

Міністерство освіти і науки України
Харківський національний університет імені В. Н. Каразіна

Навчально-науковий інститут комп'ютерних наук та штучного інтелекту

Кафедра математичного моделювання та аналізу даних

До захисту допущено

Кафедрою ІПСіТ _____ протокол № _____ від _____

завідувач кафедри _____ Володимир СТРУКОВ
(підпис) (ім'я, прізвище)

« _____ » _____ 2026 р.

Кваліфікаційна робота
здобувача першого (бакалаврського) рівня вищої освіти

Створення веб-платформи для генерації коротких відео за текстовими
запитами користувачів із сучасним інтерфейсом та інтеграцією
мультимедійних сервісів

(назва роботи)

Спеціальність (спеціалізація) 122 Комп'ютерні науки

Освітня програма Комп'ютерні науки

Виконавець



(підпис)

Вікторія Євстегнеєва

(ім'я, прізвище)

Науковий керівник



(підпис)

Ольга Маций

(ім'я, прізвище)

Харків – 2026

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
Харківський національний університет імені В. Н. Каразіна

Навчально-науковий інститут комп'ютерних наук та штучного інтелекту
Кафедра Математичного моделювання та аналізу даних
Рівень вищої освіти (освітньо-кваліфікаційний рівень) Бакалавр
Галузь знань Інформаційні технології
Спеціальність 122 Комп'ютерні науки
Освітня програма Комп'ютерні науки

ЗАТВЕРДЖУЮ
Завідувач кафедри ММ та АД
Струков Володимир Михайлович
« 13 » жовтня 2025 року

ЗАВДАННЯ НА КВАЛІФІКАЦІЙНУ РОБОТУ (ПРОЕКТ)

Євстегнєва Вікторія Денисівна
(прізвище, ім'я, по батькові студента)

1. Тема роботи: *Створення веб-платформи для генерації коротких відео за текстовими запитами користувачів із сучасним інтерфейсом та інтеграцією мультимедійних сервісів*

керівник роботи: Мацуй Ольга Борисівна,
(прізвище, ім'я, по батькові, науковий ступінь, вчене звання)

затверджені наказом по університету від “___” _____ 20__ року
№ _____

2. роботи	Строк подання студентом
-----------	-------------------------

3. Перелік питань, які потрібно розробити:

1. Проектування архітектури веб-платформи автоматизованої генерації коротких відео за текстовими запитам користувачів;
2. Розробка та реалізація автоматизованого workflow генерації відео у середовищі p8n;
3. Інтеграція великих мовних моделей (LLM), сервісів синтезу мовлення (TTS) та сервісів генерації зображень і відео;

4. Реалізація модуля генерації сценарію, сцен, промптів та формування відеоконтенту;
5. Реалізація системи збереження та обліку результатів генерації відео з використанням Google Drive та Google Sheets
6. Проведення тестування веб-платформи та оцінювання якості згенерованого відеоконтенту;
7. Розробка технічної документації та підготовка пояснювальної записки до дипломної роботи.

4. План роботи

№ з/п	Назви етапів роботи	Термін виконання етапів роботи
1	Отримання завдання та узгодження технічного завдання на розроблення веб-платформи для генерації коротких відео за текстовими запитами користувачів із сучасним інтерфейсом та інтеграцією мультимедійних сервісів.	13.10.2025 - 20.10.2025
2	Аналіз предметної області, аналогів та сучасних підходів до організації систем веб-платформи для генерації коротких відео.	14.11.2025 - 26.11.2025
3	Визначення функціональних і нефункціональних вимог до веб-платформи.	12.12.2025 - 20.12.2025
4	Дослідження технологій LLM, TTS та відеогенерації.	22.12.2025 - 06.01.2026
5	Проектування інформаційної архітектури системи, структури сторінок і навігації користувачів.	08.01.2026 - 29.01.2026
6	Розробка п8n workflow для генерації відео.	12.02.2026 - 20.02.2026
7	Інтеграція зовнішніх сервісів (Kling, ElevenLabs, Google Drive).	20.2.2026 - 25.02.2026
8	Реалізація генерації сцен, промптів та відео.	25.02.2026 - 20.03.2026
9	Реалізація озвучення (TTS) та рендерингу відео.	10.04.2026 - 15.04.2026
10	Проведення тестування веб-платформи, аналіз коректності роботи та оцінка зручності використання.	16.04.2026 - 23.04.2026
11	Оформлення пояснювальної записки, підготовка графічних матеріалів і матеріалів до захисту.	17.05.2026 - 29.05.2026

5. Дата видачі завдання 13.10.2025

Студент

Євстегнєєва В.Д.

ініціали, прізвище

підпис

Керівник роботи

Маций О.Б.

ініціали, прізвище

підпис

АНОТАЦІЯ

Пояснювальна записка до кваліфікаційної роботи містить 102 сторінки, 3 розділи, 16 таблиць, 35 джерел посилання та 5 додатків.

Кваліфікаційна робота присвячена розробці програмної системи автоматизованої генерації відеоконтенту на основі технологій штучного інтелекту. Метою роботи є створення системи автоматичного формування відеороликів у форматі 9:16 на основі текстового запиту користувача шляхом інтеграції мовних моделей, сервісів генерації відео, синтезу мовлення та інструментів автоматизації бізнес-процесів.

Об'єктом дослідження є процес автоматизованої генерації відеоконтенту із застосуванням технологій штучного інтелекту. Предметом дослідження є методи оркестрації AI-сервісів, генерації сценаріїв, відеофрагментів та мовлення, а також автоматизованого рендерингу мультимедійного контенту.

Для реалізації системи використано платформу n8n, OpenAI GPT, Google Gemini, PiAPI Kling, ElevenLabs, Telegram Bot API, Google Drive API, Google Sheets API та FFmpeg.

Практичним результатом роботи є створення програмної системи, яка забезпечує автоматичну генерацію коротких відеороликів за текстовим описом користувача через Telegram-бота. Система виконує формування сценарію, генерацію відеофрагментів, синтез мовлення, рендеринг фінального відео та збереження результатів у хмарних сховищах.

Отримані результати можуть бути використані у сфері цифрового маркетингу, туристичному бізнесі, медіавиробництві та інших сферах створення мультимедійного контенту.

Ключові слова: ШТУЧНИЙ ІНТЕЛЕКТ, ГЕНЕРАЦІЯ ВІДЕО, LLM-АГЕНТИ, N8N, TELEGRAM BOT, ELEVENLABS, ВЕРТИКАЛЬНІ ВІДЕО, АВТОМАТИЗАЦІЯ КОНТЕНТУ, AI-СИСТЕМИ.

ABSTRACT

The explanatory note to the qualification thesis consists of 102 pages, 3 chapters, 16 tables, 35 references, and 5 appendices.

The qualification thesis is devoted to the development of an automated video content generation system based on artificial intelligence technologies. The aim of the work is to develop a software system for automatic creation of 9:16 video clips based on user text requests through the integration of large language models, video generation services, speech synthesis technologies, and business process automation tools.

The object of the research is the process of automated video content generation using artificial intelligence technologies. The subject of the research includes methods of AI service orchestration, generation of video scenarios, video clips and speech, as well as automated rendering of multimedia content.

The system is implemented using the n8n platform, OpenAI GPT, Google Gemini, PiAPI Kling, ElevenLabs, Telegram Bot API, Google Drive API, Google Sheets API, and FFmpeg.

The practical result of the work is a software system that provides automatic generation of short video clips based on user text descriptions through a Telegram bot. The system performs scenario generation, video clip creation, speech synthesis, final video rendering, and storage of generated results in cloud repositories.

The obtained results can be applied in digital marketing, tourism, media production, and other fields related to multimedia content creation.

Keywords: ARTIFICIAL INTELLIGENCE, VIDEO GENERATION, LLM AGENTS, N8N, TELEGRAM BOT, ELEVENLABS, VERTICAL VIDEOS, CONTENT AUTOMATION, AI SYSTEMS.

ЗМІСТ

<u>ПЕРЕЛІК СКОРОЧЕНЬ, УМОВНИХ ПОЗНАЧЕНЬ ТА ТЕРМІНІВ</u>	8
<u>ВСТУП</u>	10
<u>РОЗДІЛ 1 АНАЛІЗ СУЧАСНИХ ТЕХНОЛОГІЙ ГЕНЕРАЦІЇ</u> <u>ВІДЕОКОНТЕНТУ ТА AI-СИСТЕМ</u>	13
<u>1.1. Особливості створення цифрового відеоконтенту в сучасних умовах</u>	13
<u>1.2. Аналіз великих мовних моделей для автоматизованої генерації сценаріїв</u> <u>та текстового контенту</u>	15
<u>1.3. Аналіз сучасних сервісів генерації відео та синтезу мовлення</u>	18
<u>1.4. Аналіз платформ автоматизації бізнес-процесів та оркестрації AI-сервісів</u>	20
<u>1.5. Порівняльний аналіз існуючих рішень та постановка задачі дослідження</u>	23
<u>1.6. Висновки до Розділу 1</u>	25
<u>РОЗДІЛ 2 ПРОЄКТУВАННЯ ТА РЕАЛІЗАЦІЯ ВЕБ-ПЛАТФОРМИ ДЛЯ</u> <u>ГЕНЕРАЦІЇ ВІДЕОКОНТЕНТУ</u>	28
<u>2.1. Концепція програмної системи та загальна архітектура рішення</u>	28
<u>2.2. Проєктування workflow генерації відео на базі платформи n8n</u>	30
<u>2.3. Інтеграція зовнішніх сервісів</u>	34
<u>2.4. Моделі даних та організація зберігання інформації</u>	39
<u>2.5. Обробка помилок та забезпечення відмовостійкості</u>	44
<u>2.6. Висновки до Розділу 2</u>	52
<u>РОЗДІЛ 3 ТЕСТУВАННЯ, ОЦІНКА ЕФЕКТИВНОСТІ ТА АНАЛІЗ</u> <u>РЕЗУЛЬТАТІВ РОБОТИ СИСТЕМИ</u>	49
<u>3.1. Методика тестування програмної системи та сценарії використання</u>	49
<u>3.2. Аналіз продуктивності системи та оцінка стабільності роботи</u>	52
<u>3.3. Економічна ефективність розробленої системи</u>	56
<u>3.4. Оцінювання якості відеоконтенту</u>	58
<u>3.5. Аналіз отриманих результатів</u>	61
<u>3.6. Висновки до Розділу 3</u>	63

<u>ВИСНОВКИ</u>	66
<u>ПЕРЕЛІК ВИКОРИСТАНИХ ДЖЕРЕЛ</u>	69
<u>ДОДАТОК А. Повний workflow платформи n8n</u>	72
<u>ДОДАТОК Б. Основні фрагменти програмного коду та конфігурацій</u>	75
<u>ДОДАТОК В. Приклади сформованих сценаріїв та промтів для генерації відео</u>	80
<u>ДОДАТОК Г. Результати тестування та приклади згенерованого контенту</u>	82
<u>ДОДАТОК Є. Політики безпеки та відмовостійкості</u>	88

ПЕРЕЛІК СКОРОЧЕНЬ, УМОВНИХ ПОЗНАЧЕНЬ, ТЕРМІНІВ

AI – Artificial Intelligence, штучний інтелект.

API – Application Programming Interface, програмний інтерфейс взаємодії.

FFmpeg – Fast Forward MPEG, програмний комплекс для обробки та рендерингу відео.

GPT – Generative Pre-trained Transformer, велика мовна модель для генерації тексту.

HTTP – HyperText Transfer Protocol, протокол передачі гіпертекстових даних.

JSON – JavaScript Object Notation, формат обміну структурованими даними.

Kling AI – сервіс генерації відеоконтенту на основі технологій штучного інтелекту.

LLM – Large Language Model, велика мовна модель.

NLP – Natural Language Processing, обробка природної мови.

n8n – платформа автоматизації бізнес-процесів та інтеграції програмних сервісів.

OpenAI – компанія-розробник великих мовних моделей та AI-сервісів.

Prompt – текстова інструкція для генеративної моделі штучного інтелекту.

REST API – Representational State Transfer Application Programming Interface, програмний інтерфейс взаємодії веб-сервісів.

SFTP – Secure File Transfer Protocol, захищений протокол передачі файлів.

TTS – Text-to-Speech, технологія синтезу мовлення з тексту.

Telegram Bot API – програмний інтерфейс для створення та керування Telegram-ботами.

UI – User Interface, користувацький інтерфейс.

URL – Uniform Resource Locator, уніфікований покажчик ресурсу.

UX – User Experience, користувацький досвід взаємодії із системою.

Webhook – механізм автоматичної передачі даних між програмними системами при виникненні події.

Workflow – послідовність автоматизованих процесів та операцій у програмній системі.

Google Drive API – програмний інтерфейс доступу до сервісу Google Drive.

Google Sheets API – програмний інтерфейс доступу до електронних таблиць Google Sheets.

ElevenLabs – сервіс синтезу мовлення на основі нейронних мереж.

Gemini – велика мовна модель компанії Google.

CSV – Comma-Separated Values, формат зберігання табличних даних.

MP3 – цифровий формат зберігання аудіофайлів.

MP4 – цифровий формат зберігання відеофайлів.

ВСТУП

Сучасний етап розвитку цифрових технологій характеризується стрімким впровадженням систем штучного інтелекту у різні сфери діяльності людини. Особливо помітним цей процес став у сфері створення мультимедійного контенту, де генеративні моделі штучного інтелекту дозволяють автоматизувати складні творчі процеси та значно скорочувати час виробництва цифрових матеріалів.

Одним із найбільш популярних форматів контенту в сучасному інформаційному просторі є короткі вертикальні відеоролики, які активно використовуються на платформах TikTok, Instagram Reels, Facebook Reels та YouTube Shorts. За останні роки саме цей формат став одним із ключових інструментів цифрового маркетингу, просування брендів, туристичних послуг, освітніх проєктів та персональних медіаресурсів. Висока залученість користувачів та швидкість поширення інформації обумовлюють постійне зростання попиту на створення великої кількості якісного відеоконтенту.

Проте процес підготовки відеороликів традиційними методами залишається досить складним та ресурсомістким. Для створення навіть короткого відео необхідно виконати декілька послідовних етапів: розробити сценарій, підготувати текстове наповнення, створити або підібрати візуальні матеріали, виконати озвучення, монтаж та фінальний рендеринг. Реалізація зазначених процесів потребує значних витрат часу, фінансових ресурсів та залучення спеціалістів різних напрямків.

Особливо актуальною дана проблема є для туристичних компаній, діяльність яких безпосередньо залежить від регулярного створення якісного рекламного та інформаційного контенту. Туристичні оператори постійно потребують нових відеоматеріалів для популяризації екскурсій, маршрутів, природних локацій та туристичних послуг. В умовах високої конкуренції швидкість створення контенту стає одним із факторів успішності бізнесу.

Аналіз сучасних наукових досліджень та програмних рішень свідчить про активний розвиток генеративних моделей штучного інтелекту. Великі

мовні моделі (Large Language Models) здатні автоматично формувати тексти, сценарії та описи, системи генерації відео дозволяють створювати відеофрагменти на основі текстових запитів, а технології синтезу мовлення забезпечують автоматичне створення якісного аудіосупроводу. Разом з тим більшість існуючих рішень орієнтовані на виконання окремих завдань та не забезпечують повністю автоматизований цикл створення готового відеоконтенту.

У зв'язку з цим актуальною науково-практичною задачею є розробка інтегрованої програмної системи, яка об'єднує можливості сучасних AI-сервісів у межах єдиного автоматизованого процесу генерації відеороликів.

Об'єктом дослідження є процес автоматизованої генерації відеоконтенту із застосуванням технологій штучного інтелекту.

Предметом дослідження є методи інтеграції великих мовних моделей, сервісів генерації відео, технологій синтезу мовлення та засобів автоматизації бізнес-процесів для створення програмної системи генерації мультимедійного контенту.

Метою роботи є розробка веб-платформи для автоматизованої генерації коротких відеороликів за текстовими запитами користувачів із використанням сучасних технологій штучного інтелекту та інтеграцією мультимедійних сервісів.

Для досягнення поставленої мети необхідно вирішити такі завдання:

- провести аналіз сучасних технологій штучного інтелекту, що використовуються для автоматизованої генерації мультимедійного контенту;
- дослідити можливості великих мовних моделей для формування сценаріїв та текстового супроводу відеороликів;
- виконати аналіз сучасних сервісів генерації відео та синтезу мовлення;
- дослідити засоби автоматизації бізнес-процесів та оркестрації AI-сервісів;
- спроектувати архітектуру веб-платформи для генерації коротких відеороликів;

- розробити автоматизований workflow генерації відеоконтенту на базі платформи n8n;
- реалізувати інтеграцію сервісів OpenAI GPT, Google Gemini, Kling AI, ElevenLabs, Google Drive та Telegram Bot API;
- реалізувати модуль генерації сценарію, сцен та промптів для створення відеоконтенту;
- реалізувати механізм автоматичного озвучення та рендерингу фінального відео;
- забезпечити збереження та облік результатів генерації за допомогою Google Drive та Google Sheets;
- провести тестування програмної системи та оцінити якість згенерованого контенту;
- проаналізувати ефективність використання розробленої системи для автоматизації процесу створення відеоконтенту.

РОЗДІЛ 1 АНАЛІЗ СУЧАСНИХ ТЕХНОЛОГІЙ ГЕНЕРАЦІЇ ВІДЕОКОНТЕНТУ ТА AI-СИСТЕМ

1.1. Особливості створення цифрового відеоконтенту в сучасних умовах

У сучасному інформаційному суспільстві цифровий контент став одним із ключових інструментів комунікації між бізнесом та споживачами. Особливу роль у цьому процесі відіграє відеоконтент, який забезпечує високий рівень залучення аудиторії та ефективне донесення інформації. Завдяки розвитку соціальних мереж та мобільних технологій короткі відеоролики стали одним із найбільш популярних форматів представлення інформації в мережі Інтернет.

Найбільшого поширення набули короткі вертикальні відео, що використовуються на платформах TikTok, Instagram Reels, Facebook Reels та YouTube Shorts. Основними перевагами такого формату є швидке сприйняття інформації користувачами, висока ефективність рекламних кампаній та можливість охоплення значної аудиторії за короткий проміжок часу. Саме тому підприємства різних сфер діяльності активно використовують відеоконтент для просування власних товарів і послуг.

Особливо актуальним використання відеоконтенту є для туристичної галузі. Туристичні компанії постійно потребують створення якісних рекламних матеріалів для демонстрації туристичних маршрутів, природних локацій, екскурсійних програм та додаткових послуг. Візуальне представлення туристичних продуктів значно підвищує зацікавленість потенційних клієнтів та позитивно впливає на прийняття рішення щодо бронювання послуг.

Базою для дослідження у даній роботі виступає компанія TRÓPICO EXCÊNTRICO – UNIPESSOAL LDA, що працює під торговою маркою Adventure Tours Madeira. Компанія здійснює діяльність на острові Мадейра (Португалія) та спеціалізується на організації активного відпочинку, туристичних екскурсій, джип-турів, багі-турів, піших маршрутів левадами, морської риболовлі, йога-турів та інших видів туристичних послуг. Крім

туристичного напрямку компанія також працює у сфері нерухомості, займаючись пошуком, проєктуванням та реалізацією будівельних об'єктів на території острова.

Для ефективного просування своїх послуг Adventure Tours Madeira активно використовує соціальні мережі та цифрові маркетингові інструменти. Щодня компанія створює та публікує фото- і відеоматеріали, що демонструють туристичні маршрути, природні пам'ятки, активний відпочинок та реальний досвід клієнтів. Постійна потреба у створенні нового контенту вимагає значних часових і фінансових ресурсів, пов'язаних із написанням сценаріїв, підготовкою візуальних матеріалів, монтажем відео та створенням озвучення.

Традиційний процес виробництва відеоконтенту включає декілька послідовних етапів: формування концепції майбутнього ролика, написання сценарію, підготовку відеоматеріалів, монтаж, озвучення, додавання ефектів та фінальний рендеринг. Кожен із цих етапів потребує використання спеціалізованого програмного забезпечення та залучення відповідних спеціалістів. У результаті створення навіть короткого рекламного відео може займати від кількох годин до декількох днів.

Суттєвий розвиток технологій штучного інтелекту дозволив автоматизувати значну частину зазначених процесів. Великі мовні моделі здатні формувати сценарії, описи сцен, рекламні тексти та інші текстові матеріали на основі коротких запитів користувача. Паралельно з цим активно розвиваються генеративні моделі створення зображень та відео, а також системи синтезу мовлення, що забезпечують автоматичне формування аудіосупроводу.

Сучасні генеративні моделі відео, зокрема Kling AI, Runway та Pika, у поєднанні з платформами автоматизації бізнес-процесів на кшталт n8n відкривають принципово нові можливості для створення цифрового контенту. Використання таких технологій дозволяє автоматизувати більшість етапів виробництва коротких відеороликів, починаючи від формування початкової ідеї та сценарію і закінчуючи створенням готового відеофайлу для подальшої

публікації в соціальних мережах.

На відміну від традиційного підходу, який потребує залучення сценаристів, дизайнерів, монтажерів та інших спеціалістів, сучасні AI-системи здатні виконувати значну частину цих завдань автоматично. Великі мовні моделі аналізують текстовий запит користувача та формують структуру майбутнього відеоролика, сервіси генерації відео створюють візуальні сцени на основі текстових описів, а системи синтезу мовлення генерують аудіосупровід природним голосом. Завдяки цьому процес створення контенту суттєво прискорюється та стає більш доступним для підприємств різного масштабу.

У межах даної роботи для компанії Adventure Tours Madeira було розроблено програмну систему автоматизованої генерації коротких вертикальних відеороликів тривалістю близько 15 секунд. Розроблене рішення забезпечує повний цикл створення контенту: від отримання текстового запиту користувача через Telegram-бот до генерації готового відеоролика із подальшим збереженням результатів та наданням посилання на створений контент. Для реалізації системи використано платформу автоматизації n8n, великі мовні моделі для генерації сценаріїв і промптів, сервіс Kling AI для створення відеофрагментів та систему ElevenLabs для синтезу мовлення.

Таким чином, автоматизація процесів створення цифрового відеоконтенту є одним із найбільш перспективних напрямків розвитку сучасних інформаційних технологій. Використання генеративного штучного інтелекту дозволяє значно підвищити швидкість створення мультимедійних матеріалів, зменшити витрати ресурсів та забезпечити стабільне виробництво якісного контенту для потреб бізнесу.

1.2 Аналіз великих мовних моделей для автоматизованої генерації сценаріїв та текстового контенту

Одним із найважливіших напрямків розвитку сучасного штучного інтелекту є створення великих мовних моделей (Large Language Models, LLM),

призначених для обробки та генерації текстової інформації природною мовою. Такі моделі навчаються на значних обсягах текстових даних та здатні виконувати широкий спектр завдань: відповідати на запитання, генерувати тексти, перекладати інформацію між мовами, створювати сценарії, аналізувати документи та формувати структуровані дані.

Основою сучасних LLM є архітектура Transformer, запропонована компанією Google у 2017 році. На відміну від попередніх підходів до обробки тексту, архітектура Transformer дозволяє ефективно враховувати контекст усієї послідовності слів, що значно покращує якість генерації тексту. Саме ця технологія стала фундаментом для створення таких моделей, як GPT, Gemini, Claude та Llama.

Серед найбільш відомих великих мовних моделей сьогодні можна виділити GPT від компанії OpenAI, Gemini від Google, Claude від Anthropic та Llama від Meta. Кожна з цих систем має власні особливості та переваги, однак усі вони здатні ефективно вирішувати задачі генерації текстового контенту на основі запитів користувача.

Для автоматизованого створення відеоконтенту великі мовні моделі мають особливе значення, оскільки саме вони забезпечують перетворення початкової ідеї користувача на структурований сценарій майбутнього відеоролика. На відміну від звичайних шаблонних систем, LLM здатні аналізувати зміст запиту, враховувати контекст та формувати логічно послідовні сценарії відповідно до заданої тематики.

У розробленій системі великі мовні моделі використовуються як інтелектуальні агенти, що виконують окремі функціональні завдання в межах автоматизованого workflow. Такий підхід дозволяє розподілити процес генерації контенту на декілька незалежних етапів та забезпечити більш стабільну якість результатів.

Першим етапом роботи є формування структури майбутнього відеоролика. Після отримання текстового запиту користувача LLM-агент аналізує зміст повідомлення та генерує набір сцен, які повинні бути

використані для побудови відео. У межах реалізованої системи кожен ролик складається з трьох сцен тривалістю близько п'яти секунд.

Для кожної сцени формується набір параметрів, серед яких основний об'єкт зображення, опис дії, характеристика оточення, особливості руху камери, освітлення, настрій сцени та інформація про переходи між окремими фрагментами відео. Такий підхід дозволяє забезпечити логічну послідовність подій та зберегти цілісність сюжету.

Наступним етапом є генерація промптів для відеомоделі Kling AI. Для цього використовується окремий агент, який на основі опису сцени формує текстовий запит у форматі, оптимізованому для генерації відео. Під час створення промптів дотримується певна структура, що включає опис головного об'єкта, виконуваної дії, навколишнього середовища, руху камери та умов освітлення.

Якість сформованих промптів безпосередньо впливає на результат генерації відео, тому правильне проєктування агентів генерації є важливим елементом усієї системи. Використання LLM дозволяє створювати більш деталізовані та змістовні описи порівняно з шаблонними алгоритмами.

Окрім генерації сцен та промптів, мовні моделі використовуються для автоматичного створення текстового супроводу готового відеоролика. На основі початкової ідеї та змісту сцен система формує заголовок (Title) і опис (Description), які можуть бути використані під час публікації контенту в соціальних мережах.

Автоматична генерація заголовків дозволяє створювати більш привабливий контент для аудиторії та підвищувати ефективність просування відеоматеріалів. При цьому враховуються ключові слова, тематика ролика та рекомендації щодо оптимізації контенту для соціальних мереж.

Додатково мовні моделі використовуються для створення тексту озвучення (Voiceover). Сформований текст передається до системи синтезу мовлення ElevenLabs, яка генерує аудіофайл природною мовою. Отриманий голосовий супровід використовується під час фінального рендерингу

відеоролика.

Використання великих мовних моделей у межах єдиного автоматизованого конвеєра дозволяє забезпечити високий рівень автоматизації процесу створення відеоконтенту та мінімізувати участь людини на всіх етапах підготовки матеріалів. Завдяки цьому значно скорочується час створення відеороликів та підвищується ефективність виробництва цифрового контенту.

1.3 Аналіз сучасних сервісів генерації відео та синтезу мовлення

Стрімкий розвиток генеративного штучного інтелекту призвів до появи великої кількості програмних сервісів, здатних автоматично створювати відеоконтент на основі текстових описів користувача. Такі системи використовують сучасні нейромережеві моделі для генерації відеоряду, анімації сцен та синтезу мовлення, що дозволяє значно спростити процес створення мультимедійного контенту.

Одним із найбільш перспективних напрямків є використання текстово-відео моделей (Text-to-Video), які формують відеофрагменти без необхідності виконання реальної відеозйомки. Користувачеві достатньо сформулювати текстовий опис бажаної сцени, після чого система автоматично створює відповідний відеоряд.

Серед найбільш відомих платформ генерації відео сьогодні можна виділити Kling AI, Runway та Pika. Кожна з них має власні особливості, переваги та обмеження.

Kling AI є однією з найсучасніших генеративних відеомоделей, розроблених компанією Kuaishou Technology. Сервіс дозволяє створювати високоякісні відеоролики на основі текстових запитів та забезпечує реалістичне відтворення рухів об'єктів, освітлення та навколишнього середовища. Однією з ключових переваг Kling AI є висока деталізація сцен і якісна анімація персонажів, що робить дану модель особливо придатною для створення рекламного та туристичного контенту.

Для інтеграції можливостей Kling AI у межах розробленої системи використовується сервіс PiAPI, який забезпечує програмний доступ до функціональності відеогенерації через API. Завдяки цьому процес створення відеофрагментів може бути повністю автоматизований та інтегрований у workflow платформи n8n.

Під час генерації відео система передає до PiAPI сформовані LLM-агентами промпти, після чого створюється задача генерації (task). Після запуску процесу відбувається періодичне опитування статусу виконання. У разі успішного завершення система отримує посилання на готовий відеофайл для подальшої обробки та рендерингу.

Іншим популярним рішенням є платформа Runway. Даний сервіс активно використовується професійними контент-мейкерами та студіями цифрового виробництва. Runway забезпечує широкий набір інструментів для генерації та редагування відео, підтримує генерацію сцен за текстовими описами та має розвинені засоби постобробки. Основною перевагою платформи є висока якість створюваного контенту, однак доступ до багатьох функцій потребує платної підписки.

Ще одним відомим сервісом є Pika. Дана система орієнтована на швидке створення коротких відеороликів для соціальних мереж. Платформа має простий інтерфейс та дозволяє генерувати анімований контент на основі текстових описів. Серед переваг Pika можна виділити простоту використання та високу швидкість генерації, проте якість відео у складних сценах часто поступається результатам Kling AI або Runway.

Порівняльний аналіз показує, що для задач автоматизованого створення коротких туристичних відеороликів найбільш доцільним є використання Kling AI, оскільки дана модель забезпечує оптимальне співвідношення між якістю генерації, реалістичністю сцен та можливістю інтеграції через API.

Окрім генерації відеоряду важливим компонентом мультимедійного контенту є озвучення. Для створення голосового супроводу широко використовуються системи синтезу мовлення (Text-to-Speech, TTS), які

дозволяють перетворювати текст у природне мовлення.

Одним із лідерів даного напрямку є сервіс ElevenLabs. Платформа використовує сучасні нейромережеві моделі для генерації голосу та підтримує велику кількість мов. Особливістю сервісу є висока природність звучання, можливість відтворення емоційних інтонацій та налаштування параметрів голосу.

У розробленій системі використовується модель `eleven_multilingual_v2`, яка забезпечує генерацію голосового супроводу для коротких відеороликів тривалістю близько 15 секунд. Після створення тексту озвучення великими мовними моделями відповідний текст передається до сервісу ElevenLabs, який формує аудіофайл у форматі MP3. Надалі аудіофайл використовується під час фінального рендерингу відеоролика.

Поєднання технологій Kling AI та ElevenLabs дозволяє реалізувати повний цикл автоматизованого створення відеоконтенту, що включає генерацію візуальної частини, синтез мовлення та підготовку готового мультимедійного продукту до публікації. Саме тому зазначені сервіси були обрані як основа розробленої програмної системи.

1.4 Аналіз платформ автоматизації бізнес-процесів та оркестрації AI-сервісів

Зі збільшенням кількості цифрових сервісів та API особливої актуальності набувають платформи автоматизації бізнес-процесів. Такі системи дозволяють інтегрувати різні програмні рішення в єдиний автоматизований workflow, забезпечуючи обмін даними між сервісами без необхідності ручного виконання операцій. У сучасних AI-проектах платформи автоматизації виконують роль центрального оркестратора, який координує взаємодію мовних моделей, сервісів генерації контенту, хмарних сховищ та зовнішніх інформаційних систем.

Однією з найвідоміших платформ автоматизації є Zapier. Дане рішення орієнтоване насамперед на бізнес-користувачів та надає можливість створення

автоматизованих сценаріїв без програмування. Платформа підтримує інтеграцію з тисячами популярних сервісів, включаючи Google Workspace, Slack, Trello, Dropbox та інші. Основною перевагою Zapier є простота використання та велика кількість готових інтеграцій. Разом з тим система має обмежені можливості кастомізації складних процесів, а використання великої кількості операцій може суттєво збільшувати вартість експлуатації.

Іншою популярною платформою є Make (раніше Integromat). На відміну від Zapier, Make надає більш гнучкі засоби побудови бізнес-процесів та підтримує складні сценарії з великою кількістю розгалужень, умов та циклів. Візуальний редактор дозволяє створювати досить складні інтеграційні рішення без написання програмного коду. Перевагою платформи є наочне відображення логіки роботи процесів та широкий набір інтеграцій. Недоліком є обмежений доступ до внутрішньої логіки системи та залежність від хмарної інфраструктури постачальника послуг.

Для автоматизації технічних процесів також широко використовується Node-RED. Платформа була розроблена компанією IBM і орієнтована на створення інтеграційних рішень, систем Інтернету речей (IoT) та серверних автоматизованих процесів. Node-RED використовує графічний підхід до проєктування потоків даних та дозволяє інтегрувати різноманітні програмні компоненти. Основною перевагою системи є відкритий вихідний код та можливість локального розгортання. Водночас для реалізації складних бізнес-процесів часто необхідне програмування власних вузлів та додаткове налаштування системи.

Окремої уваги заслуговує платформа n8n, яка останніми роками набула значної популярності серед розробників систем штучного інтелекту та автоматизації бізнес-процесів. n8n є open-source платформою автоматизації, яка поєднує переваги візуального проєктування workflow із широкими можливостями програмної кастомізації. Система підтримує інтеграцію з великою кількістю зовнішніх сервісів через API та дозволяє створювати складні автоматизовані процеси із використанням умовних переходів, циклів,

обробки помилок та паралельного виконання завдань.

Важливою перевагою n8n є можливість локального розгортання на власній інфраструктурі. Це дозволяє забезпечити більший рівень контролю над даними, підвищити безпеку системи та уникнути залежності від сторонніх хмарних сервісів. Крім того, n8n підтримує виконання користувацького JavaScript-коду, що значно розширює можливості інтеграції та обробки даних.

Для задач, пов'язаних із генеративним штучним інтелектом, n8n має вбудовані інструменти інтеграції з OpenAI, Google Gemini, Anthropic Claude та іншими мовними моделями. Також платформа дозволяє легко підключати зовнішні сервіси генерації відео, системи синтезу мовлення, бази даних, файлові сховища та месенджери.

Порівняльний аналіз розглянутих платформ показує, що Zapier та Make є зручними інструментами для швидкого створення бізнес-процесів, однак їх використання може бути обмеженим у випадку складних AI-конвеєрів із великою кількістю взаємопов'язаних компонентів. Node-RED забезпечує високу гнучкість, але потребує більш глибоких технічних знань для підтримки та розвитку системи.

У межах даної роботи як центральний оркестратор було обрано платформу n8n. Основними причинами такого вибору стали відкритий вихідний код, можливість локального розгортання, підтримка інтеграції з AI-сервісами, наявність візуального редактора workflow та широкі можливості програмної кастомізації. Крім того, n8n забезпечує ефективну реалізацію складних сценаріїв обробки даних, включаючи генерацію сцен, створення промптів, запуск задач у Kling AI, синтез мовлення через ElevenLabs, взаємодію з Google Drive, Google Sheets та Telegram Bot API.

Таким чином, проведений аналіз платформ автоматизації показав, що n8n найбільш повно відповідає вимогам розроблюваної системи автоматизованої генерації відеоконтенту та забезпечує необхідний рівень гнучкості, масштабованості та інтеграційних можливостей.

1.5 Порівняльний аналіз існуючих рішень та постановка задачі дослідження

Проведений аналіз сучасних технологій генерації відеоконтенту показав, що на ринку існує значна кількість програмних рішень, які забезпечують окремі етапи створення мультимедійного контенту. До найбільш поширених систем належать Kling AI, Runway, Pika та інші генеративні платформи. Окрім цього, активно використовуються сервіси синтезу мовлення та платформи автоматизації бізнес-процесів.

Незважаючи на високий рівень розвитку зазначених технологій, більшість існуючих рішень орієнтовані на виконання окремих функцій і не забезпечують повністю автоматизований цикл створення відеоконтенту. У більшості випадків користувачеві необхідно самостійно виконувати генерацію сценарію, створення промптів, запуск відеогенерації, підготовку озвучення, рендеринг та публікацію результатів.

Для визначення доцільності розробки власного рішення було проведено порівняльний аналіз існуючих платформ.

Таблиця 1.1 – Порівняльна характеристика існуючих рішень

Характеристика	Kling AI	Runway	Pika	Запропонована система
Генерація відео за текстом	Так	Так	Так	Так
Автоматична генерація сценарію	Ні	Ні	Ні	Так
Автоматичне створення промптів	Ні	Ні	Ні	Так

Інтеграція з Telegram	Ні	Ні	Ні	Так
Автоматичне озвучення	Частково	Частково	Ні	Так
Логування результатів	Ні	Ні	Ні	Так
Інтеграція з Google Sheets	Ні	Ні	Ні	Так
Інтеграція з Google Drive	Ні	Ні	Ні	Так
Автоматичний рендеринг	Ні	Частково	Ні	Так
Повний цикл «Ідея → Готове відео»	Ні	Ні	Ні	Так

Аналіз таблиці показує, що існуючі сервіси забезпечують якісну генерацію окремих компонентів відеоконтенту, проте не реалізують повний автоматизований процес створення готового мультимедійного продукту. Це обумовлює необхідність створення інтегрованої системи, яка об'єднуватиме можливості великих мовних моделей, сервісів генерації відео, синтезу мовлення та платформ автоматизації бізнес-процесів.

У зв'язку з цим метою роботи є розробка веб-платформи для автоматизованої генерації коротких вертикальних відеороликів за текстовими запитами користувачів із використанням сучасних AI-сервісів та інтеграцією мультимедійних компонентів.

Об'єктом дослідження є процес автоматизованої генерації

відеоконтенту із використанням технологій штучного інтелекту.

Предметом дослідження є методи інтеграції великих мовних моделей, сервісів генерації відео, систем синтезу мовлення та платформ автоматизації бізнес-процесів у межах єдиної програмної системи.

Для досягнення поставленої мети необхідно вирішити такі завдання:

- провести аналіз сучасних технологій генерації мультимедійного контенту;
- дослідити можливості великих мовних моделей для автоматичного створення сценаріїв та текстового супроводу відеороликів;
- виконати аналіз сучасних сервісів генерації відео та синтезу мовлення;
- дослідити можливості платформ автоматизації бізнес-процесів та обґрунтувати вибір програмних засобів реалізації системи;
- спроектувати архітектуру веб-платформи генерації відеоконтенту;
- реалізувати автоматизований workflow генерації відеороликів на базі платформи n8n;
- забезпечити інтеграцію сервісів OpenAI GPT, Google Gemini, Kling AI, ElevenLabs, Google Drive, Google Sheets та Telegram Bot API;
- реалізувати механізми автоматичного створення сценарію, генерації відео, синтезу мовлення та рендерингу фінального ролика;
- провести тестування розробленої системи та оцінити її ефективність;
- сформулювати рекомендації щодо подальшого розвитку та вдосконалення програмного рішення.

Таким чином, результати проведеного аналізу підтверджують актуальність розробки інтегрованої системи автоматизованої генерації відеоконтенту та визначають основні напрямки подальшого дослідження, які будуть розглянуті у наступних розділах роботи.

1.6. Висновки до розділу 1

У першому розділі дипломної роботи було проведено аналіз сучасних технологій створення цифрового відеоконтенту та досліджено можливості

використання систем штучного інтелекту для автоматизації процесів його виробництва.

Розглянуто особливості створення відеоконтенту в умовах активного розвитку соціальних мереж та платформ коротких вертикальних відео, таких як TikTok, Instagram Reels та YouTube Shorts. Встановлено, що відеоконтент є одним із найбільш ефективних інструментів цифрового маркетингу, особливо у сфері туризму, де візуальне представлення послуг суттєво впливає на зацікавленість потенційних клієнтів.

На прикладі діяльності компанії TRÓPICO EXCÊNTRICO – UNIPESSOAL LDA (Adventure Tours Madeira) було визначено проблему значних часових витрат на підготовку та публікацію відеоматеріалів для соціальних мереж. Встановлено, що автоматизація процесів створення контенту є актуальним напрямом підвищення ефективності роботи туристичних компаній.

У ході дослідження проаналізовано сучасні великі мовні моделі та визначено їх можливості щодо автоматичної генерації сценаріїв, формування сцен, створення промптів для відеогенерації, підготовки текстів озвучення та генерації метаданих для соціальних мереж. Встановлено, що використання LLM-моделей дозволяє значно скоротити час підготовки контенту та забезпечити високу якість текстових матеріалів.

Також було проведено аналіз сучасних сервісів генерації відео та синтезу мовлення. Розглянуто можливості платформ Kling AI, Runway та Pika, а також сервісу ElevenLabs. За результатами порівняльного аналізу встановлено, що для задач автоматизованої генерації коротких рекламних відеороликів найбільш доцільним є використання Kling AI у поєднанні з ElevenLabs завдяки високій якості візуального контенту та природності синтезованого мовлення.

Окрему увагу приділено аналізу платформ автоматизації бізнес-процесів. Досліджено функціональні можливості Zapier, Make, Node-RED та n8n. За результатами порівняння обґрунтовано вибір платформи n8n як

центрального елемента майбутньої системи завдяки відкритому вихідному коду, можливості локального розгортання, підтримці AI-сервісів та високій гнучкості побудови workflow.

На основі проведеного порівняльного аналізу існуючих рішень встановлено, що більшість сучасних сервісів забезпечують автоматизацію лише окремих етапів створення відеоконтенту та не реалізують повний цикл перетворення користувацької ідеї у готовий відеоролик. Це підтвердило доцільність розробки інтегрованої програмної системи, яка поєднуватиме можливості великих мовних моделей, сервісів генерації відео, систем синтезу мовлення та платформ автоматизації бізнес-процесів.

У результаті проведеного дослідження сформульовано мету, об'єкт, предмет та основні завдання дипломної роботи, що стали основою для подальшого проєктування та реалізації веб-платформи автоматизованої генерації відеоконтенту, описаної в наступному розділі.

РОЗДІЛ 2 ПРОЄКТУВАННЯ ТА РЕАЛІЗАЦІЯ ВЕБ-ПЛАТФОРМИ ДЛЯ ГЕНЕРАЦІЇ ВІДЕОКОНТЕНТУ

2.1 Концепція програмної системи та загальна архітектура рішення

Основною метою розробки програмної системи є автоматизація процесу створення коротких вертикальних відеороликів на основі текстових запитів користувачів. Запропоноване рішення реалізує повний цикл генерації відеоконтенту із використанням сучасних технологій штучного інтелекту та забезпечує мінімальне залучення користувача до процесу створення мультимедійних матеріалів.

Концепція системи базується на використанні архітектури автоматизованого AI-конвеєра, у межах якого кожен окремий компонент відповідає за виконання конкретного функціонального завдання. Центральним елементом архітектури виступає платформа автоматизації бізнес-процесів n8n, яка забезпечує координацію роботи всіх підсистем та управління потоками даних між ними.

Розроблена система дозволяє користувачу сформулювати короткий текстовий запит через Telegram-бота. На основі цього запиту автоматично створюється сценарій майбутнього відеоролика, генеруються окремі сцени, формуються промпти для відеомоделі, створюється озвучення та виконується фінальний рендеринг готового відео.

Загальна архітектура системи представлена у вигляді послідовності взаємодії основних компонентів:

Telegram → n8n → LLM-агенти → Kling AI → ElevenLabs → Render Service → Google Drive → Telegram.

На першому етапі користувач надсилає текстовий запит до Telegram-бота. Наприклад, запит може містити опис туристичної локації, екскурсії або рекламної ідеї для соціальних мереж. Telegram Bot API передає отримані дані до платформи n8n, яка запускає відповідний workflow генерації контенту.

Після отримання запиту система виконує його реєстрацію в Google

Sheets. Таблиця використовується як централізоване сховище метаданих і містить інформацію про користувацький запит, сформовані сцени, статуси генерації, посилання на створені файли та результати обробки. Такий підхід забезпечує простежуваність процесу та можливість моніторингу виконання кожного етапу.

Наступним кроком є робота інтелектуальних агентів на базі великих мовних моделей. LLM-агент аналізує зміст запиту користувача та формує структуру майбутнього відеоролика. У межах реалізованого рішення відео складається з трьох сцен тривалістю приблизно п'ять секунд кожна. Для кожної сцени визначаються основний об'єкт, виконувана дія, навколишнє середовище, параметри руху камери, освітлення та переходи між сценами.

На основі сформованих сцен окремий агент створює спеціалізовані текстові промпти для генерації відео. Структура промптів адаптована до вимог відеомоделі Kling AI та містить детальний опис майбутньої сцени. Це дозволяє підвищити якість генерації та забезпечити узгодженість окремих фрагментів відео.

Для створення відеоконтенту використовується сервіс Kling AI через програмний інтерфейс PiAPI. Для кожної сцени формується окрема задача генерації. Після запуску задачі система періодично виконує перевірку її статусу до моменту отримання готового відеофрагмента. У результаті формується набір із трьох окремих кліпів, які надалі використовуються під час фінального рендерингу.

Паралельно із генерацією відео виконується створення тексту озвучення. Для цього LLM-агент генерує короткий Voiceover-текст, який передається до сервісу синтезу мовлення ElevenLabs. У роботі використовується модель eleven_multilingual_v2, що забезпечує створення природного голосового супроводу високої якості. Отриманий аудіофайл автоматично зберігається у Google Drive та використовується під час формування фінального відеоролика.

Після завершення генерації всіх відеофрагментів та аудіодоріжки

запускається модуль рендерингу. Спеціалізований сервіс об'єднує три відеокліпи та файл озвучення в єдиний відеоролик формату 9:16, оптимізований для публікації у соціальних мережах. Під час рендерингу забезпечується синхронізація відео та аудіо, а також формування кінцевого файлу у форматі MP4.

Готовий ролик автоматично завантажується до Google Drive, де для нього створюється публічне посилання. Отримане посилання записується до Google Sheets та надсилається користувачу через Telegram-бота. Таким чином користувач отримує готовий результат без необхідності виконання будь-яких додаткових дій.

Важливою особливістю архітектури є забезпечення відмовостійкості системи. Для цього у workflow реалізовано механізми повторних спроб виконання операцій, перевірку статусів генерації, обробку помилок зовнішніх API та логування всіх етапів роботи. У випадку виникнення помилки користувач отримує відповідне повідомлення, а система фіксує причину збою для подальшого аналізу.

Запропонована архітектура має модульну структуру та може бути легко масштабована шляхом додавання нових сервісів генерації відео, альтернативних мовних моделей або додаткових каналів взаємодії з користувачами. Це забезпечує гнучкість рішення та можливість його подальшого розвитку відповідно до потреб бізнесу.

Таким чином, розроблена програмна система реалізує повністю автоматизований процес створення коротких відеороликів і поєднує можливості великих мовних моделей, сервісів генерації відео, систем синтезу мовлення та інструментів автоматизації бізнес-процесів у межах єдиної інтегрованої архітектури.

2.2 Проєктування workflow генерації відео на базі платформи n8n

Для реалізації автоматизованого процесу створення відеоконтенту було обрано платформу n8n, яка виконує роль центрального оркестратора всіх

бізнес-процесів системи. Використання n8n дозволило об'єднати в межах єдиного workflow великі мовні моделі, сервіси генерації відео, системи синтезу мовлення, хмарні сховища та Telegram-бота.

Основним принципом побудови workflow є поетапне перетворення текстового запиту користувача на готовий відеоролик. Кожен етап представлений окремим вузлом (Node), який виконує конкретну функцію та передає результат наступному компоненту системи.

Отримання запиту від користувача

Робота системи починається з вузла Telegram Trigger, який відстежує нові повідомлення від користувачів через Telegram Bot API. Для взаємодії із системою було створено Telegram-бота, який приймає текстові запити та запускає процес генерації відеоконтенту.

Після надходження повідомлення n8n автоматично отримує текст запиту, ідентифікатор користувача та службову інформацію про повідомлення. Саме цей етап є точкою входу до всієї системи.

Логування початкового запиту

Після отримання повідомлення виконується блок Search Topic та Add Query, який відповідає за реєстрацію нового запиту в Google Sheets.

У таблицю автоматично записуються:

- ідентифікатор повідомлення;
- текст запиту користувача;
- дата та час створення;
- поточний статус виконання;
- службові параметри генерації.

Використання Google Sheets дозволяє централізовано зберігати історію роботи системи та забезпечує можливість подальшого аналізу результатів генерації.

Формування структури майбутнього відео

Після збереження запиту запускається блок Shots Prompt Agent, який використовує можливості великої мовної моделі для аналізу введеної

користувачем інформації.

На основі отриманої теми система автоматично формує структуру майбутнього відеоролика. У розробленій системі використовується підхід із поділом ролика на три сцени тривалістю приблизно п'ять секунд кожна.

Для кожної сцени генеруються такі параметри:

- main_subject — головний об'єкт сцени;
- action — опис дії;
- environment — середовище виконання дії;
- camera_movement — рух камери;
- lighting_time_mood — освітлення та атмосфера;
- duration_seconds — тривалість сцени;
- continuity_notes — рекомендації щодо послідовності;
- transition_to_next — опис переходу до наступної сцени.

Після генерації структура сцен проходить додаткову обробку за допомогою вузлів Remove, Code та Set Scenes, які виконують очищення та форматування отриманих даних.

Розподіл сцен для подальшої обробки

Оскільки кожна сцена повинна генеруватися окремо, використовується вузол Split Out.

Даний компонент розділяє сформований список сцен на окремі елементи, після чого кожна сцена обробляється незалежно від інших.

Такий підхід дозволяє реалізувати паралельну генерацію відеофрагментів та скоротити загальний час виконання workflow.

Генерація промптів для відеомоделі

Після розподілу сцен запускається вузол Video Prompt Generator.

Його завданням є автоматичне створення оптимізованих текстових промптів для відеомоделі Kling AI.

Промпти формуються за стандартизованою структурою:

Subject → Action → Environment → Camera → Lighting.

Наприклад, для туристичного ролика система може сформувати промпт

такого типу:

“A tourist walking along Madeira levada, surrounded by tropical forest, cinematic drone movement, soft morning sunlight, realistic style.”

Подібний формат дозволяє отримувати більш стабільні результати генерації відео та забезпечує узгодженість між окремими сценами.

Генерація відеофрагментів у Kling AI

Після формування промтів workflow переходить до вузла Generate Videos By Kling.

Даний компонент взаємодіє з сервісом PiAPI, який забезпечує доступ до функціоналу відеогенерації Kling AI через API.

Для кожної сцени створюється окрема задача генерації відео.

Після створення задачі система отримує унікальний ідентифікатор (Task ID), який використовується для подальшого контролю виконання.

Оскільки генерація відео потребує значного часу, у workflow використовується вузол очікування має URL готового відеофрагмента та передає його до наступних етапів обробки.

Формування метаданих контенту

Паралельно з генерацією відео виконується робота вузла Title and Description Generator.

На основі змісту ролика та сформованих сцен великі мовні моделі створюють:

- назву відео (Title);
- опис відео (Description);
- ключові слова для публікації;
- заклик до дії (Call To Action).

Згенеровані дані використовуються для подальшої публікації контенту в соціальних мережах та автоматично записуються до Google Sheets.

Загальна структура workflow

У спрощеному вигляді логіка роботи системи може бути представлена таким ланцюжком:

Telegram Trigger → Search Topic → Add Query → Shots Prompt Agent → Set Scenes → Split Out → Video Prompt Generator → Generate Videos By Kling → Get Videos → Check Status → Title and Description Generator.

Зазначений workflow формує основу всієї системи автоматизованої генерації відеоконтенту та забезпечує підготовку відеофрагментів для подальшого озвучення, рендерингу та публікації результатів.

2.3 Інтеграція зовнішніх сервісів

Однією з ключових особливостей розробленої системи є інтеграція декількох незалежних сервісів у межах єдиного автоматизованого workflow. Використання зовнішніх API дозволяє реалізувати повний цикл генерації відеоконтенту без необхідності створення власних моделей штучного інтелекту або сервісів зберігання даних. У межах проєкту було інтегровано Telegram Bot API, Google Sheets API, Google Drive API, сервіс генерації відео Kling AI через PiAPI, систему синтезу мовлення ElevenLabs та SFTP-сервер для передачі мультимедійних файлів між компонентами системи.

Інтеграція з Telegram Bot API

Telegram Bot API використовується як основний канал взаємодії між користувачем та системою генерації відеоконтенту. Для реалізації проєкту було створено Telegram-бота, який приймає текстові запити користувачів та запускає процес автоматичної генерації відеороликів.

Після надсилення повідомлення користувачем Telegram передає до п8n інформацію про текст запиту, ідентифікатор користувача та службові параметри повідомлення. Отримані дані використовуються як початкові параметри для запуску workflow.

Після завершення процесу генерації система автоматично надсилає користувачу повідомлення з посиланням на готовий відеоролик. Таким чином

Telegram виступає не лише джерелом вхідних даних, а й засобом доставки результатів роботи системи.

Перевагами використання Telegram Bot API є простота інтеграції, підтримка великої кількості користувачів, можливість швидкого обміну повідомленнями та відсутність необхідності створення окремого веб-інтерфейсу.

Інтеграція з Google Sheets

Google Sheets використовується як централізоване сховище метаданих системи. Таблиця виконує роль єдиної точки зберігання інформації про всі запуски генерації відеоконтенту.

Для кожного нового запиту автоматично створюється запис, який містить:

- ідентифікатор повідомлення;
- текст запиту користувача;
- згенеровані сцени;
- промпти для відеогенерації;
- назву відео;
- опис відео;
- текст озвучення;
- посилання на створені кліпи;
- посилання на фінальний відеоролик;
- статус виконання генерації.

Використання Google Sheets дозволяє спростити моніторинг роботи системи, аналіз статистики використання та пошук можливих помилок у workflow.

Таблиця 2.1 - Структура таблиці містить такі основні поля:

Поле	Опис
Id	Ідентифікатор повідомлення Telegram
UserQuery	Текст запиту користувача
ScenePrompt1	Дані першої сцени
ScenePrompt2	Дані другої сцени
ScenePrompt3	Дані третьої сцени
Title	Назва відео
Description	Опис відео
Voiceover	Текст озвучення
Clip1_URL	Посилання на перший кліп
Clip2_URL	Посилання на другий кліп
Clip3_URL	Посилання на третій кліп
Final_URL	Посилання на фінальне відео
VideoStatus	Поточний статус
Timestamp	Дата та час створення

Використання унікального ідентифікатора повідомлення Telegram дозволяє забезпечити ідемпотентність операцій та уникнути дублювання записів.

Інтеграція з Google Drive

Google Drive використовується як основне хмарне сховище для мультимедійних файлів, створених у процесі генерації контенту.

Після завершення генерації аудіофайлів та фінального відеоролика відповідні матеріали автоматично завантажуються до визначених директорій Google Drive. Для кожного файлу створюється публічне посилання з правами лише на перегляд.

Збереження файлів у хмарному сховищі забезпечує:

- централізоване зберігання результатів;
- доступність матеріалів з будь-якого пристрою;
- можливість подальшого використання контенту;
- резервне копіювання результатів генерації.

Після створення посилання інформація автоматично записується до Google Sheets та надсилається користувачу через Telegram.

Інтеграція з PiAPI та Kling AI

Основним компонентом генерації відеоконтенту є сервіс Kling AI. Оскільки прямий доступ до моделі обмежений, у системі використовується сервіс PiAPI, який надає програмний інтерфейс для взаємодії з можливостями Kling AI.

Для кожної сцени система виконує наступну послідовність дій:

1. Формування текстового промпта.
2. Надсилання запиту до PiAPI.
3. Створення задачі генерації відео.
4. Отримання ідентифікатора задачі.
5. Перевірка статусу виконання.
6. Завантаження готового відеофрагмента.

Під час роботи використовуються три основні статуси:

- processing – відео знаходиться у процесі генерації;
- completed – генерація завершена успішно;
- failed – виникла помилка генерації.

Після успішного завершення система отримує URL відеофайлу та використовує його під час формування фінального ролика.

Використання PiAPI дозволяє автоматизувати взаємодію з Kling AI та інтегрувати відеогенерацію безпосередньо в n8n workflow.

Інтеграція з ElevenLabs

Для створення голосового супроводу використовується сервіс синтезу мовлення ElevenLabs.

Після генерації сценарію великими мовними моделями формується короткий текст озвучення (Voiceover), який надсилається до API ElevenLabs.

У роботі використовується модель:

`eleven_multilingual_v2`

Дана модель підтримує багатомовну генерацію мовлення та забезпечує високу природність звучання.

Процес роботи складається з таких етапів:

1. Генерація тексту озвучення.
2. Передача тексту до ElevenLabs API.
3. Створення аудіофайлу у форматі MP3.
4. Завантаження аудіо до Google Drive.
5. Передача аудіофайлу до модуля рендерингу.

Використання ElevenLabs дозволяє відмовитися від ручного запису голосу та значно прискорює створення контенту.

Інтеграція з SFTP-сервером

Для передачі файлів між окремими компонентами системи використовується SFTP-сервер, розгорнутий на серверній інфраструктурі Digital Ocean.

SFTP використовується як проміжне сховище для:

- відеофрагментів, створених Kling AI;
- аудіофайлів, створених ElevenLabs;
- службових файлів рендерингу;
- фінального відеоролика.

Перед запуском рендерингу система автоматично завантажує до SFTP три відеокліпи та один аудіофайл. Після завершення обробки фінальний ролик завантажується назад до Google Drive.

Використання SFTP забезпечує стабільну передачу великих мультимедійних файлів та дозволяє розділити процес генерації й рендерингу між різними компонентами інфраструктури.

Таким чином інтеграція Telegram Bot API, Google Sheets, Google Drive,

PiAPI Kling, ElevenLabs та SFTP-сервера дозволила реалізувати єдину програмну екосистему для автоматизованої генерації відеоконтенту. Взаємодія зазначених сервісів забезпечує повний цикл обробки запиту користувача — від введення ідеї до отримання готового відеоролика.

2.4 Моделі даних та організація зберігання інформації

Ефективна робота системи автоматизованої генерації відеоконтенту потребує організованого зберігання як службової інформації, так і мультимедійних файлів, що створюються на різних етапах виконання workflow. Для забезпечення простежуваності процесу, контролю статусів виконання та подальшого аналізу результатів у розробленій системі реалізовано централізовану модель зберігання даних на базі Google Sheets, Google Drive та SFTP-сервера.

Основним сховищем метаданих виступає Google Sheets, яка використовується як єдина точка зберігання інформації про всі запити користувачів та результати їх обробки. Кожен запуск workflow створює окремий запис у таблиці, що дозволяє відстежувати повний життєвий цикл генерації відеоролика.

Ключовим елементом моделі даних є запис про окремий процес генерації відео. Для кожного запиту користувача система зберігає початкову ідею, згенеровані сцени, промпти для відеомоделі, посилання на створені файли та поточний статус виконання.

Структура основної таблиці наведена в таблиці 2.2.

Таблиця 2.2 – Структура даних системи

Поле	Призначення
Id	Унікальний ідентифікатор повідомлення Telegram
UserQuery	Початковий текстовий запит користувача

ScenePrompt1	Дані першої сцени у форматі JSON
ScenePrompt2	Дані другої сцени у форматі JSON
ScenePrompt3	Дані третьої сцени у форматі JSON
Title	Згенерована назва відеоролика
Description	Згенерований опис відеоролика
Voiceover	Текст озвучення
Clip1_URL	Посилання на перший відеофрагмент
Clip2_URL	Посилання на другий відеофрагмент
Clip3_URL	Посилання на третій відеофрагмент
Final_URL	Посилання на фінальний відеоролик
Status	Поточний статус виконання
CreatedAt	Дата та час створення запису
UpdatedAt	Дата та час останнього оновлення

Поле Id використовується як ключ ідентифікації запису та відповідає значенню Telegram message_id. Такий підхід забезпечує унікальність кожного запиту та дозволяє уникнути дублювання даних при повторному виконанні workflow.

Поле UserQuery містить текстовий опис ідеї, який користувач надсилає через Telegram-бота. Саме на основі цього запиту великі мовні моделі формують структуру майбутнього відеоролика.

Поля ScenePrompt1, ScenePrompt2 та ScenePrompt3 містять опис окремих сцен відео у структурованому форматі. Для кожної сцени зберігаються параметри головного об'єкта, виконуваної дії, навколишнього середовища, освітлення, руху камери та додаткових характеристик.

Після завершення генерації система автоматично формує поля Title та

Description, які використовуються для подальшої публікації контенту у соціальних мережах.

Поля Clip1_URL, Clip2_URL та Clip3_URL містять посилання на окремі відеофрагменти, отримані від сервісу Kling AI. Після виконання рендерингу формується поле Final_URL, яке містить посилання на готовий відеоролик.

Для контролю процесу використовується поле Status, яке відображає поточний стан обробки запиту. У системі реалізовано декілька основних статусів:

- New – новий запит отримано;
- Processing – виконується генерація контенту;
- Rendering – триває формування фінального відео;
- Completed – генерацію завершено успішно;
- Failed – виникла помилка під час виконання workflow.

Використання системи статусів дозволяє контролювати роботу всіх компонентів програмного комплексу та оперативно виявляти проблемні етапи обробки.

Для зберігання мультимедійних файлів використовується комбінація Google Drive та SFTP-сервера. Такий підхід дозволяє розділити постійне та тимчасове зберігання даних.

Google Drive використовується як основне довготривале сховище результатів генерації. Після завершення обробки до Google Drive автоматично завантажуються:

- аудіофайли, створені сервісом ElevenLabs;
- фінальні відеоролики у форматі MP4;
- додаткові службові файли за необхідності.

Для кожного завантаженого файлу автоматично створюється публічне посилання з правами доступу лише на перегляд. Отримані URL-адреси записуються до Google Sheets та використовуються для передачі результатів користувачу.

SFTP-сервер використовується як проміжне сховище файлів під час

виконання рендерингу. До нього тимчасово завантажуються окремі відеофрагменти, аудіодоріжки та службові файли, необхідні для роботи сервісу обробки відео. Після завершення формування фінального ролика тимчасові файли можуть бути автоматично видалені.

Для забезпечення ефективного використання ресурсів у системі реалізовано політики зберігання даних.

Проміжні відеофрагменти та допоміжні файли, розміщені на SFTP-сервері, зберігаються протягом обмеженого періоду часу. Після завершення встановленого терміну вони автоматично видаляються для звільнення дискового простору.

Фінальні відеоролики та аудіофайли зберігаються у Google Drive протягом тривалого часу та можуть бути використані повторно. Доступ до файлів здійснюється виключно за прямим посиланням без індексації пошуковими системами.

Для підвищення надійності роботи системи вся службова інформація про процес генерації дублюється у Google Sheets. Це дозволяє відновити інформацію про виконані операції навіть у випадку тимчасової недоступності окремих компонентів інфраструктури.

Таким чином, запропонована модель даних та організація зберігання інформації забезпечують централізований облік результатів роботи системи, простежуваність усіх етапів генерації відеоконтенту, надійне зберігання мультимедійних файлів та можливість масштабування програмного рішення в майбутньому.

2.5 Обробка помилок та забезпечення відмовостійкості

Під час розробки систем автоматизованої генерації відеоконтенту особливу увагу необхідно приділяти забезпеченню стабільності роботи та коректній обробці помилок. Розроблена система взаємодіє з великою кількістю зовнішніх сервісів, зокрема Telegram Bot API, Google Sheets, Google Drive, PiAPI Kling AI, ElevenLabs та SFTP-сервером. Кожен із зазначених

компонентів може тимчасово бути недоступним або повертати помилки, що потенційно впливає на виконання workflow.

Для підвищення надійності роботи програмного рішення в архітектурі системи реалізовано комплекс механізмів контролю стану задач, повторного виконання операцій, обробки виняткових ситуацій та відновлення процесів після збоїв.

Одним із ключових елементів системи є контроль статусів задач генерації відео. Після відправлення промпта до сервісу Kling AI через PiAPI створюється окрема задача генерації відеофрагмента. Оскільки процес створення відео може тривати декілька хвилин, система не отримує результат миттєво, а періодично виконує перевірку поточного статусу задачі.

У межах реалізованого рішення використовуються такі основні статуси:

- processing – генерація відео виконується та ще не завершена;
- completed – відеофайл успішно створений і готовий до подальшої обробки;
- failed – під час генерації виникла помилка або сервіс не зміг сформувати результат.

У випадку отримання статусу processing workflow переходить у режим очікування та через визначений проміжок часу повторно виконує перевірку стану задачі. Такий підхід дозволяє уникнути надмірного навантаження на API та забезпечує коректне очікування завершення генерації.

Якщо сервіс повертає статус completed, система отримує посилання на готовий відеофрагмент та автоматично переходить до наступного етапу обробки. Після цього запускаються модулі генерації озвучення, рендерингу та збереження результатів.

У випадку виникнення статусу failed система фіксує помилку у журналі подій та оновлює інформацію в Google Sheets. Додатково користувач отримує повідомлення про неможливість завершення генерації. Такий механізм дозволяє забезпечити прозорість роботи системи та своєчасно інформувати користувачів про проблеми.

Для зменшення впливу тимчасових збоїв реалізовано механізм повторних спроб виконання операцій (retries). Повторні спроби використовуються під час взаємодії з зовнішніми API, завантаження файлів до Google Drive, передачі даних через SFTP та виконання HTTP-запитів.

У разі виникнення тимчасової помилки система автоматично виконує повторний запит через визначений інтервал часу. Такий підхід дозволяє успішно завершувати значну кількість операцій без втручання користувача, особливо у випадках короткочасної недоступності сервісів або перевищення лімітів запитів.

Для уникнення нескінченного очікування відповідей від зовнішніх систем у workflow використовуються механізми обмеження часу виконання (timeout). Для кожного етапу визначено максимально допустимий час очікування відповіді. Якщо протягом встановленого інтервалу відповідь не отримана, система завершує поточну операцію та переходить до сценарію обробки помилки.

Особливо актуальним використання timeout є під час генерації відео в Kling AI та передачі великих мультимедійних файлів через мережу. Це дозволяє уникнути блокування всього workflow через збій одного із зовнішніх сервісів.

Важливою складовою забезпечення надійності є механізм відновлення після помилок (recovery). Завдяки використанню Google Sheets як централізованого сховища метаданих система зберігає інформацію про кожен етап виконання. У випадку аварійного завершення workflow адміністратор може визначити останній успішно виконаний крок та повторно запустити процес без необхідності виконання всіх операцій з початку.

Крім того, всі ключові події автоматично фіксуються у журналі виконання. До журналу записуються статуси генерації відео, результати взаємодії з API, повідомлення про помилки та посилання на створені файли. Це дозволяє виконувати аудит роботи системи та швидко виявляти причини можливих збоїв.

Практичне тестування показало, що найбільш поширеними причинами помилок є перевищення навантаження на сервіс Kling AI, тимчасова недоступність зовнішніх API та обмеження модерації контенту. Для мінімізації впливу таких ситуацій у системі реалізовано механізми повторних спроб, додаткових затримок між запитами та автоматичного інформування користувачів.

Таким чином, реалізовані механізми контролю статусів, повторних спроб виконання операцій, обмеження часу очікування та відновлення після збоїв забезпечують високий рівень відмовостійкості системи та дозволяють підтримувати стабільну роботу процесу автоматизованої генерації відеоконтенту навіть за умов нестабільної роботи зовнішніх сервісів.

2.6. Висновки до розділу 2

У другому розділі дипломної роботи було виконано проєктування та реалізацію веб-платформи для автоматизованої генерації відеоконтенту із використанням сучасних технологій штучного інтелекту та платформ автоматизації бізнес-процесів.

На основі результатів аналізу, проведеного в першому розділі, було розроблено загальну архітектуру програмної системи, яка забезпечує повний цикл створення коротких вертикальних відеороликів — від отримання текстового запиту користувача до формування готового відеофайлу та його передачі через Telegram-бота.

У процесі проєктування було визначено основні компоненти системи та їх взаємодію. Центральним елементом архітектури обрано платформу n8n, яка виконує функції оркестрації бізнес-процесів та забезпечує інтеграцію всіх зовнішніх сервісів у межах єдиного workflow.

У межах роботи було реалізовано автоматизований workflow генерації відеоконтенту, який включає отримання користувацького запиту, формування структури сцен за допомогою великих мовних моделей, генерацію промптів для відеомоделі Kling AI, створення озвучення через сервіс ElevenLabs,

рендеринг фінального відеоролика та його збереження у хмарному сховищі.

Також було реалізовано інтеграцію із зовнішніми сервісами, зокрема Telegram Bot API, Google Sheets, Google Drive, PiAPI Kling AI, ElevenLabs та SFTP-сервером. Використання зазначених компонентів дозволило побудувати єдину інформаційну систему, що забезпечує автоматизоване виконання всіх етапів створення мультимедійного контенту.

Окрему увагу приділено проектуванню моделей даних та організації зберігання інформації. Для централізованого обліку результатів роботи використано Google Sheets, а для зберігання мультимедійних файлів — Google Drive та SFTP-сервер. Такий підхід забезпечує простежуваність процесів генерації та спрощує адміністрування системи.

Крім того, у роботі реалізовано механізми обробки помилок та забезпечення відмовостійкості. Передбачено контроль статусів виконання задач, автоматичне повторне виконання операцій, обмеження часу очікування та можливість відновлення роботи після виникнення збоїв. Це дозволяє забезпечити стабільну роботу системи навіть у випадку тимчасової недоступності зовнішніх сервісів.

Таким чином, у другому розділі було успішно вирішено задачі проектування та програмної реалізації веб-платформи автоматизованої генерації відеоконтенту. Створена система забезпечує повний цикл обробки користувацького запиту та створює основу для проведення подальшого тестування й оцінювання ефективності розробленого рішення, результати яких наведено в наступному розділі.

РОЗДІЛ 3 ТЕСТУВАННЯ, ОЦІНКА ЕФЕКТИВНОСТІ ТА АНАЛІЗ РЕЗУЛЬТАТІВ РОБОТИ СИСТЕМИ

3.1 Методика тестування програмної системи та сценарії використання

Після завершення розробки програмної системи автоматизованої генерації відеоконтенту було проведено комплексне тестування з метою оцінки працездатності реалізованого рішення, визначення часу виконання окремих етапів обробки, аналізу стабільності роботи інтегрованих сервісів та оцінки якості отриманих результатів.

Основною метою тестування було підтвердження можливості автоматичного створення коротких вертикальних відеороликів на основі текстових запитів користувачів без необхідності ручного втручання в процес генерації. Особлива увага приділялася перевірці коректності роботи workflow, стабільності взаємодії із зовнішніми сервісами та швидкості отримання кінцевого результату.

Для проведення експериментів було підготовлено набір тестових запитів, що охоплювали різні тематичні категорії туристичного контенту. Серед них були описи природних локацій острова Мадейра, туристичних маршрутів, екскурсійних програм, активного відпочинку та рекламних відеороликів для соціальних мереж.

Кожен тестовий запуск проходив повний цикл обробки:

1. Отримання текстового запиту через Telegram-бот.
2. Формування структури сцен за допомогою великих мовних моделей.
3. Генерація промптів для відеомоделі Kling AI.
4. Створення трьох відеофрагментів.
5. Генерація озвучення через ElevenLabs.
6. Рендеринг фінального відеоролика.
7. Завантаження результату до Google Drive.

користувачу через Telegram.

Для оцінювання продуктивності використовувався показник ТАТ (Turnaround Time), який характеризує повний час виконання процесу від моменту отримання запиту до передачі користувачу готового відеоролика.

У межах тестування було виконано 10 повних запусків системи. Для кожного запуску фіксувалися часові витрати на окремі етапи обробки.

Перед початком експерименту були встановлені такі параметри:

- кількість сцен — 3;
- тривалість кожної сцени — близько 5 секунд;
- формат відео — 9:16;
- генерація відео — Kling AI;
- озвучення — ElevenLabs;
- фінальний формат файлу — MP4.

Час виконання окремих допоміжних етапів був практично постійним та становив:

- Intake → Scenes — 7 секунд;
- Scenes → Prompts — 3 секунди;
- Clips → TTS — 5 секунд;
- TTS → Render — 13 секунд;
- Render → Link — 28 секунд.

Найбільший вплив на загальний час виконання має процес генерації відеофрагментів у Kling AI та фінальний рендеринг ролика.

Результати вимірювання часу виконання наведені в таблиці 3.1.

Аналіз отриманих результатів показує, що загальний час створення відеоролика знаходиться в межах від 10 до 15 хвилин. Найбільшу частину цього часу займає генерація відеофрагментів сервісом Kling AI, що є характерним для сучасних моделей генерації відео високої якості.

Таблиця 3.1 – Час виконання окремих етапів та сумарний ТАТ

Запуск	Intake→Scenes	Scene s→Prompts	Prompts→Clips	Clips→TTS	TTS→Render	Render→Link	Render	Загальний ТАТ
1	0:07	0:03	8:40	0:05	0:13	0:28	2:20	11:56
2	0:07	0:03	9:15	0:05	0:13	0:28	1:40	11:51
3	0:07	0:03	8:05	0:05	0:13	0:28	3:10	12:11
4	0:07	0:03	9:50	0:05	0:13	0:28	2:55	13:41
5	0:07	0:03	7:55	0:05	0:13	0:28	1:15	10:06
6	0:07	0:03	8:30	0:05	0:13	0:28	3:40	13:06
7	0:07	0:03	9:05	0:05	0:13	0:28	2:05	12:06
8	0:07	0:03	8:20	0:05	0:13	0:28	1:50	11:06
9	0:07	0:03	10:00	0:05	0:13	0:28	3:30	14:26
10	0:07	0:03	7:35	0:05	0:13	0:28	1:35	10:06

Середній час виконання одного повного циклу генерації становив приблизно 12 хвилин. З огляду на повну автоматизацію процесу та відсутність ручної обробки отриманий результат можна вважати задовільним для практичного використання в туристичній сфері та контент-маркетингу.

Проведене тестування підтвердило працездатність запропонованої архітектури та можливість автоматизованого створення коротких відеороликів із використанням технологій штучного інтелекту. Отримані результати стали основою для подальшого аналізу стабільності роботи системи, оцінки економічної ефективності та визначення якості згенерованого контенту.

3.2 Аналіз продуктивності системи та оцінка стабільності роботи

Окрім оцінювання часу виконання окремих етапів генерації відеоконтенту важливим критерієм ефективності системи є стабільність її роботи. Для визначення надійності розробленого програмного рішення було

проведено аналіз успішності виконання workflow під час серії тестових запусків.

Основними показниками оцінювання виступали:

- успішність завершення генерації відео;
- кількість повторних спроб виконання операцій;
- частота виникнення помилок;
- причини відмов зовнішніх сервісів;
- здатність системи автоматично відновлювати виконання після збоїв.

Під час тестування було виконано 10 повних запусків програмної системи. Для кожного запуску фіксувалися статуси генерації відео в сервісі Kling AI, кількість повторних спроб (Retry), результат роботи модуля рендерингу та можливі причини виникнення помилок.

Результати експерименту наведено в таблиці 3.2.

Таблиця 3.2 – Аналіз успішності роботи системи

Запуск	Статус Kling AI	Кількість повторів	Статус рендеринг у	Причина помилки	Примітка
1	completed	1	completed	–	Генерацію завершено успішно
2	completed	0	completed	–	Генерацію завершено успішно
3	completed	1	completed	–	Генерацію завершено успішно
4	failed	1	–	Sensitive data in prompt	Потрібне коригування промпта
5	completed	1	completed	–	Генерацію завершено успішно

6	completed	2	completed	–	Успішно після повторної спроби
7	completed	1	completed	–	Генерацію завершено успішно
8	completed	0	completed	–	Генерацію завершено успішно
9	failed	1	–	Rate limit servісу	Потрібна додаткова затримка
10	completed	1	completed	–	Генерацію завершено успішно

За результатами тестування успішно завершилися 8 із 10 запусків. Таким чином показник успішності роботи системи (Success Rate) становить:

$$SR = (8 / 10) \times 100\% = 80\%.$$

Отримане значення свідчить про достатньо високий рівень стабільності роботи програмного рішення з урахуванням використання великої кількості зовнішніх AI-сервісів та хмарних платформ.

Аналіз результатів показує, що більшість запусків завершилися успішно навіть за наявності окремих тимчасових збоїв. Це стало можливим завдяки використанню механізму повторних спроб виконання операцій. Середня кількість повторів становила приблизно один повторний запит на запуск, що свідчить про ефективність реалізованої системи відмовостійкості.

Однією з причин невдалого завершення генерації стала помилка модерації контенту в сервісі Kling AI. У четвертому запуску сформований промпт містив елементи, які були розцінені системою модерації як потенційно небажаний контент. У результаті сервіс відмовив у генерації відеофрагмента та повернув статус failed.

Інша помилка була зафіксована під час дев'ятого запуску. Причиною

збою стало перевищення допустимої кількості запитів до сервісу генерації відео. У відповідь сервіс повернув повідомлення про тимчасове перевантаження (Rate Limit Exceeded), що призвело до завершення процесу без отримання результату.

Варто зазначити, що обидві зафіксовані помилки були пов'язані не з внутрішньою логікою розробленої системи, а з особливостями роботи зовнішніх сервісів. Це підтверджує коректність реалізованого workflow та правильність побудови програмної архітектури.

Окремий аналіз показав ефективність механізмів повторного виконання операцій. У більшості випадків додатковий запит дозволяв успішно завершити генерацію без необхідності втручання користувача. Найбільша кількість повторів була зафіксована під час шостого запуску, де система виконала дві додаткові перевірки статусу задачі перед отриманням готового результату.

Реалізовані механізми контролю статусів, обмеження часу очікування та автоматичного повторного виконання операцій забезпечили стабільну роботу системи навіть за умов нестабільної відповіді зовнішніх API. Це підтверджує доцільність використання платформи n8n як центрального оркестратора процесів та правильність обраного підходу до побудови архітектури рішення.

Таким чином проведений аналіз продуктивності показав, що розроблена система забезпечує високий рівень надійності та може використовуватися для практичної автоматизації процесів створення відеоконтенту. Отримані результати також дозволили визначити напрямки подальшого вдосконалення, зокрема впровадження адаптивних алгоритмів повторних спроб та використання альтернативних сервісів генерації відео у випадку недоступності основного провайдера.

3.3 Економічна ефективність розробленої системи

Важливим критерієм оцінювання ефективності будь-якої інформаційної системи є економічна доцільність її використання. Для розробленої системи автоматизованої генерації відеоконтенту було проведено аналіз собівартості

створення одного відеоролика з урахуванням використання зовнішніх AI-сервісів.

Основними складовими витрат є:

- використання сервісу генерації відео Kling AI;
- використання великих мовних моделей для генерації сценаріїв та промптів;
- використання сервісу синтезу мовлення ElevenLabs;
- витрати на зберігання мультимедійних файлів;
- експлуатаційні витрати серверної інфраструктури.

Під час проведення тестування використовувалися безкоштовні або тестові тарифні плани окремих сервісів. Найбільшу частку витрат формувала генерація відео в Kling AI, тоді як витрати на генерацію текстового контенту та озвучення були мінімальними.

Для оцінювання економічної ефективності було виконано аналіз вартості кожного із десяти тестових запусків системи.

Результати наведено в таблиці 3.3.

Таблиця 3.3 – Собівартість генерації одного відеоролика

Запуск	LLM (USD)	Kling AI (USD)	TTS (USD)	Зберіган ня (USD)	Загальна вартість (USD)
1	0,00	1,38	0,00	0,00	1,38
2	0,00	1,38	0,00	0,00	1,38
3	0,00	1,38	0,00	0,00	1,38
4	0,00	0,92	0,00	0,00	0,92
5	0,00	1,38	0,00	0,00	1,38

6	0,00	1,38	0,00	0,00	1,38
7	0,00	1,38	0,00	0,00	1,38
8	0,00	1,38	0,00	0,00	1,38
9	0,00	0,46	0,00	0,00	0,46
10	0,00	1,38	0,00	0,00	1,38

Аналіз отриманих результатів показує, що середня вартість створення одного відеоролика становить приблизно 1,24 долара США. Найбільшу частку витрат формує сервіс Kling AI, який використовується для генерації відеофрагментів. Саме цей компонент визначає економічну складову роботи всієї системи.

Варто зазначити, що окремі тестові запуски мають нижчу собівартість через передчасне завершення генерації або неповне використання ресурсів відеосервісу. Зокрема, у четвертому та дев'ятому тестових сценаріях через виникнення помилок фактичні витрати виявилися меншими за середній показник.

Під час використання безкоштовного тарифного плану ElevenLabs витрати на синтез мовлення були відсутні. Однак після вичерпання безкоштовного ліміту орієнтовна додаткова вартість озвучення одного ролика становить близько 0,06 долара США. Навіть за таких умов загальна собівартість одного відеоролика залишається невисокою порівняно з традиційним процесом створення відеоконтенту.

Для порівняння, створення аналогічного рекламного відеоролика традиційними методами потребує залучення сценариста, монтажера та спеціаліста з озвучення, що може коштувати від кількох десятків до кількох сотень доларів залежно від складності проєкту. Використання автоматизованої системи дозволяє скоротити ці витрати у десятки разів.

Додатковою перевагою є значне зменшення часу створення контенту. Якщо традиційна підготовка короткого рекламного відео може тривати декілька годин або навіть днів, то запропонована система забезпечує отримання готового результату в середньому за 10–15 хвилин.

Таким чином проведений аналіз підтверджує економічну доцільність використання розробленої системи для автоматизованої генерації відеоконтенту. Низька собівартість одного ролика у поєднанні з високим рівнем автоматизації та швидкістю отримання результату роблять запропоноване рішення перспективним для використання в туристичній сфері, цифровому маркетингу та контент-індустрії.

3.4 Оцінювання якості відеоконтенту

Окрім технічних показників продуктивності та економічної ефективності важливим критерієм оцінювання розробленої системи є якість згенерованого відеоконтенту. Саме якість отриманих відеороликів визначає можливість практичного використання системи для створення рекламних, інформаційних та маркетингових матеріалів.

Для оцінювання якості результатів було проведено експертне опитування за участю восьми респондентів. До складу групи увійшли користувачі, які мають досвід перегляду контенту в соціальних мережах та знайомі з сучасними вимогами до коротких вертикальних відеороликів.

Кожен респондент переглянув результати генерації відеоконтенту та виконав оцінювання за трьома основними критеріями:

- Композиція – оцінка візуальної цілісності сцен, якості кадрування та загального сприйняття відеоряду;
- Відповідність ідеї – ступінь відповідності створеного контенту початковому текстовому запиту користувача;
- Плавність – якість переходів між сценами, природність рухів та загальна динаміка відео.

Оцінювання здійснювалося за п'ятибальною шкалою, де:

- 1 бал – дуже низька якість;
- 2 бали – низька якість;
- 3 бали – задовільна якість;
- 4 бали – добра якість;
- 5 балів – відмінна якість.

Результати експертного оцінювання наведені в таблиці 3.4.

Таблиця 3.4 – Оцінка якості згенерованого відеоконтенту

Запуск	Композиція	Відповідність ідеї	Плавність	Середній бал
1	4,3	4,4	4,0	4,2
2	4,5	4,6	4,3	4,5
3	4,2	4,1	3,9	4,1
4	—	—	—	—
5	4,6	4,7	4,4	4,6
6	4,1	4,0	3,8	4,0
7	4,4	4,2	4,1	4,2
8	4,3	4,5	4,2	4,3
9	—	—	—	—
10	4,5	4,4	4,3	4,4

Варто зазначити, що четвертий та дев'ятий запуски не були включені до оцінювання якості через невдале завершення процесу генерації відео. Тому аналіз проводився лише для успішно створених відеороликів.

На основі отриманих результатів було визначено середні значення за кожним критерієм:

- середня оцінка композиції – 4,36 бала;
- середня оцінка відповідності ідеї – 4,36 бала;
- середня оцінка плавності – 4,13 бала;
- загальна середня оцінка якості – 4,28 бала.

Отримані результати свідчать про достатньо високий рівень якості згенерованого контенту. Найвищі оцінки були отримані за критерієм відповідності початковій ідеї користувача. Це підтверджує ефективність використання великих мовних моделей для генерації сценаріїв та промптів, а також правильність побудови логіки формування сцен.

Дещо нижчі оцінки спостерігаються за критерієм плавності відеоряду. Основною причиною цього є незалежна генерація окремих сцен сервісом Kling AI. Незважаючи на використання механізмів забезпечення сюжетної послідовності, окремі переходи між сценами можуть виглядати менш природними порівняно з відео, створеними професійними монтажерами.

Разом з тим середній бал понад чотири пункти свідчить про те, що отриманий контент може ефективно використовуватися для публікації в соціальних мережах, рекламних кампаніях та просування туристичних послуг.

Особливо важливим є той факт, що висока якість контенту була досягнута без участі людини у процесі монтажу та створення озвучення. Усі етапи генерації виконувалися автоматично за допомогою великих мовних моделей, сервісів генерації відео та систем синтезу мовлення.

Таким чином результати експертного оцінювання підтверджують, що розроблена система здатна створювати якісний відеоконтент, який відповідає сучасним вимогам цифрового маркетингу та може використовуватися для автоматизованого виробництва коротких вертикальних відеороликів у туристичній сфері та інших галузях діяльності.

3.5 Аналіз отриманих результатів

Після проведення тестування розробленої системи було виконано комплексний аналіз отриманих результатів, що охоплює показники

продуктивності, надійності, економічної ефективності та якості згенерованого відеоконтенту. Метою аналізу є оцінка відповідності створеної системи поставленим вимогам та визначення перспектив її практичного використання.

Першим важливим показником є загальний час генерації відеоролика (TAT). Відповідно до результатів тестування, наведених у таблиці 3.1, середній час створення одного відеоролика становив близько 12 хвилин. Мінімальне значення було зафіксоване під час п'ятого та десятого запусків і становило 10 хвилин 6 секунд, тоді як максимальний показник склав 14 хвилин 26 секунд під час дев'ятого запуску.

Аналіз структури часових витрат показав, що найбільшу частину часу займає етап генерації відеофрагментів за допомогою сервісу Kling AI. У середньому цей процес триває від 8 до 10 хвилин, що становить понад 70 % від загального часу роботи системи. Інші етапи, зокрема генерація сцен, формування промптів, створення озвучення та передача файлів, виконуються значно швидше і практично не впливають на підсумковий час отримання результату.

Отримані показники свідчать про те, що навіть за використання сучасних генеративних моделей відео автоматизоване створення контенту відбувається значно швидше, ніж традиційний процес підготовки аналогічного ролика за участю сценариста, монтажера та спеціаліста з озвучення. Це підтверджує доцільність використання технологій штучного інтелекту для автоматизації виробництва коротких відеоматеріалів.

Другим важливим показником є стабільність роботи системи. Згідно з результатами тестування, успішно завершилися 8 із 10 запусків, що відповідає показнику Success Rate на рівні 80 %. При цьому всі зафіксовані помилки були пов'язані не з внутрішньою логікою системи, а з особливостями роботи зовнішніх сервісів.

У четвертому запуску виникла помилка модерації контенту в Kling AI через використання промпта, який містив елементи, що не відповідали політиці сервісу. У дев'ятому запуску було зафіксовано перевищення ліміту

запитів до API відеогенерації. Незважаючи на це, реалізовані механізми повторних спроб та контролю статусів дозволили забезпечити стабільну роботу системи в більшості випадків.

Середня кількість повторних спроб становила приблизно один повторний запит на запуск. Це свідчить про ефективність реалізованих механізмів відмовостійкості та правильність використання інструментів оркестрації в середовищі p8n.

Окремий аналіз було проведено щодо економічної ефективності системи. Результати показали, що середня собівартість створення одного відеоролика становить близько 1,24 долара США. Основну частину витрат формує сервіс Kling AI, який використовується для генерації відеофрагментів. Витрати на генерацію тексту та озвучення в межах проведеного експерименту були мінімальними завдяки використанню тестових тарифних планів.

Порівняння із традиційними підходами до створення контенту показує суттєву економію ресурсів. У випадку ручного виробництва аналогічного відеоролика необхідно залучати декількох спеціалістів та витратити значно більше часу на підготовку матеріалів. Таким чином використання автоматизованої системи дозволяє суттєво зменшити витрати та підвищити продуктивність контент-виробництва.

Важливим аспектом оцінювання стала якість створеного відеоконтенту. Результати експертного опитування продемонстрували високі середні оцінки за всіма критеріями. Загальний середній бал склав 4,28 із 5 можливих. Найкращі результати були отримані за критерієм відповідності початковій ідеї користувача, що свідчить про ефективність використання великих мовних моделей під час формування сценарію та промптів.

Середня оцінка композиції склала 4,36 бала, що підтверджує здатність відеомоделі формувати візуально привабливі сцени. Показник плавності переходів між сценами становив 4,13 бала. Незважаючи на деякі недоліки, пов'язані з незалежною генерацією окремих фрагментів, результати залишаються достатньо високими для практичного використання контенту у

соціальних мережах.

Проведене дослідження також дозволило визначити сильні сторони запропонованого рішення. До них належать повна автоматизація процесу створення відео, інтеграція сучасних AI-сервісів у межах єдиної платформи, можливість масштабування системи та низька собівартість отриманого результату.

Серед напрямків подальшого вдосконалення можна виділити покращення механізмів узгодження сцен, використання альтернативних моделей генерації відео, впровадження додаткових алгоритмів оптимізації промптів та автоматичний вибір найбільш ефективних параметрів генерації залежно від тематики контенту.

Отримані результати підтверджують досягнення поставленої мети дослідження та демонструють можливість практичного застосування розробленої системи для автоматизованого створення коротких вертикальних відеороликів. Реалізоване рішення може бути використане туристичними компаніями, маркетинговими агентствами, контент-студіями та іншими організаціями, які потребують швидкого виробництва якісного відеоконтенту із мінімальними витратами часу та ресурсів.

3.6. Висновки до розділу 3

У третьому розділі дипломної роботи було проведено комплексне тестування розробленої веб-платформи автоматизованої генерації відеоконтенту та виконано оцінювання її ефективності за основними технічними, економічними та якісними показниками.

У ході дослідження розроблено методику тестування системи та проведено серію експериментальних запусків із використанням різних користувацьких запитів. Отримані результати дозволили визначити часові характеристики роботи системи та оцінити швидкість виконання окремих етапів автоматизованого workflow.

Аналіз показників продуктивності продемонстрував, що середній час

створення одного відеоролика становить близько 12 хвилин. Найбільшу частину часу займає генерація відеофрагментів за допомогою сервісу Kling AI, тоді як інші етапи обробки виконуються практично миттєво. Отримані результати підтвердили можливість автоматизованого створення готового відеоконтенту без участі користувача після надсилання початкового запиту.

Під час дослідження також було оцінено стабільність роботи системи. За результатами десяти тестових запусків показник успішності виконання становив 80 %. Виявлені помилки були пов'язані переважно з обмеженнями зовнішніх сервісів генерації відео та не впливали на коректність роботи реалізованого програмного рішення. Реалізовані механізми повторних спроб, контролю статусів та обробки помилок забезпечили достатній рівень відмовостійкості системи.

Окрему увагу приділено аналізу економічної ефективності розробленого рішення. Встановлено, що середня собівартість створення одного відеоролика становить приблизно 1,24 долара США. Це значно менше порівняно з традиційними підходами до виробництва відеоконтенту, які потребують залучення декількох спеціалістів та значних часових витрат.

Для оцінювання якості згенерованих відеороликів було проведено експертне опитування. Результати показали, що середній бал якості контенту становить 4,28 із 5 можливих. Найвищі оцінки були отримані за показниками відповідності початковій ідеї користувача та композиції відеоматеріалів. Це підтверджує ефективність використання великих мовних моделей для формування сценаріїв і промптів, а також високу якість генерації відео за допомогою Kling AI.

Проведений аналіз результатів підтвердив досягнення поставленої мети дослідження та ефективність запропонованого підходу до автоматизації процесів створення цифрового відеоконтенту. Розроблена система продемонструвала високу функціональність, прийнятний рівень продуктивності, низьку собівартість і достатню якість результатів для практичного використання.

Таким чином, результати тестування підтверджують можливість застосування розробленої веб-платформи в туристичній сфері, цифровому маркетингу, контент-студіях та інших галузях, де існує потреба у швидкому створенні якісного відеоконтенту з використанням технологій штучного інтелекту

ВИСНОВКИ

У дипломній роботі було вирішено актуальне науково-практичне завдання автоматизації процесу створення коротких вертикальних відеороликів за допомогою сучасних технологій штучного інтелекту. Актуальність теми обумовлена стрімким зростанням популярності відеоконтенту в соціальних мережах та необхідністю скорочення часу і ресурсів, необхідних для його створення. Особливої важливості ця проблема набуває для підприємств туристичної галузі, які регулярно потребують великої кількості якісних рекламних матеріалів для просування своїх послуг.

У процесі виконання роботи було проведено аналіз сучасних технологій генерації відеоконтенту, великих мовних моделей, сервісів синтезу мовлення та платформ автоматизації бізнес-процесів. Досліджено можливості моделей GPT, Gemini та інших LLM-систем щодо автоматичного формування сценаріїв, промптів та текстового супроводу. Виконано аналіз платформ Kling AI, Runway, Pika, ElevenLabs, Zapier, Make, Node-RED та n8n, що дозволило обґрунтувати вибір технологій для реалізації програмного рішення.

На основі проведеного дослідження було спроектовано та реалізовано веб-платформу автоматизованої генерації відеоконтенту, яка забезпечує повний цикл створення відеоролика від моменту введення користувачем текстового запиту до отримання готового відеофайлу. Центральним елементом архітектури стала платформа n8n, яка забезпечує оркестрацію всіх бізнес-процесів та інтеграцію зовнішніх сервісів.

У межах роботи реалізовано інтеграцію Telegram Bot API, Google Sheets, Google Drive, сервісу генерації відео Kling AI через PiAPI, системи синтезу мовлення ElevenLabs та SFTP-сервера для обміну мультимедійними файлами. Розроблений workflow автоматично виконує формування сцен, генерацію промптів, створення відеофрагментів, озвучення, рендеринг фінального ролика та передачу результату користувачеві.

Особливістю реалізованої системи є використання великих мовних

моделей як інтелектуальних агентів для автоматичного створення структури відео, генерації описів сцен, формування промптів для відеомоделі та підготовки метаданих для публікації контенту в соціальних мережах. Це дозволило значно підвищити рівень автоматизації процесу та мінімізувати необхідність ручного втручання.

Проведене тестування підтвердило працездатність та ефективність розробленого рішення. За результатами експериментів встановлено, що середній час створення одного відеоролика становить близько 12 хвилин. Показник успішності виконання workflow досяг 80 %, що свідчить про достатньо високий рівень надійності системи з урахуванням використання зовнішніх AI-сервісів.

Аналіз економічної ефективності показав, що середня собівартість генерації одного відеоролика становить приблизно 1,24 долара США. Отримані результати демонструють значне скорочення витрат порівняно з традиційними методами створення відеоконтенту, які потребують залучення сценаристів, монтажерів та інших спеціалістів.

Оцінювання якості згенерованого контенту за участю експертів показало середній результат 4,28 бала з 5 можливих. Найвищі оцінки були отримані за показниками відповідності відео початковій ідеї користувача та загальної композиції сцен. Це підтверджує ефективність використання сучасних генеративних моделей для автоматизованого створення мультимедійного контенту.

Практична цінність роботи полягає в можливості використання розробленої системи для автоматизації створення рекламних та інформаційних відеоматеріалів у туристичній сфері. Рішення вже було апробоване на базі компанії TRÓPICO EXCÊNTRICO – UNIPESSOAL LDA (Adventure Tours Madeira), діяльність якої пов'язана з організацією туристичних послуг та активним просуванням контенту в соціальних мережах. Використання розробленої платформи дозволяє суттєво скоротити час підготовки відеоматеріалів та підвищити ефективність маркетингової

діяльності підприємства.

Перспективними напрямками подальшого розвитку роботи є інтеграція додаткових моделей генерації відео, впровадження механізмів автоматичної оптимізації промптів, використання адаптивних алгоритмів вибору AI-сервісів залежно від типу контенту, а також автоматична публікація створених відеороликів у соціальних мережах.

Таким чином, поставлена мета дипломної роботи досягнута, усі визначені завдання виконані в повному обсязі, а отримані результати підтверджують ефективність використання технологій штучного інтелекту для автоматизації процесів створення сучасного цифрового відеоконтенту.

ПЕРЕЛІК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. n8n. Документація [Електронний ресурс]. — Режим доступу: <https://docs.n8n.io>
2. 2. Telegram. Bot API [Електронний ресурс]. — Режим доступу: <https://core.telegram.org/bots/api>
3. 3. Google Developers. Google Sheets API v4 [Електронний ресурс]. — Режим доступу: <https://developers.google.com/sheets/api>
4. Google Developers. Google Drive API [Електронний ресурс]. — Режим доступу: <https://developers.google.com/drive/api>
5. PiAPI Kling. Video Generation API [Електронний ресурс]. — Режим доступу: <https://piapi.ai>
6. ElevenLabs. API Reference [Електронний ресурс]. — Режим доступу: <https://docs.elevenlabs.io>
7. OpenAI. API Documentation [Електронний ресурс]. — Режим доступу: <https://platform.openai.com/docs>
8. OpenSSH. SFTP protocol specifications [Електронний ресурс]. — Режим доступу: <https://www.openssh.com/specs.html>
9. ДСТУ 8302:2015. Бібліографічне посилання. Загальні положення та правила складання. — Київ: ДП «УкрНДНЦ», 2016. 16 с.
10. Google Developers. Drive Sharing — File permissions & links [Електронний ресурс]. — Режим доступу: <https://developers.google.com/drive/api/guides/share-files>
11. 8n. Error handling (обробка помилок у flows) [Електронний ресурс]. — Режим доступу: <https://docs.n8n.io/flows/error-handling/>
12. n8n. Credentials (керування секретами) [Електронний ресурс]. — Режим доступу: <https://docs.n8n.io/credentials/>
13. n8n. Wait node (вузол очікування) [Електронний ресурс]. — Режим доступу: <https://docs.n8n.io/integrations/builtin/core-nodes/n8n-nodes-base.wait/>

14. Google Developers. Drive API — Best practices [Електронний ресурс]. — Режим доступу: <https://developers.google.com/drive/api/guides/best-practices>
15. Google Developers. Sheets API — Limits [Електронний ресурс]. — Режим доступу: <https://developers.google.com/sheets/api/limits>
16. Telegram. Bot API — Limits/FAQ [Електронний ресурс]. — Режим доступу: <https://core.telegram.org/bots/faq#my-bot-is-hitting-limits-how-do-i-avoid-this>
17. FFmpeg. ffmpeg (CLI довідка) [Електронний ресурс]. — Режим доступу: <https://ffmpeg.org/ffmpeg.html>
18. FFmpeg. Filters (фільтри для відео/аудіо) [Електронний ресурс]. — Режим доступу: <https://ffmpeg.org/ffmpeg-filters.html>
19. Google Cloud. Exponential backoff (рекомендації) [Електронний ресурс]. — Режим доступу: <https://cloud.google.com/storage/docs/exponential-backoff>
20. IETF. RFC 9110 — HTTP Semantics [Електронний ресурс]. — Режим доступу: <https://www.rfc-editor.org/rfc/rfc9110>
21. IETF. RFC 8259 — The JSON Data Interchange Format [Електронний ресурс]. — Режим доступу: <https://www.rfc-editor.org/rfc/rfc8259>
22. IETF. RFC 4253 — The Secure Shell (SSH) Transport Layer Protocol [Електронний ресурс]. — Режим доступу: <https://www.rfc-editor.org/rfc/rfc4253>
IETF. RFC 4252 — The Secure Shell (SSH) Authentication Protocol [Електронний ресурс]. — Режим доступу: <https://www.rfc-editor.org/rfc/rfc4252>
23. Microsoft. REST API Guidelines [Електронний ресурс]. — Режим доступу: <https://github.com/microsoft/api-guidelines>
24. OWASP. Logging Cheat Sheet [Електронний ресурс]. — Режим доступу: https://cheatsheetseries.owasp.org/cheatsheets/Logging_Cheat_Sheet.html

25. OWASP. Rate Limiting Cheat Sheet [Электронный ресурс]. — Режим — доступу:
https://cheatsheetseries.owasp.org/cheatsheets/Rate_Limiting_Cheat_Sheet.html
26. Google Developers. Sheets API — Values: append [Электронный ресурс]. — Режим — доступу:
<https://developers.google.com/sheets/api/reference/rest/v4/spreadsheets.values/append>
27. Google Developers. Drive API — Files: create [Электронный ресурс]. — Режим — доступу:
<https://developers.google.com/drive/api/reference/rest/v3/files/create>
28. Google Developers. Drive API — Permissions: create [Электронный ресурс]. — Режим — доступу:
<https://developers.google.com/drive/api/reference/rest/v3/permissions/create>
29. Telegram. Bot API — sendMessage [Электронный ресурс]. — Режим доступу: <https://core.telegram.org/bots/api#sendmessage>
30. Telegram. Bot API — sendVideo [Электронный ресурс]. — Режим доступу: <https://core.telegram.org/bots/api#sendvideo>
31. ElevenLabs. Text-to-Speech API [Электронный ресурс]. — Режим доступу: <https://docs.elevenlabs.io/api-reference/text-to-speech>

ДОДАТОК А. Повний workflow платформи N8N

У додатку наведено повний workflow платформи n8n, що реалізує автоматизований процес генерації коротких вертикальних відеороликів на основі текстового запиту користувача.

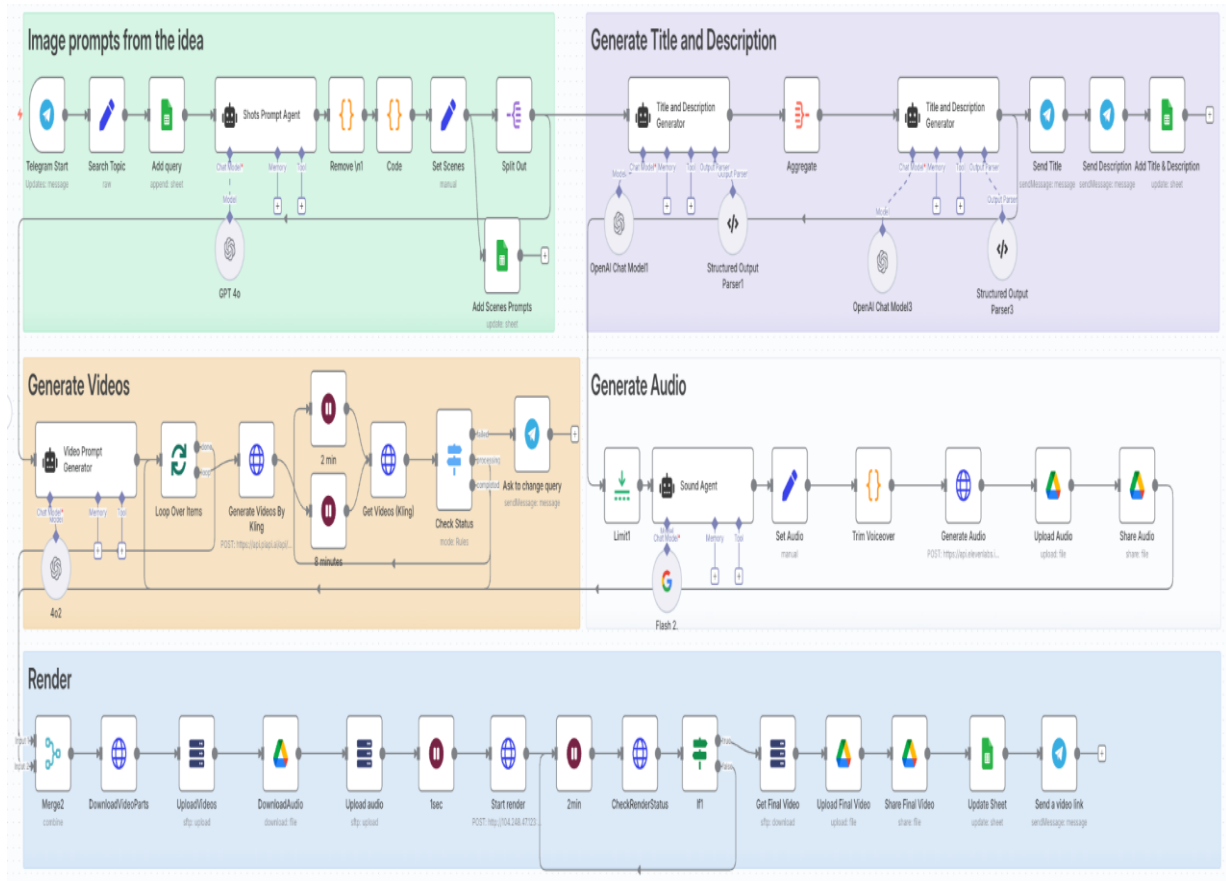


Рисунок А.1 – Загальна структура workflow платформи n8n (Повний JSON workflow.json (окремий файл)).

Workflow забезпечує автоматизоване створення відеоконтенту шляхом інтеграції Telegram Bot API, Google Sheets, Google Drive, Kling AI, ElevenLabs та сервісу рендерингу відео.

```

{
  "name": "[DEV]Video_From_Idea",
  "nodes": [
    {
      "parameters": {
        "updates": [
          "message"
        ],
        "additionalFields": {}
      },
      "type": "n8n-nodes-base.telegramTrigger",
      "typeVersion": 1.1,
      "position": [
        416,
        528
      ],
      "id": "90c5eb18-8372-43ff-8379-1e829a45415a",
      "name": "Telegram Start",
      "webhookId": "1475e981-e418-4cf8-aa00-a2bd6a1f9add",
      "credentials": {
        "telegramApi": {
          "id": "uDnbKYuPDraJNQVc",
          "name": "ContentBot"
        }
      }
    },
    {
      "parameters": {
        "mode": "raw",
        "jsonOutput": "={\n  \"query\": \"{{ $json.message.text }}\"\n}\n",
        "options": {}
      },
      "type": "n8n-nodes-base.set",
      "typeVersion": 3.4,
      "position": [
        576,
        528
      ],
      "id": "2ae83495-e6e7-44bd-8581-3677cdb323d1",
      "name": "Search Topic"
    },
    {
      "parameters": {
        "jsCode": "const parsed = JSON.parse($json.output);\nreturn\n[{\n  json: parsed\n}];\n",
        "type": "n8n-nodes-base.code",
        "typeVersion": 2,
        "position": [
          1344,
          528
        ],
      },
    }
  ]
}

```

Ключові вузли:

- Telegram Start (тригер), Shots Prompt Agent, Video Prompt Generator (формування сцен і промптів),
-
- Generate Videos By Kling (POST /api/v1/task, duration=5, aspect_ratio=9:16, mode=pro, version=2.0), Get Videos (Kling) (polling за task_id), Check Status (гілки failed/processing/completed),
-
- Sound Agent → Generate Audio (ElevenLabs eleven_multilingual_v2) → Upload Audio (Drive),
-
- DownloadVideoParts → UploadVideos (SFTP /root/render/input/video{1..3}.mp4) → Upload audio (SFTP /root/render/input/1.mp3),
-
- Start render (POST .../render, timeout до 20 хв), CheckRenderStatus (GET .../status/{jobId}), Get Final Video (SFTP /root/render/output/final.mp4),
-
- Upload Final Video → Share Final Video (Drive) → Update Sheet → Send a video link (TG).
-

Файл додатка workflow.json

ДОДАТОК Б. Основні фрагменти програмного коду та конфігурацій

Лістинг Б.1 – Фрагмент файлу workflow.json

JSON

```
{
  "name": "[DEV]Video_From_Idea",
  "nodes": [
    {
      "name": "Telegram Start",
      "type": "n8n-nodes-base.telegramTrigger"
    }
  ]
}
```

Структура артефактів і метаданих

- **GSheets (Videos by Prompt / Sheet1):** Id (message_id), UserQuery, ScenePrompt1/2/3, Title, Description, VideoFinal, VideoStatus.
- **GDrive:** Sound/1.mp3, FinalVideos/Content_YYYYMMDD_HH:mm:ss.mp4 (публічні посилання reader).

- **SFTP (DO server):** вхід /root/render/input/*, вихід /root/render/output/final.mp4.

Video Prompt Generator Execute step

Parameters Settings Docs

Source for Prompt (User Message)

Define below

Prompt (User Message)

```
Original user_request: {{ $('Search Topic').item.json.query }}
```

```
Scene: {{ $json.toJsonString() }}
```

Require Specific Output Format

☐

Options

System Message

```
You are an expert AI agent specialized in writing highly effective prompts for Kling AI video generation.
```

```
## OBJECTIVE
```

```
Convert the user's high-level idea ar
```

You are an expert AI agent specialized in writing highl...

Add Option

Рисунок Б.1 – Налаштування вузла Video Prompt Generator

Лістинг Б.2 — Конфігурація запиту до PiAPI Kling

JSON

```
{  
  
  "model": "kling",  
  
  "task_type": "video_generation",  
  
  "duration": 5,  
  
  "aspect_ratio": "9:16",  
  
  "mode": "pro",  
  
  "version": "2.0"
```

Expression

Anything inside `{{ }}` is JavaScript. [Learn more](#)

```
{
  "model": "kling",
  "task_type": "video_generation",
  "input": {
    "prompt": "{{ $json.output }}",
    "negative_prompt": "blurry, distorted, mutated, low quality,
bad anatomy, bad hands, disfigured, deformed, missing fingers,
extra limbs, fused limbs, unrealistic movement, camera instability,
artifact, grainy, poorly drawn, duplicate, watermark, text, logo",
    "cfg_scale": 0.5,
    "duration": 5,
    "aspect_ratio": "9:16",
    "camera_control": {
      "type": "simple",
      "config": {
        "horizontal": 0,
        "vertical": 0,
        "pan": 0,
        "tilt": 0,
        "roll": 0,
        "zoom": 0
      }
    },
    "mode": "pro",
    "version": "2.0"
  },
  "config": {
    "service_mode": "",
    "webhook_config": {
      "endpoint": "",
      "secret": ""
    }
  }
}
```

 Get Videos (Kling)

Execute step

ParametersSettingsDocs

Import cURL

Method

GET

URL

fx

https://api.piapi.ai/api/v1/task/{{ \$json.data.task_id }}

Authentication

Generic Credential Type

Generic Auth Type

Header Auth

Header Auth

PiAPI Kling

Send Query Parameters

Send Headers

Send Body

Options

No properties

Add option

Рисунок Б.3 – Отримання результатів генерації через PiAPI Kling

ДОДАТОК В. Приклади сформованих сценаріїв та ароматів

Лістинг В.1 – Приклад користувацького запиту

Молода сім'я чоловік з дружиною і маленькою дитиною гуляють парком, а потім пляжем біля океану.

Лістинг В.2 – Згенерована структура сцен

```
{  
  
  "main_subject": "Young family",  
  
  "action": "walking together",  
  
  "environment": "green park",  
  
  "camera_movement": "slow pan",  
  
  "duration_seconds": 5  
  
}
```

Лістинг В.3 – Kling Prompt для сцени №1

Young family walking together holding hands

in a lush green park with trees,

slow pan following the family,

warm afternoon light,

transition fades into a sandy beach.

Лістинг В.4 – Kling Prompt для сцени №2

Young family playing near the water's edge,

golden hour lighting,

child splashing happily.

Лістинг В.5 – Kling Prompt для сцени №3

Ocean waves gently crashing onto the shore,

steady zoom out,

soft evening light.

Лістинг В.6 – Negative Prompt

blurry, distorted, mutated, low quality,

bad anatomy, bad hands,

watermark, text, logo

ДОДАТОК Г. Результати тестування та приклади згенерованого контенту

Таблиця 1. TAT по етапах і сумарний TAT (10 запусків)

Запуск	Intake→Scenes	Scenes→Prompts	Prompt→Clips	Clips→TTS	TTS→Render	Render→Link	Render	TAT
1	0:07	0:03	8:40	0:05	0:13	0:28	2:20	11:56
2	0:07	0:03	9:15	0:05	0:13	0:28	1:40	11:51
3	0:07	0:03	8:05	0:05	0:13	0:28	3:10	12:11
4	0:07	0:03	9:50	0:05	0:13	0:28	2:55	13:41
5	0:07	0:03	7:55	0:05	0:13	0:28	1:15	10:06
6	0:07	0:03	8:30	0:05	0:13	0:28	3:40	13:06
7	0:07	0:03	9:05	0:05	0:13	0:28	2:05	12:06
8	0:07	0:03	8:20	0:05	0:13	0:28	1:50	11:06
9	0:07	0:03	10:00	0:05	0:13	0:28	3:30	14:26
10	0:07	0:03	7:35	0:05	0:13	0:28	1:35	10:06

Таблиця Г.2 Успішність і причини відмов (SR = **80%**; успішних 8/10; середня к-сть ретраїв $\approx$ 1/запуск)

Запуск	Kling статус	Повторів	Render статус	Причина відмови	Примітка
--------	--------------	----------	---------------	-----------------	----------

1	completed	1	completed		
2	completed	0	completed		
3	completed	1	completed		
4	failed	1	—	Kling error — sensitive data in prompt	Переформулювати промпт, прибрати чутливі дані
5	completed	1	completed		
6	completed	2	completed		
7	completed	1	completed		
8	completed	0	completed		
9	failed	1	—	Kling error — busy service (rate limit)	Додано extra wait + повтор через 2 хв
10	completed	1	completed		

Таблиця 3. Собівартість за компонентами (USD/ролик)

Запуск	LLM	Kling	TTS	Зберігання	Разом
1	0.00	1.38	0.00	0.00	1.38
2	0.00	1.38	0.00	0.00	1.38
3	0.00	1.38	0.00	0.00	1.38
4	0.00	0.92	0.00	0.00	0.92

5	0.00	1.38	0.00	0.00	1.38
6	0.00	1.38	0.00	0.00	1.38
7	0.00	1.38	0.00	0.00	1.38
8	0.00	1.38	0.00	0.00	1.38
9	0.00	0.46	0.00	0.00	0.46
10	0.00	1.38	0.00	0.00	1.38

Таблиця 4. Якість відео (опитування $n = 8$ респондентів; шкала 1–5)

Запуск	Композиція	Відповідність ідеї	Плавність	Середнє
1	4.3	4.4	4.0	4.2
2	4.5	4.6	4.3	4.5
3	4.2	4.1	3.9	4.1
4				
5	4.6	4.7	4.4	4.6
6	4.1	4.0	3.8	4.0
7	4.4	4.2	4.1	4.2
8	4.3	4.5	4.2	4.3
9				
10	4.5	4.4	4.3	4.4

Додаткові матеріали

- скріншоти Telegram-бота;
- приклади записів Google Sheets;
- посилання на готові відео;
- результати генерації.

Таблиця Д.1 – Структура таблиці Google Sheets

Поле	Тип
Id	String
UserQuery	Text
ScenePrompt1	JSON
ScenePrompt2	JSON
ScenePrompt3	JSON
Title	String
Description	Text

Voiceover	Text
Final_URL	URL
VideoStatus	Enum

Структура Google Drive

Sound/

└─ 1.mp3

FinalVideos/

└─ Content_YYYYMMDD_HH:mm:ss.mp4

Структура SFTP

/root/render/input/

/root/render/output/

Таблиця Д.2 - Схема полів таблиці Google Sheets

Поле	Тип	Приклад
Id (message_id)	string	1234567890

UserQuery	string	"Make a 15s vertical video about ..."
ScenePrompt1.. 3	json	{ main_subject: ..., duration: 5 }
Title	string	"Sunset over Kyiv"
Description	text	"A 15s vertical video..."
VideoFinal	url	https://.../final.mp4
VideoStatus	enum	Created/ToDo

ДОДАТОК Є. Політики безпеки та відмовостійкості

Основні рекомендації

- використання Credentials n8n;
- розмежування доступів;
- резервування API-ключів;
- тайм-аути HTTP-запитів;
- автоматичні повторні спроби;
- експоненціальний backoff;
- контроль статусів:
 - processing;
 - completed;
 - failed;
- логування всіх етапів генерації;
- видалення тимчасових файлів із SFTP.

Таблиця Є.1 - Ретраї за етапами

Етап	Середня к-сть повторів
Kling task create/poll	0.6
Render start/status	0.3
Drive share (audio/final)	0.1

Таблиця Є.2 – Розподіл причин відмов

Причина	К-сть	Частка
Kling — sensitive data in prompt	1	10%
Kling — busy service / rate limit	1	10%
Інші	0	0%
Разом	2	20%

Таблиця Є.3 – Кошторис інфраструктури

Сервіс	Одиниця	Обсяг	Тариф	Вартість
n8n (хостинг)	інстанс	1	15.00	15.00
SFTP (VPS/облік)	інстанс	1	6.00	6.00
GDrive (дод. сховище)	ГБ	50	0.02	1.00

Моніторинг/л огування	пакет	1	5.00	5.00
Разом				27.00

Керівник від кафедри

_____ Зінов'єв Д.В.

(підпис) (прізвище, ініціали)