.IdUser, u.Surname, u.Name, u.Patronymic, u.Email,
               sa.Name as PublicationName, sa.Type, sa.Date, sa.Quartile,
               sa.Citations, sa.Description
        FROM Users u
        LEFT JOIN ScientificActivity sa ON u.IdUser = sa.IdUser
        WHERE u.IdUser = ? AND u.IdRole = (SELECT IdRole FROM UserRole WHERE Name = 'Teacher')
        """
        return self.db.execute_query(query, (teacher_id,))

    def collect_scopus_data(self, teacher_info: Dict) -> Dict:
        scopus_data = {}
        try:
            search_url = "https://api.elsevier.com/content/search/author"
            params = {
                'query': f"authlast({teacher_info['Surname']}) and
authfirst({teacher_info['Name']})",
                'apiKey': self.scopus_api_key
            }

            response = requests.get(search_url, params=params)
            if response.status_code == 200:
                data = response.json()
                if data.get('search-results', {}).get('entry'):
                    author_id = data['search-results']['entry'][0]['dc:identifier'].split(':')[1]

                    author_url = f"https://api.elsevier.com/content/author/author_id/{author_id}"
                    author_response = requests.get(author_url,
                                                    params={'apiKey': self.scopus_api_key})

                    if author_response.status_code == 200:
                        author_data = author_response.json()
                        scopus_data = self.parse_scopus_author_data(author_data)
        except:
            pass
1]

```

```

except Exception as e:
    logger.error(f"Помилка збору даних з Scopus: {e}")

return scopus_data

def parse_scopus_author_data(self, author_data: Dict) -> Dict:
    parsed_data = {
        'publications_count': 0,
        'citations_count': 0,
        'h_index': 0,
        'documents': []
    }

    try:
        author_profile = author_data.get('author-retrieval-response', [{}])[0]
        coredata = author_profile.get('coredata', {})

        parsed_data['publications_count'] = int(coredata.get('document-count', 0))
        parsed_data['citations_count'] = int(coredata.get('citation-count', 0))
        parsed_data['h_index'] = int(coredata.get('h-index', 0))
        documents = author_profile.get('documents', {}).get('document', [])
        for doc in documents:
            document_info = {
                'title': doc.get('dc:title', ''),
                'authors': doc.get('authors', {}).get('author', []),
                'publication_date': doc.get('prism:coverDate', ''),
                'citation_count': int(doc.get('citedby-count', 0)),
                'doi': doc.get('prism:doi', ''),
                'source': doc.get('prism:publicationName', ''),
                'quartile': self.get_journal_quartile(doc.get('prism:issn', ''))
            }
            parsed_data['documents'].append(document_info)

    except Exception as e:
        logger.error(f"Помилка парсингу даних Scopus: {e}")

    return parsed_data

def collect_google_scholar_data(self, teacher_info: Dict) -> Dict:
    scholar_data = {}
    try:
        options = webdriver.ChromeOptions()
        options.add_argument('--headless')
        options.add_argument('--no-sandbox')
        options.add_argument('--disable-dev-shm-usage')

        driver = webdriver.Chrome(options=options)
        search_url =
f"{self.google_scholar_base_url}/citations?hl=en&view_op=search_authors&mauthors={teacher_info['S
urname']}&{teacher_info['Name']}"
        driver.get(search_url)
        WebDriverWait(driver, 10).until(
            EC.presence_of_element_located((By.CLASS_NAME, "gs_ai_t"))
        )

```

```

soup = BeautifulSoup(driver.page_source, 'html.parser')
author_elements = soup.find_all('div', class_='gs_ai_t')

if author_elements:
    author_element = author_elements[0]
    author_link = author_element.find('a', href=True)

    if author_link:
        author_url = f"{self.google_scholar_base_url}{author_link['href']}"
        driver.get(author_url)

        profile_soup = BeautifulSoup(driver.page_source, 'html.parser')

        scholar_data = self.parse_google_scholar_profile(profile_soup)

    driver.quit()

except Exception as e:
    logger.error(f"Помилка збору даних з Google Scholar: {e}")

return scholar_data

def parse_google_scholar_profile(self, soup: BeautifulSoup) -> Dict:
    profile_data = {
        'citations': 0,
        'h_index': 0,
        'i10_index': 0,
        'publications': []
    }

    try:
        stats_table = soup.find('table', id='gsc_rsb_st')
        if stats_table:
            rows = stats_table.find_all('tr')
            for row in rows[1:]:
                cols = row.find_all('td')
                if len(cols) >= 2:
                    metric = cols[0].text.strip()
                    value = cols[1].text.strip()

                    if 'Citations' in metric:
                        profile_data['citations'] = int(value) if value else 0
                    elif 'h-index' in metric:
                        profile_data['h_index'] = int(value) if value else 0
                    elif 'i10-index' in metric:
                        profile_data['i10_index'] = int(value) if value else 0
        publications_table = soup.find('table', id='gsc_a_t')
        if publications_table:
            publications = publications_table.find_all('tr', class_='gsc_a_tr')
            for pub in publications:
                title_element = pub.find('a', class_='gsc_a_at')
                citation_element = pub.find('a', class_='gsc_a_ac')
                year_element = pub.find('span', class_='gsc_a_h')

                if title_element and citation_element:

```

```

        publication_info = {
            'title': title_element.text.strip(),
            'citations': int(citation_element.text.strip()) if
citation_element.text.strip() else 0,
            'year': year_element.text.strip() if year_element else ''
        }
        profile_data['publications'].append(publication_info)

    except Exception as e:
        logger.error(f"Помилка парсингу профілю Google Scholar: {e}")

    return profile_data

```

```
class MetricsCalculator:
```

```

    def __init__(self, db_manager: DatabaseManager):
        self.db = db_manager

    def calculate_publications_metric(self, teacher_data: Dict) -> float:
        try:
            publications = teacher_data.get('publications', [])
            if not publications:
                return 0.0

            total_score = 0.0
            current_year = datetime.now().year

            for pub in publications:
                qf = self.get_quality_factor(pub.get('quartile', 'Q4'))
                tf = self.get_type_factor(pub.get('type', 'Стаття'))
                pub_year = pub.get('year', current_year)
                if isinstance(pub_year, str):
                    pub_year = int(pub_year) if pub_year.isdigit() else current_year

                df = np.exp(-0.1 * (current_year - pub_year))

                total_score += qf * tf * df

            return total_score / len(publications)

        except Exception as e:
            logger.error(f"Помилка розрахунку метрики публікацій: {e}")
            return 0.0

    def get_quality_factor(self, quartile: str) -> float:
        quality_factors = {
            'Q1': 1.0,
            'Q2': 0.8,
            'Q3': 0.6,
            'Q4': 0.4,
            '': 0.2
        }
        return quality_factors.get(quartile, 0.2)

```

```

def get_type_factor(self, pub_type: str) -> float:
    type_factors = {
        'Наукова стаття': 1.0,
        'Стаття': 0.8,
        'Монографія': 0.9,
        'Тези': 0.6,
        'Доповідь': 0.5,
        '': 0.4
    }
    return type_factors.get(pub_type, 0.5)

def calculate_citations_metric(self, teacher_data: Dict, max_citations: int,
max_avg_citations: float) -> float:
    try:
        total_citations = teacher_data.get('citations_count', 0)
        publications = teacher_data.get('publications', [])

        if not publications:
            avg_citations = 0.0
        else:
            citations_list = [pub.get('citations', 0) for pub in publications]
            avg_citations = sum(citations_list) / len(publications) if citations_list else
0.0

        citation_component = 0.6 * (
            np.log(1 + total_citations) / np.log(1 + max_citations)) if max_citations
> 0 else 0
        avg_citation_component = 0.4 * (avg_citations / max_avg_citations) if
max_avg_citations > 0 else 0

        return citation_component + avg_citation_component

    except Exception as e:
        logger.error(f"Помилка розрахунку метрики цитування: {e}")
        return 0.0

def calculate_h_index_metric(self, h_index: int, max_h_index: int) -> float:
    try:
        return h_index / max_h_index if max_h_index > 0 else 0.0
    except Exception as e:
        logger.error(f"Помилка розрахунку метрики h-index: {e}")
        return 0.0

def calculate_i10_metric(self, i10_index: int, max_i10_index: int) -> float:
    try:
        return i10_index / max_i10_index if max_i10_index > 0 else 0.0
    except Exception as e:
        logger.error(f"Помилка розрахунку метрики i10-index: {e}")
        return 0.0

def calculate_final_rating(self, metrics: Dict) -> float:
    try:
        weights = {
            'publications': 0.25,
            'citations': 0.35,
            'h_index': 0.25,

```

```

        'i10_index': 0.15
    }

    final_score = 0.0
    for metric, value in metrics.items():
        final_score += weights.get(metric, 0) * value
    normalized_score = min(final_score * 10, 10.0)
    return normalized_score

except Exception as e:
    logger.error(f"Помилка розрахунку фінального рейтингу: {e}")
    return 0.0

class DataAnalyzer:

    def __init__(self, db_manager: DatabaseManager):
        self.db = db_manager

    def perform_cluster_analysis(self, teachers_data: pd.DataFrame) -> Dict:
        try:
            features = teachers_data[['publications_count', 'citations_count', 'h_index',
'i10_index']].fillna(0)
            scaler = StandardScaler()
            scaled_features = scaler.fit_transform(features)
            optimal_k = self.find_optimal_clusters(scaled_features)
            kmeans = KMeans(n_clusters=optimal_k, random_state=42)
            clusters = kmeans.fit_predict(scaled_features)

            teachers_data['cluster'] = clusters
            cluster_analysis = self.analyze_clusters(teachers_data)

            return {
                'clusters': clusters,
                'cluster_centers': kmeans.cluster_centers_,
                'cluster_analysis': cluster_analysis,
                'optimal_k': optimal_k
            }

        except Exception as e:
            logger.error(f"Помилка кластерного аналізу: {e}")
            return {}

    def find_optimal_clusters(self, data: np.ndarray, max_k: int = 10) -> int:
        inertias = []
        k_range = range(1, max_k + 1)

        for k in k_range:
            kmeans = KMeans(n_clusters=k, random_state=42)
            kmeans.fit(data)
            inertias.append(kmeans.inertia_)
        optimal_k = 4
        for i in range(1, len(inertias) - 1):
            if (inertias[i - 1] - inertias[i]) / (inertias[i] - inertias[i + 1]) > 2:
                optimal_k = i + 1

```

```

        break

    return optimal_k

def analyze_clusters(self, clustered_data: pd.DataFrame) -> Dict:

    cluster_stats = {}

    for cluster in clustered_data['cluster'].unique():
        cluster_data = clustered_data[clustered_data['cluster'] == cluster]

        stats = {
            'size': len(cluster_data),
            'avg_publications': cluster_data['publications_count'].mean(),
            'avg_citations': cluster_data['citations_count'].mean(),
            'avg_h_index': cluster_data['h_index'].mean(),
            'avg_i10_index': cluster_data['i10_index'].mean(),
            'avg_rating': cluster_data['final_rating'].mean(),
            'level': self.determine_cluster_level(cluster_data)
        }

        cluster_stats[cluster] = stats

    return cluster_stats

def determine_cluster_level(self, cluster_data: pd.DataFrame) -> str:
    avg_rating = cluster_data['final_rating'].mean()

    if avg_rating >= 8.0:
        return "Відмінний"
    elif avg_rating >= 6.0:
        return "Високий"
    elif avg_rating >= 4.0:
        return "Середній"
    else:
        return "Початковий"

class RecommendationEngine:

    def __init__(self, db_manager: DatabaseManager):
        self.db = db_manager

    def generate_recommendations(self, teacher_data: Dict, metrics: Dict, final_rating: float) ->
    List[str]:
        recommendations = []

        try:
            pub_metric = metrics.get('publications', 0)
            if pub_metric < 0.4:
                recommendations.append(
                    "Рекомендується підвищити публікаційну активність. "
                    "Зосередьтеся на публікаціях у журналах з високим квантилем (Q1-Q2).")
        )
        cit_metric = metrics.get('citations', 0)

```

```

if cit_metric < 0.4:
    recommendations.append(
        "Необхідно підвищити цитованість публікацій. "
        "Розгляньте можливість участі у міжнародних конференціях та співпраці з
іншими дослідниками."
    )
    h_index = teacher_data.get('h_index', 0)
    i10_index = teacher_data.get('i10_index', 0)

if h_index < 5:
    recommendations.append(
        "Індекс Гірша нижче середнього. Рекомендується зосередитись на якості
публікацій "
        "та посилити роботу над цитуваннями існуючих робіт."
    )

if i10_index < 3:
    recommendations.append(
        "Показник i10-index потребує покращення. Необхідно створити більше впливових
робіт, "
        "які отримають понад 10 цитувань."
    )

if final_rating < 4.0:
    recommendations.append(
        "Загальний рівень наукової діяльності потребує значного покращення. "
        "Рекомендується розробити індивідуальний план розвитку наукової діяльності."
    )

elif final_rating < 6.0:
    recommendations.append(
        "Наукова діяльність на середньому рівні. Рекомендується активізувати роботу "
        "над публікаціями та посилити міжнародну співпрацю."
    )

else:
    recommendations.append(
        "Відмінні результати наукової діяльності! Рекомендується продовжувати роботу"
        "у тому ж напрямку та розглянути можливість керівництва науковими проектами."
    )

except Exception as e:
    logger.error(f"Помилка генерації рекомендацій: {e}")
    recommendations.append("На даний момент немає конкретних рекомендацій.")

return recommendations

```