

**ХЕРСОНСЬКИЙ НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ  
ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ ТА ДИЗАЙНУ  
КАФЕДРА ІНФОРМАТИКИ І КОМП'ЮТЕРНИХ НАУК**

**Пояснювальна записка**  
кваліфікаційна робота магістра  
на тему:

**Створення програмного продукту для аналізу світового ринку штучного  
інтелекту та окремих галузей за допомогою NLP**

Виконав: студент 6 курсу, групи 6КІ  
спеціальності 122 «Консолідована інформація»

Дубоносов І.О.

Керівник: Корніловська Н.В.

Рецензент: Огнева О.Є.

Факультет  
Кафедра  
Рівень вищої освіти  
Галузь підготовки  
Освітньо - професійна програма  
Спеціальність

**Кібернетики та системної інженерії  
Інформатики і комп'ютерних наук  
другий (магістерський) рівень  
12 «Інформаційні технології»  
Консолідована інформація  
122 «Комп'ютерні науки»**

### **ЗАТВЕРДЖУЮ**

Завідувач кафедри ІКН,  
професор

\_\_\_\_\_ В.І. Литвиненко  
« \_\_\_\_ » \_\_\_\_\_ 2024 року

## **ЗАВДАННЯ НА ДИПЛОМНУ РОБОТУ СТУДЕНТУ**

Дубоносова Іллі Олександра  
(прізвище, ім'я, по батькові)

**Тема роботи: «Створення програмного продукту для аналізу світового ринку штучного інтелекту та окремих галузей за допомогою NLP»**

1. Керівник роботи: Корніловська Наталія Володимирівна, к.т.н., доцент.  
затверджені наказом ХНТУ від 29.09.2023 року № 506-с.

2. Рядків подання студентом роботи

3. Вихідні дані до роботи: Способи збирання, обробки та аналізу інформації з різних джерел для надання корисних інсайтів стосовно тенденцій, перспектив розвитку, ключових гравців та можливостей для інвесторів, підприємств та дослідників.

4. Зміст розрахунково-пояснювальної записки (перелік питань, які потрібно розробити): Вступ; 1. Огляд ринку штучного інтелекту для аналізу та аргументації актуальності дослідження; 2. Аналіз способів та методів консолідації інформаційного ресурсу; 3. Методологія та інструментарій. Технологічний стек для створення програмного продукту; 4. Реалізація

програмного продукту для аналізу світового ринку штучного інтелекту та окремих галузей за допомогою NLP; Висновки.

5. Перелік графічного матеріалу (з точним зазначенням обов'язкових креслень):  
Малюнків – 24.

6. Консультанти розділів роботи.

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		завдання видав	завдання прийняв
1.	Корніловська Н.В., к.т.н., доцент кафедри інформатики і комп'ютерних наук	17.10.2023	
2.	Корніловська Н.В., к.т.н., доцент кафедри інформатики і комп'ютерних наук	1.11.2023	
3.	Корніловська Н.В., к.т.н., доцент кафедри інформатики і комп'ютерних наук	15.11.2023	

7. Дата видачі завдання 17.10.2023

## КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назва етапів дипломного роботи	Строк виконання етапів роботи	Примітка
1	Характеристика складових консолідованого інформаційного ресурсу.	17.10.2023-21.10.2023	Виконано
2	Фактори росту компаній з розробки ШІ.	21.10.2023-25.10.2023	Виконано
3	Огляд існуючих методів аналізу інформаційних ресурсів	26.10.2023-28.10.2023	Виконано
4	Аналіз основних експертних методів аналізу ринку.	28.10.2023-02.11.2023	Виконано
5	Дослідження та консолідація інформаційного ресурсу.	02.11.2023-09.11.2023	Виконано
6	Розробка та дослідження моделі збору та обробки інформації.	09.11.2023-22.11.2023	Виконано
7	Практичне застосування методів та способів дослідження ринку за допомогою NLP.	22.11.2023-08.11.2023	Виконано

Студент \_\_\_\_\_ Дубоносів І.О.  
 ( підпис ) ( прізвище та ініціали )

Керівник роботи \_\_\_\_\_ Корніловська Н.В.  
 ( підпис ) ( прізвище та ініціали )

## ПЕРЕЛІК УМОВНИХ СКОРОЧЕНЬ

Скорочення, термін, позначення	Пояснення
AI	Artificial intelligence – штучний інтелект
NLP	Natural language processing – обробка природної мови
NER	Named entity recognition – розпізнавання іменованих сутностей
API	Application Programming Interface – інтерфейс програмування застосунків
IoT	Internet of Things – інтернет речей
FE	Feature Extraction – вилучення ознак

## ЗМІСТ

РЕФЕРАТ.....	8
ABSTRACT.....	9
ВСТУП.....	10
РОЗДІЛ 1. ОГЛЯД РИНКУ ШТУЧНОГО ІНТЕЛЕКТУ ДЛЯ АНАЛІЗУ ТА АРГУМЕНТАЦІЇ АКТУАЛЬНОСТІ ДОСЛІДЖЕННЯ.....	14
1.1. Ринок штучного інтелекту.....	14
1.2. Сучасні дослідження у сфері використання nlp для аналітики компаній.....	24
1.3. Виявлення невирішених або недостатньо вирішених аспектів у літературі.....	28
ВИСНОВКИ ДО РОЗДІЛУ 1.....	31
РОЗДІЛ 2. АНАЛІЗ СПОСОБІВ ТА МЕТОДІВ КОНСОЛІДАЦІЇ ІНФОРМАЦІЙНОГО РЕСУРСУ.....	32
2.1. Поняття та типи інформаційних ресурсів.....	32
2.2. Значення інформаційних ресурсів.....	37
2.3. Форми представлення інформаційних ресурсів.....	39
2.4. Методи доступу до інформаційних ресурсів.....	42
2.5. Технології зберігання та обробки інформаційних ресурсів.....	48
ВИСНОВКИ ДО РОЗДІЛУ 2.....	51
РОЗДІЛ 3. МЕТОДОЛОГІЯ ТА ІНСТРУМЕНТАРІЙ. ТЕХНОЛОГІЧНИЙ СТЕК ДЛЯ СТВОРЕННЯ ПРОГРАМНОГО ПРОДУКТУ.....	53
3.1. Мови програмування та фреймворки.....	53
3.2. Обробка природної мови (nlp).....	60
3.3. Машинне навчання та аналіз даних.....	74
3.4. Бази даних та сховища даних.....	77
3.5. Інтеграція api та зовнішніх джерел.....	92
3.6. Бібліотеки для роботи з ai та nlp.....	95
ВИСНОВКИ ДО РОЗДІЛУ 3.....	97
РОЗДІЛ 4. РЕАЛІЗАЦІЯ ПРОГРАМНОГО ПРОДУКТУ ДЛЯ АНАЛІЗУ СВІТОВОГО РИНКУ ШТУЧНОГО ІНТЕЛЕКТУ ТА ОКРЕМИХ ГАЛУЗЕЙ ЗА ДОПОМОГОЮ NLP.....	99
4.1. Архітектурна схема програмного продукту.....	99
4.2. Опис функціональної системи.....	100
4.3. Результат.....	104

ВИСНОВКИ ДО РОЗДІЛУ 4. ....	106
ВИСНОВКИ ДО РОБОТИ. ....	107
ДОДАТОК А. КОД ПРОГРАМИ INTERNET-SNEAK.PY .....	109
ДОДАТОК Б. КОД ПРОГРАМИ MAIN.PY. ....	111
ДОДАТОК В. КОД ПРОГРАМИ A.PY.....	113
СПИСОК ЛІТЕРАТУРИ.....	114

## **Реферат.**

Дана дипломна робота присвячена розробка програмного продукту для аналізу світового ринку AI за допомогою NLP для отримання узагальненої інформації про стан ринку, а також виявляти найперспективніші AI-компанії.

Ринок штучного інтелекту (AI) характеризується бурхливим зростанням. За оцінками, до 2025 року обсяг світового ринку AI досягне 190 млрд доларів. Це зростання створює нові можливості для бізнесу та інвесторів. Однак, щоб успішно конкурувати на ринку AI, компаніям необхідно постійно аналізувати поточні тенденції та розробляти нові продукти та послуги. Це завдання ускладнюється величезним обсягом інформації, що доступна на ринку.

Розроблений програмний продукт має наукову та практичну цінність. Він може бути застосований бізнесом та інвесторами для таких завдань, як моніторинг тенденцій на ринку AI, визначення перспективних напрямків розвитку AI, оцінка конкурентоспроможності AI-компаній, виявлення потенційних партнерів та клієнтів.

Ключові слова: штучний інтелект (AI), NLP, аналіз ринку, перспектива росту, обробка інформації.



**Abstract.**

This thesis is devoted to the development of a software product for the analysis of the world AI market using NLP to obtain generalized information about the state of the market, as well as to identify the most promising AI companies.

The artificial intelligence (AI) market is characterized by rapid growth. According to estimates, by 2025 the global AI market will reach \$190 billion. This growth creates new opportunities for businesses and investors. However, to compete successfully in the AI market, companies need to constantly analyze current trends and develop new products and services. This task is complicated by the huge amount of information available on the market.

The developed software product has scientific and practical value. It can be used by businesses and investors for tasks such as monitoring trends in the AI market, determining promising directions for AI development, assessing the competitiveness of AI companies, and identifying potential partners and customers.

Keywords: artificial intelligence (AI), NLP, market analysis, growth prospects, information processing.

## **Вступ.**

### **Актуальність.**

Штучний інтелект (AI) є однією з найдинамічніших та найперспективніших галузей сучасного світу. AI-технології проникають у всі сфери людської діяльності, від виробництва та медицини до фінансів та розваг.

Ринок AI-технологій характеризується бурхливим зростанням. За оцінками, до 2025 року обсяг світового ринку AI досягне 190 млрд доларів. Це зростання створює нові можливості для бізнесу та інвесторів.

Однак, щоб успішно конкурувати на ринку AI, компаніям необхідно постійно аналізувати поточні тенденції та розробляти нові продукти та послуги. Це завдання ускладнюється величезним обсягом інформації, що доступна на ринку.

### **Мета і завдання дослідження**

Основною метою проекту є розробка програмного продукту для аналізу світового ринку AI за допомогою NLP. Продукт повинен дозволяти користувачам отримувати узагальнену інформацію про стан ринку, а також виявляти найперспективніші AI-компанії.

Для досягнення цієї мети необхідно вирішити такі завдання:

- зібрати та обробити інформацію з різних джерел, включаючи веб-сайти, новинні статті, наукові публікації та фінансові звіти;
- застосувати методи NLP для виявлення значимої інформації в зібраному масиві даних;
- розробити алгоритми для оцінки перспективності AI-компаній.

Аргументація актуальності та важливості вибору теми

- Вибір теми є актуальним з таких причин:
- зростання ринку AI створює нові можливості для бізнесу та інвесторів;
- для успішної конкуренції на ринку AI компаніям необхідно постійно аналізувати поточні тенденції;

- існуючі методи аналізу ринку AI не дозволяють повною мірою враховувати інформацію, що міститься в текстових джерелах.

Важливість вибору теми полягає в тому, що розроблений програмний продукт може стати цінним інструментом для бізнесу та інвесторів, які прагнуть успішно конкурувати на ринку AI.

### **Опис області застосування програмного продукту.**

Програмний продукт може бути застосований в таких областях:

- бізнес-аналітика;
- інвестиційний аналіз;
- маркетинг;
- розробка нових продуктів та послуг.
- Продукт може бути використаний для таких завдань:
- моніторинг тенденцій на ринку AI;
- визначення перспективних напрямків розвитку AI;
- оцінка конкурентоспроможності AI-компаній;
- виявлення потенційних партнерів та клієнтів.

**Об'єкт дослідження:** методи розробки програмного продукту за допомогою обробки природної мови (NLP). У проєкті буде розглянуто основні аспекти розробки, включаючи використання NLP для обробки текстової інформації, аналіз ключових показників ринку та виведення корисних висновків. Додатково враховую особливості взаємодії програмного продукту із світовим ринком та вплив на рішення в галузях, пов'язаних із штучним інтелектом.

**Предмет дослідження** включає в себе декілька ключових аспектів:

- Вибір технологій NLP: Розгляд та аналіз різних технологій обробки природної мови для вибору оптимальних засобів, що відповідають вимогам аналізу ринку штучного інтелекту та галузей.
- Розробка алгоритмів NLP: Створення алгоритмів для ефективного використання NLP у виявленні та аналізі ключових елементів інформації, пов'язаної з ринком штучного інтелекту та відповідними галузями.

- Інтеграція з іншими системами: Розробка механізмів інтеграції програмного продукту з іншими системами, що використовуються в галузях штучного інтелекту, для забезпечення плавного обміну та аналізу даних.
- Автоматизована консолідація інформації: Розробка засобів для автоматизованої консолідації інформації з різних джерел та її подальший аналіз за допомогою методів NLP.

Предмет дослідження також включає вивчення продуктів консолідації інформації, які можуть використовуватися для отримання звітів та даних, і обрання таких засобів, які оптимально взаємодіють з розроблюваним програмним продуктом.

**Наукова новизна одержаних результатів.** Полягає в розробці нового підходу до аналізу світового ринку штучного інтелекту за допомогою NLP. Даний підхід дозволяє:

- повною мірою враховувати інформацію, що міститься в текстових джерелах;
- виявляти приховані тенденції та закономірності на ринку AI;
- оцінювати перспективність AI-компаній.

**Практичне значення одержаних результатів** полягає в можливості їх використання бізнесом та інвесторами для таких завдань, як:

- моніторинг тенденцій на ринку AI;
- визначення перспективних напрямків розвитку AI;
- оцінка конкурентоспроможності AI-компаній;
- виявлення потенційних партнерів та клієнтів.

**Особистий внесок здобувача.** Усі результати отримано автором самостійно.

**Апробація результатів роботи.** Корніловська Н.В., Дубоносів І.О., Крупнов А.С.. Використання сучасних комп'ютерних систем та інформаційних технологій в туристичній сфері України. Сучасні комп'ютерні системи та мережі в управлінні: матеріали VI Всеукраїнської наук.-практ. Інтернет-конф. здобувачів вищої освіти та молодих вчених (30 листопада 2023 р., м.

Хмельницький, м. Херсон) / за ред. А.А. Григорової. – Херсон: Книжкове видавництво ФОП Вишемирський В. С., 2023. – С 23-26.