

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ  
ЗАПОРІЗЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ  
ІНЖЕНЕРНИЙ НАВЧАЛЬНО-НАУКОВИЙ ІНСТИТУТ ім. Ю.М. ПОТЕБНИ

Кафедра інформаційної економіки, підприємництва та фінансів  
(повна назва кафедри)

**Кваліфікаційна робота**  
другий (магістерський)  
(рівень вищої освіти)

на тему Управління освітнім процесом Університету на основі інформаційної моделі великих даних у блокчейн

Виконав: студент 2 курсу, групи 8.0512-іє  
спеціальності 051 Економіка  
(код і назва спеціальності)  
освітньої програми Інформаційна економіка  
(код і назва освітньої програми)  
спеціалізації \_\_\_\_\_  
(код і назва спеціалізації)

Горін Ярослав Владиславович  
(ініціали та прізвище)

Керівник доцент кафедри інформаційної економіки,  
підприємництва та фінансів, к.е.н. Мержинський Є.К.  
(посада, вчене звання, науковий ступінь, прізвище та ініціали)

Рецензент \_\_\_\_\_  
(посада, вчене звання, науковий ступінь, прізвище та ініціали)

Запоріжжя  
2023

**МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ**  
**ЗАПОРІЗЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ**  
**ІНЖЕНЕРНИЙ НАВЧАЛЬНО-НАУКОВИЙ**  
**ІНСТИТУТ ім. Ю.М. ПОТЕБНІ**

Кафедра Інформаційної економіки, підприємництва та фінансів

Рівень вищої освіти другий (магістерський)

Спеціальність 051 Економіка  
(код та назва)

Освітня програма Інформаційна економіка  
(код та назва)

Спеціалізація \_\_\_\_\_  
(код та назва)

**ЗАТВЕРДЖУЮ**

Завідувач кафедри д.е.н. доц.

Шапуров О.О.

«\_\_\_\_\_» \_\_\_\_\_ 20\_\_ року

**З А В Д А Н Н Я**  
**НА КВАЛІФІКАЦІЙНУ РОБОТУ СТУДЕНТОВІ**

Горін Ярослав Владиславович

(прізвище, ім'я, по батькові)

1. Тема роботи (проекту) Управління освітнім процесом Університету на основі інформаційної моделі великих даних у блокчейн

керівник роботи Мержинський Євгеній Костянтинович  
(прізвище, ім'я, по батькові, науковий ступінь, вчене звання)

затверджені наказом ЗНУ від від «01» травня 2023 року № 633-с

2. Строк подання студентом роботи 9 грудня 2023 року

3. Вихідні дані до роботи аналітичні показники маркетингової діяльності Klion Group, дані Google Analytics web-ресурса компанії Klion Group.

4. Зміст розрахунково-пояснювальної записки (перелік питань, які потрібно розробити) 1) провести аналіз сучасного стану інформаційних технологій в галузі Великих даних та застосування блокчейну; 2) розробити інформаційну модель Великих даних у блокчейн; 3) проаналізувати досвід впровадження інформаційних моделей великих даних освітніх систем; 4) запропонувати архітектуру Великих даних у блокчейн для освітнього процесу в інформаційній системі з видачі дипломів; 5) розробити електронний кабінет студента на базі технології блокчейн.

5. Перелік графічного матеріалу (з точним зазначенням обов'язкових креслень) Загальна схема структури ланцюга блоків, Джерела Big Data в освітній сфері, Узагальнена модель роботи блокчейн з видачі дипломів, Алгоритм роботи технології

блокчейн процесу видачі диплому, Схема взаємодії інструментів, що використовуються у проекті електронного кабінету студента.

#### 6. Консультанти розділів роботи

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		завдання видав	завдання прийняв
<b>1</b>	доцент, к.е.н. кафедри інформаційної економіки, підприємництва та фінансів Мержинський Є.К.	17.09	19.10
<b>2</b>	доцент, к.е.н. кафедри інформаційної економіки, підприємництва та фінансів Мержинський Є.К.	19.10	29.10
<b>3</b>	доцент, к.е.н. кафедри інформаційної економіки, підприємництва та фінансів Мержинський Є.К.	29.10	16.11

7. Дата видачі завдання \_\_\_\_\_

### КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назва етапів кваліфікаційної роботи	Строк виконання етапів роботи	Примітка
1	Призначення наукових керівників. Затвердження тем дипломних робіт	<b>01.05.2023</b>	
2	Напрацювання теоретичного матеріалу: дослідження сутності об'єкту та предмету дослідження, критичний аналіз існуючих методологічних засад, вибір та обґрунтування напрямку проведення дослідження	<b>01.09.2023</b>	
3	Апробація результатів на Міжнародних та Всеукраїнських конференціях	<b>Протягом навчального року</b>	
4	Розробка інформаційної моделі забезпечення основних елементів концептуального підходу.	<b>01.10.2023</b>	
5	Збір та систематизація статистичного та нормативного матеріалу дослідження.	<b>15.10.2023</b>	
6	Узагальнення отриманих результатів. Оформлення роботи	<b>30.10.2023</b>	
7	Надання роботи та автореферату до рецензії. Нормоконтроль	<b>25.11.2023</b>	
8	Прилюдний захист дипломної роботи на засіданні ЕК	<b>15.12.2023</b>	

Студент \_\_\_\_\_  
(підпис)

Горін Я.В.  
(ініціали та прізвище)

Керівник роботи (проекту) \_\_\_\_\_  
(підпис)

Мержинський Є.К.  
(ініціали та прізвище)

**Нормоконтроль пройдено**

Нормоконтролер \_\_\_\_\_  
(підпис)

\_\_\_\_\_ (ініціали та прізвище)

## АНОТАЦІЯ

Горін Я. В. Управління освітнім процесом Університету на основі інформаційної моделі великих даних у блокчейн.

Кваліфікаційна випускна робота для здобуття ступеня вищої освіти магістра за спеціальністю 051 – Економіка, науковий керівник Є.К. Мержинський. Запорізький національний університет Інженерний навчально-науковий інститут ім. Ю.М. Потебні. Кафедра інформаційної економіки, підприємництва та фінансів, 2023.

В роботі вирішено актуальну наукову проблему підвищення якості управління освітнім процесом за рахунок скорочення часу доступу до даних на основі розроблення інформаційної технології блокчейн для Великих даних.

Ключові слова: БЛОКЧЕЙН, ВЕЛИКІ ДАНІ, ІНФОРМАЦІЙНА МОДЕЛЬ, РОЗПОДІЛЕНІ СИСТЕМИ.

## ABSTRACT

Horin Y. Management of the University's Educational Process Based on Big Data Information Model in Blockchain.

Qualifying final work for obtaining a master's degree in higher education by specialty 051 - Economics, supervisor Y. Merzhynskiy Zaporizhzhya National University Engineering Educational and Scientific Institute named after Y.M. Potebni. Department of Information Economics, Entrepreneurship and Finance, 2023.

The work solves scientific problem of enhancing the quality of educational process management by reducing data access time through the development of blockchain-based information technology for Big Data.

Keywords: BLOCKCHAIN, BIG DATA, INFORMATION MODEL, DISTRIBUTED SYSTEMS.

## ЗМІСТ

ВСТУП	6
РОЗДІЛ 1. ТЕОРЕТИКО-МЕТОДОЛОГІЧНИЙ ОГЛЯД ТЕХНОЛОГІЇ БЛОКЧЕЙН ДЛЯ ОПРАЦЮВАННЯ ВЕЛИКИХ ДАНИХ	9
1.1. Аналіз технологій блокчейн	9
1.2. Теоретичний аналіз системи великих даних	14
1.3. Інструментарій процесу збору та обробки великих даних в управлінні освітнім процесом	19
Висновки до розділу 1	24
РОЗДІЛ 2. РОЗРОБКА ІНФОРМАЦІЙНОЇ МОДЕЛІ ВЕЛИКИХ ДАНИХ У БЛОКЧЕЙН ПРИ УПРАВЛІННІ ОСВІТНІМ ПРОЦЕСОМ УНІВЕРСИТЕТУ	25
2.1. Дослідження інформаційних моделей великих даних освітніх систем	25
2.2. Розроблення інформаційної моделі Великих даних в блокчейн	32
2.3. Структурно-логічна модель видачі дипломів з використанням блокчейн	35
Висновки до розділу 2	40
РОЗДІЛ 3. РЕАЛІЗАЦІЯ ІНФОРМАЦІЙНОЇ МОДЕЛІ ВЕЛИКИХ ДАНИХ У БЛОКЧЕЙН ПРИ УПРАВЛІННІ ОСВІТНІМ ПРОЦЕСОМ	41
3.1. Реалізація архітектури освітньої систем з використанням блокчейн-технології	41
3.2. Розробка блокчейн-застосунку електронного кабінету студента кафедри ІЕПФ ЗНУ	46
3.3. Розробка протоколів обміну даними при управлінні освітнім процесом	58
Висновки до розділу 3	62
ВИСНОВКИ	64
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ	67

## ВСТУП

Великі набори даних (Big Data) та технології розподіленого реєстру (блокчейн, blockchain) на даний момент є найбільш обговорюваними темами в інформаційних технологіях. Можливості їх застосування в різних областях, від фінансів до охорони здоров'я, активно розглядаються на заходах різних рівнів.

У великих наборах даних виникає необхідність прискорення доступу до інформації через різноманіття джерел. Це може бути вирішено за допомогою використання технології блокчейну.

Важливо відзначити основну відмінність між технологіями розподіленого реєстру та великими наборами даних: Big Data передбачає інтеграцію інформації з різних джерел, тоді як у блокчейні інформаційні ланцюги зберігаються на різних комп'ютерах. Децентралізоване зберігання та послідовний запис даних у блокчейні визначає низьку швидкість доступу до них. Ідея Big Data полягає у швидкій обробці великих обсягів інформації, чого може досягти блокчейн. Незважаючи на відмінності у технологіях, їх поєднання може принести значні переваги, зокрема, прискорення доступу до даних та підвищення якості через усунення дублікатів та суперечностей у записах.

Дані, що внесені до блокчейну, залишаються там назавжди. Тому застосування цієї технології доцільне лише для завдань, де постійне зберігання незмінної інформації є необхідним, включаючи застарілу і невикористану. Незмінність та достовірність інформаційних ланцюгів блокчейна може бути корисною для автоматизованого архівування операцій з даними, зокрема, для записів в залежності від галузей застосування.

Подібним чином поєднання Big Data і блокчейну можна використовувати в освітній діяльності, щоб важливі персональні дані студентів та абітурієнтів Університету були максимально захищені, незмінні, перевірені і неможливі до будь-яких маніпуляцій. За допомогою блокчейну

ЗВО зможуть обмінюватися достовірними відомостями зі роботодавцями, базою ЄДБО, кабінет вступника, науковими установами та іншими організаціями, які задіяні в освітній діяльності та потребують інформації. Для таких галузей, як освіта, є необхідність збереження усієї історії зміни параметрів об'єкта – здобувача, науково-педагогічного працівника тощо. Отже, задача розроблення інформаційної технології блокчейн для опрацювання Великих даних є актуальною.

Об'єктом дослідження є процеси опрацювання Великих даних у блокчейні.

Предметом дослідження є методи, моделі та інформаційні технології при управлінні освітнім процесом для опрацювання Великих даних у блокчейні.

Метою дослідження є підвищення якості управління освітнім процесом за рахунок скорочення часу доступу до даних на основі розроблення інформаційної технології блокчейн для Великих даних.

Для досягнення цієї мети поставлено та вирішено такі завдання:

- 1) провести аналіз сучасного стану інформаційних технологій в галузі Великих даних та застосування блокчейну;
- 2) розробити інформаційну модель Великих даних у блокчейн;
- 3) проаналізувати досвід впровадження інформаційних моделей великих даних освітніх систем;
- 4) запропонувати архітектуру Великих даних у блокчейн для освітнього процесу в інформаційній системі з видачі дипломів;
- 5) розробити електронний кабінет студента на базі технології блокчейн.

Методи дослідження. У процесі розробки моделі Великих даних у блокчейн використано математичну статистику, теорію множин. Для побудови алгоритмів запису у блокчейн використано теорію інформації, хеш-функції, теорію адгоритмів. Для побудови методу перевірки якості внесених даних використано теорію кодування.

Наукова новизна одержаних результатів. У кваліфікаційній роботі вирішено актуальну наукову проблему підвищення якості управління освітнім процесом за рахунок скорочення часу доступу до даних на основі розроблення інформаційної технології блокчейн для Великих даних. Найбільш суттєві наукові результати полягають у такому:

дістала подальшого розвитку:

розроблена архітектуру Великих даних у блокчейн для освітнього процесу в інформаційній системі з видачі дипломів. Розроблену структурно-логічну модель можна використовувати, як конструктор для оптимізації даних, які є незмінними;

удосконалено:

електронний кабінет студента на базі технології блокчейн, який при впровадженні в освітній процес дозволить зменшити кількість корупції та іншого роду недоброчестностей під час проходження тестів та виконання індивідуальних завдань.

Практичне та теоретичне значення дослідження полягає в тому, що запропонований підхід до управління освітнім процесом Університету на основі інформаційної моделі великих даних у блокчейн дає змогу зберігати дані такого розміру, що певні реляційні бази даних (тестування відбувалося на MySQL) навіть на потужних серверах, не змогли б виконати операції CRUD (Create, Update, Delete) в час, зазначений у вимогах якостей Кодда.

Структура та обсяг кваліфікаційної роботи. Магістерська робота складається зі вступу, трьох розділів, висновків, списку використаних літературних джерел. Загальний обсяг роботи - 71 сторінка, 4 таблиці, 34 рисунки, 45 джерел літератури.



## РОЗДІЛ 1

# ТЕОРЕТИКО-МЕТОДОЛОГІЧНИЙ ОГЛЯД ТЕХНОЛОГІЇ БЛОКЧЕЙН ДЛЯ ОПРАЦЮВАННЯ ВЕЛИКИХ ДАНИХ

### 1.1. Аналіз технологій блокчейн

«Blockchain – розподілена децентралізована мережа, яка забезпечує незмінність, конфіденційність, безпеку та прозорість. Не існує центрального органу для перевірки транзакцій, проте кожна транзакція в блокчейні вважається повністю захищеною та перевіреною» [1].

Наше існування тісно пов'язане із фінансами, інформацією та документами. Це змушує нас устанавлювати зв'язки з різноманітними посередниками, які надають нам фінанси, документи та дані. Вони перевіряють їхню валідність, засвідчують їхню автентичність, надають копії, впевнюються в достовірності копій і так далі.

«Причому нас змушують довіряти цим посередникам, хоча раз у раз несумлінні банкіри ховаються з грошима клієнтів, нотаріуси заднім числом підробляють заповіти і договори, співробітники державних органів і комерційних організацій користуються службовим становищем у зловмисних цілях» [2].

«Технологія blockchain якраз і вирішує всі ці проблеми. Blockchain – ланцюжок блоків даних, де кожен блок пов'язаний з попереднім. Блок містить у собі набір записів. А нові блоки завжди додаються строго в кінець ланцюжка» [3-4].

«Система побудована на трьох принципах» [4]:

- розподіленість;
- відкритість;
- захищеність.

Усі учасники системи blockchain об'єднуються у мережу серверів (вузлів), де кожен вузол містить свою копію інформації blockchain. Зазвичай

це є повна дублікат блоків, але теоретично можна зберігати лише необхідні дані на конкретному комп'ютері.

Завдяки цьому зупинити, виключити або зламати blockchain практично неможливо, оскільки для цього треба вимкнути або зламати всі вузли [5-6]. Поки є хоч один користувач, blockchain буде існувати. Кожен новий користувач розширює і зміцнює цю мережу. Причому всі вузли рівноправні, там немає організаторів, модераторів, контролерів і менеджерів. Кожен відповідає за себе сам.

«Усі дані блоки в blockchain і їх вміст є завжди відкритими та для усіх доступними. Можна легко прочитати будь-який блок і побачити всі записи в ньому. Також можна подивитися повний ланцюжок і відстежити зміну інформації. Таким чином, усі дані в blockchain легко перевіряються, з цього випливає, що не обов'язково довіряти іншим учасникам мережі, адже самому завжди можна їх перевірити і отримати гарантовано достовірну відповідь» [6].

«Для захисту даних в blockchain використовуються різноманітні види шифрування. Завдяки цьому можна отримати здавалося б несумісне – відкритість і достовірність даних при повній недовірі до решти учасників, а також можливо, навіть при їх зловмисному умислі» [7-8].

«Усі дані blockchain зберігаються на вузлах користувачів blockchain мережі. Усі користувачі мережі мають рівні права і обов'язки, узагалі кажучи, можуть робити все, що завгодно, в тому числі безуспішно намагатися обманути інших користувачів та намагатись змінити ланцюг блоків. Заборонити цього їм ніхто не може, тому що всі знаходяться в рівних умовах, мають рівні права, і в однаковій мірі можуть виконувати чи навіть порушувати свої обов'язки» [7].

Отже, учаснику мережі blockchain не потрібні жодні посередники, такі як банки, державні органи, аудитори, контролери, страхові компанії або реєстратори. Тут немає потреби в зверненні за дозволом, жодна зайнята

позиція не призначає додаткових прав або можливостей, авторитет не має більшого захисту, ніж власна особиста обізнаність.

Вступивши до мережі blockchain, учасник приєднується до інших комп'ютерів мережі для обміну даними, такими як блоки та транзакції. Важливо, що ця мережа ніяк не прив'язана до географії, тобто користувач з Запоріжжя може одночасно підключитися до користувачів з Києва, Мілану і Вашингтону. І це, до речі, також захищає його від будь-яких регіональних особливостей.

Учасники системи розділені на дві категорії: звичайні учасники, які генерують транзакції з даними, і гравці, які виконують видобуток блоків для подальшого їх включення в ланцюг блоків. Формування блоку – це вкрай витратний і складний процес, і не кожен має бажання або можливість цим займатися.

Звичайні учасники створюють та розповсюджують по мережі записи, наприклад, "учасник з ключем А передає 10 монет учаснику з ключем Б" або "особа з ключем А відправила 100 монет з ключем Х". Усі записи відкриті, але зашифровані. Якщо вам відомий ключ відправника, ви можете переглядати входи та виходи, а також перевіряти, чи є монети на відповідній адресі, але ім'я відправника залишається анонімним, якщо він сам не розкриє цю інформацію в публічному доступі. Крім того, одна особа може мати кілька ключів, тому навіть знаючи ключ А, неможливо визначити загальну кількість монет.

Далі, майнери збирають записи, проводять їх перевірку і включають в блоки, які потім поширюються по мережі. Звичайні учасники отримують ці блоки та зберігають їх для коректного створення своїх записів та достовірної перевірки інших записів.

Поки новий запис не включений у жоден блок, його вважають недостовірним, і він залишається у мемпулі. Будь-який учасник мережі може використовувати його на свій ризик, оскільки існує ймовірність, що це некоректний або навіть підроблений запис. Є можливість, що його відкинуть,

тобто відхилять ланцюг з цим записом, а сам запис визначають як "сирота". Таким чином, учасники зазвичай просто пересилають новий запис, очікуючи, що він дійде до майнера, який додасть його в блок. Лише після того, як запис збережено у блоку, його можна вважати перевіреним, коректним і неможливим для відміни. Таким чином, працює технологія blockchain, завдяки консенсусному протоколу, який є необхідною частиною будь-якої мережі blockchain.

«Алгоритм консенсусу – це процедура, за допомогою якої всі однолітки мережі Blockchain досягають спільної згоди щодо поточного стану розподіленої книги» [9]. «Таким чином, консенсусні алгоритми досягають надійності в мережі Blockchain та встановлюють довіру між невідомими однолітками у розподіленому обчислювальному середовищі. По суті, консенсус-протокол є гарантом, що кожен новий блок з даними, який додається до блокчейну, є єдиною версією істини, яка є узгодженою усіма вузлами в блокчейні» [10]. Протокол консенсусу в системі Blockchain складається з конкретних завдань, таких як досягнення узгодженості, співпраця, забезпечення рівних прав для кожного вузла та обов'язкова участь кожного вузла в процесі консенсусу. З цього випливає, що алгоритм консенсусу націлено на досягнення загальної узгодженості, яка приносить вигоди всій системі.

«Сам blockchain складається з блоків і найцікавіше те, як ті блоки генеруються при видобуванні (mining)» [11-12]. Блок в blockchain складається з заголовка і тіла (рис. 1.1).

Тіло кожного блоку просто представляє собою перелік записів. Блоки в мережі blockchain взаємодіють через використання ключів, оскільки в заголовку кожного блоку зберігається ключ попереднього блоку. Це важливе та технічно обґрунтоване рішення, яке забезпечує безпеку blockchain. Кожен ключ блоку розрахований на дані всього блоку та ключі попереднього блоку. Це означає, що в ключі будь-якого блоку закодовані не лише дані цього конкретного блоку, але й всі попередні блоки. Навіть найменша зміна у

даних будь-якого блоку призводить до повної зміни його ключа, і, відповідно, вимагає зміни ключів для всіх наступних блоків.

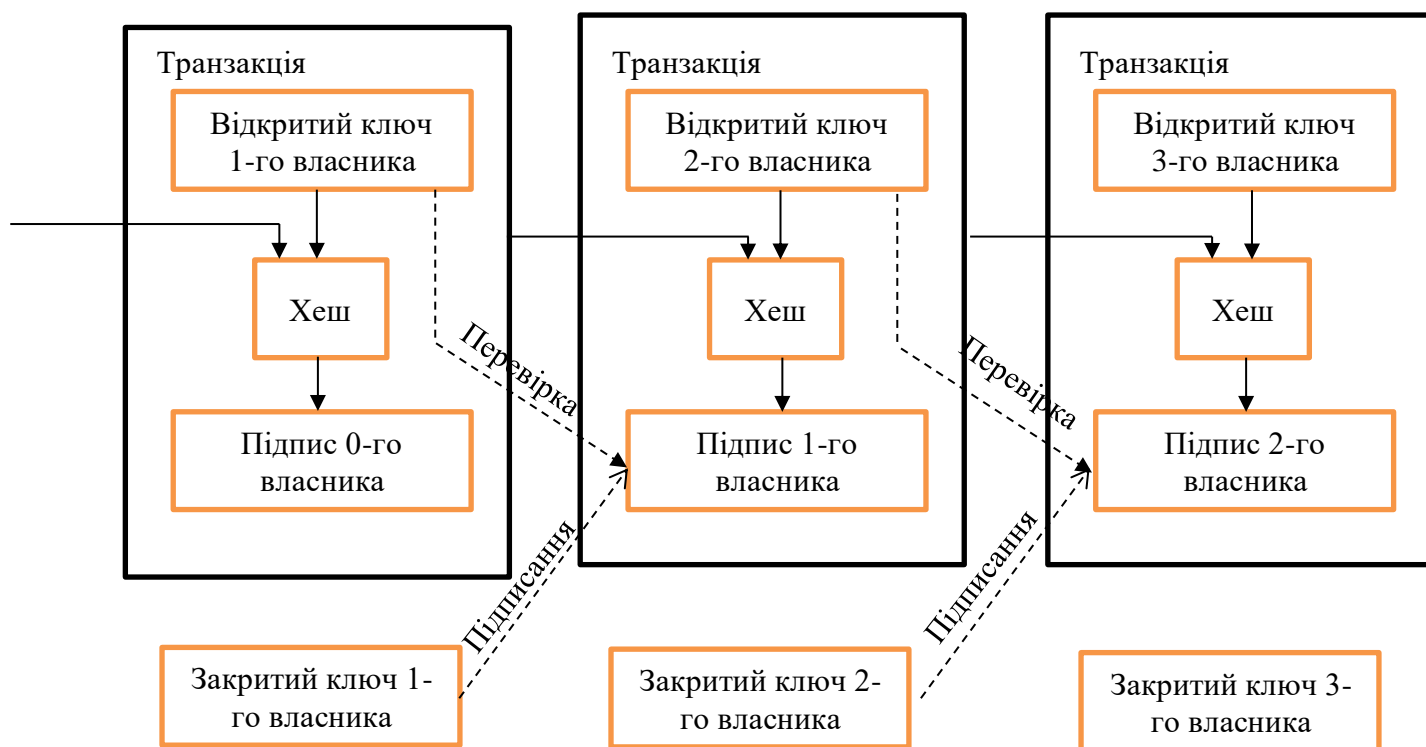


Рисунок 1.1 – Загальна схема структури ланцюга блоків

Отже, при огляді усього ланцюга блоків і ключів можна легко перевірити точність будь-яких даних, зокрема, переконатися, чи правильна послідовність блоків, чи відсутній пропущений блок – можливо, що вставлений новий блок всередину ланцюга, і також – чи відповідає ключ блоку збереженим у ньому даним. Тобто, завжди можна перевірити дані в конкретному блоку.

Майнер – це такий самий учасник blockchain-мережі, як і всі інші. Однак, крім перевірки та поширення даних, він також займається створенням нових блоків.

Отримавши нові записи від інших учасників мережі, ці записи накопичуються в мемпулі (пул пам'яті), після чого майнер об'єднує їх, формує заголовок майбутнього блоку і обчислює ключ блоку. Допустимо, після першого обчислення ключ виглядає так:

“885b7ceafdsfs348e290d6a9f244f8883419f5c69df53df3d04d749a2cfa50d70e1”.

Але, відповідно до встановлених правил, ключ має починатися з десяти нулів. Щоб отримати новий ключ, необхідно змінити вихідні дані. Для досягнення цього в заголовку блоку передбачено спеціальне поле, відоме як "nonce". Під час початкового обчислення його значення становить 0. Таким чином, майнер змінює це значення на 1 і знову розраховує ключ. Тепер ключ повністю змінився і став рівним:

“hsdfkh3hk3440d91428d7433c3d9ede5648a08e961c5e31425830284da1028b0e”.

Отже, знову розпочинається не з нулів. В цьому випадку майнер збільшує nonce до 2 і проводить перерахунок ключа. Для знаходження відповідного значення ключа майнерові доводиться здійснити трильйони перерахунків. Коли відповідний ключ знайдений, майнер зберігає блок і розсилає його іншим учасникам мережі. Тепер всі дані в блоку підтверджені і захищені ключем, який вельми важко підробити. Більше того, як зазначено вище, у ключі блоку закодований ключ попереднього блоку, що робить його підроблення майже неможливим.

За створення кожного блоку майнери отримують винагороду. Той, хто першим знаходить ключ і створює блок, заробляє, отримуючи системну нагороду. Решта учасників не отримують жодного винагородження. Таким чином, складний процес розрахунку ключів, безумовно, ускладнює створення блоку, але водночас робить підробку блоків майже неможливою.

## 1.2. Теоретичний аналіз системи великих даних

«Великі дані – це різноманітні дані, які надходять з постійно зростаючою швидкістю і обсяг яких постійно зростає. Таким чином, три основні властивості великих даних – це різноманітність, висока швидкість надходження і великий обсяг» [13].

«Аналітика великих даних – це процес аналізу великих і складних джерел даних для виявлення тенденцій, моделей поведінки клієнтів і ринкових переваг, що допомагає приймати більш ефективні бізнес-рішення. Складність аналізу великих даних вимагає нових аналітичних інструментів, таких як аналітика прогнозів, машинне навчання, потокова аналітика, і такі методи, як аналіз в базі даних і в кластері.

Визначаються великі дані зазвичай чотирма V» [14-15]:

- обсяг (Volume): великі обсяги даних;
- різноманітність (Variety): багато різних форм даних, неструктурованих і структурованих;
- швидкість (Velocity): частота вхідних даних;
- правдивість (Veracity): достовірність даних.

До вже перелічених факторів іноді ще додають:

- складність (Value): різний його рівень, наприклад, відомості про користувачів соцмереж та інформацію про транзакції в банківській системі;
- змінність (Variability): сплески і спади даних, які вимагають певних технологій для обробки.

«Окрім величезного обсягу даних, складність зібраних даних створює проблеми в управлінні даними, їх інтеграції та аналізі. Але компанії, що об'єднують неструктуровані джерела даних, такі як вміст соціальних мереж, з існуючими структурованими даними, такими як транзакції, можуть додавати контекст і генерувати нові, а часто і багатші ідеї» [16-17].

Безперечно, великі обсяги інформації визначають підвищену швидкість прийому даних із різноманітних джерел, таких як сенсори, мобільні пристрої, потоки веб-кліків і транзакції, що вимагає аналізу в режимі реального часу. Організації, які мають можливість здобути вигоду від миттєвості подій для уникнення відмов обладнання, рекомендацій покупок, виявлення кредитних карткових шахраїв та інших аспектів, швидко виходять на лідерські позиції у своїх галузях.

Аналітика заснована на перетворенні даних у конкретні дії, які збільшують цінність для організації. Але зростання обсягів структурованих і неструктурованих даних, відомих як великі дані, радикально змінює природу аналітики.

Хоча великі дані розширюють можливості для бізнесу, вони також народжують нові виклики у сферах збору, зберігання та доступу до інформації.

Цей переворот ставить перед нами значні завдання щодо зберігання даних і створює нові вимоги до аналітичного програмного забезпечення. Водночас це надає потужні можливості для виявлення та реалізації передових стратегій для досягнення конкурентної переваги.

Упровадження цих можливостей потребує двох складових: технологічного потенціалу для збору та зберігання великих обсягів даних і нові інструменти для перетворення даних у розуміння та, в кінцевому рахунку, у їхню цінність.

«Аналітика великих даних може об'єднувати дані в стані спокою (традиційні структуровані дані) з даними в русі (дані в реальному часі), щоб виявляти можливості й використовувати їх у даний момент» [18].

Великі об'єми даних вже є неот'ємною реальністю для більшості корпорацій, а обсяг та складність цих великих даних можуть викликати враження нестримності. Організації повинні приділяти увагу створенню можливостей як для аналізу стаціонарних, так і для динамічних даних. Великі дані представляють собою розширений набір потужних можливостей. Хоча точний сценарій використання великих даних через рік, три або п'ять років ще не відомий, необхідність перетворення їх на конкурентну перевагу означає, що жодна компанія не може відкладати прийняття стратегічних рішень.

Отримання користі від великих інвестицій у дані передбачає ефективне їх використання. Пошук цінної інформації, яка збільшить рентабельність організаційних інвестицій, аналогічний вибору голки в стогні



сіна. Саме тому багато компаній скаржаться на низький коефіцієнт прибутковості великих даних.

«Щоб реалізувати потенціал епохи великих даних і мінімізувати її ризики, підприємствам потрібна уніфікована архітектура даних, а також програмне забезпечення для аналізу та візуалізації даних. Переваги рішення для аналізу великих даних:

- Платформи для аналізу великих даних, що дозволяє користувачам виявляти невидимі тенденції і закономірності у великих і складних наборах даних, які сприяють більш швидкій ідентифікації стратегічних можливостей і загроз.

- Єдиний погляд на бізнес – завдяки уніфікованій архітектурі даних, надає компаніям дуже узгоджене і всеосяжне вікно даних, що підвищує ефективність прийняття рішень і дозволяє користувачам працювати з найточнішою і своєчасною інформацією.

- Найшвидший час для дій – аналітика великих даних підвищує продуктивність прийняття рішень, дозволяючи будь-якому співробітнику компанії передбачити ситуації і можливості, поставити актуальні і своєчасні запитання та отримувати відповіді, які ведуть до рішучих дій» [19].

Розглянемо основні можливості аналітики великих даних:

- Розширені аналітичні та обчислювальні методи, включаючи техніки машинного навчання, є інструментами науки про дані та статистичних розрахунків. Ці інструменти використовують великі обсяги історичних даних для отримання нових знань і виявлення закономірностей. Машинне навчання дозволяє створювати та оптимізувати потужні алгоритми, які можуть удосконалювати бізнес-процеси та підвищувати цінність підприємства.

- Автоматизована аналітика в режимі реального часу, використовуючи аналітичні та прогностичні моделі, може бути використана для оперативної обробки даних. З використанням інтерфейсу візуальної розробки для швидкого створення та впровадження потокових додатків можна допомогти операційним системам в оцінці даних, надсиланні повідомлень та

ефективному виконанні завдань для прийняття своєчасних рішень відповідно до контексту.

- Важливою складовою є візуалізація даних, яка використовує просту статистику та вбудовані коннектори для швидкого імпорту даних в інтуїтивно зрозумілі панелі моніторингу. Це надає можливість бізнес-користувачам аналізувати обширні джерела даних, приймати рішення на підставі реальних даних і постійно користуватися інформаційними панелями, що відповідають потребам бізнес-середовища.

- Використання віртуалізації даних надає сучасний підхід до управління інформацією, що дозволяє користувачам отримувати доступ, об'єднувати, трансформувати і доставляти різні набори даних з надзвичайною швидкістю і ефективністю [20]. В результаті використання віртуалізації даних користувачі мають доступ не лише до актуальних, але й до найновіших даних, які легко знаходити, використовувати і розуміти.

- Ефективне керування ресурсами інформації забезпечує постійний доступ, надійну доставку, ефективне управління та забезпечення безпеки даних відповідно до потреб організації. Це досягається за допомогою інструментів, таких як управління основними даними, віртуалізація даних, каталогізація інформації, а також можливості самообслуговування для підготовки та обробки даних.

- Виявлення даних самообслуговування – це високоефективне аналітичне рішення для даних, яке надає можливість користувачам у всій організації проводити дослідження даних і отримувати відповіді, не залежно від потреби в спеціалізованому та глибокому моделюванні даних. Це зменшує відношення до ІТ та високоресурсних бізнес-аналітичних (ВІ) рішень і значно прискорює процеси прийняття рішень.

«Загальні джерела даних для аналітики великих даних:

- Платформи великих даних;
- Транзакційні дані;
- IoT дані датчиків;

- Соціальні медіа;
- Інтернет / Онлайн дані;
- Дані мобільних пристроїв;
- Історичні дані та дані в реальному часі;
- Торгова точка (PoS);
- Геопозиційні дані;
- Текстові дані» [21-23].

Завдяки таким обширним наборам даних можна: аналізувати повний спектр клієнтів, забезпечувати захист від шахрайства, підвищувати ефективність процесів ціноутворення, визначати ефективність роботи, розробляти рекомендаційні сервіси, проводити аналіз соціальних мереж, прогнозувати погоду і багато іншого.

### 1.3. Інструментарій процесу збору та обробки великих даних в управлінні освітнім процесом

Освітня сфера в умовах глобалізації суспільства зазнає постійних кардинальних змін, пов'язаних із впровадженням інформатизації та розвитком інформаційно-комунікаційних технологій (ІКТ), швидким нарощуванням обсягу знань і стрімкою динамікою змін, необхідних на ринку праці. У контексті актуальних тенденцій у розвитку освіти в контексті цифрової освіти, виникає необхідність розгляду специфіки та особливостей впровадження технології обробки великих обсягів даних в блокчейн. Зазначена технологія може бути використана для вирішення як традиційних педагогічних завдань різних класів та рівнів, так і для інноваційних викликів.

Наразі роботи в галузі використання блокчейн в освіті знаходяться на початковому етапі. Провідні ЗВО, експерименти яких лягли в основу звіту Єврокомісії, включають Массачусетський технологічний інститут, Відкритий університет Великобританії, Університет НОРД та кілька освітніх закладів в Берліні. Згідно з висновками доповіді, успішна реалізація запропонованих

сценаріїв залежить від ініціативи країн-учасниць ЄС у сфері регулювання та стандартизації блокчейн. Автори пропонують державним органам підтримувати інноваційні приватні компанії та створити експертні комітети для впровадження блокчейн-рішень. Зазначається, що ключовою перевагою технології блокчейн в освіті є забезпечення надійності і безпеки збору та зберігання інформації, дозволяючи одночасно включати різноманітні дані у самі записи.

Наприклад, використовуючи технологію блокчейн, можна зберігати інформацію про іспити, отримані дипломи та сертифікати, а також деталі про те, хто і коли їх проводив чи видавав. Таким чином, традиційний паперовий документ втрачає свою унікальність, оскільки будь-хто здатен перевірити його автентичність та отримати завірену копію безпосередньо, не вдаючись до архівів відповідної організації. «Елементами зберігання можуть бути не тільки дипломи і атестати про закінчення навчання, але і відомості про закінчення онлайн- курсів, здачі контрольних робіт» [24].

Ключовою частиною навчального процесу є остаточне оцінювання та атестація, які включають іспити, кваліфікаційні проекти та інші форми навчальних заходів. Під час цих етапів учні демонструють свої навчальні досягнення, такі як знання, вміння, навички та здібності. Щоб оцінити, як технологія розподіленого реєстру може бути застосована в освіті, слід розглянути й обмежити різні сфери навчання, які можна дослідити в контексті перспектив впровадження цієї технології. Спочатку слід провести розрізнення між університетською та відкритою освітою, що включає сучасні масові відкриті онлайн курси (МООС, Massive Open Online Courses) та інші. «По-друге, у рамках університетського навчання має сенс розглядати безпосередньо освітній процес і супроводжуючі його адміністративно-управлінські аспекти з одного боку, а з іншого боку – результати освітнього процесу, документи про освіту або сертифікати про прослухані курси, укупі з урахуванням даних документів за межами окремого університету» [25].

Оптимальним контекстом для впровадження технології розподіленого реєстру в галузь освіти виступають платформи для масових онлайн-курсів. Зовсім нещодавно був представлений проект DISCIPLINA 3, розроблений командою вчених. Цей проект представляє собою блокчейн-платформу, яка об'єднує студентів, навчальні заклади, викладачів і, в подальшому, потенційних роботодавців. Використовуючи переваги розподіленого реєстру, розробники прагнуть вирішити завдання створення єдиного реєстру для обліку успішності студентів на різних курсах, які пропонуються різними навчальними закладами. Також метою є створення об'єктивної та достовірної системи рейтингування студентів, що спростить пошук роботодавцями відповідних кандидатів з відповідними знаннями для конкретних посад.

В такій відкритій системі оцінювання і рейтингування можливе оцінювання будь-чого - відгуки і рейтинги окремих курсів чи програм, репутація університетів і роботодавців, які користуються цим реєстром. Проблему обробки особистих даних вирішується шляхом фіксації в блоках лише анонімізованих хешів, тоді як конкретні досягнення студентів зберігаються в їх особистих облікових записах. Доступ для роботодавців до системи передбачається платним.

Токени, які функціонують у блокчейні DISCIPLINA, мають за мету використовуватися для розрахунків за платні курси та за участь у формуванні блоків у ланцюгу. Аналогічна логіка може бути застосована й в традиційному університетському навчальному процесі, принаймні в рамках навчальних закладів, що беруть участь у Болонському процесі. Основною його особливістю є впроваджена система кредитів, які отримуються за вивчення конкретних дисциплін. «Певна кількість набраних кредитів за певними предметів дозволяє студенту отримати ступінь бакалавра, магістра або доктора наук» [26].

Уніфікованість також визначається спрощенням процедури нострифікації (визнання диплома країни при вступі до навчального закладу іншої країни) під час вступу до іншого університету для отримання

наступного ступеня освіти або при переході з одного вищого навчального закладу до іншого. Проблема сумісності кредитів за однією й тією ж дисципліною в різних університетах, однак, залишається високою складністю, оскільки, навіть при повному відповідності вивченим темам у межах конкретного предмету, якість викладання та кваліфікація викладача можуть значно варіюватися. «Частина описаного процесу, котрий залежить від безпосередньо викладача, можливо значно спростити і автоматизувати, що і пропонують автори доповіді» [27]. Автори планують створити розподілений реєстр на блокчейні, до якого доступ може бути здійснений за допомогою механізму реєстрації та ідентифікації осіб. Передбачається, що основною одиницею інформації в реєстрі будуть досягнення студентів, які, з метою збереження конфіденційності персональних даних, мають бути зашифровані.

Якщо студент отримує оцінку за іспит, йому автоматично присвоюється унікальний ідентифікаційний номер для кожної дисципліни. Цей номер пов'язаний із закритим ключем, який забезпечує доступ до відповідного облікового запису студента. В метаданих "залікової" транзакції міститься інформація про університет, де був іспит, кількість кредитів, години на освоєння, теми інші важливі дані. При виявленні помилок організація створює нову транзакцію для коригування балів з вказівкою причини. Всі ці метадані служать основою для створення смарт-контрактів.

Наприклад, якщо дві освітні організації укладають угоду про взаємоприйняття кредитів за певними дисциплінами при переході студентів, це реєструється смарт-контрактом. Таким чином, при переведенні студента і складанні індивідуального навчального плану автоматично враховуються пройдені дисципліни в іншому навчальному закладі за рахунок перевірки відповідності вимог, встановлених смарт-контрактами в такій мережі.

У перспективі побудова такої системи на базі технології розподіленого реєстру може значно спростити перехід студентів, збільшити мобільність населення в межах країни і за її межами, фіксуючи всі

досягнення студентів у надійному реєстрі, що виключає підробку. Процедура видачі дипломів також стане більш прозорою, оскільки освоєння всіх дисциплін завжди можна буде відстежити і перевірити.

В наш час Google Академія (Google Scholar) використовується як основне джерело показників цитування для оцінки значущості публікацій чи внесків дослідника. Проте механізми цього сервісу не є прозорими, і немає впевненості, чи алгоритм знаходить всі цитати та індексує всі роботи автора. Створення аналогічного сервісу на базі розподіленого реєстру може допомогти подолати обмеження існуючих сервісів, забезпечуючи об'єктивну оцінку академічної діяльності.

В загальному вигляді процес збору та обробки інформації в освітній діяльності виглядає наступним чином:

- аналітична програма отримує завдання;
- система збирає необхідну інформацію, одночасно підготовлюючи її: видаляє нерелевантне, очищує від сміття, декодує;
- вибирається модель або алгоритм для аналізу;
- програма вчиться алгоритму і аналізує знайдені закономірності.

Найчастіше дані зберігають у "озері даних" (data lake). При цьому вони зберігаються у різних форматах і різній структурності:

- рядки і колонки зі структурованими базами даних;
- CSV, XML, JSON-файли, логи - напівструктуровані;
- документи, поштові повідомлення - неструктуровані PDF-файли;
- аудіо, відео та бінарні зображення.

Для зберігання та обробки наданої інформації в data lake використовують різноманітні інструменти та технології. Декілька з них включають:

Nadoop – це потужна платформа управління даними, яка може включати один чи кілька кластерів. Використовується для ефективної обробки, зберігання та аналізу об'ємних нереляційних даних, таких як файли журналів, записи інтернет-трафіку, дані IoT-датчиків, JSON-об'єкти,

повідомлення у соцмережах, зображення та інше. Також, це комплекс інструментів і бібліотек для розробки та виконання розподілених програм на кластерах вузлів.

Storm – це фреймворк, розроблений на Clojure, призначений для обробки інформації в режимі реального часу.

Data lake – термін "Озеро даних" вказує на зберігання не лише програмної платформи, такої як Hadoop, але і на кластерах серверів для зберігання та обробки даних. Воно може включати засоби інтеграції з джерелами та споживачами інформації, системи підготовки даних та іноді інструменти машинного навчання. Також, "озеро даних" може легко масштабуватися до тисяч серверів без зупинки кластера.

MapReduce – це фреймворк для паралельних обчислень над великими обсягами даних (до декількох петабайт), розроблений Google.

NoSQL – це підходить до роботи з різномірними даними та вирішує проблеми масштабованості та доступності через атомарність та консистентність даних.

## Висновки по розділу 1.

В першому розділі було проведено аналіз технології блокчейн, подано поняття великих даних, описані методи, які лежать в основі блокчейн технології.

Встановлено, що головними перевагами технології блокчейн є висока швидкість операцій, безпека, можливість повноцінного контролю даних, децентралізація, багатофункціональність, універсальність, відкритість, публічність; простота, зручність і доступність. Головні недоліки: необхідність у стійкому криптографічному захисті, необхідність в інструментах передачі, користування і зберігання інформації, відсутність належного юридичного врегулювання і гарантій, недостатня поширеність, зацікавленість кіберзлочинності до нових технологій.



Досліджені інструменти процесу збору та обробки великих даних в організації та управлінні освітнім процесом. Наведені приклади можливих сфер застосування технології блокчейну в освітній діяльності ЗВО. Наведені приклади успішних проектів впровадження технологій блокчейн з обробки великих даних в освітній діяльності.

## РОЗДІЛ 2

### РОЗРОБКА ІНФОРМАЦІЙНОЇ МОДЕЛІ ВЕЛИКИХ ДАНИХ У БЛОКЧЕЙН ПРИ УПРАВЛІННІ ОСВІТНІМ ПРОЦЕСОМ УНІВЕРСИТЕТУ

#### 2.1 Дослідження інформаційних моделей великих даних освітніх систем

Великі дані освітніх систем можна знайти в різних джерелах такі як студентські інформаційні системи (наприклад Moodle ЗНУ <https://moodle.znu.edu.ua/>, АІС «Абітурієнт»), адміністративні системи (СИНАП ЗНУ <https://scientific-rating.znu.edu.ua/>), системи управління навчанням (АСУ «МКР», система «Деканат», Moodle) та бібліотечні інформаційні системи (Наукова бібліотека ЗНУ).

Нові розробки в галузі освіти та застосування інформаційних технологій разом з Інтернет-технологіями призвели до зростання онлайн-освіти. ЗВО все частіше пропонують і надають онлайн-навчання, що призводить до великого обсягу та наявності освітніх сховищ даних та інструментів їх обробки. Крім того, число студентів і курсів, які пропонуються на масових відкритих онлайн-курсах, дедалі зростають та стають різноманітними, що призводить до збільшення кількості даних для аналітики.

Існує також все більше різних сортів і формати аудіо, відео, текстової інформації та зображень, крім даних у реляційних базах даних ЗВО. На рис. 2.1 показано графічне представлення областей досліджень і джерел даних Big Data в освітній сфері.

Набори освітніх даних можна розглядати з двох аспектів [28]:

- набори даних, безпосередньо пов'язані з освітньою інформацією, що містить освітні ресурси, інституційні дані та освітні показники;
- набори даних з різних доменів, які можна використовувати в освітніх установах. Прикладом освітніх наборів даних є DBpedia [29],

Freebase [30] і GeoNames [31]. Дані в цих освітніх наборах даних можна використовувати для Data mining доступного освітнього контенту, відкриття нової інформації, яка може допомогти в освітніх практиках, і підключення локальних наборів даних до хмари.

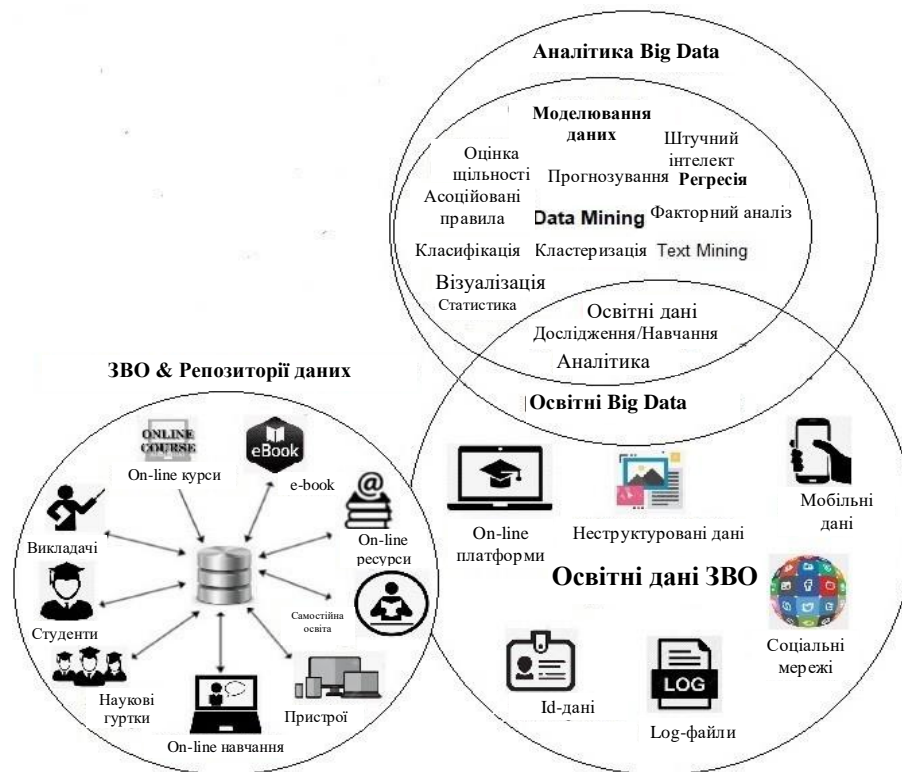


Рисунок 2.1 – Джерела Big Data в освітній сфері

\*Розроблено автором на основі [32]

Дослідники в [32] запропонували розподілену архітектуру для обробки інформації великих освітніх даних. Автори використовують цю архітектуру для прогнозування успішності студентів із застосуванням кількох рівнів аналітики у вигляді пласту. На рис. 2.2 показано запропоновану архітектуру, яка складається з трьох рівнів:

- рівень доступу до даних;
- рівень зберігання даних;
- рівень обробки даних.

Рівень доступу до даних складається з усіх джерел даних, необхідних механізмів обробки інформації, таких як журнали студентів, записи студентів

та історичні дані, а також мобільний додаток студента, який може генерувати дані на основі діяльності студента.

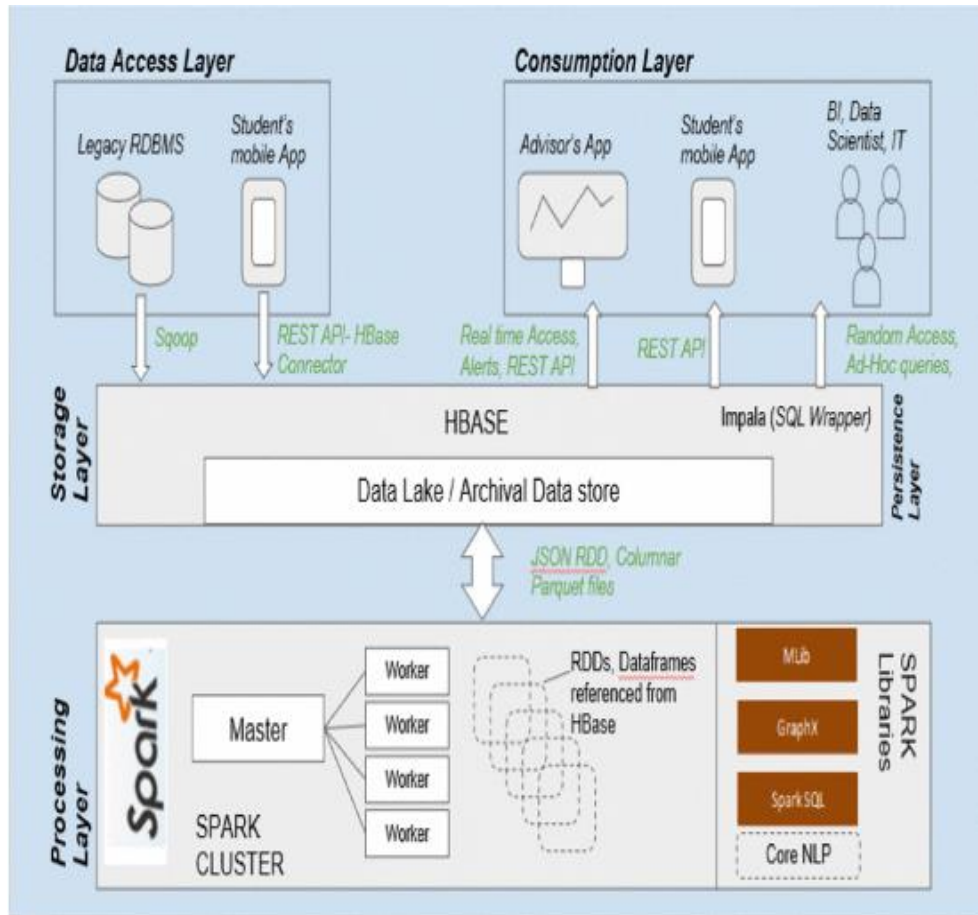


Рисунок 2.2 – Розподілена архітектура для обробки інформації великих освітніх даних за [32]

Джерела даних підключаються до рівня зберігання (HBASE) за допомогою Sqoop і REST API-HBase Connector. Другий рівень — це рівень зберігання даних, який складається з розподіленого сховища HBase та HDFS. Третій рівень — це рівень обробки, який виконує фільтрацію даних та прогнозу аналітику. Цей рівень використовує кластер Spark. На цьому рівні функції були перетворені у формати Spark Resilient Distributed Data (RDD) для виконання прогнозової аналітики. Процедури прогнозного моделювання були виконані за допомогою процесу ансамблевого моделювання.

Автори в [33] запропонували архітектуру під назвою «Concept Definition for Big Data Architecture in the Education System». Ця архітектура складається з п'яти рівнів:

- джерела даних;
- обробка великих даних;
- сховище даних;
- інструменти аналізу даних;
- звітність.

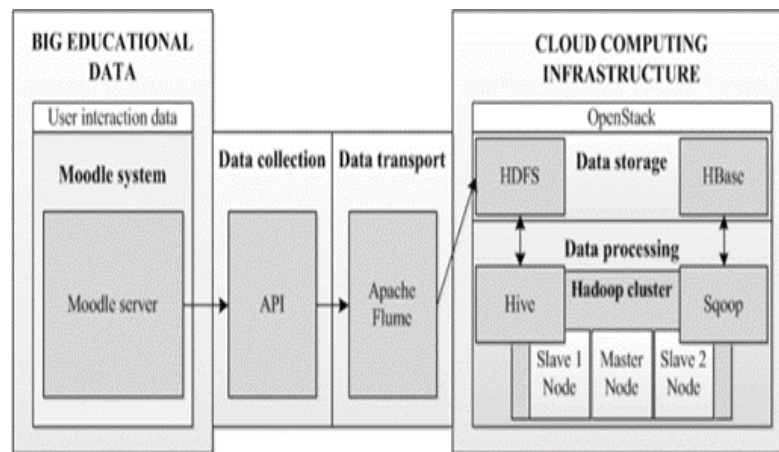


Рисунок 2.3 – Розподілена архітектура для обробки інформації великих освітніх даних за [33]

На рівні джерел даних, дані можуть зберігатися в традиційних базах даних SQL (наприклад, класичні реляційні БД) або базах даних NoSQL (наприклад, дані з соціальних мереж). На рівні обробки великих даних використовується Apache Hadoop. Третій рівень — це рівень Data Warehouse, який не залежить від технології та постачальника для створення кубів даних. На рівні інтелектуального аналізу даних використовуються інструменти IBM SPSS або SAS. Найвищий рівень (рівень звітності) виконує функцію аналізу отриманих даних для різних типів користувачів (наприклад, викладачів, адмінперсоналу, інших зацікавлених сторін для Університету). Програмне забезпечення Cognos від IBM використовується для функціональних можливостей зі створення звітів і презентацій.

Автори в [34] представили інформаційну модель великих даних освітніх систем з використанням Spark. Як показано на рис. 2.4, різні дані доставляються в HDFS відповідно до кожного атрибута.

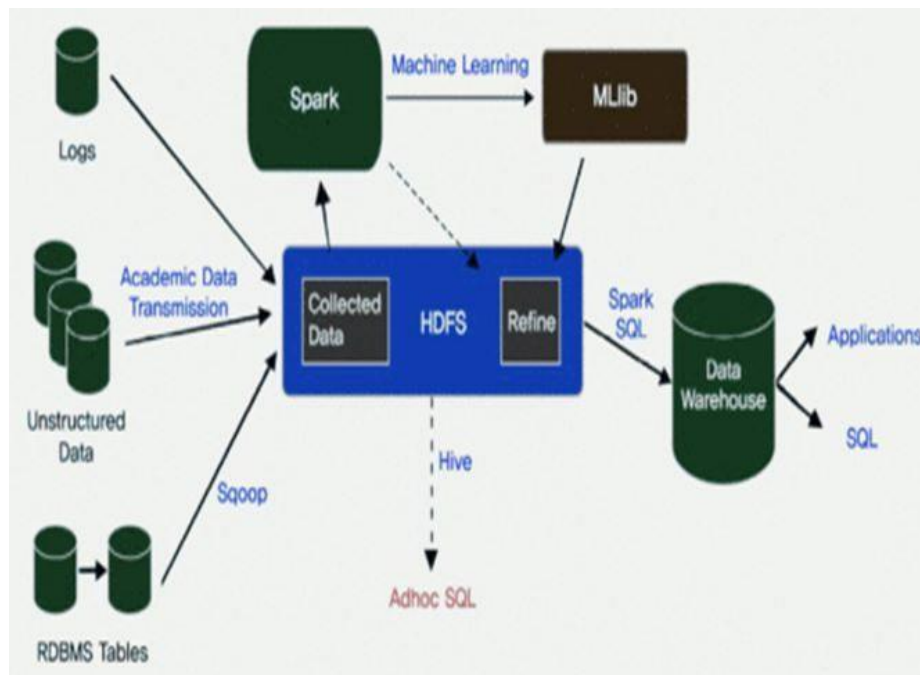


Рисунок 2.4 – Інформаційна модель великих даних освітніх систем за [34]

Структуровані дані передаються з RDBMS до HDFS за допомогою SQL-to-Hadoop (Sqoop). Серед цих зібраних даних важливим елементом є дані лекцій, а алгоритм FP-Growth виконується за допомогою MLlib, бібліотеки машинного навчання Spark. Отримані дані можуть бути використані для виявлення закономірностей, щодо знань лекційного матеріалу, якими володіють здобувачі ЗВО взяті за рік і семестр. Ці шаблони містять інформацію про патерни для бажаних лекцій студентів, і на основі цього створює систему рекомендацій, яка демонструє лекції студентам. Крім того, використовуючи дані, зібрані з датчиків інформації про відвідуваність аудиторії та інформацію про вхід у аудиторію, можна визначити кількість студентів, передбачити щільність та контролювати температуру в приміщенні та в цілому в Університеті шляхом аналізу зразків.

Автори в [35] запропонували інформаційну модель великих даних освітніх екосистем електронного навчання. Вона складається з п'яти модулів, як показано на рис. 2.5: модуль збору; транспортний модуль; модуль зберігання; обчислювальний модуль; і сервісний модуль.

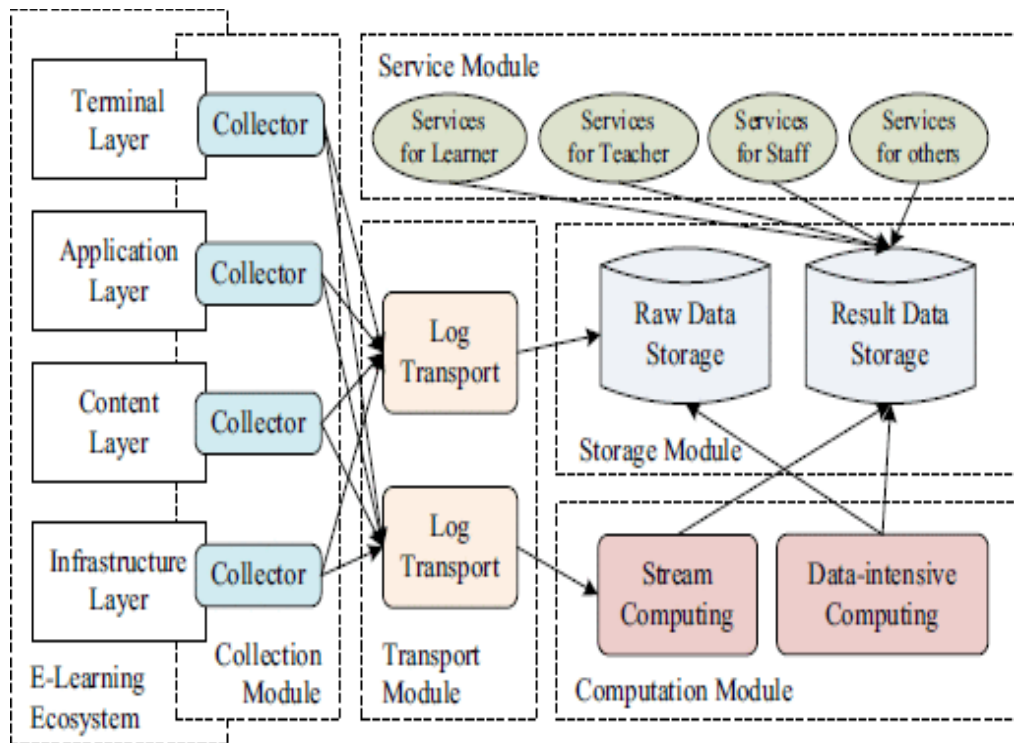


Рисунок 2.5 – Інформаційна модель великих даних освітніх систем за [35]

Окремо є модуль колекцій, який містить колектори, розподілені в кожному шарі. Кожен збирач записує дані журналу, створений різними об'єктами, і нормалізує зібрані дані. Транспортний модуль передає зібрані дані до журналу, там, де вони потрібні. Модуль зберігання включає дві категорії систем зберігання. Першою системою зберігання є система Raw Data Storage, яка зберігає історичні дані для майбутнього видобутку та аналізу даних. Другою є система зберігання даних результатів, яка має можливість швидкого доступу до даних для забезпечення високої продуктивності, вводу-виводу для потокових обчислень у модулі обчислення та модулі обслуговування.

Обчислювальний модуль містить дві обчислювальні структури. Перша — це інтенсивна обчислювальна структура, яка використовується для аналізу масивних необроблених даних, щоб отримати цінну інформацію, а друга — це структура поточкових обчислень, яка використовується для роботи з усіма даними, що надходять, у режимі реального часу. Обчислювальні інфраструктури повинні підтримувати паралельні обчислення, щоб гарантувати низьку затримку процесу аналізу та підвищити ефективність обчислень. Сервісний модуль зчитує необхідні дані із системи зберігання даних результатів для всіх об'єктів і ролей на кожному рівні екосистеми електронного навчання.

Дослідники в [36] запропонували інформаційну модель великих даних освітньої системи, розгорнутої як платформу Hadoop, щоб покращити навчальний процес. Інформаційну модель інтегрована з системою управління навчанням (LMS) на платформі Moodle. Інформаційну модель розгорнуто в інфраструктурі електронного навчання лабораторії. На рис. 2.6. показано реалізовану інфраструктуру електронного навчання Hadoop.

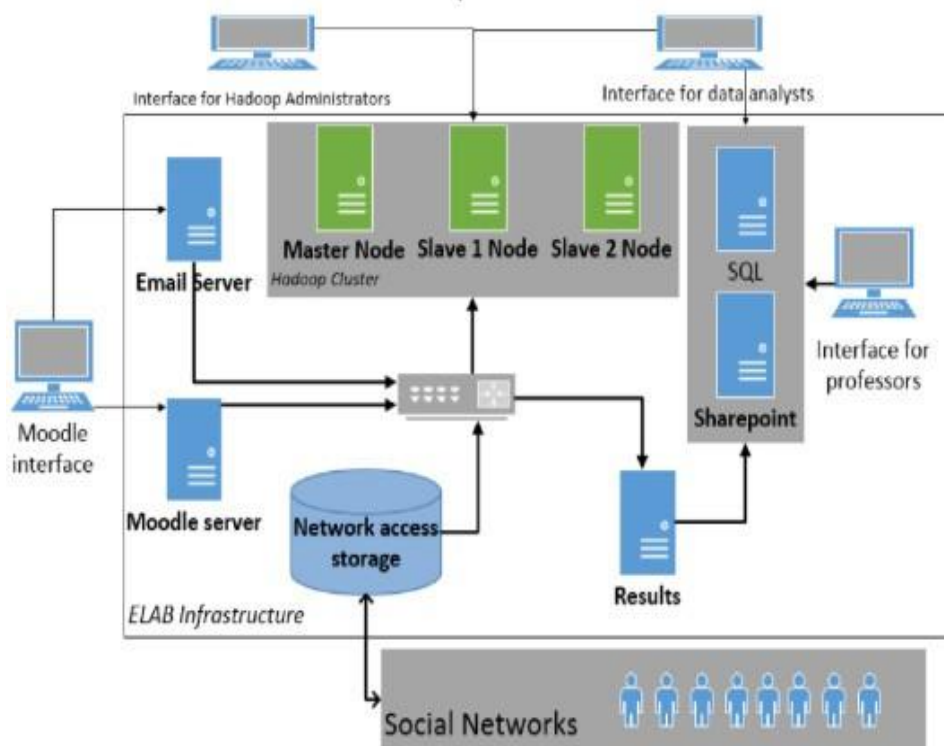


Рисунок 2.6 – Інформаційна модель великих даних освітніх систем за [36]



Кластер Hadoop містить три вузли (головний вузол, підлеглий вузол 1 і підлеглий вузол 2). Кластер Hadoop також підключений до сервера електронної пошти, сервера Moodle. Мережеве сховище даних і кластер Sharepoint через сервер результатів. Кластер взаємодіє з іншими компонентами за допомогою протоколу TCP/IP, і всі дані передаються через інфраструктуру Ethernet.

## 2.2. Розроблення інформаційної моделі Великих даних в блокчейн

Великі дані в блокчейн – це послідовні дані, подані у великому обсязі, які відправляють клієнти на анонсер, що підключені до мережі:

$$\mathbf{Block} ::= \langle \mathit{BlockHead}, \mathit{Data} \rangle, \quad (2.1)$$

де  $\mathit{BlockHead}$  гарантує цілісність  $\mathit{Data}$ .

$\mathbf{BlockHead} ::= \langle \mathit{PreviousBlockHash}, \mathit{DataHash}, \mathit{Moment}, \mathit{Difficulty}, \mathit{Timestamp} \rangle$   
 $\mathit{PreviousBlockHash}$  – це хеш-значення попереднього блоку заголовку.  
 $\mathit{DataHash}$  – хеш-значення  $\mathit{Data}$  у цьому блоці.  $\mathit{Moment}$  – значення, яке визначається шляхом випадкової перевірки.

$$\mathit{Zero}(\mathit{Hash}(\mathit{PreviousBlockHash} \parallel \mathit{DataHash} \parallel \mathit{Timestamp} \parallel \mathit{Moment})) \geq \mathit{Difficulty}, \quad (2.2)$$

де  $\mathit{Zero}(\blacksquare)$  – функція, яка повертає кількість лівих послідовних нулів у вхідному рядку в байтах;

$\mathit{Difficulty}$  – ціле число, яке визначає вимогу щодо кількості послідовних нулів у голові результату хешування.

$\mathit{Timestamp}$  – час поточного блоку.

Наприклад,  $\mathit{Difficulty}=4$  означає, що перших чотири байта вихідного хешу дорівнюють нулеві, отже, перших 4 байти з  $\mathit{Hash}(\blacksquare)$  є нулями.

$$\mathbf{Data} ::= \langle \mathbf{Metadata}, \mathbf{BData} \rangle, \quad (2.3)$$

де *Metadata* – кортеж дескриптора для *BData*, *BData* – обов’язковий кортеж з записом даних. Приклад запису для відображення неструктурованих даних поданий у Таблиці 2.1.

Таблиця 2.1 – Структура даних, які отримує анонсер з клієнта

Назва	Значення
osvita_id	Ключ hash для ідентифікації роздачі
osvita_ip	IP адреса в форматі ip2long
osvita_port	Інформація про порт клієнта
osvita_peer	Інформація про піра
osvita_md5	Унікальний хеш ключа, IP, порта в md5
osvita_sha1	Унікальний хеш ключа, IP, порта в sha1
osvita_biteup	Інформація про роздане, в байтах
osvita_bitedown	Інформація про скачане, в байтах
osvita_biteleft	Залишилось скачати, в байтах
osvita_update	Мітка часу останнього онлайну піра
osvita_hour	Мітка часу зі зміщенням на 1 годину
osvita_ipv46	Інформація про IP, звичайний формат

*Metadata* складається з таких даних як:

$$\mathbf{Metadata} ::= \langle \mathbf{LSource}, \mathbf{DataType} \rangle, \quad (2.4)$$

де *LSource* – джерела даних, а *DataType* – мітка для структурованих, напівструктурованих чи неструктурованих даних.

$$\mathbf{PreviousBlockHash} ::= \mathbf{Hash}(\mathbf{BlockHead}), \quad (2.5)$$

де *BlockHead* – заголовок з попереднього блоку. Тобто хеш-значення попереднього заголовку вбудовується в наступний блок. Його також можна розглядати як ланку двох сусідніх заголовків.

$$\mathbf{DataHash} ::= \mathbf{Hash}(\mathbf{BData1} \parallel \mathbf{BData2} \parallel \dots \parallel \mathbf{BData}_n). \quad (2.6)$$

Отже, проблема, яка вирішується за допомогою побудованої моделі – це наявність такої великої кількості даних, що стандартні реляційні бази

даних не витримують навантаження. Для прикладу в роботі було обрано базу даних в MS SQL, оскільки така СУБД – одна з найчастіших СУБД для web-ресурсів. Проводилися дослідження і в MS Access, суттєвої різниці не спостерігалось.

Через таке високе навантаження MS SQL не справляється з такою кількістю запитів, навіть на потужних серверах. Для цього використано розроблену модель даних, запис даних у Blockchain і їх вибірка через RPC-API з blockchain демона.

Оскільки дані є статичними і рідко змінюються, міняючи лише при цьому параметри, можна використовувати блокчейн для їх запису і читання.

osvita\_id – унікальний хеш торрент роздачі. Складається зі вмісту даних, перетворених в словник, і взятий з нього хеш sha1.

У цей хеш входять:

- ITEM: розмір і шлях до файла.
- Ім'я: Ім'я для пошуку.
- Довжина pieces: довжина одного фрагмента даних.
- Фрагменти: SHA1 Hash з кожного фрагменту торрента.
- Приватний: прапор обмеженого доступу.

Потім зазначену стрічку кодуємо в sha1 і отримуємо на виході: 9BEE8699E68396A4CC02D57D256A831F53969DA2. Це osvita\_id даного торрент файла, тоді зазначений параметр буде основним для створення закритого ключа доступу до відкритого ключа на блокчейні.

Далі відбувається запис Великих даних у блокчейн, який складається з таких етапів:

1. Структуризація даних, вхідні дані сортуються та автоматично очищуються від повторів.
2. Групування за змістом та довжиною даних. Далі здійснюється перетворення даних в hex вигляд і створення raw транзакції на основі hex даних.

3. Виконується перевірка на існування даних у блокчейн системі. Якщо такий запис вже існує, то він не записується, а записується мітка вказівник, де знаходяться зазначені дані.

4. Здійснюється запис у блокчейн за допомогою raw функцій перетворення даних в транзакцію.

Для того, щоб записувати дані в блокчейн, необхідні монети. Емісію даних монет створюватимуть майнери, вони в свою чергу будуть підтримувати мережу, формувати нові блоки з транзакціями та отримувати винагороду на seed адресу для наступних транзакцій.

Seed адреса зі свого боку виступає якорем для наступних транзакцій (рис. 2.7).

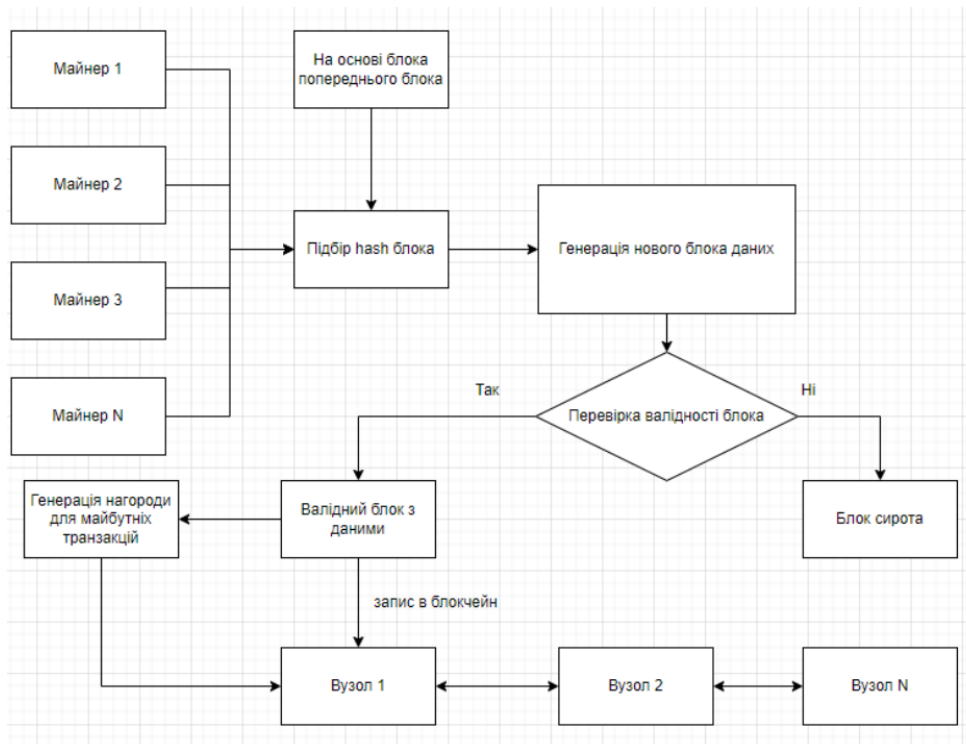


Рисунок 2.7 – Принцип записування в блокчейн

Процедура такого запису передбачає наступне: на першому етапі проводиться пошук унікальних `osvita_md5` контрольних сум. На другому – здійснюється пошук `osvita_sha1` контрольних сум. Якщо суми не мають повторів, тоді дані пишуться в блокчейн.

### 2.3. Структурно-логічна модель видачі дипломів з використанням блокчейн

Більшість ЗВО ведуть записи у паперовій формі або зберігають їх у ексклюзивних базах даних, таких як система «Деканат» в ЗНУ. При цьому доступ до структурованих даних має лише певний адміністративний персонал ЗВО. В основному ці установи мають свої спеціалізовані системи для зберігання повних даних студентів. Щоб забезпечити безпроблемну інформацію про записи студентів на всіх рівнях, тобто студент, заклад та роботодавці, організації працюють над розробкою системи управління сертифікатами на основі блокчейну. Це призводить до автоматизації процесів видачі паперового сертифікату та надсилання запиту на перевірку сертифіката через різні вузли блокчейну.

Облікові дані на паперовій основі видача передбачає час і схильність до помилок та шахрайства. Ця автоматизація може призвести до завершення роботи в режимі реального часу, а дані зберігаються в незмінній інфраструктурі. Для студентів це просто зберігання облікових даних протягом усього життя з можливістю контролювати право власності. З іншого боку, роботодавці можуть зменшити вимогу верифікації третьою стороною та мінімальний ризик отримання фальшивих заявок на сертифікацію.

У дослідженні «Блокчейн в освіті» [37] Об'єднаного дослідницького центру Європейської комісії є попередній огляд технології і її потенційних застосувань в Європі. Це демонструє, наскільки серйозно технологія сприймається в освітньому просторі. Одним із застосувань технології блокчейн описаних у статті є видача цифрових сертифікатів та дипломів.

Наприкінці захисту, оголошується оцінка всіх робіт. Диплом підписується завідувачем кафедри, головою комісії, членами комісії і секретарем. Всі документи повинні бути завірені печаткою навчального закладу. Диплом видається студенту особисто або за заявою висилається

поштою, рекомендованим відправленням. Ця заява зберігається в особовій справі випускника, а також там зберігається і копія виданого диплома. Всі документи, є бланками суворої звітності і враховуються за спеціальним реєстром. Для обліку всіх виданих дипломів, в освітньому закладі ведеться книга реєстрації, листи якої нумеруються, а сама книга прошнуровується і на ній ставиться печатка навчального закладу та зазначається кількість аркушів. Книга так само зберігається як документ суворої звітності.

Для вирішення проблем з шахрайством у сфері підробки документів, і проблемою зберігання документів, було запропоновано впровадити технологію блокчейн. Узагальнена модель роботи блокчейн з видачі дипломів представлена на рисунку 2.8

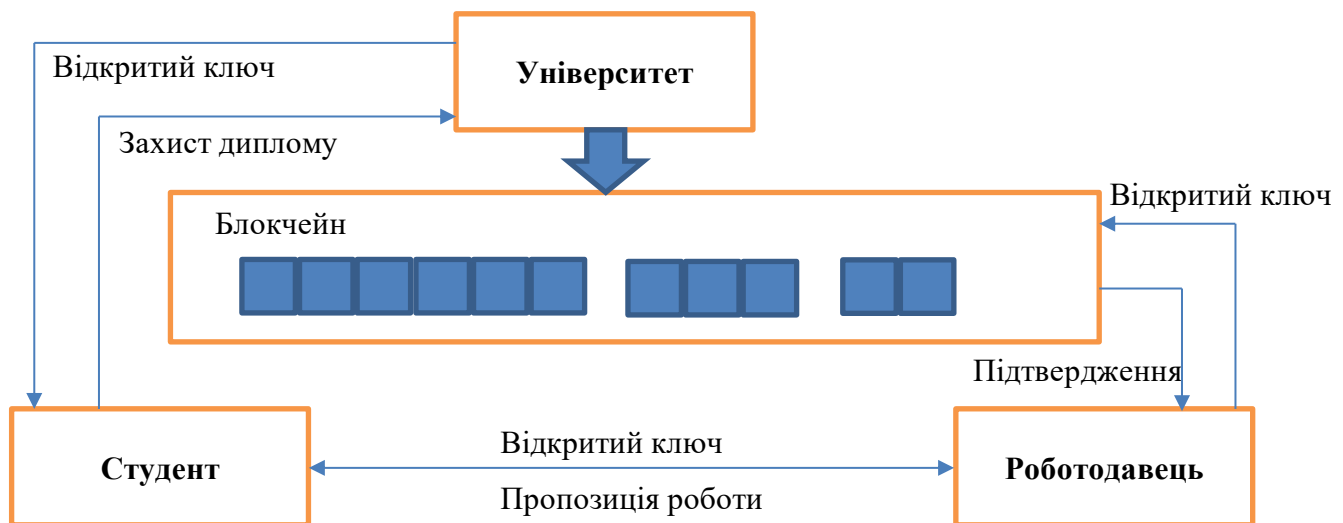


Рисунок 2.8 – Узагальнена модель роботи блокчейн з видачі дипломів

В даній моделі, ЗВО, що випускають цифрові дипломи, будуть використовувати єдиний блокчейн для їх зберігання. Унікальні дипломи, підписані приватним ключем, будуть надаватися безпосередньо роботодавцям. Таким чином, перевірка справжності диплома вимагає тільки порівняння з хешем, що зберігаються в ланцюжку блоків. Це вирішить проблему з шахрайством у сфері підробки документів і проблему збереження документів.

В першу чергу, створюється цифровий файл, який містить основну інформацію, такі як назва університету і одержувача диплома, дату видачі, посвідчення. Потім, університет підписує зміст диплома з використанням закритого ключа, до якого має доступ тільки освітня організація. Дані підтверджуються мережевим вузлом і передаються в мережу. Запис приєднується до блоку. Університет створює хеш файл облікових даних - рядок букв і цифр, які можуть використовуватися для перевірки того, щоб ніхто не порушив зміст диплома. Існує тільки одна можлива комбінація букв і цифр, яка відповідає цифровому файлу, і будь-яка зміна файлу призведе до іншого хешу. Потім, університет знову використовує свій приватний ключ для створення запису в блокчейне, в якій говориться, що освітня організація видала певний сертифікат певній особі на певну дату. Після чого, випускнику передається відкритий ключ.

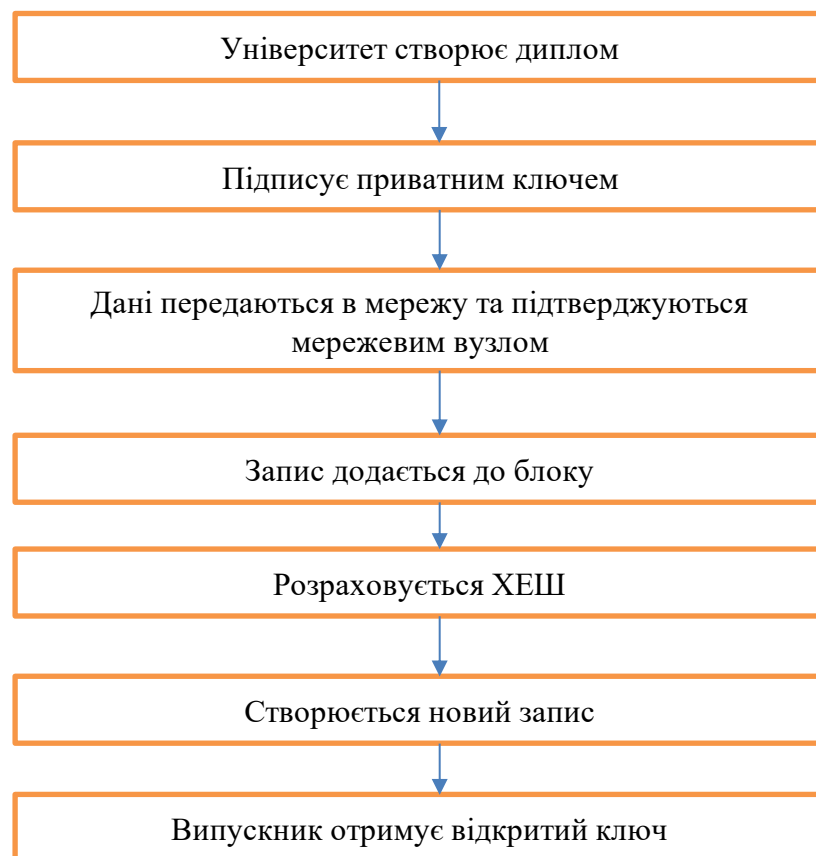


Рисунок 2.9 – Алгоритм роботи технології блокчейн процесу видачі диплому

Таким чином, користувач може перевірити, кому був виданий диплом, ким і для перевірки вмісту самого диплома [38]. Переваги перед поточним станом – докази сертифікатів будуть зберігатися повністю, надійно і в постійному блокчейні. Таким чином, навіть якщо установи, що видали сертифікати, повинні були закрити або якщо вся система освіти звалилася, ці сертифікати все ще перевіряються щодо записів, що зберігаються в блокчейні.

Крім того, як тільки установи видають диплом, їм не потрібно витрачати додаткові ресурси, щоб підтвердити дійсність цього документа третім особам, так як вони зможуть безпосередньо перевіряти дипломи у вигляді ідентифікації ланцюжка блокчейн.

Архітектура системи побудована з декількох модулів:

- Модуля запису в блокчейн, модуля читання з блокчейну,
- Модуля перевірки інформації та очистка даних від повторів,
- Модуля для тимчасових даних та їх опрацювання тощо.

На рисунку 2.10 приведена діаграма класів, що відповідає загальному вигляду архітектури системи. Основні модулі обведені червоним.

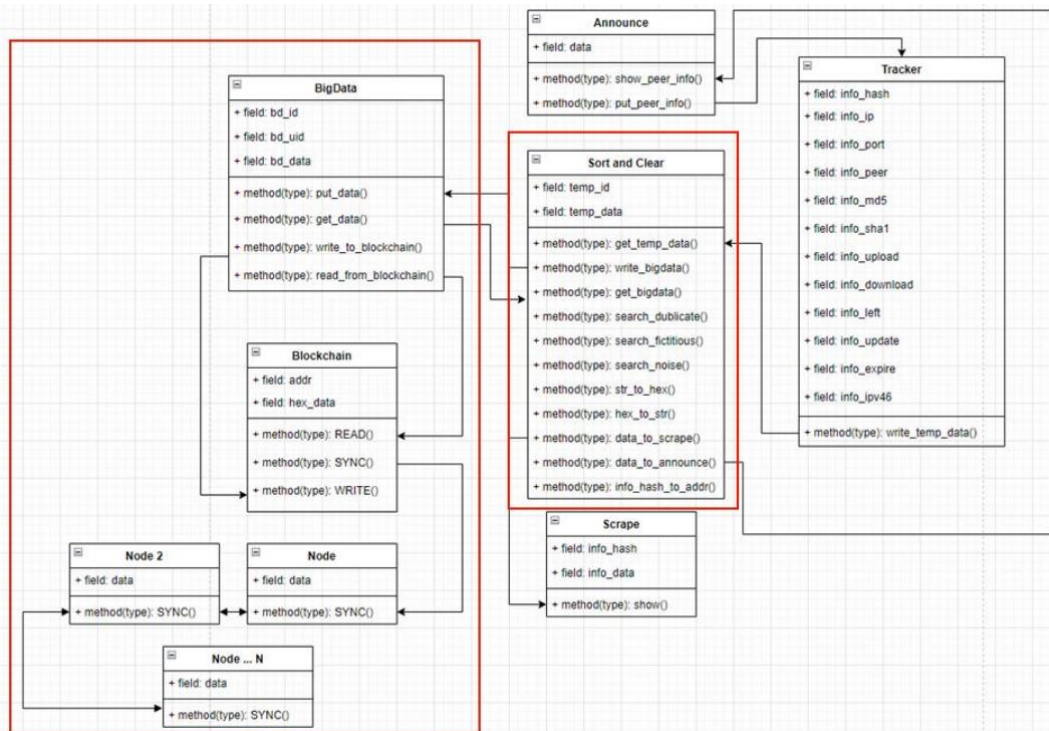


Рисунок 2.10 – Діаграма класів архітектури модулів



Особливість такої архітектури полягає в тому, що вона ділиться на основні модулі для опрацювання даних у блокчейні. Тому, в принципі, дану систему можна використовувати для різних проєктів з будь-якими даними. Для спрощення побудови проєкту архітектуру було поділено на 6 модулів:

- Announce – модуль для обробки даних, які поступають від клієнта.
- Scrape – модуль для віддачі даних клієнту.
- Tracker – модуль для опрацювання тимчасових вхідних даних.
- Sort and Clear – основний операційний модуль для опрацювання вхідних і вихідних даних, сам модуль містить у собі різноманітну кількість функцій для роботи системи.
- BigData – модуль для опрацювання великих даних, а саме запису та читання їх.
- Blockchain – модуль для реплікації по нодах, побудований на основі блокчейну з алгоритмом Scrypt, де ідентифікатором даних виступає адреса.

Розроблена модель Великих даних у блокчейн та її реалізація дає змогу зберігати дані такого розміру, що певні реляційні бази даних, наприклад MySQL, навіть на потужних серверах, не змогли б виконати операції CRUD в час, зазначений у вимогах якостей Кодда.

## Висновки по розділу 2.

У другому розділі кваліфікаційної роботи проаналізовано дослідницькі тенденції у сфері розробки інформаційних моделей обробки великих освітніх даних, включаючи джерела даних, збір даних, технологічні аспекти, інструментарій та аналітику даних. Розглянуті переваги та недоліки джерел для введення в системи великих освітніх даних, включаючи системи управління навчанням (LMS), відкриті освітні ресурси (OER), MOOC, соціальні медіа та пов'язані дані.

Проведений детальний аналіз методів хешування даних у блокчейні зі всіх можливих, описано та проаналізовано алгоритми консенсуса, необхідних для роботи з блокчейном. Розроблено модель Великих даних для їх

послідовного опрацювання, розроблений та продемонстрований, а також описаний метод перевірки якості внесених даних.

Розроблено архітектуру Великих даних у блокчейн для освітнього процесу, яка використана в інформаційній системі з видачі дипломів. Розроблену систему можна використовувати, як конструктор для оптимізації даних, які є незмінними, наприклад, логи для статистики тощо.

## РОЗДІЛ 3

### РЕАЛІЗАЦІЯ ІНФОРМАЦІЙНОЇ МОДЕЛІ ВЕЛИКИХ ДАНИХ У БЛОКЧЕЙН ПРИ УПРАВЛІННІ ОСВІТНІМ ПРОЦЕСОМ

#### 3.1. Реалізація архітектури освітньої систем з використанням блокчейн- технології

Блокчейн дозволяє створити систему, в якій дані про здобувачів, навчальні матеріали і науково-педагогічних працівників зберігаються на різних вузлах мережі. Це робить інформацію більш доступною і захищеною від змін і втрат.

Для прикладу при реалізація архітектури освітньої систем обрано торрент анонсер, що містив мільйони записів, і для його оптимізації розроблено систему на основі блокчейну, щоб зменшувалося навантаження при величезній кількості запитів на сервер. При використанні класичної реляційної бази даних MySQL маємо проблему, а саме: оператори INSERT, SELECT при величезній кількості запитів в секунду не справлялися з обробкою даних, що створювало затримку і чергу у вибірці чи додаванні даних, що в свою чергу приводило до збільшення навантаження на сервер [39-41]. Для вирішення зазначеної проблеми та зменшення навантаження на сервер було спроектовано систему і обробник для опрацювання даних у блокчейні.

Схему роботи частини системи, коли при величезній кількості запитів, зменшується навантаження між серверами, зображено на рисунку 3.1, та безпосередньо результат зменшення навантаження продемонстрований на рисунках 3.2-3.3.

Система може автоматично масштабуватися, додаючи нові сервери або ресурси відповідно до поточного навантаження. Це дозволяє підтримувати стабільну роботу системи навіть при зростаючому обсязі запитів.

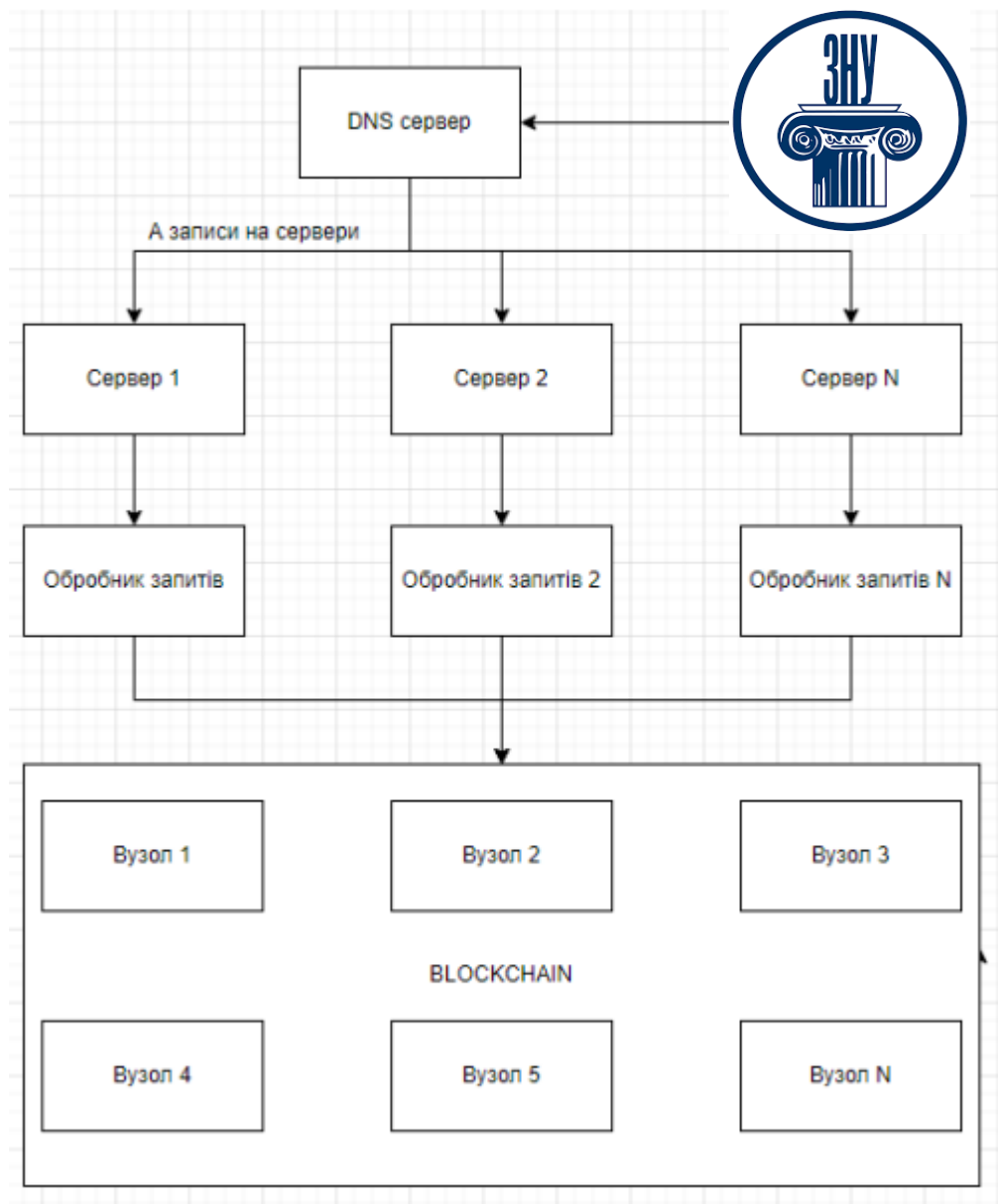


Рисунок 3.1 – Схема роботи системи

Демонстрація переключення анонсера по принципу, коли роль бази даних виконує MySQL (зображено на рисунку 3.2) та Blockchain (зображено на рисунку 3.3). Різниця очевидна, також видно, як кардинально спадає навантаженість на сервер (втричі).

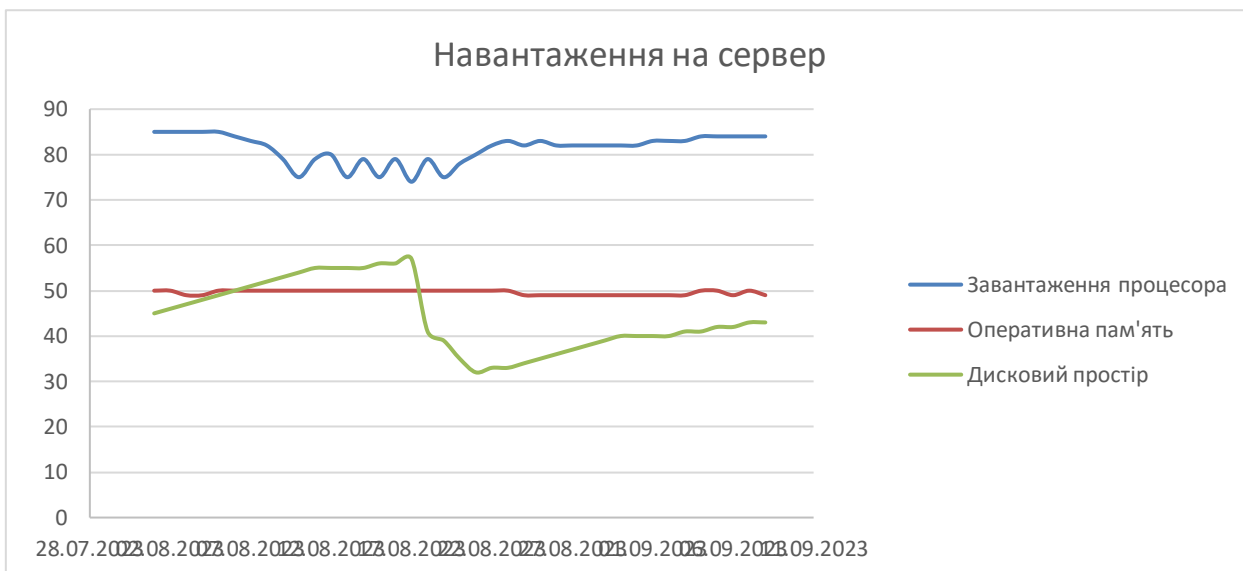


Рисунок 3.2 – Навантаження на сервер, до введення в експлуатацію



Рисунок 3.3 – Навантаження на сервер, після введення в експлуатацію

Розроблену систему можна використовувати, як конструктор для оптимізації даних, які є незмінними, наприклад логи для статистики тощо. Як ідентифікатор може виступати адреса (відкритий ключ), а транзакції – хешовані мета-дані.

Навантаження на сервер при використанні MySQL бази даних для збереження незмінних важливих даних може варіюватися в залежності від різних факторів, таких як обсяг даних, кількість запитів, інтенсивність обробки даних і інше. На рис. 3.4. представлені результати навантаження на сервер при використанні MySQL бази даних.

```

root@brown:~
top - 21:01:57 up 646 days, 14:39, 1 user, load average: 9.66, 10.83, 10.59
Tasks: 219 total, 3 running, 214 sleeping, 0 stopped, 2 zombie
%Cpu(s): 46.0 us, 16.7 sy, 0.0 ni, 35.7 id, 0.0 wa, 0.0 hi, 1.6 si, 0.0 st
KiB Mem : 16141720 total, 1211456 free, 5948180 used, 8982084 buff/cache
KiB Swap: 8387580 total, 0 free, 8387580 used. 9044716 avail Mem

  PID USER      PR  NI   VIRT   RES   SHR  S  %CPU  %MEM     TIME+ COMMAND
 11416 root        20   0  10.8g   2.2g  1416  S  193.8  14.4  21438.41 qemu-kvm
 11918 apache     20   0   53924  10488  2204  R  100.0   0.1  22116:08 nginx
 14499 mysql      20   0 4831996 233768 4120  S   50.0   1.4 165981:14 mysqld
  1228 apache     20   0  457424  12436  3392  S    6.2   0.1   0:04.55 httpd
  3414 apache     20   0  457424  11868  3336  S    6.2   0.1   0:01.72 httpd
  3862 apache     20   0  457424  11844  3304  S    6.2   0.1   0:01.06 httpd
  4405 apache     20   0  457424  12120  3384  S    6.2   0.1   0:03.48 httpd
  5758 apache     20   0  457424  11904  3332  S    6.2   0.1   0:01.69 httpd
  6144 apache     20   0  457424  12136  3388  S    6.2   0.1   0:03.46 httpd
  9075 apache     20   0  457424  12132  3384  S    6.2   0.1   0:05.13 httpd
  9181 apache     20   0  457680  12500  3428  S    6.2   0.1   0:10.19 httpd
 13347 apache     20   0  457424  11868  3332  S    6.2   0.1   0:02.12 httpd
 14461 apache     20   0  457424  11852  3304  S    6.2   0.1   0:00.77 httpd
 14462 apache     20   0  457424  11856  3304  S    6.2   0.1   0:00.77 httpd
 16706 apache     20   0  457156  11784  3304  S    6.2   0.1   0:00.70 httpd
 18737 apache     20   0  457424  12288  3400  S    6.2   0.1   0:05.89 httpd
 18783 apache     20   0  457424  12524  3496  S    6.2   0.1   0:05.93 httpd
 19742 apache     20   0  457424  12552  3500  S    6.2   0.1   0:08.48 httpd
 22577 apache     20   0  457424  12212  3444  S    6.2   0.1   0:04.00 httpd
 24543 apache     20   0  457424  12116  3384  S    6.2   0.1   0:03.28 httpd
 27945 apache     20   0  457412  11876  3332  S    6.2   0.1   0:01.89 httpd
 28057 apache     20   0  457424  11808  3296  S    6.2   0.1   0:00.43 httpd
 31109 apache     20   0  457424  11800  3304  S    6.2   0.1   0:01.18 httpd

```

Рисунок 3.4 – Навантаження на сервер при використанні MySQL бази даних для збереження незмінних важливих даних

Навантаження на сервер при використанні Blockchain як бази даних для збереження незмінних важливих даних у блоках представлено на рис. 3.5.

```

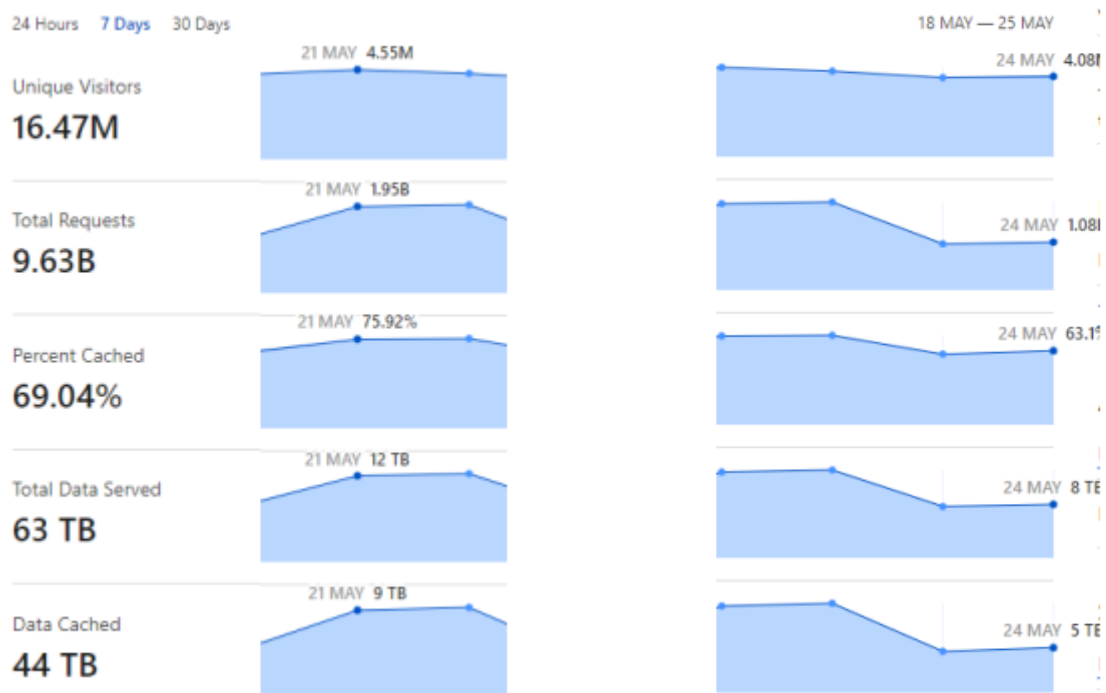
top - 21:02:48 up 91 days, 21:27, 1 user, load average: 0.37, 0.37, 0.40
Tasks: 187 total, 1 running, 186 sleeping, 0 stopped, 0 zombie
%Cpu(s): 4.6 us, 1.1 sy, 0.0 ni, 92.8 id, 1.3 wa, 0.0 hi, 0.2 si, 0.0 st
KiB Mem : 16327680 total, 14539580 used, 1788100 free, 333324 buffers
KiB Swap: 8387580 total, 0 used, 8387580 free. 12586184 cached Mem

  PID USER      PR  NI   VIRT   RES   SHR  S  %CPU  %MEM     TIME+ COMMAND
 1527 mysql      20   0 2429444 280280 11680  S    7.3   1.7 12331:45 mysqld
 2908 www-data   20   0   51980  25164  4524  S    4.3   0.2 435:04.19 nginx
32654 www-data   20   0  386360  12148  2536  S    0.3   0.1   0:17.98 apache2
    1 root        20   0   29684   5460  3040  S    0.0   0.0   0:57.97 systemd

```

Рисунок 3.5 – Навантаження на сервер при використанні Blockchain як бази даних для збереження незмінних важливих даних у блоках

Приведені дані з сервісу Cloudflare статистики запитів на анонсер, який покритий проксі сервером, до і після введення в експлуатацію:



**Кількість запитів на анонсер до роботи системи**

**Кількість запитів на анонсер після запуску роботи системи**

Рисунок 3.6 – Результати сервісу Cloudflare статистики запитів

Зниження запитів майже удвічі зумовлене тим, що падає навантаження на процесор сервера при такій схемі роботи і на сервері, який обробляє запити, не виникає проблеми з keep-timeout, де також знижується можливість втрати даних.

Розроблена модель Великих даних у блокчейн та її реалізація дає змогу зберігати дані такого розміру, що певні реляційні бази даних, наприклад MySQL, навіть на потужних серверах, не змогли б виконати операції CRUD (Create, Update, Delete) в час, зазначений у вимогах якостей Кодда (12 с) [42].

### 3.2. Розробка блокчейн-застосунку електронного кабінету студента кафедри ІЕПФ ЗНУ

Для апробації розробленої моделі Великих даних у блокчейн досліджено студентську систему електронних кабінетів. Типовий функціонал для електронних кабінетів студентів – авторизація, перегляд оцінок, перегляд файлів та завдань та проходження тестів. У даному прототипі було реалізовано такі функції:

- Авторизація, що реалізується за допомогою криптографічно-захищеного алгоритму;
- Перегляд журналу оцінок, що зберігаються на блокчейні;
- Проходження тестів;
- Запис результатів тестів на блокчейн.

Можливі дії користувача у прототипі електронного кабінету представлені у UML Use Case діаграмі на рисунку 3.7. .

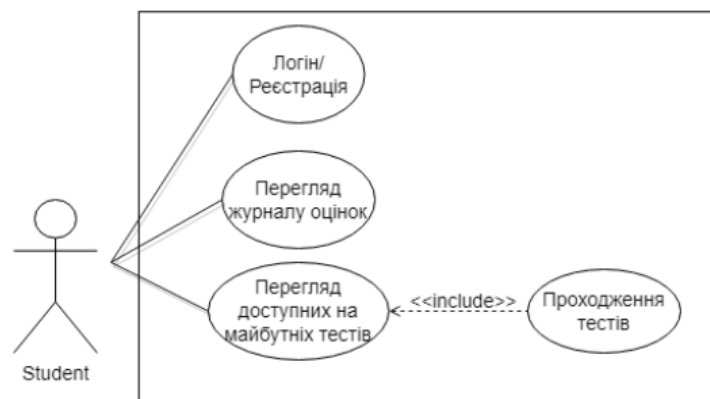


Рисунок 3.7 – UML Use Case діаграма реалізованого проєкту

З перерахованих у попередніх розділах інструментів, фреймворків та мов програмування був обраний наступний стек для розробки децентралізованого електронного кабінету:

- Мова програмування смарт-контрактів: Solidity;
- Середовище для налаштування, запуску та тестування смарт-контрактів: Remix Project;
- Програма для налаштування локального блокчейну: Ganache;



- Інструменти для взаємодії веб-застосунку з блокчейном: MetaMask та бібліотека Web3.js;

- Фреймворк для створення користувацького інтерфейсу: React.js.

Серед достатньо великої кількості мов програмування, за допомогою яких можна писати смарт-контракти для взаємодії з блокчейном та користувачами, Solidity – перша мова з якою доступне написання смарт-контрактів, і створювалася вона саме для цього. Інші мови, не маючи вбудованої підтримки такого функціоналу, використовують різні бібліотеки, кількість і якість яких не завжди може задовольнити потреби розробника. У середовищі для розробки контрактів Remix Project зібрані всі потрібні для розробки та налаштування смарт-контрактів без підключення додаткових бібліотек, на відміну від аналогів. Також у вже створений проєкт з контрактами, скомпільованими за допомогою Remix Project легко інтегруються вже наявні веб-проєкти або їх шаблони. Це дозволяє в одному редакторі одночасно розробляти смарт-контракти, автоматично тестувати їх за допомогою зручного інтерфейсу середовища та розробляти веб-проєкт і майже миттєво бачити зміни у користувацькому інтерфейсі та тестувати функції мануально.

Оскільки було обрано середовище Remix Project, до проєкту також було додано програму Ganache для взаємодії зі смарт-контрактами. Ganache – локальний блокчейн, призначений для розробки проєктів і тестування їх у режимі, близькому до продуктової версії. За допомогою зручного користувацького інтерфейсу, програма дозволяє бачити зміни у локальному блокчейні, передивлятися список транзакцій і детальні дані кожної транзакції.

Для підключення до блокчейну у веб-браузері, у проєкті використовується інструмент MetaMask. Це розширення Google Chrome на сьогоднішній день найпопулярніший інструмент для використання блокчейну у веб-проєктах, що має зручний інтерфейс та наділений найбільшою функціональністю через малої кількості аналогів.

Для взаємодії з блокчейном після підключення користувача до мережі, використовується бібліотека Web3.js. Ця бібліотека, як і MetaMask, можна сказати, є єдиним стабільним та безпечним інструментом для вирішення цього роду задач.

Користувацький інтерфейс був розроблений за допомогою фреймворку для написання веб-застосунків Vue.js. Через надмірну гнучкість, що може бути проблемою при реалізації інших веб-застосунків, цей фреймворк полегшує додавання у веб-проект інструментів для взаємодії з блокчейном, пришвидшуючи розробку та пошук помилок.

Налаштування середовища розробки – важливий етап у створенні проєкту. Для початку налаштування, на комп'ютері повинна бути встановлена платформа для виконання JavaScript-коду – Node.js. Це дозволить виконувати код не тільки на клієнтській частині застосунку, у браузері, а і для налаштування, завантаження додаткових бібліотек та збірки усього проєкту.

Після встановлення Node.js потрібно завантажити середовище для розробки контрактів Remix Project з офіційного сайту. Тепер можна створити порожню директорію, та відкрити її як нове робоче середовище. Після цього, Remix Project генерує шаблон проєкту з директоріями для написання контрактів, директоріями, де будуть зкомпільовані смарт-контракти, і директорія для написання тестів[11]. Архітектура шаблону зображена на рисунку 3.8.

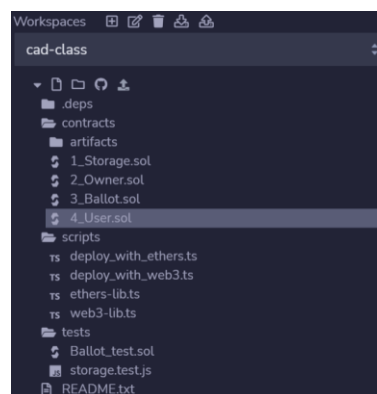


Рисунок 3.8 – Файлова структура щойно створеного проєкту з Remix Project

Структура утвореного проєкту: директорії «contracts», «.deps», «scripts», «test», а також файлу «README.txt, у якому зберігається інструкція для компіляції контрактів, їх тестування та завантаження у блокчейн. Папка «.deps» зберігає у собі файли конфігурації, у яких параметри налаштування середовища розробки смарт-контрактів, а саме: версія компілятора мови Solidity, адреса локальної мережі з якою будуть взаємодіяти смарт-контракти. У папці «contracts» зберігаються файли .sol, що представляють собою смарт-контракти. Папка «scripts» вміщує у собі JavaScript-файли, які описують хід компіляції та створення JSON-структури двійкового інтерфейсу написаних смарт-контрактів. Папка «test» містить файли для тестування контрактів[11].

Далі для того, щоб мати змогу виконувати смарт-контракти та бачити їх виконання не у результатах тестів, бачити зміни, транзакції та блоки, що створюються контрактами, треба завантажити програму Ganache[12]. Після встановлення, треба запустити програму Ця програма має зручний користувацький інтерфейс, де основними вкладками є «Акаунти»(рис. 3.9), на якій зображено 10 тестових акаунтів для взаємодії зі смарт-контрактом у якості користувачів, «Транзакції»(рис. 3.10), на якій зображений список транзакцій у рамках обраної мережі, та вкладка «Контракти»(рис. 3.11), що містить інформацію, яка зберігається у смарт-контракті та список транзакцій.

ADDRESS	BALANCE	TX COUNT	INDEX
0x81D0687D791a672C9c46Fa1fE8E5EaB50A6220e6	99.94 ETH	8	0
0xdb33632b630bC2aa3F4a88aaF5fC47bfaE050ad1	99.99 ETH	5	1
0x53089A4457b6497dC5D56a0b9b0E8d0a14fd65b7	100.00 ETH	0	2
0xE2346Cf7D610fA30532900eD5c61B8582D380bA9	100.00 ETH	0	3

Рисунок 3.9 – Вкладка «Акаунти»

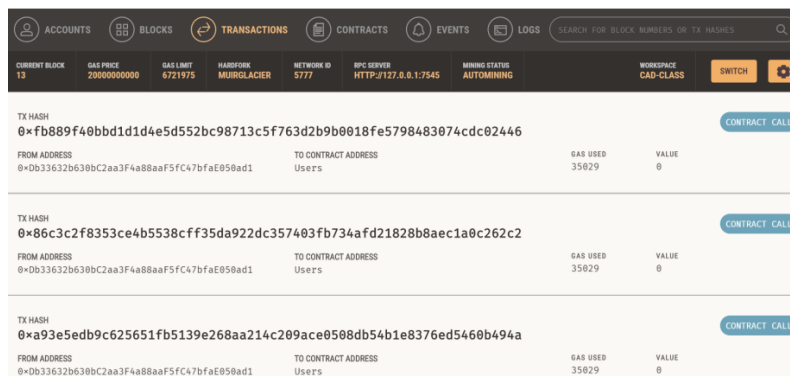


Рисунок 3.10. – Вкладка «Транзакції»

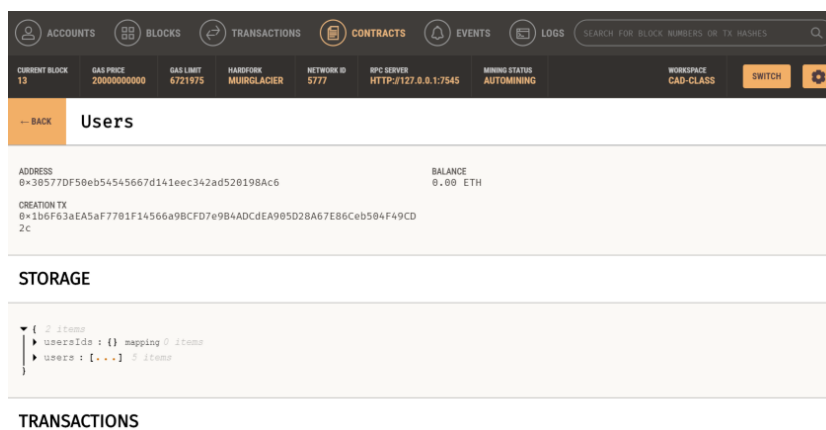


Рисунок 3.11 – Вкладка «Контракти»

Наступний крок: встановлення розширення MetaMask у веб-браузері Google Chrome та підключення акаунту до локальної мережі. На рисунку 3.12 зображено інтерфейс розширення, що містить назву мережі, до якої користувач бажає підключитися («my custom rpc»), ім'я акаунту («Горін») та його тестовий баланс (100 CPAY).

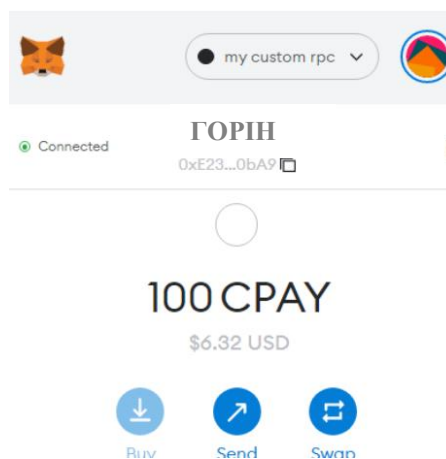
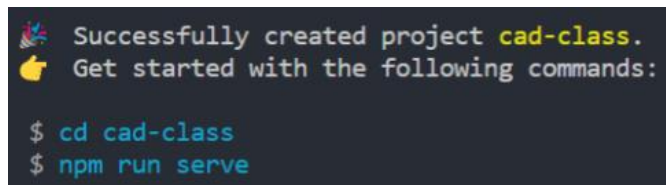


Рисунок 3.12 – Інтерфейс MetaMask

На цьому етапі частина завершується етап налаштування блокчейн-частини застосунку. Наступний етап включає в себе налаштування веб-проєкту за допомогою фреймворку React.js та підключення бібліотеки Web3.js для того, щоб взаємодіяти з налаштованою частиною блокчейн-застосунку.

Для створення шаблону проєкту на React.js, потрібно встановити цей фреймворк собі на комп'ютер за допомогою команди «`npm create-react-app my-app`», перейти до директорії з налаштованим раніше проєктом Remix Project та ввести команду «`react create`». Консольні повідомлення, що свідчать про успішну установку фреймворку та успішне ініціалізування шаблону проєкту представлено на рис.3.13.



```

Successfully created project cad-class.
Get started with the following commands:

$ cd cad-class
$ npm run serve
  
```

Рисунок 3.13 – Успішна ініціалізація шаблону web-проєкту

Останній крок – встановлення та налаштування бібліотеки Web3.js для взаємодії веб-проєкту з блокчейном. Для цього потрібно ввести команду «`npm install web3@0.20.1`», створити файл за шляхом «`src/mixins/viewMixin.js`», де реалізувати функцію, зображену на рисунку 3.14.



```

/**
 * Initialize the Web3 instance.
 *
 * @param {string} addressUrl Provider address URL like http://127.0.0.1:7545
 * @return {Promise}
 */
init(addressUrl) {
  return new Promise((resolve, reject) => {
    if (window.ethereum) {
      window.web3 = new Web3(ethereum);
      try {
        ethereum.enable()
          .then(() => {
            this.setWeb3(window.web3, addressUrl);
            resolve(window.web3);
          })
          .catch(error => reject(error));
      } catch (error) {
        reject(error);
      }
    }
  });
}
  
```

Рисунок 3.14 – Функція підключення веб-проєкту до блокчейн-мережі через розширення MetaMask

Результатом виклику функції буде підключення до розширення MetaMask, який у свою чергу підключений до тестової мережі Ganache, де зберігаються та виконуються контракти, скомпільовані середовищем Remix Project. Для більш чіткого розуміння взаємодії усіх інструментів, на рисунку 3.15 зображені зв'язки між усіма інструментами.

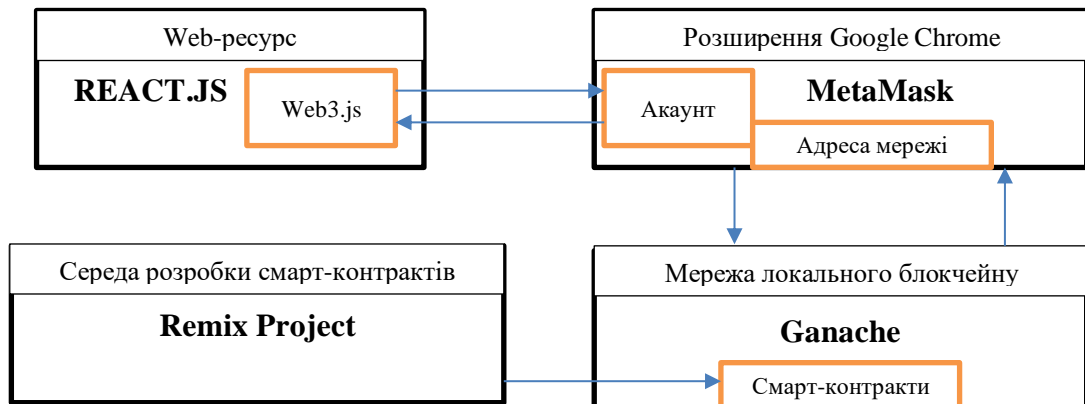


Рисунок 3.15 – Схема взаємодії інструментів, що використовуються у проєкті електронного кабінету студента

В результаті роботи у налаштованому середовищі було розроблено прототип електронного кабінету здобувача освіти кафедри Інформаційної економіки, підприємництва та фінансів ЗНУ. Застосунок має наступні сторінки: «Реєстрація», «Особистий кабінет», «Журнал оцінок», «Пройти тест» та сторінка тесту. Для огляду цих сторінок та тестування роботи написаних смарт-контрактів, пройдемося за варіантами, описаними у вищезазначеній UML Use Case діаграмі. Перше, що бачить здобувач при відкритті сайту електронного кабінету – сторінка реєстрації (рис. 3.16). На сторінці реєстрації можна побачити форму, де студент має ввести ПІБ, свою кафедру та групу.

Після заповнення полів та натискання на кнопку «Створити профіль», формуються дані транзакції та спроба записати її в блокчейн. Це робиться за допомогою налаштованої бібліотеки Web3.js, яка через MetaMask робить запит до локальної мережі на створення транзакції.

Пройти тест Журнал оцінок [Зареєструватися](#) Connected

## Реєстрація

ПІБ

Кафедра

Група

[Створити профіль](#)

Рисунок 3.16 – Сторінка реєстрації

Оскільки застосунок побудований на блокчейні Ethereum, правила запису транзакцій працюють за алгоритмом PoW, тобто за запису транзакції потрібно сплатити комісію, хоч і обчислення у локальному середовищі проводяться на комп'ютері користувача. Таким чином Ganache емулює роботу глобального блокчейну. Як тільки дані транзакції сформовані, клієнтська частина відправляє запит на запис і розширення MetaMask просить підтвердження сплати комісії за запит на локальний блокчейн для запису транзакції (рис. 3.17).

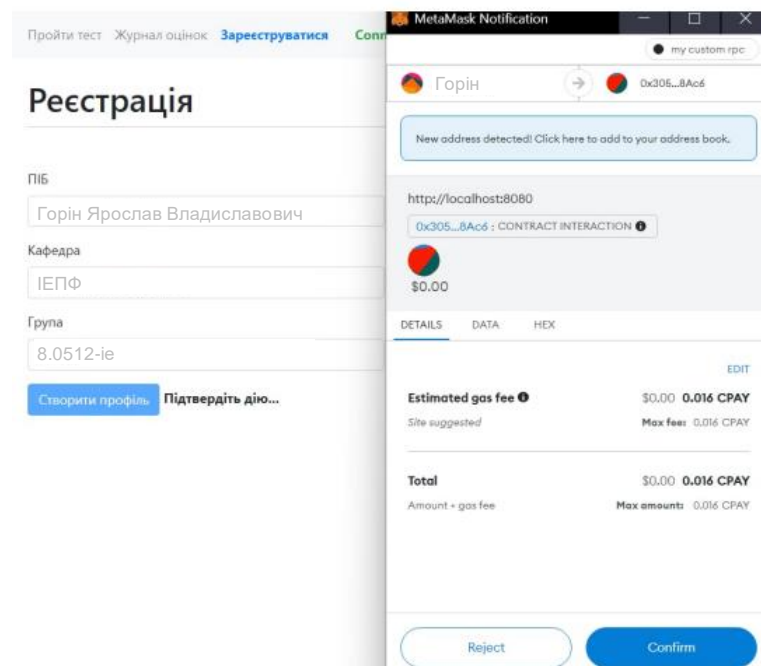


Рисунок 3.17 – Підтвердження відправки запиту на реєстрацію

Якщо на акаунті студента достатньо токенів для сплати комісії, то робиться спроба запису транзакції у блокчейн. Якщо запит на створення проходить успішно, то застосунок відправляє студента на сторінку «Профіль», а смарт-контракт записує транзакцію у блокчейн. Перевірити це можна у програмі Ganache у вкладці «Транзакції», обравши останню транзакцію (рис. 3.18). У даних транзакції можна побачити детальну інформацію щодо неї. Ця інформація включає в себе захешовані дані транзакції, розмір комісії, контракт, що записав цю транзакцію, функцію, за допомогою якої цю транзакцію було записано та аргументи цієї функції.

The screenshot displays the Ganache interface with the 'TRANSACTIONS' tab selected. The top navigation bar includes 'ACCOUNTS', 'BLOCKS', 'TRANSACTIONS', 'CONTRACTS', 'EVENTS', and 'LOGS'. Below the navigation bar, a status bar shows network information: 'CURRENT BLOCK 15', 'GAS PRICE 2000000000', 'GAS LIMIT 8721975', 'HARDFORK MURGLACIER', 'NETWORK ID 5777', 'RPC SERVER HTTP://127.0.0.1:7545', 'MINING STATUS AUTOMINING', and 'WORKSPACE CAD-CLASS'. The main area shows transaction details for block 15, including 'VALUE 0.00 ETH', 'GAS USED 23887', 'GAS PRICE 2000000000', 'GAS LIMIT 800000', and 'MINED IN BLOCK 15'. The 'TX DATA' section contains a long hexadecimal hash. Below this, the 'CONTRACT' section is expanded to show details for the 'Users' contract, including the 'FUNCTION registerUser(\_userName: string, \_department: string, \_group: string)'.

Рисунок 3.18 – Дані транзакції, що свідчать про реєстрацію здобувача освіти

Далі студент може переглянути свої оцінки по предметах на сторінці «Журнал оцінок» (рис. 3.19). На цій сторінці зображена таблиця оцінок студента. Кожен рядок має інформацію про назву дисципліни, назву теми тесту по якому виставлено оцінку, дату оцінки та адресу акаунту викладача.

Переглядати доступні тести можна на сторінці «Пройти тест». На цій сторінці зображена таблиця майбутніх тестів, які можна пройти у конкретну дату. Кожен рядок таблиці має у собі назву предмету, назву теми тесту, що







Методи смарт контракту, що розроблені у ході виконання програмної реалізації представлені в таблиці 3.1.

Таблиця 3.1 – Методи смарт контракту

Метод	Опис
get_temp_data()	Видобуток тимчасових даних з бази даних
write_bigdata()	Запис даних у тимчасову таблицю великих даних
get_bigdata()	Читання даних із тимчасової таблиці великих даних
search_duplicate()	Пошук та очистка дублікатів даних
search_fictitious()	Пошук та очистка даних від фіктивних даних
search_noise()	Пошук та очистка даних від шуму
str_to_hex()	Перетворення стрічкових даних в hex вигляд
hex_to_str()	Перетворення даних з hex вигляду в стрічкові дані
data_to_scrape()	Перетворення даних в bencode дані про сідів
data_to_announce()	Перетворення даних в bencode дані про пірів
nfo_hash_to_addr()	Перетворення info_hash в адресу на блокчейн мережі
write_to_blockchain()	Запис великих даних у блокчейн через RPC API
read_from_blockchain()	Читання великих даних із блокчейну через RPC API

Таким чином, було визначено функціонал прототипу електронного кабінету кафедри ІЕПФ, перераховані обрані технології та обґрунтований їх вибір. Також детально розглянута послідовність налаштування проєкту для розробки застосунків для веб-браузера що мають можливість взаємодіяти з блокчейном. За наведеними результатами роботи застосунку можна зробити висновок, що описаний функціонал працює правильно.

На відміну від централізованих застосунків, що дозволяють впроваджувати у освітній процес електронні кабінети та тестування, розроблений децентралізований більш прозорий як для студентів, так і для викладачів. З боку студентів перевагою застосунку є відображення наявності тестів до початку проведення, відправлення відповідей до тесту та моментальне отримання результатів напряму з блокчейну, без посередників. Запис результатів також відбувається на пряму, тому при впровадженні таких систем у освітній процес, кількість корупції та іншого роду недоброчестностей зменшується. Викладачеві застосунок дозволяє полегшити процес ведення журналу оцінок, бо за допомогою методів смарт-контракту така необхідність майже зникає.

### 3.3. Розробка протоколів обміну даними при управлінні освітнім процесом

Для обміну даних базовий протокол буде використовувати базовий протокол прикладного рівня HTTPS [43-44]. Для отримання даних буде використовуватись REST API через метод GET з випадковим ключем для забезпечення конфіденційності передачі даних [44]. Структура URL адреса буде наступною `http://1337.abcvg.info/announce` – для прийому і віддачі даних з торрент клієнта, а також: `http://1337.abcvg.info/scrape` для отримання додаткових даних на рахунок пірів. Схема роботи такої системи, зображена на рис. 3.23.

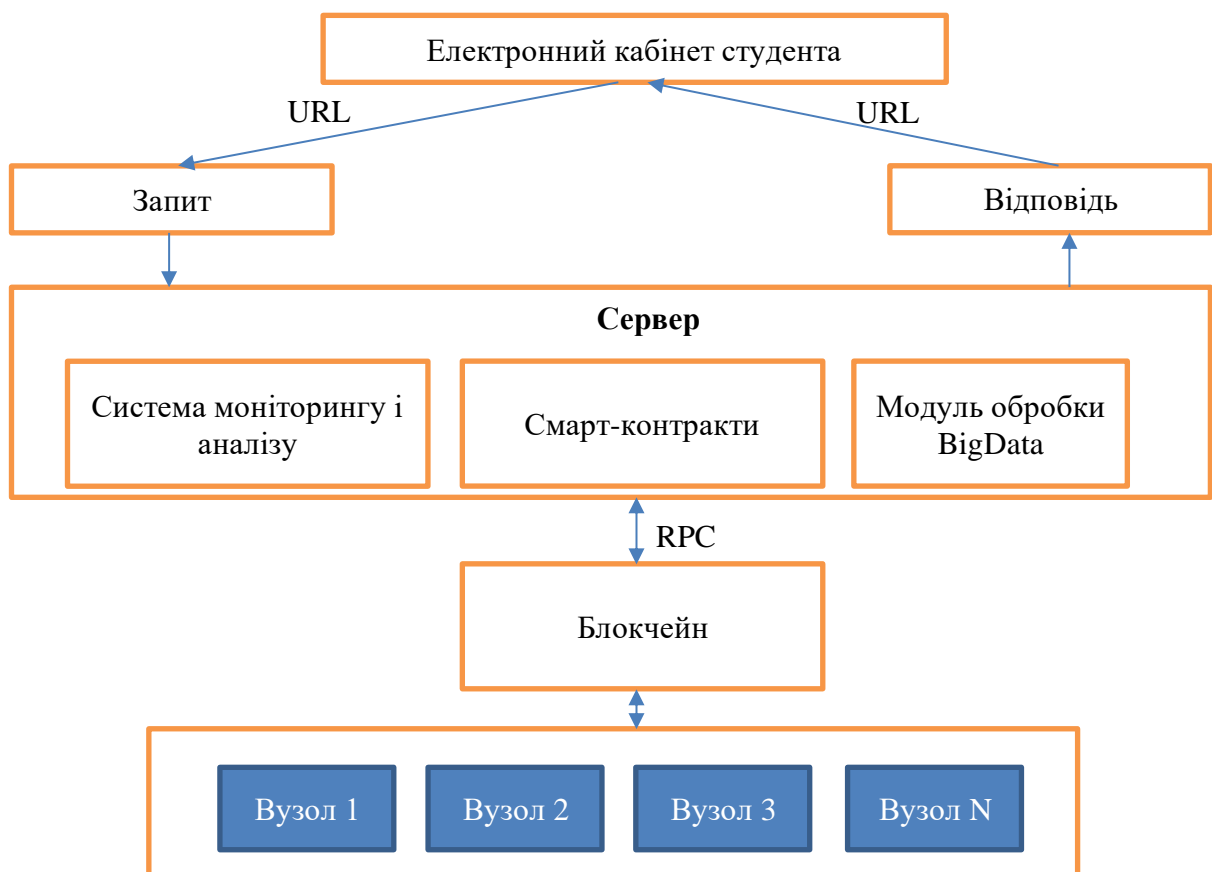


Рисунок 3.23 – Схема роботи протоколів обміну даними системи

Система розроблена на мові програмування JavaScript з використанням бази даних MySQL для запису тимчасових даних та

Blockchain системою, на основі алгоритму script для запису постійних незмінних даних, які дублюються між вузлами, серверами для майнінгу і створення емісії монет для майбутніх транзакцій.

Призначення елементів такої схеми представлено в таблиці 3.2.

Таблиця 3.2 – Схема роботи всього процесу

Елемент системи	Опис
Електронний кабінет студента	Клієнт, який передає дані на сервер
Запит	REST API запити
Відповідь	REST API відповіді
Сервер	Сервер для опрацювання всіх даних
Модуль обробки BigData	Модуль для опрацювання даних
Блокчейн	Блокчейн використовується як реплікаційна ДБ
Смарт-контракти	Програмне забезпечення та інструменти Solidity і фреймворки розробки, такі як Truffle або Hardhat
Система моніторингу і аналізу	Системи моніторингу такі як, Grafana, і засоби аналізу даних, такі як ELK (Elasticsearch, Logstash, Kibana)

Під час реалізації протоколу обміну даними виникла необхідність в оцінці якості даних та дії щодо його підвищення. Це пов'язано з тим, що оцінка якості даних є необхідним етапом будь-якого аналітичного проекту в освітній діяльності, оскільки аналітичні алгоритми або не зможуть працювати з неякісними даними, або будуть давати некоректні результати.

Приведення вихідних «сирих» даних у відповідність до необхідних критеріїв якості є найважливішим завданням Data Mining і утворює цілий напрямок, який називають передобробкою. Основними проблемами, що викликають зниження якості даних, зазвичай вважають [45]:

1. Недостовірність джерел даних: Коли джерела даних мають низьку достовірність або в них є помилки, це може призвести до введення невірних або неточних даних у систему. Наприклад, помилки вводу даних або недостовірні дані від джерел можуть спричинити проблеми з якістю даних.

2. Відсутність стандартів і правил: Наявність чітких стандартів і правил для зберігання і обробки даних може бути важливою для забезпечення якості даних. Відсутність стандартів може вести до різних форматів даних, дублювання або неправильного форматування.

3. Дублювання даних: Наявність дубльованих записів або дублювання даних у системі може викликати конфлікти і розбіжності в даних, що впливає на їхню якість.

4. Помилки вводу даних: Людські помилки при введенні даних можуть бути поширеними причинами невірних або неточних даних. Це може включати в себе орфографічні помилки, неправильні дані і т. д.

5. Відсутність забезпечення якості даних: Недостатній контроль за якістю даних і відсутність процесів виправлення помилок можуть призвести до зниження якості даних у часі.

6. Застарілість даних: Якщо дані не оновлюються регулярно, вони можуть застаріти і втратити актуальність, що також може вплинути на їхню якість.

7. Недостатня інтеграція даних: Якщо системи та джерела даних погано інтегровані між собою, це може призвести до розбіжностей і конфліктів даних.

8. Відсутність якісного контролю: Недостатній контроль за якістю даних під час їх зберігання, передачі і обробки може призвести до зниження якості даних.

У роботі проблемами, через які знижується якість даних, вважатимемо:

1. Дублікати.
2. Неповні дані.
3. Некоректні формати та подання даних (відрізняється від описаних в метаданих).
4. Фіктивні дані.
5. Суперечливі дані (козії хеш-функції).
6. Дані з порушеною структурою.

Профайлінг даних — один із найпоширеніших методів перевірки якості даних та виявлення проблем у Data Mining. Профайлінг виконується автоматично відповідно до деякого заздалегідь налаштованого сценарію на

основі аналізу інформації про структуру даних [45]. У процесі профайлінгу перевіряються поля джерела даних на відповідність до заданих обмежень. Якщо параметри полів задовільняють обмеженням, дані вважаються відповідними необхідному рівню якості, в іншому випадку необхідно вживати заходів для приведення параметрів до відповідних обмежень. Так як дані про пірів у даній дисертаційній роботі є динамічними з торрент-клієнтів, то перевірка їх якості полягає у перевірці співпадіння їх типів.

Таким чином вирішується проблема:

- неспівпадіння форматів даних,
- неповних даних,
- порушення структури даних.

Якщо формат не зазначений у метаданих блоку або результат порожній, чи чогось не вистачає, наприклад, клієнт повернув через параметр port замість цифр випадкові знаки, то даний запис відхиляється і не додається в Blockchain [42].

Таблиця 3.3 – Вигляд даних за їх типом

Змінна	Тип даних
osvita_hash	sha1 hash
osvita_ip	integer
osvita_port	integer
osvita_peer	integer
osvita_md5	md5 hash (info_ip + info_hash + info_port + info_peer)
osvita_sha1	md5 hash (info_ip + info_hash + info_port + info_peer)
osvita_upload	integer
info_download	integer
osvita_download	integer
osvita_update	tatestamp to timestamp
osvita_expire	integer
osvita_ipv46	string

Для цього типу даних важливі дані тільки osvita\_hash, osvita\_port, osvita\_peer, а також osvita\_ip. Таке подвійне хешування знижує ймовірність появи суперечливості даних (колізій хеш-функції).

Контрольна сума для перевірки цілісності та якості даних береться з параметрів: `osvita_ip`, `osvita_port`, `osvita_peer`, через те, що вони є статичними і відтворюють унікальність даних. Такі параметри, як-от `osvita_upload`, `osvita_download`, `osvita_left`, `osvita_update`, `osvita_expire` є динамічними.

Обробник приймає ці дані, записує в тимчасову таблицю тимчасової бази даних SQL, знаходить їх хеш. Якщо хеш вже існує, тоді вони не записуються в блокчейн і просто викидаються для наступного запису. Також це стосується хеш даних порожній, тоді означає, що дані поступили браковані, тому їх не потрібно записувати. Це дає змогу вирішувати проблему дублікатів значень. Є особливо важливим для Великих даних, де одні і ті самі дані можуть надходити з різних джерел.

Основна проблема таких систем – високонавантаженість, і тому інколи дані є неповними [9]. Щоб такого уникнути треба також проводити перевірку на тип даних, чи він є статичним. Якщо тип не співпадає з тим, що отримує обробник, то дані просто ігноруються і видаляються, навіть не записуються. Такий метод дозволяє уникнути проблем, коли навмисно хтось пробує записати неіснуючі дані чи сам клієнт відправляє неправдиві дані. Наприклад, коли скачане по факту одне і вираховується в `integer` форматі, а клієнт відправляє тип даних у форматі `float` з плаваючою крапкою.

### Висновки до розділу 3.

У третьому розділі була розроблена архітектура системи, описаний та розроблений протокол для обміну даних, спроектована для майбутньої розробки архітектура освітньої системи з використання блокчейн-технології для бази даних незмінних даних. Запис Великих даних у блокчейн на тестовій системі дозволив знизити кількість запитів майже удвічі. Це зумовлене тим, що падає навантаження на процесор сервера при такій схемі роботи і на сервері, який обробляє запити, не виникає проблеми з `keep-timeout`, де також знижується можливість втрати даних. Експериментально



доведено зменшення кількості запитів до системи (втричі) через підвищення якості даних, а саме через зменшення кількості дублікатів.

Розроблено прототип електронного кабінету студента кафедри ІЕПФ ЗНУ на базі блокчейн технологій, перераховані обрані технології та обґрунтований їх вибір. Детально розглянута послідовність налаштування деяких освітніх процесів для розробки застосунків для веб-браузера що мають можливість взаємодіяти з блокчейном. За наведеними результатами роботи застосунку зроблено висновок, що описаний функціонал працює правильно.

На відміну від централізованих застосунків, таких як Moodle, що дозволяють впроваджувати у освітній процес електронні кабінети та тестування, розроблений децентралізований більш прозорий як для студентів, так і для викладачів. З боку студентів перевагою застосунку є відображення наявності тестів до початку проведення, відправлення відповідей до тесту та миттєве отримання результатів напряму з блокчейну, без посередників. Запис результатів також відбувається на пряму, тому при впровадженні таких систем у освітній процес, кількість корупції та іншого роду недоброчестностей зменшується. Викладачеві застосунок дозволяє полегшити процес ведення журналу оцінок, бо за допомогою методів смарт-контракту така необхідність майже зникає.

## ВИСНОВКИ

У результаті проведеного дослідження вирішено актуальну проблему підвищення якості управління освітнім процесом за рахунок скорочення часу доступу до даних на основі розроблення інформаційної технології блокчейн для Великих даних. На прикладі кафедри інформаційної економіки, підприємництва та фінансів ЗНУ було удосконалено електронний кабінет студента на базі технології блокчейн, який при впровадженні в освітній процес дозволить зменшити кількість корупції та іншого роду недоброхочестостей під час проходження тестів та виконання індивідуальних завдань, та розроблено архітектуру Великих даних у блокчейн для освітнього процесу в інформаційній системі з видачі дипломів, яку можна використовувати, як конструктор для оптимізації даних, які є незмінними;

Основними результатами кваліфікаційної роботи є:

Проведено аналіз технології блокчейн, подано поняття великих даних, описані методи, які лежать в основі блокчейн технології. Встановлено, що головними перевагами технології блокчейн є висока швидкість операцій, безпека, можливість повноцінного контролю даних, децентралізація, багатофункціональність, універсальність, відкритість, публічність; простота, зручність і доступність. Головні недоліки: необхідність у стійкому криптографічному захисті, необхідність в інструментах передачі, користування і зберігання інформації, відсутність належного юридичного врегулювання і гарантій, недостатня поширеність, зацікавленість кіберзлочинності до нових технологій.

Досліджені інструменти процесу збору та обробки великих даних в організації та управлінні освітнім процесом. Наведені приклади можливих сфер застосування технології блокчейну в освітній діяльності ЗВО. Наведені приклади успішних проектів впровадження технологій блокчейн з обробки великих даних в освітній діяльності.

Проаналізовано дослідницькі тенденції у сфері розробки

інформаційних моделей обробки великих освітніх даних, включаючи джерела даних, збір даних, технологічні аспекти, інструментарій та аналітику даних. Розглянуті переваги та недоліки джерел для введення в системи великих освітніх даних, включаючи системи управління навчанням (LMS), відкриті освітні ресурси (OER), MOOC, соціальні медіа та пов'язані дані.

Проведений детальний аналіз методів хешування даних у блокчейні зі всіх можливих, описано та проаналізовано алгоритми консенсуса, необхідних для роботи з блокчейном. Розроблено модель Великих даних для їх послідовного опрацювання, розроблений та продемонстрований, а також описаний метод перевірки якості внесених даних.

Розроблено архітектуру Великих даних у блокчейн для освітнього процесу, яка використана в інформаційній системі з видачі дипломів. Розроблену систему можна використовувати, як конструктор для оптимізації даних, які є незмінними, наприклад, логи для статистики тощо.

Розроблена архітектура системи, описаний та розроблений протокол для обміну даних, спроектована для майбутньої розробки архітектура освітньої системи з використання блокчейн-технології для бази даних незмінних даних. Запис Великих даних у блокчейн на тестовій системі дозволив знизити кількість запитів майже удвічі. Це зумовлене тим, що падає навантаження на процесор сервера при такій схемі роботи і на сервері, який обробляє запити, не виникає проблеми з keep-timeout, де також знижується можливість втрати даних. Експериментально доведено зменшення кількості запитів до системи (втричі) через підвищення якості даних, а саме через зменшення кількості дублікатів.

Розроблено прототип електронного кабінету студента кафедри ІЕПФ ЗНУ на базі блокчейн технологій, перераховані обрані технології та обґрунтований їх вибір. Детально розглянута послідовність налаштування деяких освітніх процесів для розробки застосунків для веб-браузера що мають можливість взаємодіяти з блокчейном. За наведеними результатами

роботи застосунку зроблено висновок, що описаний функціонал працює правильно.

На відміну від централізованих застосунків, таких як Moodle, що дозволяють впроваджувати у освітній процес електронні кабінети та тестування, розроблений децентралізований більш прозорий як для студентів, так і для викладачів. З боку студентів перевагою застосунку є відображення наявності тестів до початку проведення, відправлення відповідей до тесту та моментальне отримання результатів напряму з блокчейну, без посередників. Запис результатів також відбувається на пряму, тому при впровадженні таких систем у освітній процес, кількість корупції та іншого роду недоброчестностей зменшується. Викладачеві застосунок дозволяє полегшити процес ведення журналу оцінок, бо за допомогою методів смарт-контракту така необхідність майже зникає.

## СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Khan, Mohammed. BigData Analytics Techniques to Obtain Valuable Knowledge. Indian Journal of Science and Technology. №11. 2018. URL: 10.17485/ijst/2018/v11i14/120977 (дата звернення: 16.10.2023).
2. Пічкурова З.В. Особливості розвитку технології великих даних у глобальній цифровій економіці / З.В. Пічкурова // Modern international relations: topical problems of theory and practice: collective monograph / Faculty of International Relations of the National Aviation University; under general editorship of W. Welskop, Y.O. Voloshin. – Lodz: Wydawnictwo Naukowe Wyższej Szkoły Biznesu i Nauko Zdrowiu w Łodzi, 2021. P. 178-182.
3. Zimmerman P. Blockchain structure and cryptocurrency prices // Staff working Paper / Bank of England. No 855, February 2020. P.1-75. URL: <https://www.bankofengland.co.uk/-/media/boe/files/working-paper/2020/blockchain-structure-and-cryptocurrency-prices.pdf>
4. Belotti M. et al. A vademecum on blockchain technologies: When, which, and how // IEEE Communications Surveys & Tutorials. Vol. 21 (4). 2019. P. 3796-3838. URL: <https://ieeexplore.ieee.org/document/8760539>.
5. Elrom E. Blockchain Nodes // The Blockchain Developer: Practical Guide for Designing, Implementing, Publishing, Testing, and Securing Distributed Blockchain-based Projects / Elad Elrom. – Apress Berkeley CA, 2019. – Ch. 2. – P. 31-72. URL: <https://link.springer.com/article/10.1007/s12599-017-0467-3>
6. Вергелюк Ю. Потенціал використання блокчейн технологій на фінансовому ринку. Економіка та суспільство. №38. 2022. URL: <https://doi.org/10.32782/2524-0072/2022-38-15>
7. Zhang R., Xue R., Liu L. Security and privacy on blockchain // ACM Computing Surveys (CSUR). Vol. 52 (3), art. No 51, 2019. P. 30. URL: <https://dl.acm.org/doi/pdf/10.1145/3316481>
8. Демченко О. Е. Розробка методу використання технологій Big Data та data mining в інтелектуальних системах обробки неструктурованих даних /

О. Е. Демченко // Радіоелектроніка та молодь в ХХІ столітті : матеріали 24 Міжнар. молодіж. форуму, 7-9 квіт. 2020 р. Харків : ХНУРЕ, 2020. Т. 5. С. 142–143.

9. Lashkari B., Musilek P. A comprehensive review of blockchain consensus mechanisms // IEEE Access. Vol.9, 2021. P. 43620-43652. URL: <https://ieeexplore.ieee.org/abstract/document/9376868>

10. Blockchain Development Trends 2021. URL: [https://outlierventures.io/wp-content/uploads/2021/02/OV-Blockchain-Dev-Q1-2021-\\_v7-1.pdf](https://outlierventures.io/wp-content/uploads/2021/02/OV-Blockchain-Dev-Q1-2021-_v7-1.pdf)

11. Islam N. et al. Is blockchain mining profitable in the long run? // IEEE Transactions on Engineering Management. 2021. P. 1-14. URL: <https://ieeexplore.ieee.org/abstract/document/9325951>.

12. Моторин Р. М. Використання технологій Big Data для розвитку вищої освіти. К.: КНТЕУ, 2019. С. 42-46.

13. Yaqoob I. et al. Big data: From beginning to future // International Journal of Information Management. Vol. 36, Issue 6, 2016. P. 1231-1247. URL: <https://www.sciencedirect.com/science/article/abs/pii/S0268401216304753>

14. Болюбаш Н. М. Задачі та методи інтелектуального аналізу освітніх даних. Миколаїв: ЧНУ ім. Петра Могили, 2019. С. 26-29

15. Wu X. et al. Data mining with big data // IEEE transactions on knowledge and data engineering. Vol. 26, Issue 1. P. 2013. 97-107. URL: <https://ieeexplore.ieee.org/abstract/document/6547630>.

16. Rodríguez-Mazahua L. et al. A general perspective of Big Data: applications, tools, challenges and trends // The Journal of Supercomputing. Vol. 72, 2016. P. 3073-3113.

17. Яцько О., Ватаманіца Е., Горський М. Особливості застосування OLAP-моделювання в освіті. Опт-ел. інф-енерг. техн., вип. 44, 2023. с. 5–12.

18. Rathore M. M. U. et al. Real-time big data analytical architecture for remote sensing application // IEEE journal of selected topics in applied earth observations and remote sensing. Vol. 8, Issue 10, 2015. P. 4610-4621.

19. Shi Y. *Advances in big data analytics: theory, algorithms and practices.* – Springer Nature, 2022. 728 p. URL: <https://link.springer.com/book/10.1007/978-981-16-3607-3>
20. Mathivanan S., Jayagopal P. A big data virtualization role in agriculture: a comprehensive review. *Walailak Journal of Science and Technology (WJST)*. Vol. 16, No 2, 2019. P. 55-70.
21. Rajaraman V. Big data analytics // *Resonance*. – 2016. – Vol. 21, No 8. – P. 695-716. URL: <https://link.springer.com/article/10.1007/s12045-016-0376-7>
22. Заєць С. В., Хантіль А. К. Big Data у вітчизняних маркетингових дослідженнях: роль, можливості, тенденції. Нові джерела та методи поширення даних у статистиці: матеріали XVIII Міжнародної науково-практичної конференції з нагоди Дня працівників статистики. Київ: «Інформаційно-аналітичне агентство», 2020. С. 22-28.
23. Кірей К. О. Розвиток і трансформація поняття «Big Data». *Вісник Черкаського державного технологічного університету*. 2019. №1. С. 33-40.
24. Improving Online Education Using Big Data Technologies website. URL: <https://www.intechopen.com/chapters/68814> (дата звернення: 15.09.2023).
25. Лопушинський І.П. «Цифровізація» освіти в контексті розвитку інформаційного суспільства в Україні. *Педагогічний альманах*. 2018. Вип. 37. С. 46–55.
26. Блінов В.І. Основні ідеї дидактичної концепції цифрової професійної освіти і навчання. Харків: Вид-во «Перо», 2019. 24 с.
27. Філіппова В.Д. Форсайт-технологія як інструмент формування і реалізації державної політики в галузі педагогічної освіти. *Теорія та практика державного управління і місцевого самоврядування*. No1. 2020. URL: <https://bit.ly/3MePIY0>.
28. Alblawi A. S., Alhamed A. A., “Big data and learning analytics in higher education: Demystifying variety, acquisition, storage, NLP and analytics,” in *Proc. IEEE Conf. Big Data Analytics (ICBDA)*, Nov. 2017, pp. 124–129.

29. S. Auer, C. Bizer, G. Kobilarov, J. Lehmann, R. Cyganiak, and Z. Ives, “Dbpedia: A nucleus for a Web of open data,” In *The Semantic Web*. Berlin, Germany: Springer, 2007, pp. 722–735.
30. K. Bollacker, R. Cook, and P. Tufts, “Freebase: A shared database of structured general human knowledge,” in *Proc. AAAI vol. 7*, Jul. 2007, pp. 1962–1963.
31. T. Rebele, F. Suchanek, J. Hoffart, J. Biega, E. Kuzey, and G. Weikum, “YAGO: A multilingual knowledge base from wikipedia, wordnet, and geonames,” in *Proc. Int. Semantic Web Conf.* Cham, Switzerland: Springer, Oct. 2016, pp. 177–185.
32. R. Ferguson and D. Clow, “Where is the evidence?: A call to action for learning analytics,” in *Proc. 7th Int. Learn. Analytics Knowl. Conf.*, Mar. 2017, pp. 56–65.
33. P. Michalik, J. Stofa, and I. Zolotova, “Concept definition for big data architecture in the education system,” in *Proc. IEEE 12th Int. Symp. Appl. Mach. Intell. Informat. (SAMI)*, Jan. 2014, pp. 331–334.
34. M.-S. Lee, E. Kim, C.-S. Nam, and D.-R. Shin, “Design of educational big data application using spark,” in *Proc. 19th Int. Conf. Adv. Commun. Technol. (ICACT)*, Feb. 2017, pp. 355–357.
35. Q. Zheng, H. He, T. Ma, N. Xue, B. Li, and B. Dong, “Big log analysis for E-Learning ecosystem,” in *Proc. IEEE 11th Int. Conf. e-Bus. Eng.*, Nov. 2014, pp. 258–263.
36. D. Marjanovic, M. Milovanovic, and B. Radenkovic, “Hadoop infrastructure for education,” in *Proc. 14th Int. Symp. New Bus. Models Sustain. Competitiveness*, 2014, pp. 365–370.
37. *Blockchain in Education* Alexander Grech Anthony F. Camilleri. Editor: Andreia Inamorato dos Santos 2017, URL:[https://publications.jrc.ec.europa.eu/repository/bitstream/JRC108255/jrc108255\\_blockchain\\_in\\_education\(1\).pdf](https://publications.jrc.ec.europa.eu/repository/bitstream/JRC108255/jrc108255_blockchain_in_education(1).pdf)



38. Поляков Н. Е. Впровадження технології блокчейн в освіту: зарубіжний досвід / Н. Є. Поляков, А. В. Солодов // Управління соціально-економічними системами: теорія, методологія, практика: збірник статей III Міжнародної науково-практичної конференції. - П: МЦНС «Наука і Просвещение». Ч. 2, 2017. С. 100-104.

39. Khrystynets, N., Miskevych, O., & Mazurenko, V. Технології Blockchain для оптимізації процесів документообігу // Computer-integrated technologies: education, science, production. – 2020. – № 40. – P. 153-157.

40. Tariq N. et al. The security of big data in fog-enabled IoT applications including blockchain: A survey // Sensors. – 2019. – Т. 19. – №. 8. – С. 1788. URL: <https://www.mdpi.com/1424-8220/19/8/1788>

41. Li J. et al. Blockchain-based public auditing for big data in cloud storage // Information Processing & Management. – 2020. – Т. 57. – №. 6. – С. 102382. URL: <https://doi.org/10.1016/j.ipm.2020.102382>

42. Кучковський В. В. Алгоритми консенсуса блокчейн систем // Вісник Хмельницького національного університету. Серія: Технічні науки. № 3 (297), 2021. С. 30–33.

43. Кучковський В. В. Алгоритми консенсусу // Trends in science and practice of today : abstracts of XXVIII International scientific and practical conference (Ankara, Turkey; June 01 – 04, 2021), 2021. С. 502–505.

44. Кучковський В. В. Змішані алгоритми консенсусу // Trends in science and practice of today : abstracts of XXVIII International scientific and practical conference (Ankara, Turkey; June 01 – 04, 2021), 2021. С. 506–507.

45. Hand D. J. Principles of data mining // Drug safety. – 2007. – Vol. 30, Issue 7. – P. 621-622. URL: <https://link.springer.com/article/10.2165/00002018-200730070-0001>.