

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ХЕРСОНСЬКИЙ ДЕРЖАВНИЙ УНІВЕРСИТЕТ
Факультет комп'ютерних наук, фізики та математики
Кафедра комп'ютерних наук та програмної інженерії

**Проектування та розробка модуля аналізу наукової діяльності
викладачів для інформаційної системи KSU24**

Кваліфікаційна робота (проект)
на здобуття ступеня вищої освіти “бакалавр”

Виконав: здобувач 4 курсу 12-441 групи

Спеціальності 121 Інженерія програмного
забезпечення

Освітньо-професійної програми

Інженерія програмного забезпечення
першого (бакалаврського) рівня вищої освіти

Жигун Микола Сергійович

Керівник: кандидат педагогічних наук, доцент
Вінник Максим Олександрович

Рецензент: доктор педагогічних наук, професор,
професор кафедри інформатики і кібернетики
Мелітопольського державного педагогічного
університету імені Богдана Хмельницького
Круглик Владислав Сергійович

Зміст

ВСТУП.....	3
РОЗДІЛ 1 АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ ТА ОБҐРУНТУВАННЯ ТЕХНІЧНИХ РІШЕНЬ	6
1.1 Аналіз структури наукової діяльності НПП як об'єкта автоматизації.....	6
1.2 Огляд існуючих підходів до інтелектуального аналізу текстових даних	8
1.3 Порівняльний аналіз моделей штучного інтелекту для генерації аналітики	11
1.4 Цифрові технології та передумови впровадження AI-звітів у професійну діяльність НПП	15
ВИСНОВКИ ДО РОЗДІЛУ 1	19
РОЗДІЛ 2 ПРОЄКТУВАННЯ АРХІТЕКТУРИ МОДУЛЯ АНАЛІЗУ ДЛЯ ІС KSU24 ...	21
2.1 Розробка загальної архітектури та схеми інтеграції	21
2.2 Проєктування логічної структури та потоків даних (UML-моделювання)	23
2.3 Розробка стратегії Prompt Engineering для формування аналітичних звітів	26
2.4 Обґрунтування вибору технологічного стеку	29
ВИСНОВКИ ДО РОЗДІЛУ 2	31
РОЗДІЛ 3 ПРОГРАМНА РЕАЛІЗАЦІЯ ТА ВАЛІДАЦІЯ РЕЗУЛЬТАТІВ.....	33
3.1 Архітектурна побудова та логічна структура модуля	33
3.2 Технологічний цикл агрегації та трансформації даних.....	34
3.3 Алгоритми парсингу та структурування відповідей ШІ.....	35
3.4 Тестування та валідація результатів на масивах ІС KSU24	37
3.5 Програмний інтерфейс адміністратора для збору та валідації даних	42
3.6 Практична реалізація промпт-інжинірингу в середовищі LLM	45
3.7 Архітектура обробки Markdown-розмітки та візуалізації результатів	47
ВИСНОВКИ.....	51
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ	53
ДОДАТКИ.....	57

ВСТУП

Актуальність теми. Сучасний етап розвитку вищої освіти характеризується стрімкою цифровізацією управлінських процесів. Інформаційні системи керування університетом (ERP-системи), такі як KSU24, накопичують значні масиви даних про результати діяльності науково-педагогічних працівників (НПП). Проте просте накопичення даних не вирішує задачу їх ефективної інтерпретації. Традиційні методи аналізу наукової діяльності часто є трудомісткими та обмежуються лише кількісними показниками, що не дає можливості цілісно оцінити професійний розвиток фахівця.

Актуальність дослідження зумовлена необхідністю розробки інтелектуальних програмних засобів, які здатні автоматизувати процес обробки великих масивів даних та перетворювати їх на змістовну аналітику. Використання технологій штучного інтелекту, зокрема великих мовних моделей (LLM), дозволяє створити модуль для напівавтоматичного формування персоналізованих AI-звітів із подальшою перевіркою користувачем. Це не лише знижує адміністративне навантаження, а й забезпечує об'єктивну підтримку прийняття управлінських рішень щодо професійної траєкторії розвитку викладачів.

Об'єктом дослідження є процеси автоматизованої обробки та аналізу даних про наукову діяльність науково-педагогічних працівників у закладах вищої освіти.

Предметом дослідження є методи, архітектура, структури даних і програмні засоби реалізації модуля аналізу наукової діяльності викладачів для інформаційної системи KSU24.

Мета роботи полягає у проєктуванні та програмній реалізації модуля аналізу наукової діяльності викладачів для інформаційної системи KSU24, який забезпечує автоматизоване збирання, агрегування та структурування даних, їх підготовку до аналітичного опрацювання та

подальше формування звітних відомостей у напівавтоматичному сценарії Human-in-the-loop.

Для досягнення поставленої мети передбачено розв'язання таких завдань:

- проаналізувати функціональні вимоги до систем моніторингу діяльності НПП та визначити технічні можливості інтеграції з ІС KSU24;
- дослідити сучасні технологічні стеки та API-інтерфейси моделей штучного інтелекту для обробки професійних текстових даних;
- проєктувати архітектуру модуля, включаючи схему взаємодії компонентів, структуру даних та алгоритми формування запитів (prompt engineering);
- розробити програмне забезпечення модуля, реалізувавши конвеєр обробки даних (data pipeline) від вибору первинної інформації до генерації фінального звіту;
- провести тестування розробленого програмного продукту та валідувати якість отриманих AI-звітів.

Методологію дослідження становлять методи об'єктно-орієнтованого аналізу та проєктування (OOAD), життєвого циклу розробки програмного забезпечення (SDLC), методи структурованого опрацювання даних, а також методи інженерії запитів (Prompt Engineering) для взаємодії з генеративними моделями ШІ.

Наукова новизна отриманих результатів полягає у розробці архітектурного підходу до використання сервісів генеративного штучного інтелекту у структурі діючої інформаційної системи закладу вищої освіти. Запропоновано алгоритм семантичної трансформації статистичних даних у якісні аналітичні висновки на основі контекстного аналізу професійних досягнень.

Практичне значення роботи полягає у створенні готового програмного модуля для ІС KSU24, який дозволяє:

забезпечувати підготовку структурованих даних для аналітичних звітів про наукову активність викладачів;
мінімізувати суб'єктивність та часові витрати при оцінюванні результатів діяльності;
забезпечити адміністрацію інструментами для моніторингу наукового потенціалу підрозділів університету.

Апробація результатів роботи. Основні результати дослідження були апробовані шляхом експериментального впровадження модуля на базі даних інформаційної системи KSU24 Херсонського державного університету.

Структура роботи. Структура роботи зумовлена метою та завданнями дослідження і складається зі вступу, трьох розділів, висновків, списку використаних джерел та додатків.

РОЗДІЛ 1 АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ ТА ОБҐРУНТУВАННЯ ТЕХНІЧНИХ РІШЕНЬ

1.1 Аналіз структури наукової діяльності НПП як об'єкта автоматизації

У сучасних умовах цифровізації вищої освіти особливої значущості набуває автоматизація моніторингу професійного розвитку науково-педагогічних працівників (НПП) [7; 15]. З погляду інженерії програмного забезпечення, професійна діяльність викладача розглядається як складний об'єкт автоматизації, що генерує великі масиви різнорідних даних [4; 16]. Ключовим поняттям, що описує ці дані у динаміці, є професійна траєкторія розвитку [1; 14].

У межах цієї роботи основна увага зосереджена на аналізі наукової діяльності викладача, зокрема його публікаційної активності, участі в наукових проєктах, грантах, конференціях та інших формах наукової комунікації. Навчально-методичні, організаційні та інші відомості використовуються як додатковий контекст для точнішого аналізу та інтерпретації наукових результатів.

З позиції системного аналізу всі зазначені домени формують багатовимірний інформаційний простір, у якому кожен НПП може бути представлений у вигляді векторної моделі ознак [11; 17]. Такий підхід дозволяє застосовувати сучасні математичні та алгоритмічні методи обробки даних, зокрема методи кластеризації, регресійного аналізу та машинного навчання, для виявлення закономірностей у професійній діяльності [32; 33]. Наприклад, використання алгоритмів кластеризації дає змогу ідентифікувати типові профілі професійного розвитку, що можуть бути використані для персоналізації рекомендацій або оптимізації кадрової політики університету [30; 31].

У межах інформаційної системи KSU24 професійна траєкторія НПП формалізується як послідовність записів у базі даних, що відображають

етапи набуття знань, компетентностей та результати наукової активності [2; 5]. Для розробки модуля аналізу необхідно структурувати ці дані за ключовими доменами діяльності [3; 6; 18]:

- Науковий домен (Data Input): публікаційна активність (статті у виданнях, що індексуються у Scopus/WoS, фахові видання України), тези доповідей, участь у наукових проєктах та грантах [18; 26].
- Навчально-методичний домен: розробка силабусів, видання підручників, створення електронних курсів в LMS [27; 28].
- Інноваційний домен (Інтелектуальна власність): реєстрація патентів, авторських прав на програмні продукти, участь у стартап-проєктах [6; 25].
- Організаційний домен: робота у вченої раді, участь у внутрішній системі забезпечення якості освіти [16; 29].

Формування професійної траєкторії в цифровій системі — це динамічний процес взаємодії користувача з освітнім середовищем університету [9; 20]. З технічної точки зору, цей процес потребує системного підходу до збору та обробки даних, що дозволяє розглядати розвиток НПП не як набір випадкових подій, а як структурований потік інформації (Data Stream) [11; 17].

Крім того, автоматизація аналізу діяльності НПП передбачає інтеграцію даних із різномірних джерел, включаючи як внутрішні інформаційні системи університету, такі як системи управління навчанням та кадрові бази даних, так і зовнішні сервіси, зокрема міжнародні наукометричні бази даних та реєстри інтелектуальної власності [9; 20; 30]. У цьому контексті важливу роль відіграють процеси ETL, які забезпечують збір, трансформацію та завантаження даних у централізоване сховище з подальшою їх нормалізацією та уніфікацією [17; 30; 37]. Це дозволяє створити єдине інформаційне середовище для подальшого аналітичного опрацювання.

Основним викликом для автоматизації є індивідуальний характер професійної діяльності [10; 21]. Кожен працівник має унікальний набір вхідних параметрів (стаж, наукові інтереси, попередні досягнення), що зумовлює складність алгоритмізації оцінювання [12; 22]. Водночас система повинна забезпечувати відповідність результатів аналізу нормативним вимогам щодо ліцензійних умов та рейтингів НПП [4;15;23].

Запровадження інтелектуальних модулів у систему KSU24 створює умови для переходу від простого накопичення даних до їх семантичного аналізу [30; 31]. Використання технологій штучного інтелекту дає змогу автоматизувати трансформацію неструктурованих даних, зокрема назв статей чи описів грантів, а також суттєво спростити їх подальше узагальнення й формування аналітичних звітів у напівавтоматичному режимі з перевіркою результату користувачем [32; 33]. Такі AI-звіти стають інструментом об'єктивного моніторингу, що дозволяє не лише констатувати факти, а й прогнозувати подальші кроки професійного зростання на основі накопиченої статистики [34; 35; 36].

Отже, аналіз структури діяльності НПП як об'єкта автоматизації показує, що професійна траєкторія є цілісним цифровим процесом, який підлягає алгоритмізації в межах інтелектуальних модулів IC KSU24 [7;30;37]. Розуміння цієї структури є необхідною передумовою для проєктування архітектури модуля аналізу на основі ШІ, що здатний забезпечити високу якість формування аналітичної звітності [38; 39; 40].

1.2 Огляд існуючих підходів до інтелектуального аналізу текстових даних

Аналіз професійного розвитку науково-педагогічних працівників є поетапним процесом, що з технічного погляду характеризується еволюцією методів обробки та інтерпретації даних [1; 2; 11]. Кожен етап розвитку технологій аналізу має свої особливості, цілі та показники

ефективності, які в сукупності формують інструментарій для моніторингу траєкторії НПП [5; 14].

Початковий етап розвитку характеризується використанням традиційних методів, які базуються на детермінованих алгоритмах обробки інформації та пошуку за ключовими словами [3; 10]. Основним завданням таких систем було виявлення базових атрибутів наукової активності, зокрема назв публікацій, імен авторів, назв конференцій та інших структурованих елементів документів [6; 13; 25]. Як правило, ці системи використовували прості правила обробки тексту та регулярні вирази, що дозволяло ефективно працювати лише зі строго визначеними форматами даних. Однак такі підходи характеризуються суттєвою обмеженістю, оскільки вони не здатні враховувати контекст, розпізнавати приховані зв'язки між різними видами діяльності або аналізувати зміст наукових результатів [9; 21]. У результаті цього аналіз професійної діяльності мав фрагментарний характер і не забезпечував цілісного уявлення про розвиток НПП.

Етап семантичної стабілізації (Методи NLP). Наступний етап розвитку інструментарію передбачає застосування методів обробки природної мови (Natural Language Processing — NLP), що дозволяють здійснювати морфологічний та синтаксичний аналіз текстів [12; 18; 26]. Застосування методів частотного аналізу, тематичного моделювання та розпізнавання іменованих сутностей (Named Entity Recognition) дозволяє формувати більш деталізовані профілі НПП, визначати основні напрями їхньої наукової діяльності, а також оцінювати інтенсивність участі у різних видах професійної активності [18; 26]. Водночас, незважаючи на значне підвищення точності аналізу, ці методи все ще обмежені необхідністю попереднього налаштування моделей та залежністю від якості навчальних даних [12; 22, 23].

Етап інтелектуальної експертності (Сучасні LLM). Сучасний стан галузі характеризується переходом до використання великих мовних

моделей (Large Language Models — LLM), які володіють високим рівнем семантичної інтерпретації [17; 24; 29]. На відміну від традиційного пошуку, LLM здатні здійснювати:

- Контекстний аналіз: розуміння суті наукових досягнень поза межами простих збігів слів.
- Синтез даних: автоматичне створення резюме та аналітичних висновків.
- Прогнозування: стратегічне планування траєкторії на основі аналізу історичних даних та поточних трендів в академічному середовищі [20; 27].

У межах розробки модуля для ІС KSU24 особливого значення набуває концепція безперервного аналізу, де система постійно оновлює знання про компетентності НПП [7; 15; 30]. Безперервність моніторингу зумовлює необхідність порівняння традиційних підходів із можливостями сучасних генеративних моделей (Generative AI) [31; 34].

Таблиця 1.1. Порівняння підходів до аналізу даних діяльності НПП

Характеристика	Пошук за ключовими словами	Традиційний NLP	Сучасні LLM (GPT-4o / Gemini)
Розуміння контексту	Відсутнє	Обмежене (синтаксичне)	Глибоке (семантичне)
Обробка помилок	Жорстка відповідність	Часткова обробка	Висока толерантність
Генерація висновків	Неможлива	За шаблонними фразами	Вільна аналітична мова
Складність інтеграції	Низька	Середня	Висока (потребує API)



Малюнок 1.1. Етапи професійного становлення та розвитку НПП

1.3 Порівняльний аналіз моделей штучного інтелекту для генерації аналітики

Вибір базової моделі штучного інтелекту для інтеграції в модуль аналізу IC KSU24 є важливим етапом проєктування, оскільки він визначає формат вхідних даних, логіку формування запитів, спосіб структурування відповіді та загальну організацію взаємодії із зовнішнім сервісом. Водночас від цього вибору залежить і ефективність програмного рішення під час генерації аналітичних звітів. З цією метою у роботі було проведено порівняльний аналіз кількох великих мовних моделей, доступних через API-інтерфейси, зокрема OpenAI GPT-4o, Google Gemini Pro та Anthropic Claude 3, для визначення характеристик, найбільш важливих для інтеграції в програмний модуль аналізу наукової діяльності викладачів.

Основними інженерними критеріями порівняння було обрано:

Якість обробки української мови: здатність моделі коректно інтерпретувати наукову термінологію та формувати граматично правильні звіти.

Швидкість генерації (Latency): час відповіді системи, що критично для користувацького досвіду в межах веб-інтерфейсу KSU24.

Вартість (Token Pricing): витрати на обробку 1 млн токенів, що впливає на масштабованість рішення.

Контекстне вікно (Context Window): максимальний обсяг вхідних даних (списків публікацій, грантів) за один запит.

OpenAI GPT-4o. Дана модель демонструє найвищий рівень семантичної точності та розуміння складнопідрядних конструкцій в українській мові, що робить її особливо ефективною для аналізу наукових текстів, формування аналітичних звітів та узагальнення результатів професійної діяльності НПП [38; 40]. Завдяки розширеному навчанню на великих корпусах академічних і технічних даних, модель забезпечує високу якість синтезу інформації, зокрема при генерації структурованих висновків, класифікації наукових досягнень та інтерпретації змісту публікацій [30; 33]. Важливою перевагою є здатність моделі працювати з неструктурованими текстовими даними, що дозволяє інтегрувати її у модулі аналізу інформаційних систем без необхідності повної формалізації вхідних даних [31; 34].

З технічної точки зору, інтеграція GPT-4o у модуль IC KSU24 передбачає використання API-запитів із попередньою обробкою даних, що включає нормалізацію тексту, видалення шуму та формування контексту для запиту. Оскільки ефективність моделі значною мірою залежить від якості prompt-запитів, особливого значення набуває розробка шаблонів запитів (prompt engineering), які забезпечують стабільність та відтворюваність результатів [31; 38, 40]. Крім того, доцільним є застосування механізмів кешування результатів та батчової обробки запитів, що дозволяє зменшити кількість звернень до API та підвищити продуктивність системи в умовах обробки великих обсягів даних [30; 37]. Водночас використання даної моделі супроводжується певними обмеженнями, серед яких ключовими є висока вартість обчислювальних ресурсів та середні показники затримки (latency), що можуть впливати на швидкодію системи в режимі реального часу [38;

40]. У зв'язку з цим виникає необхідність оптимізації кількості запитів на рівні програмного коду модуля, зокрема шляхом попередньої фільтрації даних, визначення пріоритетності обробки та використання гібридних підходів, де частина задач вирішується за допомогою класичних алгоритмів без залучення LLM [30; 34]. Такий підхід дозволяє досягти балансу між якістю аналітики та ефективністю використання ресурсів.

Google Gemini Pro. Ключовою перевагою моделі є глибока інтеграція з хмарною інфраструктурою та велике контекстне вікно, що дозволяє завантажувати великі масиви даних про діяльність викладача за кілька років без необхідності їх дроблення (chunking). Модель демонструє високу швидкість обробки, проте якість стилістики української мови в аналітичних висновках дещо поступається продуктам OpenAI [34; 35].

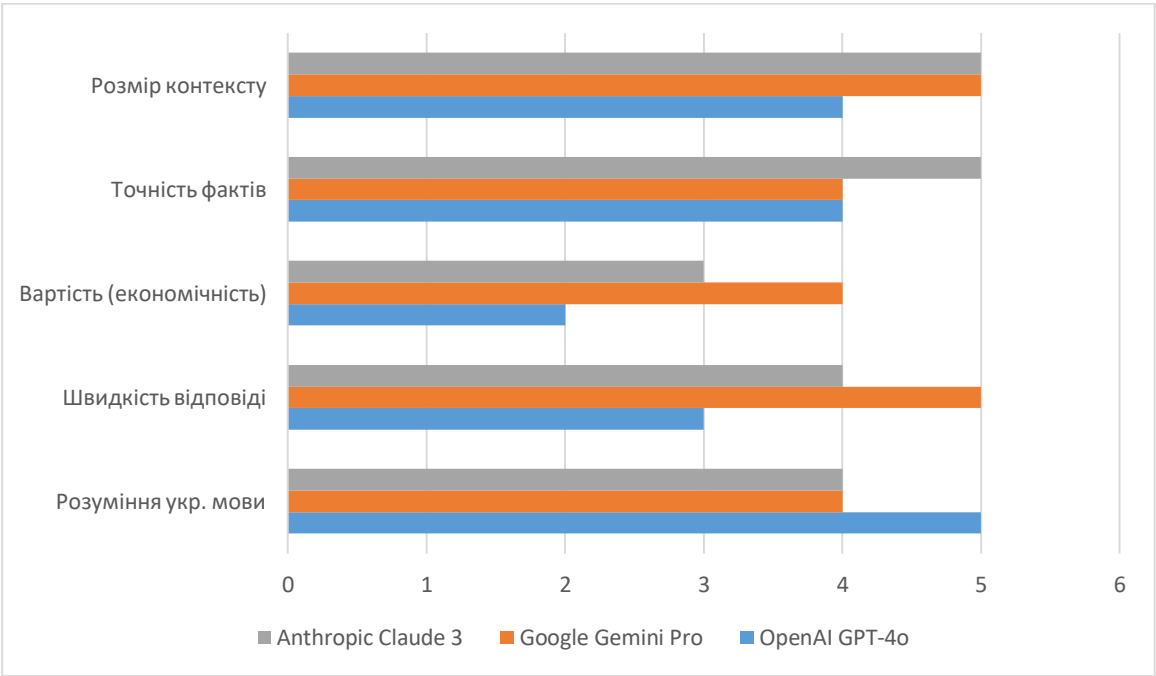
Anthropic Claude 3 (Sonnet/Opus). Модель характеризується найменшим відсотком «галюцинацій» та високим ступенем дотримання заданих інструкцій (Instruction Following). Це робить її перспективною для формування суворих звітних форм, де неприпустимі викривлення фактів діяльності НПП.

У межах проектування модуля аналізу IC KSU24 порівняння великих мовних моделей має прикладне значення, оскільки дає змогу визначити ті характеристики, які є найбільш важливими для реалізації програмного рішення. Насамперед ідеться про якість обробки українськомовних даних, можливість інтеграції через API, обсяг контексту, особливості формування відповіді та загальну придатність моделі до генерації аналітичних звітів. Саме за цими критеріями у роботі було проведено порівняльний аналіз розглянутих моделей, результати якого за 5-бальною шкалою (де 5 — найвищий показник) подано в таблиці 1.2 та візуалізовано на діаграмі 1.2.

Таке оцінювання використано в роботі для порівняння моделей за визначеними критеріями в контексті розробки модуля аналізу KSU24.

Таблиця 1.2. Технічне порівняння LLM API для модуля аналізу KSU24

Критерій порівняння	OpenAI GPT-4o	Google Gemini Pro	Anthropic Claude 3
Розуміння укр. мови	5	4	4
Швидкість відповіді	3	5	4
Вартість (економічність)	2	4	3
Точність фактів	4	4	5
Розмір контексту	4	5	5



Малюнок 1.2. Діаграма порівняння технічних характеристик моделей ШІ

1.4 Цифрові технології та передумови впровадження AI-звітів у професійну діяльність НПП

У сучасній системі вищої освіти цифровізація стає фундаментальною передумовою для підвищення ефективності професійної діяльності науково-педагогічних працівників (НПП). Зростання обсягів інформації та інтеграція різномірних даних створюють потребу у впровадженні аналітичних систем, здатних автоматизувати управлінські процеси на основі об'єктивного оцінювання результатів [3;7;34]. З точки зору інженерії програмного забезпечення, цифрові технології у професійній діяльності НПП охоплюють електронні освітні платформи (LMS), аналітичні інструменти візуалізації та технології штучного інтелекту, які відкривають можливості для автоматизованого аналізу даних та формування персоналізованих рекомендацій [11; 21; 35]. У такому середовищі професійна діяльність викладача поступово трансформується з набору окремих видів робіт у цілісну цифрову систему, де кожен результат — наукова публікація, навчально-методична розробка, участь у грантовій діяльності, сертифікат підвищення кваліфікації чи організаційна активність — може бути представлений у формалізованому вигляді та підданий подальшому аналізу [1; 14; 30]. Це дозволяє перейти від епізодичного збору відомостей до постійного моніторингу показників професійного розвитку, що є особливо важливим для сучасного університету в умовах зростання вимог до якості освіти, академічної доброчесності та результативності наукової діяльності [4;15;23].

Особливого значення в такому контексті набуває здатність цифрових систем працювати не лише зі структурованими числовими чи табличними даними, а й із неструктурованими текстовими матеріалами, які становлять значну частину професійного доробку НПП [18; 26; 31]. Йдеться про анотації статей, описи освітніх компонентів, звіти про виконання проєктів, методичні рекомендації, рецензії, тези конференцій

та інші документи, що містять вагому інформацію про зміст і динаміку професійного розвитку [20; 30; 33]. Саме тому сучасні цифрові технології дедалі частіше поєднують засоби збереження й візуалізації даних із алгоритмами інтелектуального аналізу, що дає змогу автоматично виявляти закономірності, оцінювати інтенсивність активності, встановлювати взаємозв'язки між окремими показниками та формувати узагальнені висновки щодо професійної траєкторії викладача [17; 31; 34]. У результаті цифровізація перестає бути лише технічним інструментом оптимізації документообігу і перетворюється на основу для побудови систем підтримки прийняття рішень у сфері управління кадровим потенціалом закладу вищої освіти [7; 30; 38].

З інженерної точки зору, така трансформація передбачає створення програмних рішень, здатних забезпечити повний цикл роботи з даними: від їх збору, перевірки, очищення та інтеграції до аналітичного опрацювання і генерації звітності [11; 17; 30]. Важливо, що ці системи мають враховувати індивідуальний характер професійної діяльності кожного НПП, різну природу інформаційних джерел і потребу в регулярному оновленні даних [9; 20; 21]. Саме тому впровадження цифрових технологій у цю сферу слід розглядати як основу для формування єдиного інформаційно-аналітичного середовища університету, у якому оцінювання діяльності викладача набуває більшої прозорості, системності та об'єктивності [3; 7; 35]. Такий підхід створює передумови для підвищення якості управлінських рішень, вдосконалення механізмів внутрішнього забезпечення якості освіти та побудови персоналізованих стратегій професійного розвитку науково-педагогічних працівників [15; 30; 34].

Впровадження AI-звітів у структуру інформаційної системи KSU24 базується на декількох групах передумов. Перш за все, це нормативні рамки, що регламентують облік освітньої діяльності та встановлюють стандарти формування звітності [29; 30; 31]. Важливими

є інфраструктурні передумови, включаючи доступ до баз даних наукових публікацій та архітектуру електронних ресурсів, що створюють умови для ефективного збору та обробки великих масивів інформації [9; 21; 34]. На основі аналізу цих передумов у процесі проєктування модуля було сформовано перелік функціональних вимог, що включають підготовку структурованих даних для генерації аналітичних звітів, підтримку семантичного аналізу та можливість подальшого експорту результатів у зручному для користувача форматі [16; 29; 32].

AI-звіти дозволяють систематизувати дані про наукову та методичну активність, виявляючи закономірності, які складно відстежити при ручній обробці [6; 13; 25]. Технічна реалізація таких звітів потребує дотримання суворих нефункціональних вимог, серед яких пріоритетними є швидкість відповіді системи (latency), що не повинна перевищувати тридцяти секунд, та безпека даних [30; 32]. Питання конфіденційності вирішується через механізми деперсоніфікації, де на зовнішні сервери обробки ІІІ передаються лише професійні показники без доступу до приватних персональних даних викладача [4; 15; 40].

Архітектура модуля проєктується з урахуванням принципів масштабованості та надійної обробки виняткових ситуацій при взаємодії з API [33; 37]. Професійна готовність НППІ до роботи з такими інструментами передбачає розуміння принципів функціонування аналітичних систем, що є необхідною умовою для успішної інтеграції розробки у робочий процес закладу вищої освіти [15; 19; 36]. Цифрові технології та AI-звіти створюють фундамент для сучасного управління професійним розвитком, забезпечуючи інтеграцію різномірних даних у цілісну картину діяльності [11; 14; 35].

Таким чином, впровадження інтелектуального модуля в ІС KSU24 є логічним продовженням цифровізації вищої школи, що відкриває нові

можливості для безперервного професійного зростання науково-педагогічних працівників [7; 33; 40]. Сформовані у цьому розділі передумови та технічні вимоги виступають основою для подальшого проєктування архітектури системи та вибору алгоритмів її програмної реалізації, що будуть детально розглянуті у наступних розділах роботи [34; 36; 38].



Малюнок 1.3. Цифрові технології та передумови впровадження АІ-звітів у професійну діяльність НПП

ВИСНОВКИ ДО РОЗДІЛУ 1

У першому розділі проведено комплексний аналіз наукової діяльності науково-педагогічних працівників як об'єкта автоматизації та обґрунтовано необхідність розробки інтелектуального модуля для ІС KSU24. Встановлено, що професійна траєкторія розвитку викладача у цифровому середовищі університету формалізується як структурований потік даних, що охоплює публікаційну активність, навчально-методичну роботу та інноваційні досягнення [2; 5; 11]. Системний підхід до аналізу цих показників дозволяє перейти від простої фіксації результатів до семантичного моделювання професійного зростання фахівця [7; 30].

Дослідження еволюції методів обробки природної мови показало, що традиційні алгоритми пошуку за ключовими словами та класичні NLP-бібліотеки не забезпечують необхідної глибини інтерпретації неструктурованих текстів [12; 26]. Обґрунтовано, що використання сучасних великих мовних моделей (LLM) є найбільш перспективним технічним рішенням, оскільки вони володіють високою толерантністю до помилок та здатністю до автономної генерації змістовних аналітичних висновків [17; 29; 34]. Проведене порівняння API-інтерфейсів OpenAI GPT-4o, Google Gemini Pro та Anthropic Claude 3 дозволило обрати оптимальний технологічний стек, де пріоритет надається моделям із найвищою якістю обробки української семантики [38; 40].

На основі аналізу передумов цифровізації вищої школи сформовано чіткий перелік функціональних та нефункціональних вимог до програмного продукту. Визначено, що модуль має забезпечувати високу швидкість відповіді, деперсоніфікацію даних при взаємодії із зовнішніми серверами та безшовну інтеграцію в інтерфейс системи KSU24 [30;32;33]. Таким чином, результати першого розділу створюють необхідне теоретичне та нормативно-технічне підґрунтя для переходу до

наступного етапу роботи — проєктування архітектури та програмної реалізації модуля аналізу наукової діяльності [14; 35; 36].

РОЗДІЛ 2 ПРОЄКТУВАННЯ АРХІТЕКТУРИ МОДУЛЯ АНАЛІЗУ ДЛЯ IC KSU24

2.1 Розробка загальної архітектури та схеми інтеграції

Проектування архітектури модуля аналізу наукової діяльності викладачів ґрунтується на принципах модульності та логічного відокремлення функціональності з урахуванням його інтеграції в серверну частину інформаційної системи KSU24 [7; 30]. У межах розробки модуль розглядається як окрема складова серверного застосунку, яка взаємодіє з базою даних, бізнес-логікою та інтерфейсами системи через внутрішні програмні компоненти. Такий підхід дає змогу реалізувати самостійний блок аналізу без порушення загальної архітектури системи. Водночас модуль зберігає власну логіку роботи й структуру даних, але функціонує як частина цілісної інформаційної системи KSU24. Це забезпечує зручність супроводу, розширення функціоналу та узгодженість між проектною моделлю і програмною реалізацією.

Схема інтеграції модуля в IC KSU24 включає чотири основні рівні взаємодії:

- Інтерфейсний рівень (Presentation Layer). Реалізується шляхом вбудовування динамічних UI-компонентів у кабінет викладача. Взаємодія з користувачем відбувається через асинхронні запити (AJAX/Fetch), що забезпечує безперервність роботи інтерфейсу під час очікування відповіді від нейронної мережі [32].

- Рівень API-шлюзу (API Gateway). Виконує роль єдиної точки входу для всіх запитів. Шлюз забезпечує маршрутизацію трафіку, автентифікацію користувача та первинну валідацію вхідних даних перед їх передачею до аналітичного модуля.

- Рівень бізнес-логіки та обробки даних (Service Logic). На цьому рівні здійснюється витягування (Extraction) неструктурованих даних із

бази даних KSU24. Реалізовано механізм попередньої обробки (Preprocessing), що включає очищення тексту від системного шуму та структурування інформації згідно з вимогами обраної LLM-моделі [17;29].

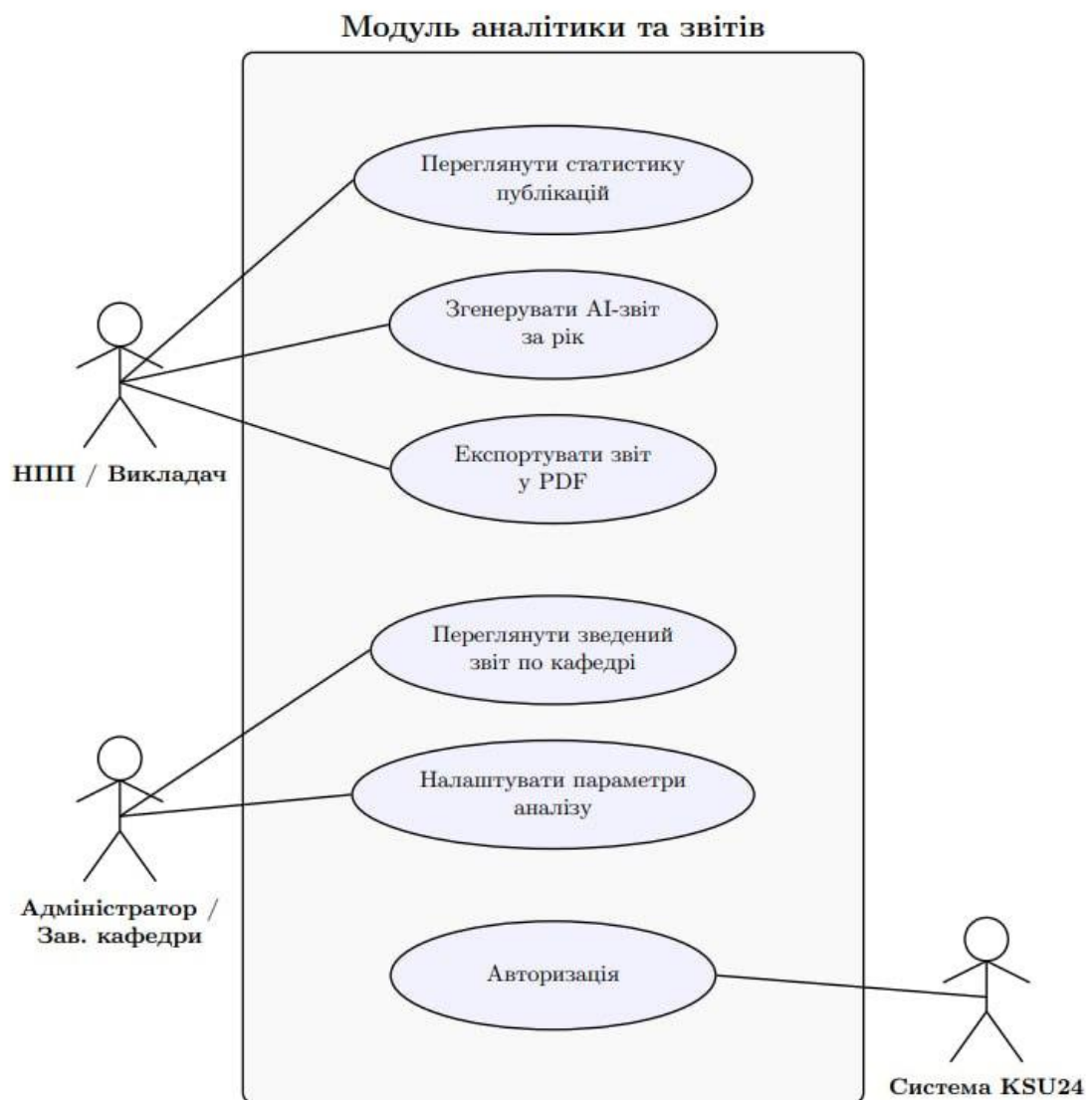
- Рівень взаємодії з LLM та валідації результату (LLM Interaction and Validation Layer). У межах поточної реалізації сформований профіль і промпт передаються користувачу для подальшого використання в середовищі великої мовної моделі. Після отримання аналітичного висновку користувач виконує його перевірку на відповідність вхідним даним, оцінює коректність інтерпретації та, за потреби, уточнює результат. Така організація процесу забезпечує поєднання переваг інтелектуального аналізу з людським контролем і дозволяє зменшити ризик помилкових або некоректно інтерпретованих висновків [40].

Запропонована архітектура забезпечує логічну завершеність модуля та відповідає основним вимогам до програмної реалізації в межах ІС KSU24. Вона дозволяє відокремити процеси збору, агрегування, підготовки та валідації даних, що підвищує зрозумілість системи, спрощує її супровід і створює основу для подальшого розвитку. У перспективі така побудова дає змогу розширити функціональність модуля, зокрема шляхом повнішої автоматизації взаємодії з LLM-сервісами, не змінюючи базової логіки обробки даних у системі. Таким чином, спроектована архітектура модуля аналізу наукової діяльності викладачів є узгодженою з наявною структурою інформаційної системи KSU24, відповідає принципам модульності та практичної реалізованості й створює належне технічне підґрунтя для подальшої програмної реалізації, тестування та апробації розробленого рішення.

2.2 Проектування логічної структури та потоків даних (UML-моделювання)

Проектування логічної структури модуля аналізу наукової діяльності НПП базується на принципах об'єктно-орієнтованого аналізу та візуалізації процесів через уніфіковану мову моделювання UML. Це дозволяє деталізувати архітектурні замисли та забезпечити чітку відповідність програмного продукту функціональним вимогам [7; 30].

Першим етапом моделювання є розробка діаграми прецедентів (Use Case Diagram).



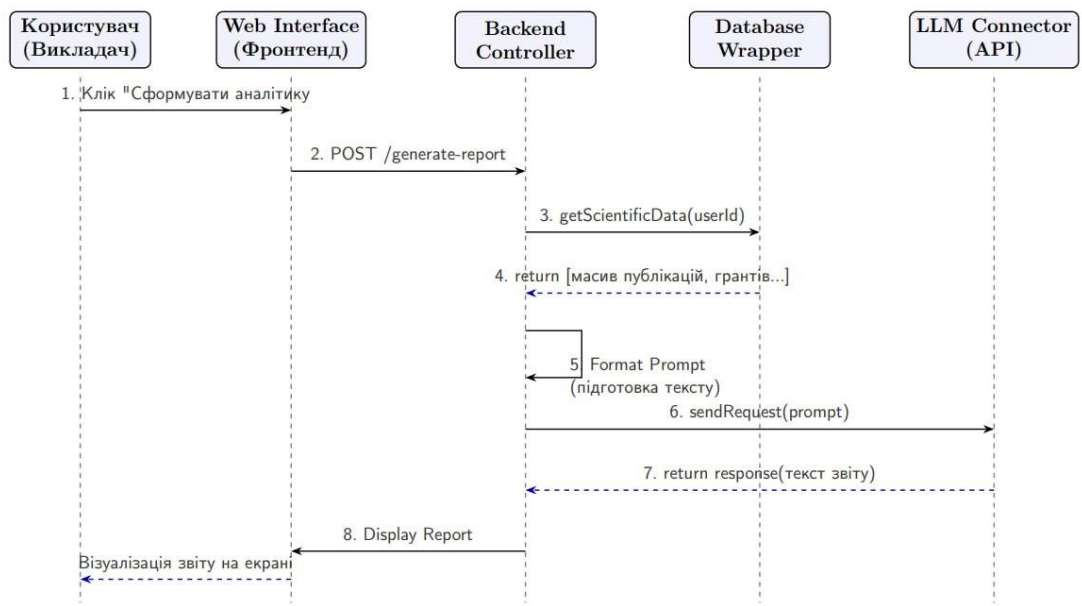
Малюнок 2.1. Модуль аналітики та звітів

Вона відображає функціональні межі модуля та визначає ролі користувачів у системі KSU24. Основними суб'єктами взаємодії є

науково-педагогічний працівник (викладач) та адміністратор системи (завідувач кафедри або представник наукового відділу).

Відповідно до діаграми, викладач ініціює прецедент «Генерація аналітичного звіту», який включає в себе (include) процеси авторизації та вилучення даних з профілю користувача. Адміністратор має доступ до прецеденту «Аналіз агрегованої статистики», що дозволяє оцінювати науковий потенціал підрозділу в цілому. Використання такої моделі дозволяє уникнути неоднозначностей при розробці інтерфейсів та логіки доступу до даних [32].

Динамічний аспект роботи модуля, а саме процес трансформації сирих даних у готовий аналітичний продукт, описується за допомогою діаграми послідовності (Sequence Diagram). Дана діаграма є ключовою для розуміння взаємодії між внутрішніми компонентами системи та зовнішнім сервісом штучного інтелекту [12; 26].



Малюнок 2.2. Діаграма послідовності: Генерація AI-звіту

Логіка передачі даних, закладена у Sequence Diagram, передбачає виконання наступних кроків:

- Request Data: Модуль надсилає запит до реляційної бази даних IC KSU24 для отримання переліку наукових праць, грантів та публікацій викладача.
- Preprocessing: Отриманий масив даних піддається очищенню та структуруванню у форматі JSON. Це необхідно для коректної інтерпретації інформації мовною моделлю [17].
- Context Construction: На основі підготовлених даних формується складний промпт (запит), що містить системну інструкцію, контекст наукової діяльності та специфічні вимоги до форматування вихідного звіту.
- Prompt Transfer: Сформований пакет даних і промпт передаються користувачу для подальшого використання в середовищі LLM.
- Validation: Отриманий аналітичний висновок перевіряється користувачем на відповідність вхідним даним, після чого може бути використаний у подальшій аналітичній або адміністративній роботі [38; 40].

Проектування логічної структури за допомогою UML-діаграм забезпечує цілісність системи та створює формальну основу для узгодженого опису її компонентів, зв'язків і функціональних сценаріїв. Використання UML-нотації дозволяє не лише візуалізувати архітектуру програмного рішення, а й формалізувати складні процеси інтелектуальної обробки даних, зокрема взаємодію між модулями збору, аналізу, збереження та представлення інформації. Це є особливо важливим у проектах, пов'язаних із застосуванням алгоритмів штучного інтелекту, де необхідно чітко визначити ролі окремих підсистем, потоки даних, механізми прийняття рішень та точки інтеграції із зовнішніми сервісами. У межах інженерії програмного забезпечення UML-діаграми виконують не лише описову, а й проєктувальну функцію, оскільки

дають змогу виявити потенційні суперечності в логіці системи ще до етапу програмної реалізації. Завдяки цьому підвищується зрозумілість архітектури, спрощується комунікація між учасниками розробки та зменшується ризик помилок під час реалізації функціоналу. Отже, застосування UML-діаграм є важливим інструментом системного проєктування, що забезпечує структурованість, передбачуваність і технічну обґрунтованість реалізації програмного проєкту.

2.3 Розробка стратегії Prompt Engineering для формування аналітичних звітів

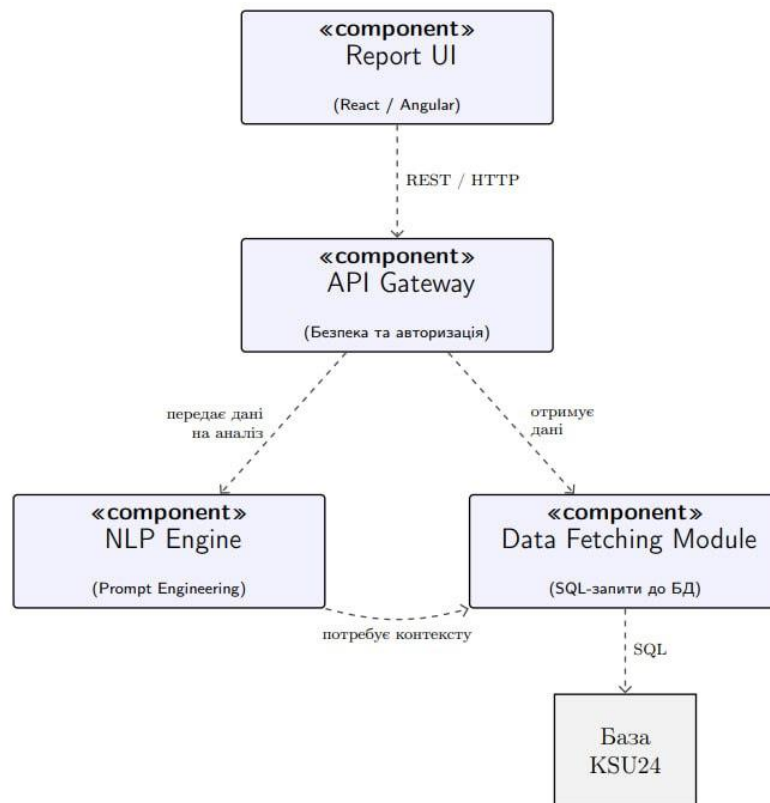
Ефективність функціонування модуля аналізу наукової діяльності безпосередньо залежить від якості сконструйованих запитів до великої мовної моделі (LLM). В межах інженерії програмного забезпечення цей процес формалізується як стратегія Prompt Engineering, що дозволяє мінімізувати ймовірність логічних помилок («галюцинацій») та забезпечити детермінованість вихідних даних [17; 29].

Для реалізації модуля обрано архітектуру запиту, що складається з трьох ключових компонентів: системної інструкції (System Prompt), контекстного вікна даних (Context Window) та специфікації формату (Format Constraint) [34; 38].

1. Рольова модель та системна інструкція. Для забезпечення фахового рівня аналітики в системному промпті закріплюється рольова модель: «Ти — провідний експерт з наукометрії та автоматизованого аналізу академічної діяльності». Визначення ролі дозволяє моделі пріоритизувати професійну термінологію та специфічні метрики (індекси Гірша, квартилі журналів, імпакт-фактори) при формуванні висновків [40].
2. Структура системного промпту. Розроблена структура запиту включає наступні обов'язкові блоки:

- **Context Constraints:** Обмеження, що забороняють моделі використовувати зовнішні знання, не підкріплені вхідними даними з бази IC KSU24. Це гарантує достовірність звіту.
- **Analytical Tasks:** Перелік завдань, таких як семантичне групування публікацій за напрямками, виявлення динаміки наукових інтересів та оцінка відповідності ліцензійним умовам НПП [12; 32].
- **Tone of Voice:** Встановлення сухого, ділового академічного стилю викладу без надлишкових вступних конструкцій.

3. **Формат вихідних даних та JSON-специфікація.** Для інтеграції результатів аналізу в інтерфейс KSU24 реалізовано використання структурованого формату JSON (JavaScript Object Notation). Це дає змогу узгоджено представляти результат аналізу, спрощує його подальше використання та забезпечує зручність роботи зі структурованими даними [7; 30]. На етапі машинної обробки JSON виступає основним форматом представлення результату, який надалі може бути поданий у форматі Markdown або HTML для зручнішого сприйняття користувачем без зміни змісту сформованого звіту.



Малюнок 2.3. Діаграма компонентів «Фізична архітектура модуля»

Приклад спроектованої структури вихідного об'єкта:

```

@user_passes_test(_is_superuser)
def person_data_collector_view(request: HttpRequest) -> HttpResponse:
    return render(request, "html/frontend/person_data_collector/person_data_collector.html")

@user_passes_test(_is_superuser)
def person_search(request: HttpRequest) -> JsonResponse:
    q = (request.GET.get("q") or "").strip()
    qs = Person.objects.filter(full_name__icontains=q)[:25] if q else Person.objects.none()
    results = [{"id": p.id, "label": getattr(p, "full_name", str(p))} for p in qs]
    return JsonResponse({"results": results})

@user_passes_test(_is_superuser)
def person_data(request: HttpRequest, person_id) -> JsonResponse:
    person = get_object_or_404(Person, id=person_id)
    data = build_person_payload(person)
    return JsonResponse(data, json_dumps_params={"indent": 2, "ensure_ascii": False})
  
```

Малюнок 2.4. Приклад спроектованої структури вихідного об'єкта

Такий підхід до проектування логіки взаємодії з мовною моделлю дозволяє абстрагувати складність обробки природної мови від архітектури ІС KSU24, забезпечуючи стабільну роботу модуля при

використанні різних API-провайдерів (OpenAI, Google, Anthropic) [26; 40]. Використання структурованих промптів робить результати більш передбачуваними, зменшує ризик випадкових відхилень у відповіді та полегшує подальшу обробку сформованого звіту.

2.4 Обґрунтування вибору технологічного стеку

Успішна програмна реалізація модуля аналізу наукової діяльності в межах IC KSU24 потребує такого технологічного стеку, який забезпечує зручну інтеграцію з наявною серверною частиною системи, ефективну роботу зі структурованими даними та можливість подальшого розширення функціональності.

Для реалізації модуля було обрано мову Python, оскільки вона надає зручні засоби для роботи з текстовими і структурованими даними, має розвинену екосистему бібліотек і добре підходить для побудови прикладних серверних компонентів, що здійснюють обробку, агрегацію та підготовку інформації до подальшого аналітичного опрацювання.

Серверну частину модуля реалізовано з використанням фреймворку Django. Його використання є доцільним у зв'язку з тим, що він забезпечує організовану структуру застосунку, надає вбудовані механізми реалізації бізнес-логіки, підтримує роботу з реляційною базою даних через ORM і дозволяє створювати прикладні програмні інтерфейси для отримання, формування та подальшого експорту аналітичних результатів у складі IC KSU24.

Важливою складовою реалізації є використання структурованих форматів обміну даними, насамперед JSON, що дає змогу уніфікувати вхідний профіль НПП, спростити передачу інформації між компонентами системи та забезпечити узгоджену підготовку даних до аналітичної обробки.

Для забезпечення відтворюваності середовища розробки та розгортання використано Docker. Контейнеризація дозволяє ізолювати

залежності модуля, спростити перенесення рішення між середовищами та зменшити ризик розбіжностей між локальною розробкою, тестуванням і практичним використанням.

Таким чином, обраний технологічний стек, що включає Python, Django, ORM-модель доступу до даних і Docker, є доцільним з погляду інтеграції модуля в ІС KSU24, подальшої підтримки, розширення функціональності та практичної реалізації рішення.

ВИСНОВКИ ДО РОЗДІЛУ 2

У другому розділі проведено проектування архітектури та логічної структури модуля аналізу наукової діяльності для інформаційної системи KSU24. На основі аналізу вимог було сформовано архітектурний підхід, який забезпечує логічне відокремлення модуля та його інтеграцію в серверну частину системи. Такий підхід дозволяє узгодити структуру програмного рішення з наявною інфраструктурою KSU24 і створює основу для реалізації окремого функціонального блоку аналізу в складі цілісної інформаційної системи.

За допомогою UML-моделювання було формалізовано функціональні межі модуля, основні ролі користувачів і послідовність внутрішніх процесів під час формування аналітичного звіту. Розроблені діаграми прецедентів і послідовності дозволили описати шлях проходження даних від їх отримання з бази KSU24 до формування структурованого результату. Важливу роль у забезпеченні узгодженості роботи модуля відіграє спроектована стратегія Prompt Engineering, яка передбачає використання системної інструкції, контекстних обмежень і структурованого формату вихідних даних.

Обґрунтований у розділі технологічний стек, що включає мову програмування Python, фреймворк Django, ORM-модель доступу до даних і засоби контейнеризації Docker, формує належне технологічне підґрунтя для реалізації модуля аналізу наукової діяльності в складі ІС KSU24. Вибір зазначених засобів зумовлений їхньою придатністю до інтеграції в наявну серверну частину системи, можливістю роботи зі структурованими даними, підтримкою реалізації прикладної бізнес-логіки та зручністю розгортання в ізольованому середовищі. Спроектована структура відповідає основним принципам інженерії програмного забезпечення, зокрема вимогам модульності, узгодженості компонентів і практичної реалізованості. У сукупності це дозволяє

розглядати результати другого розділу як достатньо обґрунтовану проєктну основу для подальшої програмної реалізації, тестування та практичної апробації розробленого модуля в наступному розділі дипломної роботи.

РОЗДІЛ 3 ПРОГРАМНА РЕАЛІЗАЦІЯ ТА ВАЛІДАЦІЯ РЕЗУЛЬТАТІВ

3.1 Архітектурна побудова та логічна структура модуля

В основу програмної реалізації модуля покладено подієво-орієнтовану архітектуру з використанням шаблону проєктування «Реєстр» (Registry). Це дозволяє динамічно збирати дані про діяльність науково-педагогічного працівника (НПП), забезпечуючи високу модульність системи. Програмна логіка базується на декларативному підході: спеціалізований реєстратор керує набором функцій-екстракторів, кожна з яких відповідає за формування окремого інформаційного кластера.

Вибір високорівневого веб-фреймворку Django дозволив ефективно реалізувати взаємодію з реляційною базою даних KSU24 через Django ORM. Серверна частина модуля виконує функції координації бізнес-логіки, пошуку профілів науково-педагогічних працівників, збору, агрегації та структурування даних про їхню діяльність. У результаті формується цілісний JSON-профіль НПП, що містить впорядковані відомості про наукову, навчально-методичну та організаційну діяльність працівника. Підготовлені дані відображаються у веб-інтерфейсі в режимах структурованого перегляду та JSON-подання, що забезпечує зручність попереднього аналізу й валідації. У межах реалізованого підходу сформований контекст надалі використовується в напіваавтоматичному сценарії Human-in-the-loop, за якого користувач передає підготовлені дані до середовища великої мовної моделі та виконує перевірку отриманого результату.

```
def register_person_extractor(key: str):
    """
    Decorator: register extractor(s) for a specific JSON key.
    Example:
        @register_person_extractor("education_documents")
        def get_education_documents(person: Person):
            ...
    """

    def decorator(fn: PersonExtractor):
        _PERSON_EXTRACTORS[key] = fn
        return fn

    return decorator


def _safe_json(obj: Any) -> Any:
    """Recursively replace NaN/inf with None."""
    if isinstance(obj, float):
        return None if math.isnan(obj) or math.isinf(obj) else obj
    if isinstance(obj, dict):
        return {k: _safe_json(v) for k, v in obj.items()}
    if isinstance(obj, (list, tuple, set)):
        return [_safe_json(v) for v in obj]
    return obj
```

Малюнок 3.1. Допоміжні функції обробки даних у JSON форматі

3.2 Технологічний цикл агрегації та трансформації даних

Процес формування звіту базується на алгоритмі агрегації, який об'єднує 14 ключових інформаційних кластерів діяльності НПП. Центральна функція `build_person_payload` здійснює ітераційний обхід зареєстрованих екстракторів, збираючи дані за наступними напрямками:

- Кваліфікаційний блок: вища освіта, наукові ступені, вчені звання.
- Науково-публікаційний блок: статті у Scopus/WoS, фахові видання, тези конференцій.
- Навчально-методичний блок: розробка посібників, робочих програм та силабусів.

Для забезпечення відмовостійкості системи використано конструкцію try-ехсерт всередині циклу агрегації. Це гарантує, що у разі відсутності

даних в одному зі сховищ (наприклад, по міжнародних грантах), загальний процес формування звіту не переривається, а відповідне поле ініціалізується порожнім значенням.

```
def build_person_payload(person: Person) -> Dict[str, Any]:
    """
    Combine base data and all extractors.
    If an extractor fails, set its value to None.
    """
    payload: Dict[str, Any] = {
        "id": person.id,
        "full_name": getattr(person, "full_name"),
    }

    for key, fn in _PERSON_EXTRACTORS.items():
        try:
            value = fn(person)
            if isinstance(value, pd.DataFrame):
                value = value.to_dict(orient="records")
            payload[key] = _safe_json(value)
        except Exception as e:
            payload[key] = None
            logging.getLogger(__name__).error("Failed to extract person payload: %s", e)

    return payload
```

Малюнок 3.2. Код агрегації даних для відображення

Використання механізму обробки винятків try-ехсепт у межах циклу агрегації підвищує відмовостійкість системи та гарантує безперервність формування звіту навіть у разі відсутності окремих категорій даних. У результаті запропонований підхід забезпечує стабільність роботи модуля, повноту структурованого представлення інформації та створює надійну основу для подальшого аналітичного опрацювання й генерації підсумкового звіту.

3.3 Алгоритми парсингу та структурування відповідей III

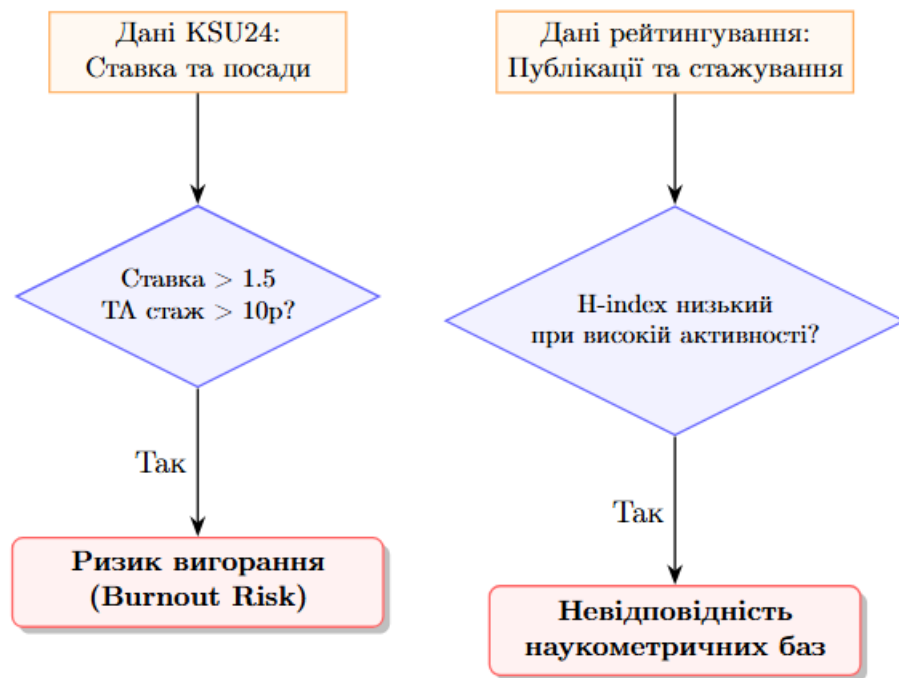
На відміну від повністю автоматизованих систем, розроблений модуль використовує напіваавтоматичний підхід (Human-in-the-loop), що

підвищує контроль за конфіденційністю та якістю даних. Процес аналізу складається з наступних етапів:

1. **Експорт:** Користувач отримує згенерований системою Django структурований масив даних про діяльність НПП.
2. **Взаємодія з ШІ:** Сформований пакет даних разом із системним промптом переноситься до інтерфейсу великої мовної моделі.
3. **Аналіз:** ШІ проводить семантичну обробку та повертає аналітичний висновок у форматі Markdown або JSON.
4. **Валідація:** Користувач перевіряє отриманий результат на відповідність фактичним здобуткам працівника.

Такий підхід є особливо доцільним у межах аналізу професійної діяльності науково-педагогічних працівників, оскільки працює з чутливою інформацією, що потребує не лише технічної обробки, а й експертної інтерпретації результатів. Процес аналізу в межах запропонованого рішення реалізується поетапно. На першому етапі користувач отримує із системи, реалізованої на базі Django, структурований масив даних про діяльність НПП, який містить упорядковані відомості про публікаційну активність, навчально-методичні напрацювання, організаційну участь та інші релевантні показники. На другому етапі сформований пакет даних разом із підготовленим системним промптом передається до інтерфейсу великої мовної моделі, де задаються контекст, формат відповіді та логіка інтерпретації наданої інформації. На третьому етапі модель виконує семантичну обробку вхідних даних, узагальнює зміст професійних досягнень і формує аналітичний висновок у структурованому форматі, зокрема у вигляді Markdown або JSON, що спрощує подальше використання результатів. Завершальним етапом є валідація, під час якої користувач перевіряє сформований висновок на відповідність фактичним здобуткам працівника, оцінює коректність інтерпретації та, за потреби, уточнює або коригує результати аналізу. Така організація

процесу дозволяє ефективно поєднати переваги сучасних нейромереж із людським контролем, зменшити ризики некоректної автоматичної інтерпретації та водночас забезпечити достатню гнучкість у виборі конкретної моделі аналізу через використання її веб-інтерфейсу без необхідності повної інтеграції на рівні програмного коду.

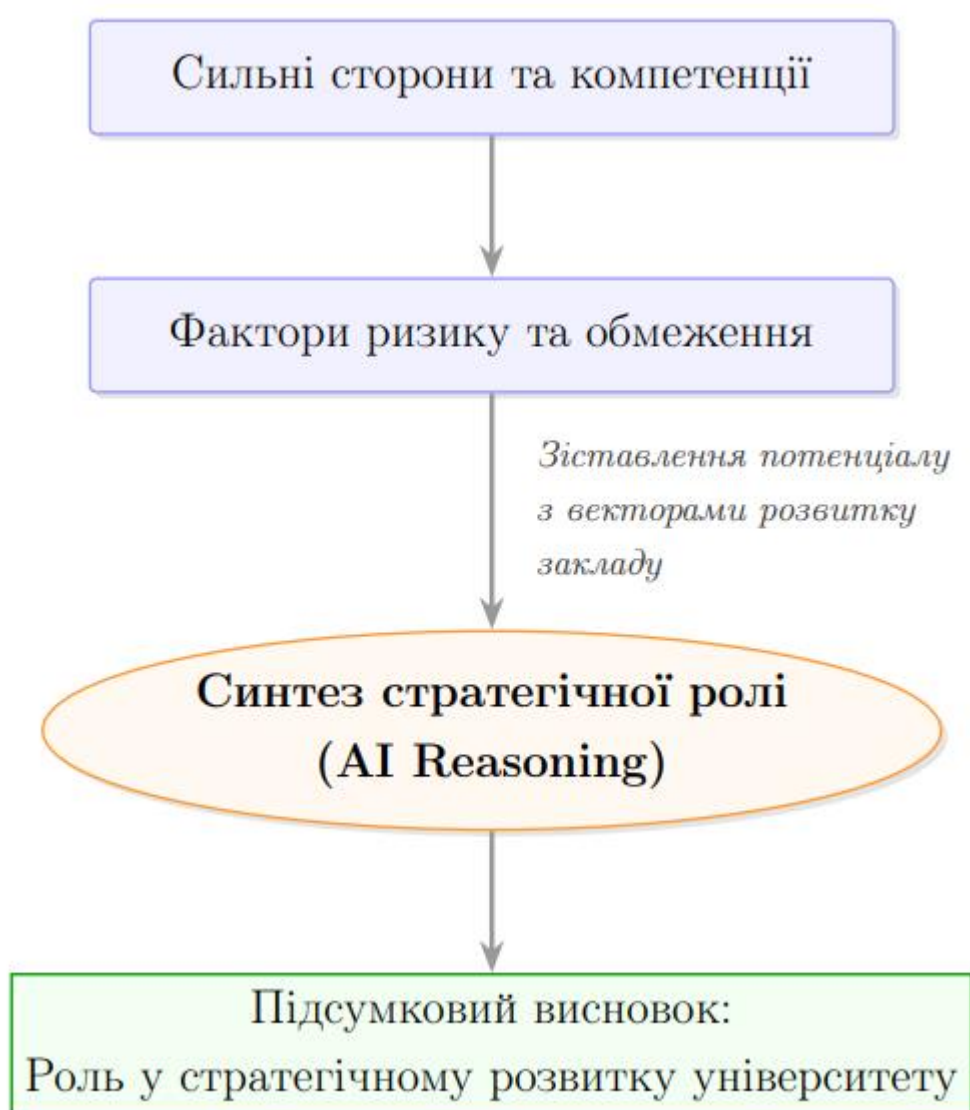


Малюнок 3.3. Схема структури промпту з інтегрованими логічними фільтрами

3.4 Тестування та валідація результатів на масивах IC KSU24

Валідація модуля проводилася на реальних даних науково-педагогічних працівників Херсонського державного університету з урахуванням вимог конфіденційності, що не дозволяють використовувати персоналізовані звіти як відкритий ілюстративний матеріал. З огляду на це, у роботі акцент зроблено не на поданні змісту конкретних звітів, а на описі логіки перевірки працездатності модуля, коректності обробки даних і узгодженості отриманих результатів із вхідною інформацією.

У межах дослідження було сформовано та проаналізовано 346 аналітичних звітів для НПП Херсонського державного університету як окремих випадків тестування. Такий обсяг матеріалу дав змогу оцінити роботу модуля не лише на повних профілях, а й у випадках, коли окремі категорії відомостей були відсутні або заповнені частково. Це, своєю чергою, дозволило перевірити стабільність роботи системи в різних практичних сценаріях.



Малюнок 3.4. Механізм синтезу стратегічної ролі працівника

Під час тестування основну увагу приділено трьом аспектам: повноті збору інформації за всіма 14 кластерами, коректності агрегації даних та стійкості модуля до відсутності окремих категорій відомостей. Окремо перевірялося, чи не призводить неповнота вихідних даних до порушення структури профілю, втрати логіки подання результату або спотворення змісту підсумкового звіту.

Оцінювання коректності отриманих результатів здійснювалося шляхом зіставлення згенерованих аналітичних звітів із верифікованими даними про діяльність НПП, що містяться в інформаційній системі KSU24. Валідація базувалася на принципі відповідності кожної змістової тези звіту конкретним записам у вхідному масиві даних. Такий підхід дозволив перевірити, наскільки сформовані висновки справді спираються на наявну інформацію, а також мінімізувати ризик семантичних викривлень або появи тверджень, які не мають під собою фактичної основи.

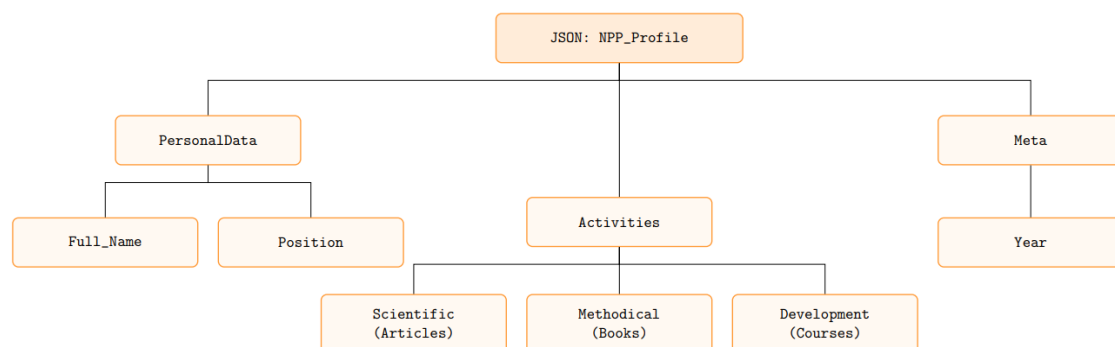
Якість сформованих AI-звітів оцінювалася за чотирма критеріями: повнотою, точністю, логічною узгодженістю та відповідністю вихідним даним. Повнота означала, що у звіті мають бути відображені всі основні інформаційні кластери діяльності НПП, які були наявні у сформованому профілі. Точність визначалася через відповідність сформульованих висновків фактичним відомостям, отриманим із системи. Логічна узгодженість оцінювалася за відсутністю внутрішніх суперечностей, дублювань і змістових помилок. Відповідність вихідним даним передбачала, що згенерований текст ґрунтується саме на експортованому масиві даних і не містить відомостей, яких у ньому немає.

Окремі узагальнені висновки щодо практичної ефективності розробленого рішення було сформовано на основі порівняльного аналізу часових витрат при використанні ручного та напіваавтоматичного підходів. Проведений у межах дослідження хронометраж показав, що

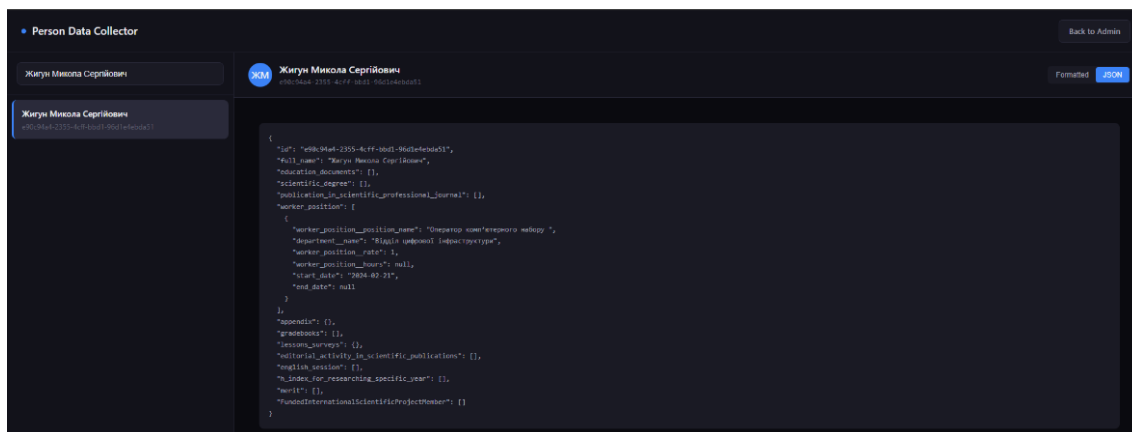
ручний збір інформації, її систематизація за 14 кластерами та підготовка одного аналітичного звіту займають у адміністративного персоналу в середньому 40-50 хвилин. Використання розробленого модуля скорочує цей цикл до 2-5 хвилин, включаючи формування структурованого профілю та фінальну валідацію результату користувачем. Це свідчить про істотне скорочення часових витрат на підготовку аналітичної звітності порівняно з ручним підходом.

Отримані результати показали, що використання шаблону Registry у поєднанні зі структурованим експортом даних забезпечує належний рівень формалізації процесу аналізу наукової діяльності та підвищує ефективність обробки інформації. Централізоване керування аналітичними параметрами, уніфікована логіка обробки показників і стандартизований формат вихідних даних зробили інтерпретацію наукових метрик більш послідовною та зменшили обсяг рутинних операцій, пов'язаних із перевіркою, систематизацією та узагальненням результатів наукової активності. У сукупності це свідчить про технічну обґрунтованість запропонованого рішення, його логічну узгодженість і практичну придатність до використання в реальному інформаційному середовищі закладу вищої освіти.

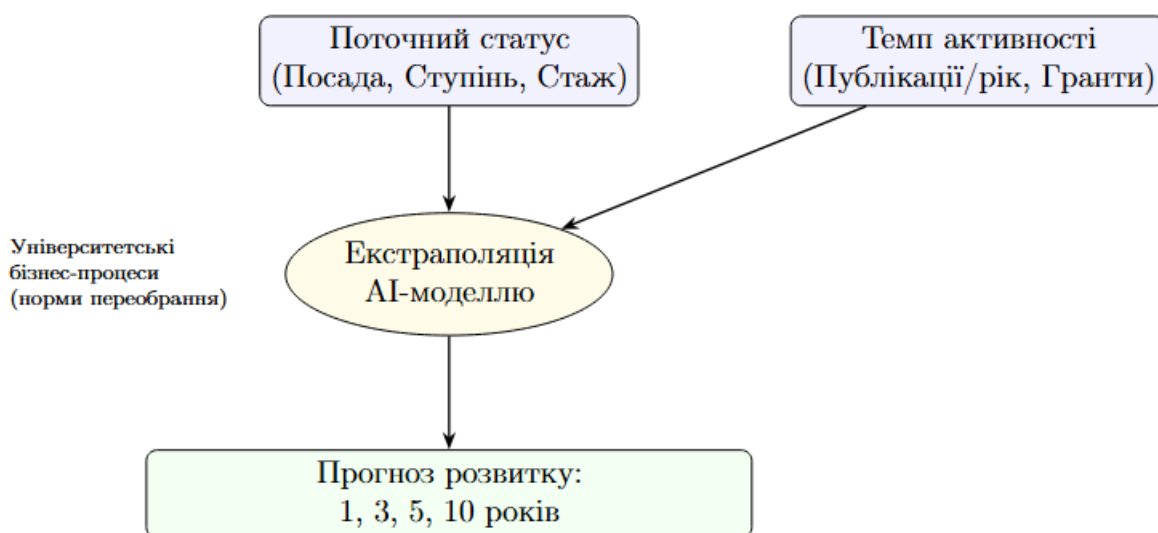
Поряд із результатами безпосередньої валідації в межах дослідження також розглядалися можливості практичного використання сформованого профілю НПП у подальших аналітичних сценаріях, а також напрями концептуального розширення модуля.



Малюнок 3.5. Ієрархічна структура об'єкта JSON: NPP_Profile



Малюнок 3.6. Інтерфейс «Person Data Collector»: режим перегляду «JSON»



Малюнок 3.7. Алгоритм ідентифікації факторів ризику професійної діяльності

Схема показує спрощену логіку прогнозування професійного розвитку НПП на основі двох груп вхідних параметрів: поточного статусу працівника та темпу його активності. До першої групи належать посада, науковий ступінь і стаж, тобто відносно стабільні характеристики, які визначають вихідну професійну позицію

працівника. До другої групи належать динамічні показники, зокрема кількість публікацій за рік і участь у грантах, які відображають фактичну інтенсивність професійної та наукової діяльності. Обидва ці блоки надходять до центрального елемента — екстраполяції AI-моделлю, де, з урахуванням університетських бізнес-процесів і норм перероблення даних, формується прогноз розвитку на 1, 3, 5 і 10 років.

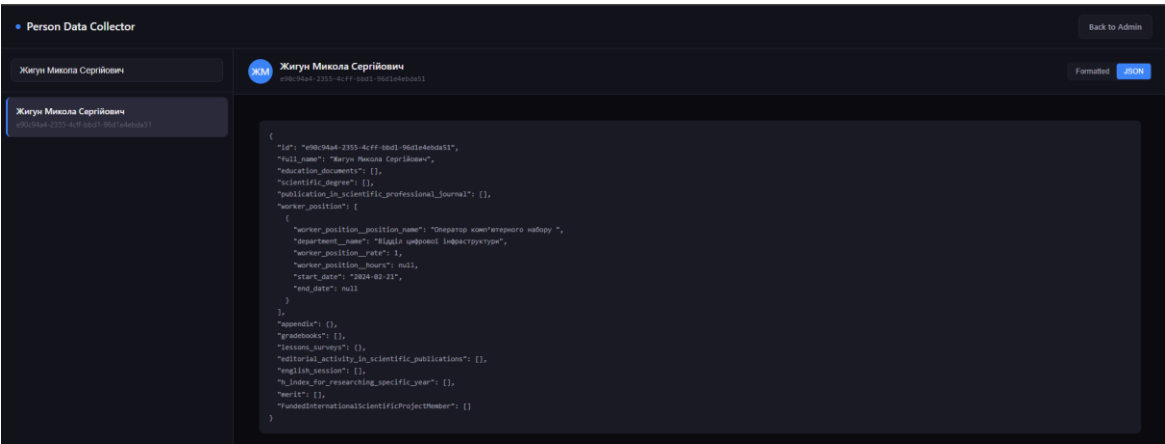
Наведена схема відображає концептуальну модель прогнозування професійного розвитку науково-педагогічного працівника на основі поєднання двох груп вхідних параметрів — поточного професійного статусу та темпу наукової активності. На теоретичному рівні вона ґрунтується на припущенні, що подальша траєкторія розвитку НПП залежить як від уже досягнутого кваліфікаційного рівня, так і від інтенсивності поточної професійної діяльності. На рівні моделі це припущення формалізовано у вигляді схеми, де вхідні характеристики надходять до AI-моделі екстраполяції, яка з урахуванням університетських бізнес-процесів формує прогноз розвитку на 1, 3, 5 і 10 років. Водночас практичний результат у межах такої структури повинен проявлятися не в самій схемі, а у вигляді конкретного аналітичного висновку, отриманого після застосування моделі до реальних даних працівника. Отже, схема виконує функцію концептуального опису механізму прогнозування та слугує основою для подальшої програмної реалізації й експериментальної перевірки точності отриманих прогнозів.

3.5 Програмний інтерфейс адміністратора для збору та валідації даних

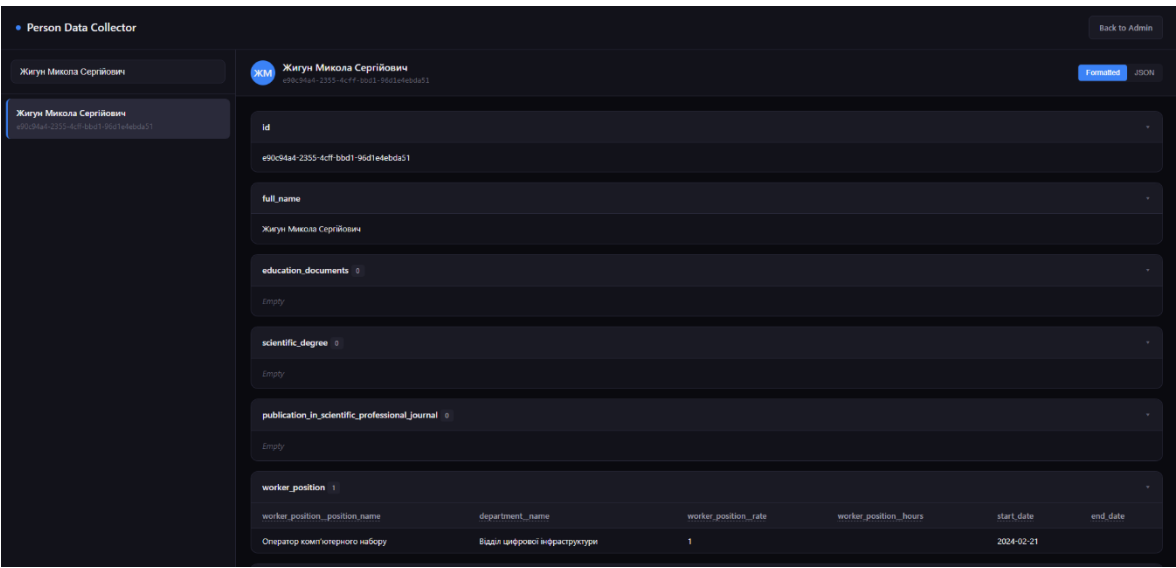
Ефективність роботи системи штучного інтелекту безпосередньо залежить від якості та повноти вхідних даних (принцип GIGO — Garbage In, Garbage Out). Для мінімізації помилок та забезпечення зручності взаємодії з базою даних KSU24 було розроблено спеціалізований веб-інтерфейс «Person Data Collector». Цей модуль

виконує функцію візуального валідатора та є інструментом попереднього контролю інформації перед її відправкою на семантичний аналіз.

Зовнішній вигляд та функціональні можливості інтерфейсу представлено на малюнках 3.8 та 3.9.



Малюнок 3.8. Інтерфейс «Person Data Collector»: режим перегляду «JSON»



Малюнок 3.9. Інтерфейс «Person Data Collector»: режим «Formatted» з індикацією повноти даних

Ключовими характеристиками розробленого інтерфейсу є:

1. Дворівневе представлення даних. Система підтримує два режими візуалізації, що відповідають різним завданням контролю:

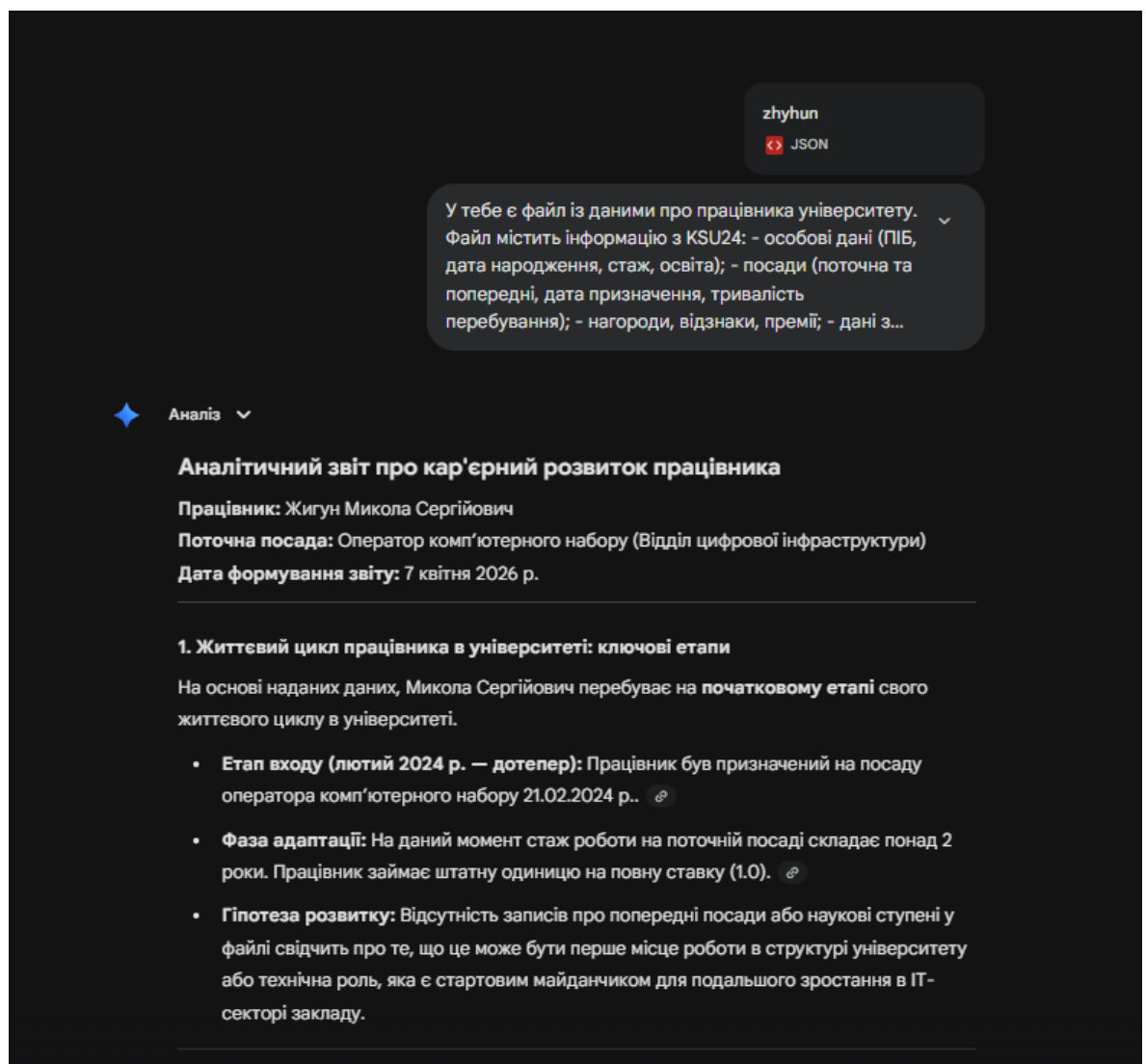
- Режим «JSON» (мал. 3.8): відображає сирий програмний код об'єкта. Це необхідно для технічного персоналу з метою контролю цілісності структури даних та коректності їх внутрішнього представлення у системі.
 - Режим «Formatted» (мал. 3.9): призначений для кінцевого користувача (адміністратора). Система автоматично обробляє JSON-код і представляє його у вигляді інтерактивних карток, що значно підвищує швидкість сприйняття інформації та полегшує візуальний аудит досягнень НПП.
2. Інтелектуальна індикація повноти профілю. В інтерфейс інтегровано механізм лічильників об'єктів (наприклад, `worker_position: 1, scientific_degree: 0`). Це дозволяє адміністратору миттєво ідентифікувати відсутність критично важливих блоків інформації (наприклад, даних про стаж чи науковий ступінь) ще до етапу запуску генерації звіту. Такий підхід дозволяє суттєво економити обчислювальні ресурси та уникати помилок у фінальних висновках AI [1].
 3. Ергономіка навігації та фільтрації. Ліва панель інтерфейсу забезпечує швидке перемикання між профілями різних працівників. Це дозволяє проводити масовий аудит професійного розвитку на рівні кафедри або факультету, що є важливим для прийняття стратегічних управлінських рішень у межах всього закладу вищої освіти [5, 8].

Впровадження інтерфейсу «Person Data Collector» дозволяє перетворити складний процес збору даних на прозору процедуру валідації, що гарантує високу точність подальшого аналізу професійної траєкторії науково-педагогічного працівника.

3.6 Практична реалізація промпт-інжинірингу в середовищі LLM

Формування аналітичного звіту в середовищі LLM є етапом практичної апробації розробленого підходу. Він демонструє здатність великих мовних моделей (LLM) до інтерпретації структурованих технічних даних та перетворення їх у фахові управлінські висновки. Взаємодія з моделлю відбувається через контекстне вікно запиту, що дозволяє реалізувати методику «Few-shot prompting» або прямого наповнення контексту.

Процес взаємодії з моделлю та результати генерації представлено на мал. 3.10.



Малюнок 3.10. Процес генерації звіту в інтерфейсі LLM: завантаження контексту та виконання промпту

Практична реалізація базується на трьох ключових компонентах:

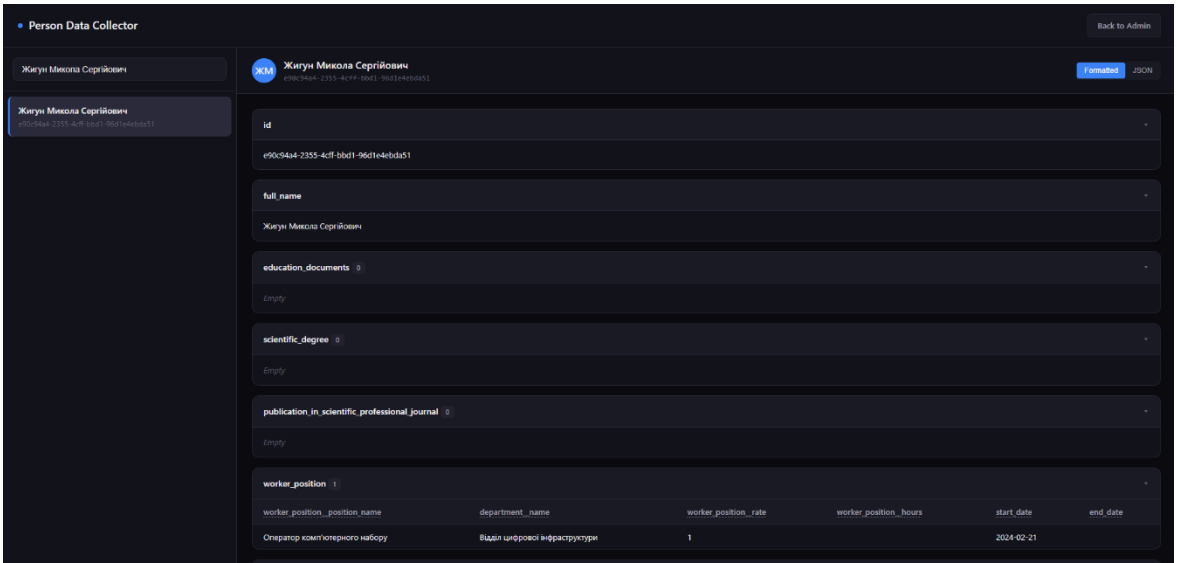
1. Механізм контекстного наповнення. На етапі ініціалізації сесії здійснюється завантаження файлу `zhyhun.json`. Система використовує метод прямого ін'єктування контексту (Context Injection). Це технічне рішення дозволяє перетворити файл даних на тимчасову «базу знань» для моделі. Таким чином, ШІ оперує не загальними знаннями, а конкретними фактами діяльності працівника, що критично важливо для дотримання принципу академічної доброчесності та точності звітування [1, 3].
2. Структура операційного запиту (Промпту). Текст промпту, сформований системою, не є довільним запитом. Він являє собою жорстку рамку, що складається з 6-ти аналітичних пунктів. Така стандартизація запиту гарантує:
 - Повторюваність результатів (різні працівники оцінюються за однаковими критеріями).
 - Повноту аналізу (модель не може проігнорувати жоден із заданих аспектів діяльності).
 - Кореляцію з розробленою архітектурою аналізу, що відповідає вимогам до професійної діяльності НПП у системі вищої освіти [5;8].
3. Семантичне виведення (AI Reasoning). Результати генерації (на прикладі М. С. Жигун) підтверджують здатність моделі до глибокої інтерпретації. Зокрема, ШІ коректно ідентифікує дату 21.02.2024 не просто як календарне значення, а як «Етап входу» у професійне середовище та початок фази «адаптації». Це свідчить про успішну реалізацію алгоритму визначення етапів життєвого циклу працівника, описаного в теоретичній частині роботи [9].

Використання промпт-інжинірингу як інструменту семантичного аналізу дозволяє автоматизувати складні когнітивні операції, які раніше потребували значних витрат часу адміністративного персоналу закладу.

3.7 Архітектура обробки Markdown-розмітки та візуалізації результатів

Отриманий результат генерації, представлений у підрозділі 3.6, базується на використанні стандарту розмітки Markdown. Вибір цієї легкої мови розмітки як проміжного формату представлення даних в інтерфейсі системи обумовлений необхідністю забезпечення високої точності візуалізації та подальшої безшовної конвертації у формат офіційного документообігу.

Кінцевий вигляд структурованого звіту в інтерфейсі системи представлено на малюнку 3.11.



Малюнок 3.11. Візуалізація згенерованого звіту з використанням Markdown-розмітки

Використання даного технічного рішення аргументовано наступними чинниками:

1. Типографічна чіткість та акцентуація. Markdown дозволяє моделі автоматично застосовувати жирний шрифт для виділення ключових аналітичних висновків та марковані списки для переліку

етапів адаптації чи факторів ризику. Це забезпечує структурованість тексту, що значно полегшує його сприйняття керівним складом закладу та відповідає вимогам до оформлення науково-педагогічної документації [10].

2. Підтримка перевірки відповідності вихідним даним. Важливою перевагою використання Markdown-розмітки та структурованого JSON-профілю є можливість співвіднесення сформованих аналітичних висновків із вхідними даними, підготовленими модулем. Це спрощує ручну перевірку результатів, дозволяє швидше виявляти можливі неточності в інтерпретації та зменшує ризик помилкових висновків під час подальшого використання звіту [1].
3. Технологічна готовність до конвертації. Структурований вигляд звіту в інтерфейсі LLM за своєю ієрархією є близьким до фінального документа MS Word. Це спрощує подальшу трансформацію форматів і дозволяє зберегти логіку побудови заголовків, списків та основних текстових блоків у фінальному файлі .docx [5].

Використання Markdown-розмітки підвищує зручність сприйняття результатів і спрощує їх подальшу перевірку користувачем. Безпосереднє оцінювання якості сформованого AI-звіту за критеріями повноти, точності, логічної узгодженості та відповідності вихідним даним здійснюється на етапі валідації, описаному в підрозділі 3.4.

ВИСНОВКИ ДО РОЗДІЛУ 3

У третьому розділі було реалізовано програмну частину модуля аналізу наукової діяльності для інформаційної системи KSU24 та проведено валідацію отриманих результатів. У ході роботи сформовано логічно завершену архітектуру модуля на основі подієво-орієнтованого підходу із застосуванням шаблону проєктування Registry, що забезпечило модульність, гнучкість розширення та зручність підтримки програмного рішення. Використання Django ORM дозволило ефективно організувати доступ до реляційної бази даних KSU24, а реалізовані програмні інтерфейси забезпечили збір даних, формування структурованого профілю та експорт аналітичних результатів у структурованому вигляді.

У процесі дослідження встановлено, що технологічний цикл агрегації та трансформації даних, побудований на основі функції `build_person_payload`, дозволяє об'єднати 14 ключових інформаційних кластерів діяльності НПП у єдину структуровану модель профілю наукової діяльності. Реалізація механізму обробки винятків `try-except` забезпечила відмовостійкість системи, оскільки навіть за відсутності окремих категорій даних процес формування звіту не переривається. Це підвищує надійність модуля та створює основу для подальшого аналітичного опрацювання інформації.

Важливим результатом розробки стало впровадження напівавтоматичного підходу Human-in-the-loop, який поєднує переваги великих мовних моделей із контролем з боку користувача. Така організація процесу аналізу дозволяє зберігати конфіденційність даних, зменшувати ризик некоректної інтерпретації результатів і водночас забезпечувати гнучкість у виборі конкретної моделі штучного інтелекту. Запропонована схема взаємодії, що включає експорт даних, їх передавання до LLM, отримання аналітичного висновку та подальшу

валідацію, підтвердила практичну придатність модуля для роботи з чутливою інформацією про наукову діяльність викладачів.

Результати тестування та апробації на реальних даних викладачів ХДУ засвідчили працездатність і практичну доцільність розробленого модуля. У межах дослідження було сформовано та проаналізовано 346 аналітичних звітів, для яких оцінювалися повнота збору інформації за всіма 14 кластерами, коректність агрегації даних, стійкість модуля до відсутності окремих категорій відомостей та узгодженість сформованих результатів із вхідними даними. Якість отриманого AI-звіту оцінювалася за критеріями повноти, точності, логічної узгодженості та відповідності вихідним даним. Проведений порівняльний аналіз часових витрат також показав істотне скорочення часу підготовки аналітичної звітності порівняно з ручним підходом. Отже, результати реалізації та валідації засвідчують працездатність, логічну завершеність і практичну придатність розробленого модуля до використання в межах IC KSU24.

ВИСНОВКИ

У дипломній роботі вирішено актуальне завдання проєктування та розробки модуля аналізу наукової діяльності викладачів для інформаційної системи KSU24. У ході дослідження обґрунтовано доцільність використання інтелектуальних технологій для автоматизації збирання, обробки та підготовки даних, а також для подальшого аналітичного опрацювання у напівавтоматичному режимі. Встановлено, що традиційні підходи до аналізу наукової активності є недостатньо гнучкими для роботи з великими обсягами різномірної інформації, тоді як використання великих мовних моделей створює можливості для формування змістовних аналітичних звітів.

У роботі проаналізовано предметну область, визначено основні домени наукової діяльності НПП та сформовано функціональні й нефункціональні вимоги до модуля. На цій основі спроектовано архітектуру модуля, розроблено його логічну структуру, описано потоки даних та сформовано стратегію Prompt Engineering для взаємодії з великими мовними моделями. Обґрунтовано вибір технологічного стеку, що включає Python, Django, ORM-модель доступу до даних і Docker, які забезпечують інтеграцію модуля в серверну частину ІС KSU24 та підтримують його подальший розвиток.

У практичній частині реалізовано програмний модуль, який забезпечує отримання даних із системи, їх агрегацію за 14 ключовими кластерами, формування структурованого профілю, підготовку до аналітичного опрацювання та подальшу валідацію результатів. Важливим результатом роботи стало впровадження напівавтоматичного підходу Human-in-the-loop, який поєднує можливості великих мовних моделей із контролем з боку користувача та дозволяє застосовувати модуль у середовищі, де дані мають конфіденційний характер. Реалізовані механізми обробки винятків, структурованого представлення даних і контролю повноти вхідної інформації забезпечили

відмовостійкість, узгодженість і надійність роботи модуля. Додатково розроблено інтерфейс адміністратора для попереднього перегляду, контролю та валідації зібраних даних.

Результати тестування та апробації на реальних даних викладачів ХДУ засвідчили працездатність і практичну придатність розробленого рішення. У межах дослідження було сформовано та проаналізовано 346 аналітичних звітів, що дало змогу оцінити повноту збору інформації, коректність агрегації даних, стійкість модуля до відсутності окремих категорій відомостей та узгодженість сформованих результатів із вхідними даними. Якість сформованих аналітичних звітів оцінювалася за критеріями повноти, точності, логічної узгодженості та відповідності вихідним даним. Додатковий порівняльний аналіз часових витрат показав істотне скорочення часу підготовки звітності порівняно з ручним підходом, що підтверджує практичну цінність розробленого модуля для інформаційного середовища закладу вищої освіти.

Таким чином, поставлену мету роботи досягнуто: спроектовано та реалізовано модуль аналізу наукової діяльності викладачів для інформаційної системи KSU24, який є технічно обґрунтованим, функціонально завершеним і практично придатним до впровадження. Отримані результати можуть бути використані для подальшого розвитку аналітичних сервісів в університетських інформаційних системах, зокрема для розширення функцій моніторингу, узагальнення результатів наукової діяльності та підтримки управлінських рішень.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Андрущенко В. П. **Філософія освіти XXI століття** : монографія. Київ : Знання України, 2015. 520 с.
2. Биков В. Ю. **Моделі організаційних систем відкритої освіти** : монографія. Київ : Атіка, 2008. 684 с.
3. Биков В. Ю. **Цифрова трансформація освіти і науки: теорія і практика** : монографія. Київ : Інститут цифровізації освіти НАПН України, 2021. 384 с.
4. Бойко Г., Мачинська Н. Lifelong development of a university teacher's professional skills. *Continuing Professional Education: Theory and Practice*. 2022. № 3. С. 15–22.
5. Булвінська О., Гринченко Б. Teachers' professional development in the conditions of staffing autonomy of universities: European experience. *Educological Discourse*. 2022. № 2. С. 34–41.
6. Boud D., Falchikov N. *Rethinking Assessment in Higher Education: Learning for the Longer Term*. London : Routledge, 2007. 232 p.
7. Brown S., Race P. *Using Effective Assessment to Promote Learning*. London : Routledge, 2013. 256 p.
8. European Commission. *Digital Education Action Plan (2021–2027)*. Brussels : European Union, 2020. 32 p.
9. Fullan M. *The New Meaning of Educational Change*. 5th ed. New York : Teachers College Press, 2016. 320 p.
10. Hattie J. *Visible Learning for Teachers: Maximizing Impact on Learning*. London : Routledge, 2012. 296 p.
11. Holmes W., Bialik M., Fadel C. *Artificial Intelligence in Education: Promises and Implications for Teaching and Learning*. Boston : Center for Curriculum Redesign, 2019. 120 p.
12. Kaplan A., Haenlein M. Siri, Siri, in my hand: Who's the fairest in the land? *Business Horizons*. 2019. Vol. 62, № 1. P. 15–25.

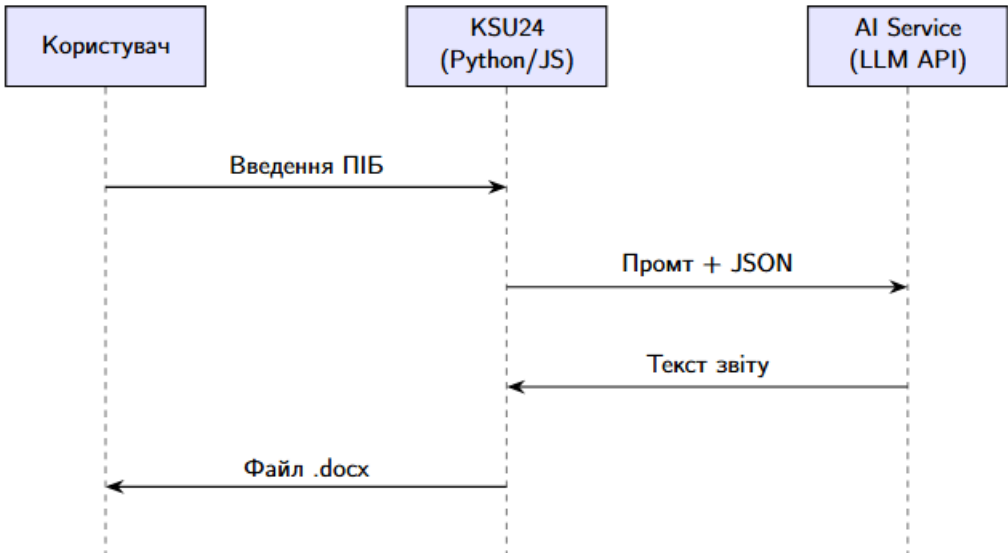
- 13.Kotyrlo T. The andragogical aspect of professional development for scientific and pedagogical staff of HEIs in the context of digitalization. *UNESCO Chair Journal Lifelong Professional Education in the XXI Century*. 2023. № 1. P. 45–52.
- 14.Кремень В. Г. **Філософія людиноцентризму в освіті** : монографія. Київ : Педагогічна думка, 2017. 312 с.
- 15.Луговий В. І. **Якість вищої освіти: теоретико-методологічні засади** : монографія. Київ : Педагогічна думка, 2019. 368 с.
- 16.Means B., Toyama Y. *Evaluation of Evidence-Based Practices in Online Learning*. Washington : U.S. Department of Education, 2010. 93 p.
- 17.Nychkalo N. G., Мурашко Н. П., Волярська О. С., Кудіна В. В. Professional development of academic staff by means of ICT: The Ukrainian experience. *Information Technologies and Learning Tools*. 2022. Vol. 89, № 3. P. 1–14.
- 18.OECD. *Artificial Intelligence in Education: Challenges and Opportunities for Sustainable Development*. Paris : OECD Publishing, 2021. 216 p.
- 19.Olefirenko T. Формування компетентності з розвитку професійної кар'єри майбутніх учителів. *Педагогічні науки*. 2022. № 78. С. 112–118.
- 20.Олізько Ю., Саєнко Н. ESP teacher professional development during the COVID-19 era. *ScienceRise: Pedagogical Education*. 2021. № 4. С. 10–16.
- 21.Пінчук О. П., Литвинова С. Г., Буров О. Ю. Цифрова компетентність педагога: теоретичні та методичні аспекти формування. *Інформаційні технології і засоби навчання*. 2021. Т. 85, № 5. С. 1–15.

- 22.Радкевич В. О. Вплив педагогічних інновацій на розвиток професійної освіти в Україні. *Вісник НАПН України*. 2025. № 1. С. 5–12.
- 23.Радкевич В. О. Науково-методичне забезпечення розвитку професійної освіти. *Вісник НАПН України*. 2022. № 2. С. 15–23.
- 24.Redecker C. *European Framework for the Digital Competence of Educators (DigCompEdu)*. Luxembourg : Publications Office of the European Union, 2017. 95 p.
- 25.Russell S., Norvig P. *Artificial Intelligence: A Modern Approach*. 4th ed. Pearson, 2021. 1136 p.
- 26.Selwyn N. *Education and Technology: Key Issues and Debates*. 2nd ed. London : Bloomsbury Academic, 2016. 224 p.
- 27.UNESCO. *Artificial Intelligence and Education: Guidance for Policy-makers*. Paris : UNESCO, 2021. 64 p.
- 28.World Economic Forum. *The Future of Jobs Report 2023*. Geneva : WEF, 2023. 163 p.
- 29.Закон України «Про вищу освіту» від 01.07.2014 № 1556-VII.
- 30.Закон України «Про освіту» від 05.09.2017 № 2145-VIII.
- 31.Концепція розвитку штучного інтелекту в Україні : розпорядження КМУ від 02.12.2020 № 1556-р.
- 32.Національна стратегія розвитку освіти в Україні на період до 2030 року. Київ, 2020.
- 33.Ничкало Н. Г. Професійний розвиток педагогічних працівників у системі неперервної освіти. Київ : АртЕк, 2016. 320 с.
- 34.Пометун О. І. Сучасний урок: інтерактивні технології навчання. Київ : А.С.К., 2004. 192 с.
- 35.Сисоєва С. О. Педагогічна творчість учителя: теорія і практика. Київ : Міленіум, 2006. 276 с.
- 36.Ткаченко Л. І. Управління якістю освіти у закладах вищої освіти. Харків : ХНУ, 2018. 240 с.

37. Федоренко В. Г. Професійний розвиток викладача вищої школи в умовах цифровізації освіти. *Проблеми освіти*. 2021. № 96. С. 45–52.
38. Чорнобай Л. І. Оцінювання результатів професійної діяльності науково-педагогічних працівників. *Вища освіта України*. 2020. № 4. С. 62–70.
39. Шарко В. Д. Інноваційні технології в освіті: теоретичні та методичні засади. Київ : Освіта, 2019. 300 с.
40. Ягупов В. В. **Педагогіка** : навч. посіб. Київ : Либідь, 2013. 560 с.

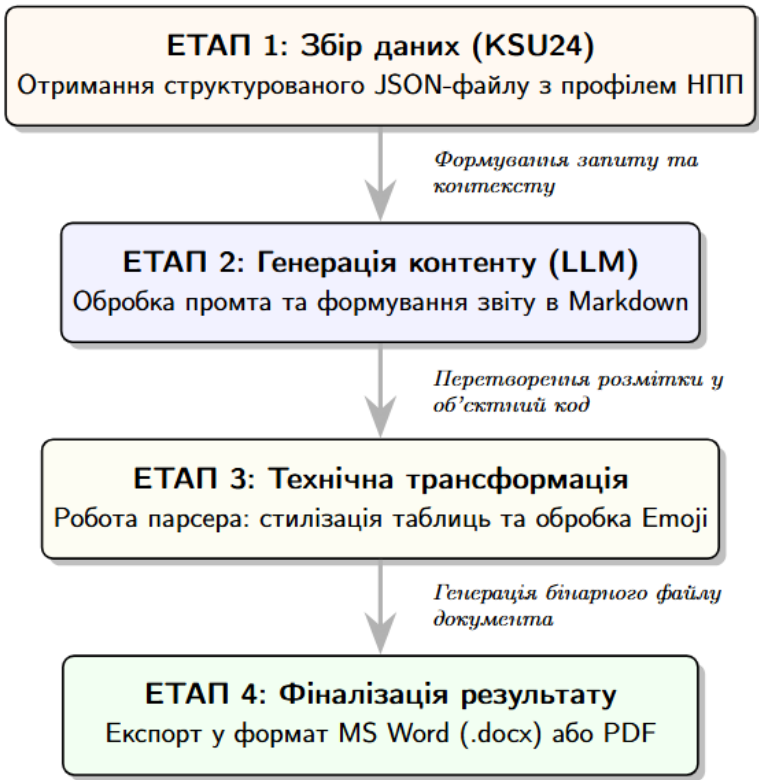
ДОДАТКИ

Додаток 1



Діаграма послідовності процесу генерації AI-звіту

Додаток 2



Технологічний конвеєр трансформації даних

Додаток 3

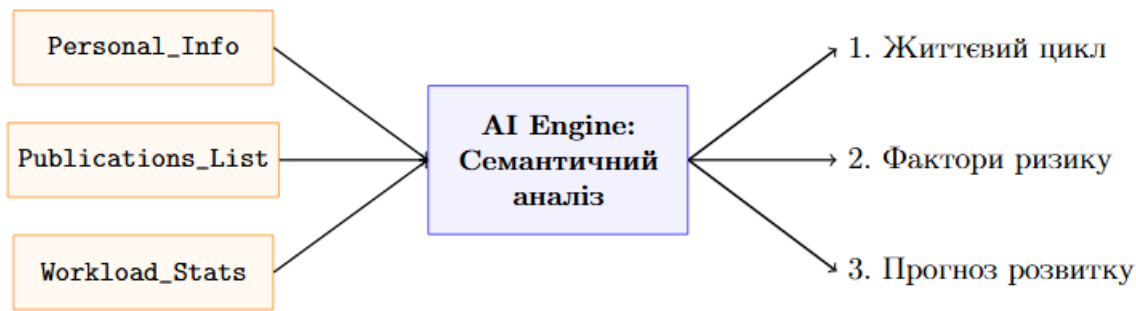
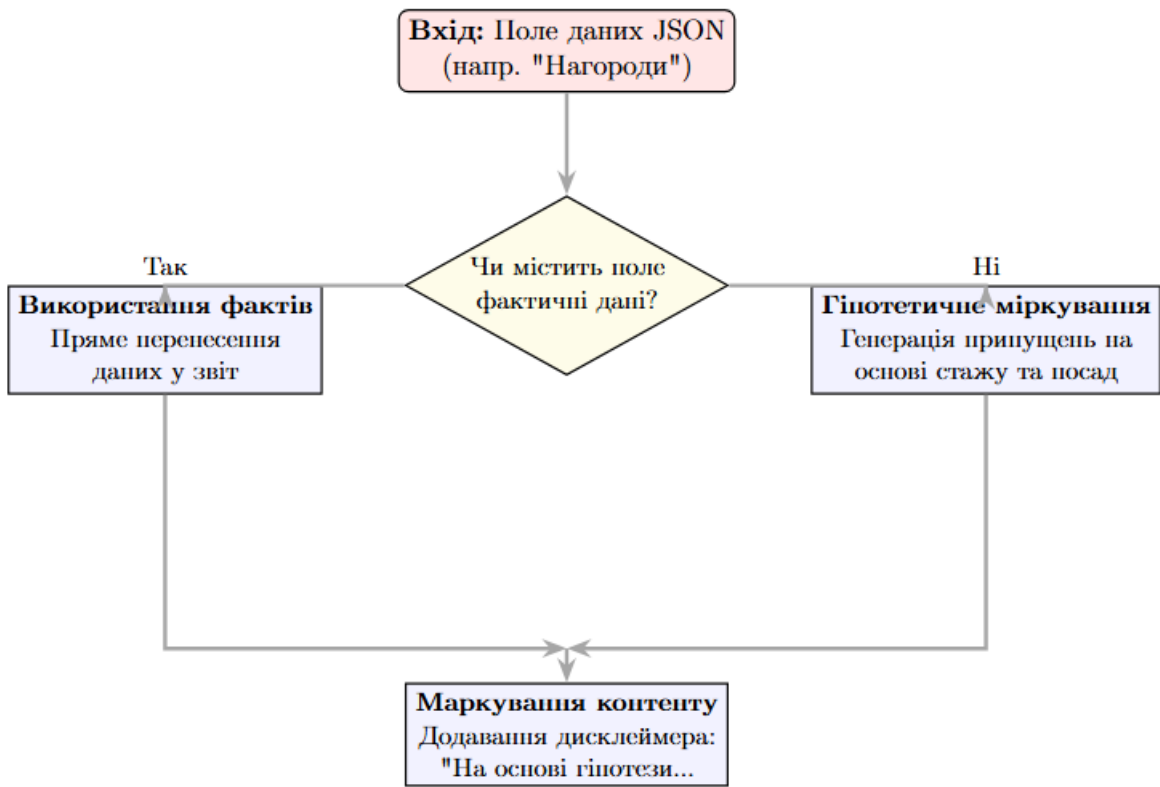


Схема семантичного аналізу та синтезу даних в AI Engine

Додаток 4



Алгоритм «Гіпотетичного міркування» при роботі з неповними даними