

Міністерство освіти і науки України
Одеський національний технологічний університет
Кафедра комп'ютерної інженерії



**ПОЯСНЮВАЛЬНА ЗАПИСКА
ДО КВАЛІФІКАЦІЙНОЇ РОБОТИ**

на тему Дослідження ефективності методів збору даних
(назва кваліфікаційної роботи згідно наказу ОНТУ)
про ціни на товари та послуги

Здобувача Гвоздьова А.В.
(прізвище, ініціали)

2 курсу 556а групи

Керівники: к.т.н., ст. викл. Ненов О.Л.
(посада, прізвище та ініціали)

ст. викл. Слушина Н.В.
(посада, прізвище та ініціали)

Консультанти: _____
(посада, прізвище та ініціали)

д.е.н., проф. Басюркіна Н.Й.
(посада, прізвище та ініціали)

Кваліфікаційна робота допускається до захисту

Рішення кафедри від 30.11 2023 р., протокол № 3

Завідувач кафедри комп. інженерії _____ Сергій АРТЕМЕНКО
(назва кафедри) (підпис) (Ім'я ПРІЗВИЩЕ)

Одеса – 2023 рік

ОДЕСЬКИЙ НАЦІОНАЛЬНИЙ ТЕХНОЛОГІЧНИЙ УНІВЕРСИТЕТ

Факультет комп'ютерної інженерії, програмування та кіберзахисту
Кафедра комп'ютерної інженерії
Ступінь вищої освіти магістр
Спеціальність 123 «Комп'ютерна інженерія»
Освітня програма Спеціалізовані комп'ютерні системи

ЗАТВЕРДЖУЮ

Зав. кафедри комп'ютерної інженерії
Сергій АРТЕМЕНКО
« 30 » листопада 2022 року

ЗАВДАННЯ НА КВАЛІФІКАЦІЙНУ РОБОТУ ЗДОБУВАЧА

Гвоздьова Артема Вадимовича

1. Тема роботи Дослідження ефективності методів збору даних про ціни на товари та послуги

Затверджена наказом університету від « 14 » серпня 2023 р., наказ № 415-03

2 Термін здачі здобувачем закінченої роботи 28 листопада 2023 р.

3. Вихідні дані роботи

1. Дослідження методів 2. Рекомендації з застосування методів 3. Текстовий редактор Microsoft Word 4. Середовище розробки PyCharm 2023 5. Мова розробки Python

4. Перелік питань, які потрібно розробити

1. Вступ. 2. Дослідження предметної області. 3. Технічне проектування та моделювання Системи. 4. Проектування і програмна реалізація системи. 5. Загальні висновки. 6. Економічні розрахунки. 7. Охорона праці.

5. Перелік графічного матеріалу (з точним зазначенням обов'язкових креслень)

Слайд 1. Вступ. Слайд 2. Мета роботи. Слайд 3. Огляд існуючих рішень. Слайд 4. Постановка задачі. Слайд 5. Основні етапи роботи парсера. Слайд 6. Визначення методів збору даних. Слайд 7 Визначення джерел інформації для збору даних. Слайд 8. Оцінка швидкості методів. Слайд 9. Оцінка методів за розробкою та підтримкою. Слайд 10. Специфіка категорій. Слайд 11. Оптимізація і вибір методу. Слайд 12. Основні економічні показники. Слайд 13. Загальні висновки

6. Консультанти по роботі, із зазначенням розділів роботи, що стосуються їх

Розділ	Консультант	Підпис, дата	
		Завдання видав	Завдання прийняв
<i>Економіка</i>	<i>д.е.н., проф. Басюркіна Н.Й.</i>		
<i>Охорона праці</i>	<i>к.т.н., доц. Шестопалов С.В.</i>		
<i>Нормоконтроль</i>	<i>ст. викл. Жуковецька С.Л.</i>		

7. Дата видачі завдання 30.11.2022

Керівники

Олексій НСНОВ

Наталя СЛУШНА

Завдання прийняв до виконання

Артем ГВОЗДЬОВ

КАЛЕНДАРНИЙ ПЛАН

№	Назва етапів кваліфікаційної роботи	Термін виконання етапів роботи	Примітка
1.	<i>Дослідження предметної області</i>	<i>26.12.2022</i>	
2.	<i>Дослідження існуючих аналогів</i>	<i>30.01.2023</i>	
3.	<i>Дослідження методів реалізації</i>	<i>28.02.2023</i>	
4.	<i>Проектування</i>	<i>15.08.2023</i>	
5.	<i>Розробка демонстраційної версії ПЗ</i>	<i>27.10.2023</i>	
6.	<i>Підготовка техніко-економічної частини</i>	<i>15.11.2023</i>	
7.	<i>Підготовка розділу охорони праці</i>	<i>15.11.2023</i>	
8.	<i>Оформлення пояснювальної записки</i>	<i>27.11.2023</i>	
9.	<i>Оформлення графічної частини та лістингу</i>	<i>27.11.2023</i>	

Здобувач-дипломник Артем ГВОЗДЬОВ

Керівники роботи Олексій НСНОВ

Наталя СЛУШНА

Несу відповідальність за ідентичність електронного та друкованого варіантів кваліфікаційної роботи, даю згоду на обробку персональних даних та не заперечую проти розміщення кваліфікаційної роботи на офіційних web-ресурсах ОНТУ.

Підтверджую, що в кваліфікаційній роботі відсутні порушення норм академічної доброчесності.

Здобувач-дипломник Артем ГВОЗДЬОВ

АНОТАЦІЯ

Дипломна робота спрямована на дослідження та порівняння різних методів збору даних про ціни на товари та послуги з метою визначення їхньої ефективності. Зі зростанням конкуренції і змінами на ринку, набуття точної та актуальної інформації про ціни стає важливим завданням для підприємств, аналітиків та урядових органів.

У роботі проводиться огляд різних методів збору даних, включаючи опитування, спостереження, аналіз веб-ресурсів, та автоматизовані системи. Проводяться експерименти та аналіз даних, щоб визначити переваги та недоліки кожного методу, а також виявити фактори, які впливають на їхню ефективність.

Результати цього дослідження можуть бути використані як підстава для вибору оптимальних методів збору даних про ціни в різних галузях та умовах. Вони також надають корисні рекомендації для бізнесу, аналітиків та урядових інституцій, які прагнуть збирати точну та надійну інформацію про ціни на товари та послуги.

Ця робота важлива для розвитку ефективних стратегій ціноутворення, конкурентоспроможності підприємств та покращення аналізу ринків. Вона також відкриває додаткові можливості для подальших досліджень у сфері збору даних та аналізу ринкових цін.

Ключові слова: збір даних, ціни на товари і послуги, методи збору даних, ефективність методів, аналіз ринкових цін.

ABSTRACT

This thesis aims to investigate and compare different methods of collecting price data for goods and services in order to determine their effectiveness. With increasing competition and changes in the market, acquiring accurate and up-to-date price information is becoming an important task for businesses, analysts, and government agencies.

This paper reviews various methods of data collection, including surveys, observations, web crawling, and automated systems. Experiments and data analysis are conducted to determine the advantages and disadvantages of each method, as well as to identify factors that affect their effectiveness.

The results of this study can be used as a basis for choosing the best methods for collecting price data in different industries and conditions. They also provide useful recommendations for businesses, analysts, and government agencies seeking to collect accurate and reliable information on prices for goods and services.

This work is important for the development of effective pricing strategies, the competitiveness of enterprises, and improved market analysis. It also opens up additional opportunities for further research in the field of data collection and market price analysis.

Keywords: *data collection, prices for goods and services, data collection methods, efficiency of methods, market price analysis.*

ЗМІСТ

	стор.
ВСТУП.....	9
РОЗДІЛ 1 ДОСЛІДЖЕННЯ ПРЕДМЕТНОЇ ОБЛАСТІ.....	11
1.1 Огляд основ ціноутворення продуктів загального користування	11
1.2 Огляд існуючих рішень для збору та аналізу цін товарів	17
1.2.1 <i>Hotline.ua</i> (сайт)	18
1.2.2 <i>Ekatalog</i> (сайт + застосунок)	19
1.2.3 <i>Costless</i> (сайт + застосунок)	20
1.3 Мета дослідження, постановка задачі, об'єкт та предмет дослідження, технічне завдання	21
1.3.1 Мета дослідження.....	21
1.3.2 Постановка задачі.....	22
1.3.3 Об'єкт дослідження	23
1.3.4 Предмет дослідження	23
1.3.5 Технічне завдання.....	24
1.3.6 Методи дослідження	24
Висновок до першого розділу	26
РОЗДІЛ 2 ТЕХНІЧНЕ ПРОЄКТУВАННЯ ТА МОДЕЛЮВАННЯ СИСТЕМИ. 27	
2.1 Дослідження методів збору даних.....	27
2.1.1 Ефективність методів	27
2.1.2 Специфіка категорій.....	29
2.1.3 Оптимізація і вибір методу	30
2.1.4 Послідовність використання методів	30
2.1.5 Рекомендації	31
2.2 Принцип роботи парсерів	32

					<i>КРМ.КІ.1. 415-03.1.4</i>		
Змн.	Арк.	№ докум.	Підпис	Дата			
Розробив		Артем ГВОЗДЬОВ			Літ.	Арк.	Акрушів
Перевірів		Наталія СЛУШНА				6	105
Рецензент		Марина БУЛЮК			<i>Дослідження ефективності методів збору даних про ціни на товари та послуги</i>		
Нормоконтроль		Світлана Жуковецька					
Затвердив		Сергій АРТЕМЕНКО					
					<i>гр. 556а, ОНТУ</i>		

2.3 Сфери використання парсерів	33
2.4 Етапи парсинга	37
2.4 Захист від конкурентного парсинга	39
2.5 Етична сторона парсингу	40
Висновок до другого розділу.....	41
РОЗДІЛ 3 ПРОЄКТУВАННЯ І ПРОГРАМНА РЕАЛІЗАЦІЯ СИСТЕМИ.....	43
3.1 Визначення методів збору даних та джерела інформації.....	43
3.2 Збір даних	45
3.3 Опис програмного комплексу для програміста	47
3.3.1 Збір даних за допомогою <i>API</i>	47
3.3.2 Веб-скрапінг.....	50
3.2.3 Аналіз <i>HTTP</i> -запитів і відповідей.....	53
Висновки до третього розділу	54
РОЗДІЛ 4 ЕКОНОМІЧНІ РОЗРАХУНКИ.....	55
4.1 Організаційно-економічне та маркетингове обґрунтування проекту	55
4.2 Економічні розрахунки проекту	59
4.2.1 Розрахунок трудомісткості розробки методом сітьового планування	59
4.2.2 Визначення трудомісткості розробки програмного продукту.....	63
4.2.3 Розрахунок ціни програмного продукту.....	66
Висновок до четвертого розділу.....	76
РОЗДІЛ 5 ОХОРОНА ПРАЦІ.....	77
5.1 Вимоги до особистого робочого місця працівника	77
5.2 Рішення з розміщенням проектованого обладнання	78
5.3 Електробезпека	78
5.4 Пожежна профілактика	82
5.5 Виробнича санітарія	84
Висновки до п'ятого розділу	88
ЗАГАЛЬНІ ВИСНОВКИ	90

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ	91
ДОДАТКИ	94
Додаток А Текст формування бази даних та збір даних	94
Додаток Б Графічний матеріал	99

					<i>КРМ.КІ.1.415-03.1.4</i>	<i>Арк.</i>
<i>Змн.</i>	<i>Арк.</i>	<i>№ докум.</i>	<i>Підпис</i>	<i>Дата</i>		8

ВСТУП

Сучасне господарське середовище характеризується постійними змінами і зростаючою конкуренцією на ринках. Один із ключових аспектів успішного функціонування підприємств, аналізу ринку та встановлення цінової політики полягає в зборі та використанні інформації про ціни на товари та послуги. Точна та актуальна інформація про ціни є важливою складовою прийняття рішень, розробці маркетингових стратегій та визначенні конкурентоспроможності на ринку.

Ця дипломна робота присвячена дослідженню ефективності методів збору даних про ціни на товари та послуги. Питання збору і аналізу інформації про ціни мають вирішальне значення для бізнесу, аналітиків, урядових органів та споживачів. Це важливий інструмент для визначення стратегій ціноутворення, аналізу ринків та визначення факторів, що впливають на ринкову динаміку.

Метою цієї роботи є визначення переваг та недоліків різних методів збору даних про ціни, а також виявлення факторів, які впливають на їхню ефективність. Ми пропонуємо дослідити різні методи збору даних, такі як опитування, спостереження, аналіз веб-ресурсів та використання автоматизованих систем, та оцінити їхню придатність для конкретних завдань та умов.

Ця робота є актуальною в умовах зростаючої конкуренції на ринках і постійних змін у споживчих вимогах. Результати дослідження можуть бути корисними для бізнесу, аналітиків та урядових інституцій, які стежать за ринковими тенденціями і регулюють цінову політику.

У наступних розділах роботи ми докладніше розглянемо методологію дослідження, результати, їх обговорення та рекомендації для практичного застосування.

					КРМ.КІ.1.415-03.1.4	Арк.
						9
Змн.	Арк.	№ докум.	Підпис	Дата		

Зазначимо мету, об'єкт, предмет та задачі, які необхідно вирішити в ході кваліфікаційної роботи.

Метою роботи є проведення детального аналізу та дослідження різноманітних методів збору цін на товари та послуги з метою визначення їх ефективності та застосовності у сучасних умовах.

Об'єктом дослідження є процеси, пов'язані зі збором цін на товари та послуги в різних галузях економіки.

Предметом дослідження є різні методи збору цін, які застосовуються для отримання інформації про вартість товарів та послуг.

Основними задачами, які необхідно вирішити в ході роботи є:

1. Проведення аналізу сучасних підходів до збору цін на товари та послуги.
2. Дослідження різноманітних методів збору цін та їх порівняльний аналіз.
3. Розробка рекомендацій щодо оптимального вибору методів збору цін у конкретних умовах.
4. Визначення переваг та обмежень використання різних методів збору цін.

Дослідження цих аспектів сприятиме розвитку та удосконаленню методології збору цін на товари та послуги, що має практичне значення для економічних суб'єктів та дослідників у даній галузі.

Дослідження цих аспектів сприятиме розвитку та удосконаленню методології збору цін на товари та послуги, що має практичне значення для економічних суб'єктів та дослідників у даній галузі.

					<i>KPM.KI.1.415-03.1.4</i>	Арк.
						10
Змн.	Арк.	№ докум.	Підпис	Дата		

РОЗДІЛ 1

ДОСЛІДЖЕННЯ ПРЕДМЕТНОЇ ОБЛАСТІ

1.1 Огляд основ ціноутворення продуктів загального користування

Дослідження ефективності методів збору інформації про ціни на товари та послуги містить аналіз цінової політики підприємств на плановий період, що ґрунтується на принципах збільшення обсягу продажів, підвищення прибутку, утримання поточного сегмента ринку та дотримання рівня витрат і прибутку, досягнутого на початок планового періоду.

На формування цін найбільше впливають покупці, конкуренти і витрати. Під час установа цін менеджери повинні дивитися на продукцію з точки зору споживачів. На сьогоднішній день ринок харчових продуктів в Україні різноманітний і насичений, але завжди виникає питання про відповідність ціни та якості продукції. Якість продукції залежить від технічного обладнання підприємства, кваліфікації працівників і відповідності стандартам сировинних ресурсів, і вона перебуває в тісному зв'язку з безпосередніми витратами виробництва, які можна розглядати як змінні, маржинальні витрати.

Знання технології конкуруючих фірм, їхніх виробничих потужностей, виробничих площ, логістики та стратегії розвитку полегшує встановлення цін на власну продукцію.

Стратегія розвитку підприємства в умовах ринкової економіки втілюється в розробленні стратегічних, перспективних і поточних планів (бюджетування).

Основою формування ціни має бути нормативна (планова) собівартість продукції за плановою номенклатурою, асортиментом і нормами на початок планового періоду з використанням розрахунку витрат методом "величини покриття". При цьому вихідна база повинна характеризувати рівень витрат і прибутку, вже досягнутий на початок планового періоду. Це означає, що ефект від заходів попереднього періоду повністю реалізується в плановому періоді і вже врахований у нормах, які використовуються при формуванні цін.

					КРМ.КІ.1.415-03.1.4	Арк.
						11
Змн.	Арк.	№ докум.	Підпис	Дата		

У харчовій промисловості використовують такі основні підходи до формування ціни на продукцію підприємств: встановлення ціни, яку можуть заплатити покупці; використання інформації про ціни конкурентів; використання інформації про собівартість продукції; використання інформації про витрати підприємства. Дослідження взаємозв'язку витрат, обсягу та прибутку дає змогу встановити мінімальну ціну, яка в короткостроковій перспективі забезпечить досягнення запланованого прибутку.

Найпоширенішими методами встановлення ціни є метод ціноутворення на основі собівартості продукції та витрат підприємства.

Економічно обґрунтованим способом ціноутворення на основі витрат є метод "витрати плюс". Якщо ринок не ставить жорстких умов, то менеджер під час установавання ціни має орієнтуватися на такі формули: виробничі змінні витрати плюс надбавка; загальні змінні витрати плюс надбавка; виробнича собівартість плюс надбавка; повна собівартість плюс надбавка.

Націнку можна розрахувати на основі: змінних виробничих витрат; повної виробничої собівартості; загальних змінних витрат; повних витрат.

Моніторинг цін – це частина бізнес-процесу онлайн-продажів, який використовується для спостереження за цінами на товари, послуги, знижки, сервісні пропозиції та акції. Він є потужним інструментом як для виробників товарів, так і для роздрібних продавців та їхніх клієнтів.

Моніторинг цін конкурентів – це регулярне відстеження цін з метою вивчення стратегії ціноутворення та маркетингової політики конкурентів. Дані моніторингу цін конкурентів використовують для формування власної стратегії ціноутворення, яка приносить компанії прийнятний прибуток.

Маркетинговий моніторинг цін – це системне і регулярне спостереження за станом цін на ринку з метою їхньої оцінки, аналізу тенденцій і розуміння конкурентного оточення. Цей вид моніторингу необхідний для ефективного ціноутворення. Результати маркетингового моніторингу цін дають змогу швидко коригувати оптові ціни компанії та давати рекомендації щодо роздрібних цін як в онлайн, так і в офлайн.

					<i>KPM.KI.1.415-03.1.4</i>	Арк.
						12
<i>Змн.</i>	<i>Арк.</i>	<i>№ докум.</i>	<i>Підпис</i>	<i>Дата</i>		

Моніторинг цін конкурентів в Інтернеті – це регулярне і системне спостереження за цінами онлайн з метою оцінки, вивчення тенденцій і розуміння конкурентного оточення. Збір даних про ціни необхідний для ефективного ціноутворення. Моніторинг цін конкурентів в Інтернеті дає змогу швидко коригувати оптові та роздрібні ціни компаній. На основі моніторингу цін в Інтернеті формується рекомендована роздрібна ціна для партнерів і клієнтів.

Аналіз моніторингу цін – це процес ухвалення рішень про ціноутворення компанії на основі даних про ціни конкурентів, пропозиції постачальників, розуміння собівартості та цілей ринкових процесів і внутрішніх цілей компанії. Для управління цим процесом використовують складні математичні моделі та машинне навчання на основі штучного інтелекту.

Моніторинг цін конкурентів інтернет-магазинів – це обов'язкова процедура для кожного учасника електронної комерції. Вона охоплює вибір майданчиків, що конкурують з інтернет-магазином, у різних товарних категоріях і систематичне порівняння цін свого магазину з цінами конкурентів. Мета цієї процедури – коригування власних цін магазину з метою отримання максимального прибутку від продажу товарів.

Система моніторингу цін в інтернеті може бути важливим інструментом для багатьох бізнесів і компаній з різних галузей. Ось декілька можливих сценаріїв, коли така система може знадобитися:

1. Прогнозування кон'юнктури ринку: За допомогою моніторингу цін в інтернеті компанії можуть аналізувати тенденції зміни цін на свої товари або послуги, що дозволяє їм передбачати майбутні зміни на ринку та адаптувати свою стратегію відповідно.
2. Аналіз конкурентів: Система моніторингу цін дозволяє компаніям відстежувати ціни конкурентів і розуміти їхні цінові стратегії. Це може бути корисним для того, щоб визначити свою власну цінову політику і знайти конкурентні переваги.

					<i>KPM.KI.1.415-03.1.4</i>	<i>Арк.</i>
<i>Змн.</i>	<i>Арк.</i>	<i>№ докум.</i>	<i>Підпис</i>	<i>Дата</i>		13

3. Керування цінами: Компанії можуть використовувати моніторинг цін для активного керування своїми цінами. Вони можуть швидко реагувати на зміни на ринку, регулювати ціни в залежності від попиту та пропозиції, а також проводити *A/B*-тести для визначення оптимальних цін.
4. Пошук нових можливостей: Аналіз моніторингу цін може допомогти компаніям виявити нові ринки або споживачів зі зростаючим попитом на певні товари або послуги.
5. Маркетингова стратегія: Знання цін на ринку дозволяє компаніям розробляти більш ефективні маркетингові стратегії, включаючи акції, знижки, рекламу і промо-кампанії, які враховують конкурентність цін.
6. Контроль над постачальниками: Компанії можуть використовувати дані з моніторингу цін для перевірки, чи дотримуються їхні постачальники обіцяних цін, і вимагати від них виконання угод.

Отже, система моніторингу цін в інтернеті може допомогти компаніям бути більш конкурентоздатними, ефективно управляти своєю ціновою політикою і забезпечувати максимальну вигоду від змін на ринку.

Збір даних про ціни може проводитися з різних джерел, залежно від конкретних потреб компанії:

1. Власний веб-скрапінг: Компанія може створити власний веб-скрапер, який автоматично відвідує веб-сайти конкурентів та інших джерел і збирає інформацію про ціни на їхні товари або послуги. Важливо бути впевненим, що це відповідає законам та правилам використання інтернет-ресурсів.
2. Постачальники даних: Деякі компанії спеціалізуються на зборі та наданні даних про ціни на ринку. Це може бути інформація від третіх сторін, які мають доступ до цінової інформації від різних джерел. Цей підхід може зекономити час і ресурси.
3. *API* інтернет-магазинів і платформ: Деякі інтернет-магазини та онлайн-платформи надають *API* (інтерфейси програмування застосунків), які

					<i>KPM.KI.1.415-03.1.4</i>	Арк.
						14
Змн.	Арк.	№ докум.	Підпис	Дата		

дозволяють компаніям отримувати доступ до даних про ціни на їхній платформі. Це дозволяє отримувати актуальну інформацію без необхідності скрапінгу веб-сайтів.

4. Публічні джерела: Деякі цінові дані можуть бути доступні в публічних джерелах, таких як офіційні веб-сайти компаній, гіпермаркетів, державних структур або інших джерел. Це може бути менш надійним джерелом, але в деяких випадках може бути корисним.

Якщо ви обираєте метод збору даних про ціни, важливо враховувати правові обмеження, етичні аспекти та можливість отримання надійної та актуальної інформації.

Аналіз отриманих даних про ціни може бути різноманітним і включати в себе наступні етапи:

1. Порівняння цін: Порівнюйте ціни на ваші товари або послуги з цінами конкурентів і власними цінами на різних ринках або платформах. Визначте, чи ви є конкурентоспроможними.
2. Аналіз тенденцій: Вивчайте тенденції зміни цін на ринку. Це допоможе передбачити майбутні зміни і виробити стратегію.
3. Визначення оптимальної цінової стратегії: Виходячи з отриманих даних, розробіть оптимальну цінову стратегію. Враховуйте рівень попиту, сезонність, конкурентний тиск і багато інших факторів.
4. Контроль за власними цінами: Підтримуйте актуальність своїх цін та впевніться, що вони відповідають вашій стратегії.
5. Розробка маркетингових кампаній: Використовуйте дані про ціни для створення ефективних маркетингових кампаній, таких як знижки, акції та реклама.

Для аналізу цінових даних можна використовувати різні інструменти, від електронних таблиць і баз даних до програм аналітики даних та спеціалізованих програм для моніторингу цін.

Система моніторингу цін має численні переваги для бізнесу:

					<i>KPM.KI.1.415-03.1.4</i>	<i>Арк.</i>
<i>Змн.</i>	<i>Арк.</i>	<i>№ докум.</i>	<i>Підпис</i>	<i>Дата</i>		15

1. Покращення стратегії ціноутворення: Вона допомагає компанії визначити оптимальні ціни для максимізації прибутку або ринкової долі.
2. Адаптація до змін на ринку: Компанія може швидко реагувати на зміни цін і конкуренції на ринку.
3. Підвищення конкурентоспроможності: Це дозволяє бути більш конкурентоздатним і привертати більше клієнтів.
4. Ефективність маркетингу: Дані з моніторингу цін можуть бути використані для створення більш ефективних маркетингових кампаній.
5. Можливості для росту прибутку: Виходячи з аналізу даних про ціни, компанія може знайти можливості для підвищення прибутку.

Якщо правильно використовувати систему моніторингу цін, це може стати потужним інструментом для оптимізації бізнесу і досягнення успіху на ринку.

Отже, який необхідний моніторинг цін в онлайн маркетингу?

1. Моніторинг цін в інтернет-магазинах.
2. Моніторинг акцій та спецпропозицій інтернет-магазинів.
3. Моніторинг цін на прайс-агрегаторах.
4. Моніторинг цін та знижок на сайтах конкурентів.
5. Моніторинг цін на маркетплейс.
6. Моніторинг цін у рекламних оголошеннях у пошуковій видачі.
7. Моніторинг цін на різних дошках оголошень.
8. Моніторинг нецінових методів просування та розповсюдження товарів клієнтами та конкурентами.
9. Моніторинг відгуків про ціни споживачів продукції компанії.

Плюси використання прайс-агрегаторів для порівняння цін для бізнесу:

На прайс-агрегаторах рекламуються як топові гравці ринку, так і невеликі інтернет-магазини, що дозволяє, не використовуючи великих фінансових вкладень у моніторинг, визначити загальну тенденцію за цінами на ринку,

					<i>KPM.KI.1.415-03.1.4</i>	<i>Арк.</i>
						16
<i>Змн.</i>	<i>Арк.</i>	<i>№ докум.</i>	<i>Підпис</i>	<i>Дата</i>		

визначити демпінгуючі інтернет-магазини та проаналізувати позиціонування своїх конкурентів.

Мінуси: на прайс-агрегаторах рекламується частина інтернет-магазинів на ринку, тому результат порівняння цін на цьому ресурсі як «середня температура по лікарні» може бути спотвореним.

Плюс для споживача: одразу можна побачити низьку ціну.

Мінус: низьку ціну зазвичай пропонують інтернет-магазини, що реалізують неофіційний товар. І тут покупець може отримати фіскальний чек, і навіть гарантійне обслуговування у разі пошкодження товару.

Серед маркетплейсів для вирішення задач порівняння цін будуть корисні великі гравці з мільйонами пропозицій. В Україні лідируючі позиції займають *Rozetka* та *Prom.ua*. Один і той же товар тут можна знайти у кількох продавців та вибрати найбільш прийнятний за ціною для кожного покупця варіант.

Для підприємців, які планують працювати з маркетплейсами, буде зайвим порівняти ціну власного товару з аналогічним товаром продавців на цих торгових майданчиках.

Порівняння цін можна проводити вручну або звернутися до спеціалізованих сервісів. Для компаній з невеликою кількістю товару можливий і ручний моніторинг, але за наявності широкого асортименту ефективніше цей процес автоматизувати.

1.2 Огляд існуючих рішень для збору та аналізу цін товарів

З настанням епохи комп'ютерних технологій та автоматизації процес пошуку вигідних пропозицій став набагато зручнішим. З використанням зручних інтерфейсів і доступного представлення інформації кінцевому користувачеві, збір, аналіз і вибір найвигідніших пропозицій і товарів за критеріями ціна-якість-функціональність стали значно простішими і доступнішими.

Давайте розглянемо приклади таких інструментів.

					<i>KPM.KI.1.415-03.1.4</i>	Арк.
						17
Змн.	Арк.	№ докум.	Підпис	Дата		

1.2.1 Hotline.ua (сайт)

Hotline.ua – український онлайн-сервіс для вибору товарів і порівняння цін (рис.1.1).



Рис. 1.1 – Логотип *Hotline.ua*

Його історія наглядно показує, як з приходом комп'ютерних технологій змінювалось представлення інформації для користувачів:

У 2011 році на *Hotline.ua* покупцям було доступно більше 2 млн пропозицій від усіх провідних і активних інтернет-магазинів України (більше 1400 магазинів). Станом на липень 2012 року на сайті розміщувалось близько 2 млн пропозицій від 1700 інтернет-магазинів.

Ресурс було створено у жовтні 1992 року як каталог прайс-листів на комп'ютерну техніку і спочатку поширювався за допомогою модемного доступу. У 1993 р. з'явилась його паперова версія, яка щотижня виходила до кінця 2005 року .

З 1 січня 2006 почав свою роботу інтернет-проект hotline.ua.

В березні 2014 року відкрито україномовну версію сайту. Щоденна аудиторія сайту складає 100 тис. відвідувачів, щомісячна аудиторія –1,5 млн.

Hotline дозволяє знайти нон-фуд (Нон-фуд – промислові товари, які є невід'ємною частиною нашого повсякденного життя. У категорію товарів нон-фуд входить ряд товарних груп, до яких можна віднести: іграшки, товари для інтер'єру, спортивні товари, канцелярське приладдя, косметичні засоби і товари побутової хімії.) товари, порівняти їх ціни у різних магазинах та характеристики.

Данні до *Hotline* потрапляють через так званий «товарний фід».

Товарний фід – це файл у структурованому форматі, який містить детальну інформацію про товари магазину. Інформація про товари та послуги надається державною мовою, згідно Законодавства України. Підготовка

					КРМ.КІ.1.415-03.1.4	Арк.
						18
Змн.	Арк.	№ докум.	Підпис	Дата		

якісного товарного фіду є ключовою умовою успішного розміщення магазину на *hotline*.

Товарний фід може бути сформований в одному з наступних форматів:

1. *Hotline XML* - структурований формат, який дозволяє найбільш повно описати асортимент, умови продажу та доставки товарів магазину.
2. *Yandex YML* - дозволяє описати асортимент, але має обмежені можливості у відображенні умов продажу та доставки товарів, оскільки в роботу беруться тільки ті теги й атрибути, які мають аналоги в стандарті *Hotline*.
3. *CSV, XLS* (формат електронної таблиці) - не підходить для завантаження товарів з розділу "Одяг для спорту і відпочинку", для завантаження фідів, які містять товари з цього розділу, потрібно використовувати тільки формат *XML (YML)*.
4. *TXT* (текстовий формат з роздільниками).

1.2.2 *Ekatalog* (сайт + застосунок)

E-Katalog – багатофункціональний сервіс пошуку товарів в інтернет-магазинах і порівняння цін. Він охоплює найрізноманітніші категорії товарів: електроніка, комп'ютери, побутова техніка, автотовари, обладнання для ремонту та будівництва, туристичне спорядження, дитячі товари і багато іншого (нон-фуд) (рис.1.2).



Рис. 1.2 – Логотип *E-Katalog*

Основна мета – допомогти покупцеві швидко і зручно знайти найвигіднішу пропозицію. Для тих, хто визначається з вибором, в кожному розділі є підбір за параметрами і можливість порівняти товари між собою. Доступний і зручний текстовий пошук, що дозволяє шукати як потрібні розділи, так і конкретні товари за назвою. А на сторінці кожної моделі є

					КРМ.КІ.1.415-03.1.4	Арк.
						19
Змн.	Арк.	№ докум.	Підпис	Дата		

докладна інформація, яка допоможе прийняти рішення: опис, технічні характеристики, фото і відео, корисні посилання та відгуки. Там же знаходиться блок «Де купити?» зі списком інтернет-магазинів, цінами і прямими посиланнями на сторінку покупки.

До системи *E-Katalog* підключено більше 3000 магазинів, дані по яких постійно оновлюються.

Як видно з опису, *E-Katalog* дуже схожий на *Hotline*, проте основна їх різниця – територія, яку вони охоплюють і кількість магазинів, з якими співпрацюють. Дивлячись на цифри, здається, що всі мережі, з якими працює *E-Katalog* є в *Hotline*, проте це не завжди так. *E-Katalog* же допомагає більшій кількості користувачів за рахунок того, що працює не тільки в Україні, а й в інших країнах – сусідах (Польща, Болгарія тощо).

Данні до *Ekatalog* потрапляють приблизно так само, як і до *Hotline*, проте з деякими відмінностями:

1. *XML* – вкрай рекомендований, оскільки несе актуальну інформацію про товари магазину. Також можлива обробка форматів Хотлайн (*Hotline XML*).
2. *CSV* – якщо підготувати прайс-лист у форматі *XML* не виходить, то можна надати у форматі *CSV* (Текстовий файл із розділювачами). Такий формат можна отримати з електронних таблиць *Excel*.

1.2.3 *Costless* (сайт + застосунок)

Costless – застосунок, що дозволяє порівнювати ціни на окремі продукти або цілі продуктові кошики (переважно фуд-онлі), визначаючи оптимальний для вас магазин поблизу (рис. 1.3).

Він може:

1. Позбавити вас стопки дисконтних карт: можна завантажити дисконтні картки в *Costless* та користуватися ними разом із близькими у будь-якому місці.
2. Повідомляти про свіжі акційні пропозиції від магазинів вашого міста.

					КРМ.КІ.1.415-03.1.4	Арк.
						20
Змн.	Арк.	№ докум.	Підпис	Дата		

3. Створювати нові списки покупок, використовувати старі, ділитися ними з близькими, що спрощує сімейні покупки.



Рис. 1.3 – Логотип *Costless*

Онлайн-сервіс збирає дані за цінами та акціями щодо найбільших українських онлайн-рітейлерів і торговельних мереж, таких як АТБ, Новус, МЕТРО, Ашан, Сільпо, Фора, Епіцентр і т.д. З його допомогою можна проводити моніторинг цін на товари конкурентів онлайн із заданою періодичністю. Отримані дані можна завантажувати у вигляді *Excel* або *pdf* звіту. У режимі реального часу можна дізнатися про найпопулярніші товари та акції.

За інформацією на їх офіційному сайті вони збирають дані як через документи, що їм надають рітейлери (*XML*, *CSV*, тощо), так і самостійно, за допомогою API торговельних мереж та інших інструментів веб-скрапінгу з відкритих джерел.

1.3 Мета дослідження, постановка задачі, об'єкт та предмет дослідження, технічне завдання

1.3.1 Мета дослідження

Мета дослідження ефективності методів збору даних про ціни на товари та послуги полягає в оцінці та порівнянні різних методів, використовуваних для збору інформації про ціни на товари та послуги в сучасному бізнес-середовищі.

Основні завдання дослідження можуть включати:

1. Аналіз існуючих методів збору даних: Ретельний огляд та аналіз різних методів, що використовуються для збору інформації про ціни, таких як опитування споживачів, моніторинг інтернет-ресурсів, аналіз даних з точок продажу тощо.

					<i>KPM.KI.1.415-03.1.4</i>	Арк.
						21
Змн.	Арк.	№ докум.	Підпис	Дата		

2. Оцінка точності та достовірності даних: Визначення, наскільки точні та надійні дані, отримані за допомогою різних методів, враховуючи можливі систематичні та випадкові помилки.
3. Визначення ефективності в залежності від галузі та масштабу бізнесу: Встановлення, які методи найбільш ефективні в конкретних галузях чи для певних розмірів підприємств.
4. Аналіз вартості та часу: Оцінка витрат часу та ресурсів, необхідних для впровадження та підтримання різних методів збору даних.
5. Розробка рекомендацій: На основі отриманих результатів розробка рекомендацій щодо вибору оптимального методу чи комбінації методів для певного типу бізнесу чи завдання.
6. Прогнозування тенденцій: Визначення можливих майбутніх тенденцій у сфері збору даних про ціни та розробка стратегій, які можуть бути актуальними у майбутньому.
7. Дослідження має на меті внести вагому вклад у розвиток практик збору інформації про ціни, що стане корисним для бізнес-спільноти та дослідників у галузі маркетингу та економіки.

1.3.2 Постановка задачі

Для досягнення вказаної мети необхідно вирішити наступні завдання:

Аналіз існуючих методів:

1. Ретельно проаналізувати різні методи збору даних про ціни на товари та послуги.
2. Вивчити переваги та недоліки кожного методу.
3. Визначити, які методи є найбільш популярними та широко використовуються в сучасному бізнес-середовищі.

Оцінка точності та достовірності даних:

1. Визначити метрики точності та надійності для оцінки якості даних, зібраних за допомогою кожного методу.

					<i>KPM.KI.1.415-03.1.4</i>	Арк.
						22
Змн.	Арк.	№ докум.	Підпис	Дата		

2. Провести порівняльний аналіз результатів, визначивши, наскільки вони відповідають дійсним цінам на ринку.

Визначення ефективності в залежності від галузі та масштабу бізнесу:

1. Вивчити, як різні методи впливають на ефективність збору даних в різних галузях.
2. Визначити, чи існують певні особливості ефективності методів для малих, середніх та великих підприємств.

Аналіз вартості та часу:

1. Визначити витрати, пов'язані з впровадженням та підтриманням кожного методу.
2. Порівняти витрати відповідно до етапів використання кожного методу.

Розробка рекомендацій – на основі отриманих результатів розробити рекомендації для бізнес-середовища щодо вибору оптимального методу чи їх комбінації.

Прогнозування тенденцій – визначити можливі тенденції у розвитку методів збору даних про ціни та розробити стратегії для їх врахування в майбутньому.

1.3.3 Об'єкт дослідження

1. Моніторинг інтернет-ресурсів: Аналіз та порівняння ефективності збору даних через моніторинг цін на веб-сайтах та електронних платформах.
2. Аналіз даних з точок продажу: Як техніки, пов'язані з обробкою та аналізом даних з точок продажу, можуть впливати на результати.
3. Інші методи: Врахування будь-яких інших методів, які використовуються в сучасних практиках для отримання інформації про ціни на товари та послуги.

1.3.4 Предмет дослідження

Предметом дослідження є відповіді на наступні запитання:

					<i>KPM.KI.1.415-03.1.4</i>	Арк.
						23
Змн.	Арк.	№ докум.	Підпис	Дата		

1. Які методи збору даних є найбільш поширеними в сучасних умовах?
2. Як точні та достовірні дані, отримані різними методами?
3. Які впливи можуть мати різні методи на результати дослідження цінової політики?
4. Чи існують певні методи, які ефективніше використовувати в певних галузях чи для різних видів товарів та послуг?

1.3.5 Технічне завдання

1. Здійснити огляд наукової літератури та методичних підходів щодо збору цінової інформації.
2. Визначити перелік джерел цінових даних для дослідження, включаючи онлайн-маркетплейси, веб-сайти виробників та роздрібних продавців.
3. Розробити план збору даних, включаючи методи, інструменти та технології, які будуть використовуватися.
4. Зібрати цінові дані з різних джерел та засобів, враховуючи різні види товарів та послуг.
5. Провести аналіз та порівняння зібраних даних, враховуючи їхню актуальність, точність і достовірність.
6. Розробити методичні рекомендації щодо вибору оптимального методу збору цінової інформації для певних сегментів ринку.
7. Підготувати звіт з результатами дослідження та рекомендаціями для подальших дій в галузі маркетингу та ціноутворення.
8. Забезпечити належну обробку та візуалізацію даних для найкращого їхнього розуміння та використання.

1.3.6 Методи дослідження

В рамках даної роботи необхідно ретельно розглянути різні методи збору цінової інформації та провести їх порівняльний аналіз. Нижче наведено детальний опис методів дослідження, які можна використовувати для цієї роботи.

					<i>KPM.KI.1.415-03.1.4</i>	<i>Арк.</i>
						24
<i>Змн.</i>	<i>Арк.</i>	<i>№ докум.</i>	<i>Підпис</i>	<i>Дата</i>		

Аналіз літературних джерел та наукових публікацій:

1. Провести пошук та аналіз наукових статей, книг, підручників та інших літературних джерел, що стосуються методів збору цінової інформації.
2. Визначити основні підходи та техніки, які використовуються в даному дослідженні.

Огляд існуючих онлайн ресурсів:

1. Переглянути і проаналізувати онлайн ресурси, такі як веб-сайти магазинів, маркетплейсів, виробників та інших джерел цінової інформації.
2. Визначити, які ресурси є найбільш актуальними та надійними для збору даних.

Збір цінових даних:

1. Розробити план збору даних, включаючи визначення обсягу та періодичності збору.
2. Використовувати інструменти автоматизації для збору цінової інформації з онлайн ресурсів.

Аналіз та порівняння даних:

1. Обробити зібрані дані, включаючи їхню структуру та очищення від дублікатів та помилок.
2. Провести статистичний аналіз та порівняння цінових даних, отриманих різними методами збору.
3. Визначити рівень точності, достовірності та актуальності кожного методу.

Оцінка ефективності методів:

1. Визначити переваги та недоліки кожного методу збору цінової інформації.
2. Оцінити витрати часу, ресурсів та коштів, які необхідно витратити на кожний метод.
3. Врахувати практичну застосовність результатів для бізнесу та прийняття управлінських рішень.

					<i>KPM.KI.1.415-03.1.4</i>	<i>Арк.</i>
<i>Змн.</i>	<i>Арк.</i>	<i>№ докум.</i>	<i>Підпис</i>	<i>Дата</i>		25

Розробка рекомендацій:

1. Висунути рекомендації щодо вибору оптимального методу збору цінової інформації в залежності від конкретних умов ринку та виду товарів чи послуг.
2. Врахувати можливості автоматизації та вдосконалення обраного методу.

Підготовка звіту:

1. Підготувати детальний звіт з результатами дослідження, включаючи опис методології, отримані дані, висновки та рекомендації.
2. Візуалізувати дані за допомогою графіків та таблиць для зрозумілості та ілюстрації результатів.

Ці методи дослідження допоможуть детально вивчити та порівняти різні способи збору цінової інформації та визначити їхню ефективність для подальшого використання в бізнес-середовищі.

Висновок до першого розділу

У першому розділі проведений аналіз предметної області методів і засобів для збору даних про ціни на товари та послуги дозволили конкретизувати загальну постановку завдання на дипломну роботу, обґрунтувати доцільність дослідження цих методів та рекомендацій щодо їх використання. Також були обрані основні показники, згідно яких буде проводитись оцінка методів в рамках дослідження. Здійснено постановку задачі з зазначенням бажаного результату.

					<i>KPM.KI.1.415-03.1.4</i>	Арк.
						26
Змн.	Арк.	№ докум.	Підпис	Дата		

РОЗДІЛ 2

ТЕХНІЧНЕ ПРОЕКТУВАННЯ ТА МОДЕЛЮВАННЯ СИСТЕМИ

2.1 Дослідження методів збору даних

В кожному кодї використовуються мітки, що відслідковують час, який витрачається на обробку інформації. Протягом виконання парсингу вони записувались у файл і тепер їх можна проаналізувати.

2.1.1 Ефективність методів

Веб-скрапінг:

1. Завантаження та аналіз *HTML*-сторінок: Веб-скрапінг включає в себе процес завантаження *HTML*-сторінок з веб-сайту та подальший аналіз їх вмісту. Це може займати значний час, особливо якщо сторінки великі або складні.
2. Специфіка сайту: Складні структури сторінок, використання *JavaScript* або інші фактори можуть роблять веб-скрапінг більш часоємним.
3. Потреба у парсингу *HTML*: веб-скрапінг вимагає парсингу *HTML*-коду для видобування необхідної інформації, що є менш ефективним порівняно з іншими методами, які отримують структуровані дані.

API:

1. Швидкість отримання даних: *API* надають структуровані дані, що знижує час на обробку та аналіз. Вони призначені для ефективної передачі інформації між сервером та клієнтом.
2. Обмеження *API*: час може залежати від обмежень самого *API*, таких як обмеження швидкості відповіді або обмеження кількості запитів.

HTTP-запити:

1. Комбінація різних методів: *HTTP*-запити можуть використовувати різні методи, включаючи як *API*, так і веб-скрапінг, залежно від конкретної реалізації та потреб дослідження, що є плюсом.

					КРМ.КІ.1.415-03.1.4	Арк.
						27
Змн.	Арк.	№ докум.	Підпис	Дата		

2. Час на обробку відповіді. Час відправлення запиту та час отримання відповіді може бути значущим фактором. Він може залежати від швидкості сервера, який обробляє запит, його навантаження, та загального обсягу даних.

Висновок: веб-скрапінг часто є найчасоємкішим методом через потребу в обробці *HTML*-сторінок та можливість виникнення проблем при аналізі складних структур. *API* надає швидший та більш ефективний спосіб отримання структурованих даних. *HTTP*-запити можуть поєднувати обидва методи, але його швидкість також залежить від різних факторів, таких як швидкість сервера та обмеження *API* (рис. 2.1, рис.2.2).

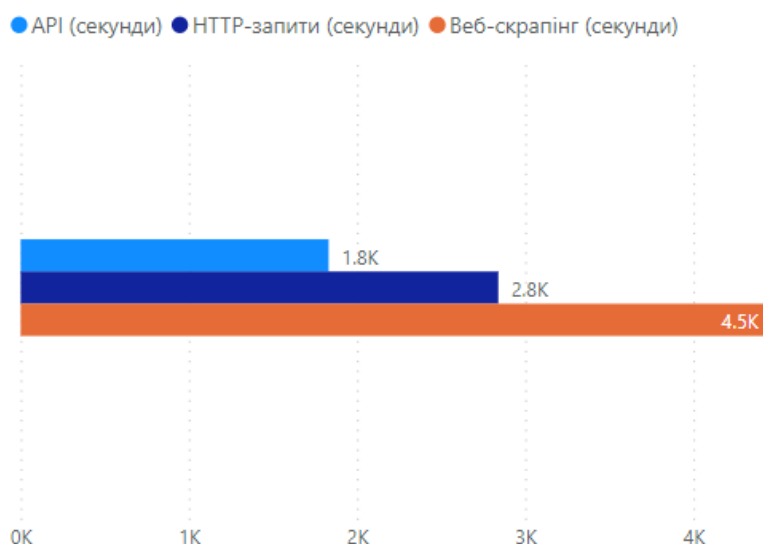


Рис. 2.1 – Порівняння загального часу витраченого на збір даних

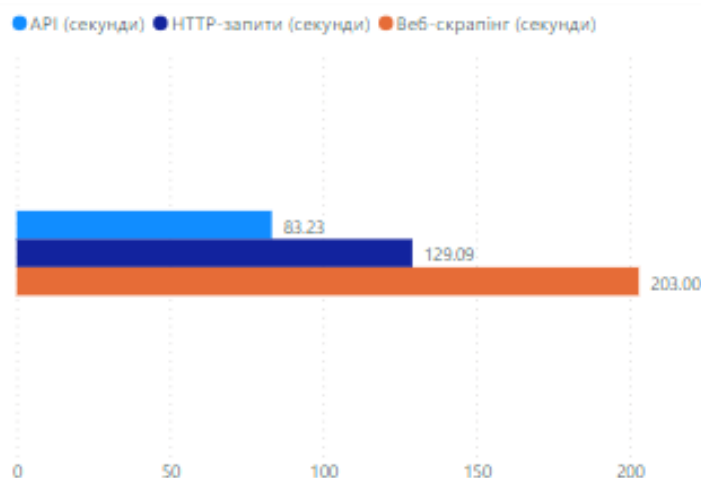


Рис. 2.2 – порівняння середнього часу, витраченого на збір даних по категорії

2.1.2 Специфіка категорій

Різна кількість даних. Різні категорії товарів можуть мати різну кількість доступної інформації. Наприклад, категорія "Кулінарія" може містити багато різних товарів з додатковою інформацією, що може впливати на час парсингу.

Складні структури сторінок. Різні категорії можуть використовувати різні структури сторінок та розміщення даних. Структура сторінок може варіюватися від простої до складної, що впливає на час, необхідний для аналізу.

Величина фотографій та інших мультимедійних даних. Деякі категорії можуть включати багато зображень, відео або інших мультимедійних даних, що також може вплинути на час парсингу, особливо при використанні веб-скрапінгу.

Категорія	API (секунди)	HTTP-запити (секунди)	Веб-скрапінг (секунди)
18+	29	31	59
Алкоголь	118	179	241
Бакалія	63	121	183
Гігієна	58	120	182
Дитячі товари	116	181	239
Для дому	119	177	244
Для тварин	57	122	177
Заморожена продукція	62	118	243
Кава, чай	59	121	181
Квіти, товари для саду та городу	62	57	121
Консерви, соуси, приправи	118	179	245
Кулінарія	121	239	365
Молочні продукти та яйця	61	59	121
М'ясо	178	122	235
М'ясо-ковбасні вироби	62	118	185
Напої	61	119	178
Риба	122	238	355
Сири	59	122	175
Снеки	63	58	123
Солодощі	57	123	179
Фрукти, овочі, соління	128	175	310
Хліб та хлібобулочні вироби	58	61	125
Total	1,831	2,840	4,466

Рис. 2.3 – порівняння часу, витраченого на збір даних в залежності від категорії та методу збору

Загальна ідея полягає в тому, що різні категорії мають свої унікальні особливості, які можуть впливати на час та ефективність процесу парсингу. Наприклад, категорії з більшою кількістю та різноманітністю даних, а також

					КРМ.КІ.1.415-03.1.4	Арк.
						29
Змн.	Арк.	№ докум.	Підпис	Дата		

складнішими структурами сторінок, можуть вимагати більше часу для аналізу та обробки. Оптимізація методів парсингу може враховувати ці специфікації для максимальної ефективності в конкретному контексті дослідження (рис.2.3).

2.1.3 Оптимізація і вибір методу

Комбінація методів. Одним з підходів до оптимізації може бути використання комбінації різних методів парсингу для різних завдань. Наприклад, використання *API* для отримання загальної інформації, а веб-скрапінгу для детального аналізу.

Оптимізація запитів. Для використання *API* може бути важливим оптимізувати запити, вибираючи лише необхідні дані та обмежуючи кількість запитів для зменшення часу відповіді.

Кешування даних. Збереження результатів попередніх запитів у кеші може допомогти уникнути повторних запитів на одні й ті самі дані, що прискорює процес.

Асинхронний парсинг. Використання асинхронних запитів або паралельного парсингу може покращити ефективність, дозволяючи виконувати декілька операцій одночасно.

Оптимізація методу парсингу важлива для максимальної продуктивності та швидкості збору даних. Вибір методу повинен враховувати особливості конкретного дослідження, вимоги до швидкості та обсягу даних, які необхідно зібрати. Комбінування різних методів та використання оптимізаційних стратегій може сприяти ефективному використанню ресурсів при парсингу великих обсягів даних.

2.1.4 Послідовність використання методів

Загальна стратегія. Оптимальний порядок використання методів може варіюватися в залежності від конкретних потреб дослідження. Зазвичай, дослідники починають з використання найбільш швидкого та ефективного методу, такого як *API*, для отримання загальної інформації.

					<i>KPM.KI.1.415-03.1.4</i>	Арк.
						30
Змн.	Арк.	№ докум.	Підпис	Дата		

Веб-скрапінг для деталей. Після отримання загальних даних за допомогою *API*, веб-скрапінг може використовуватися для детального аналізу сторінок та отримання специфічних даних, які можуть бути недоступні через *API*.

Оптимізація запитів. Якщо потрібно використовувати *HTTP*-запити, оптимізація запитів може включати в себе вибір конкретних категорій для оптимального розподілу ресурсів та обмеження кількості запитів.

Ітеративний процес. Послідовність може бути ітеративною, де результати одного методу використовуються для покращення вибору та оптимізації іншого методу.

Правильний порядок використання методів може бути визначений експериментально та залежить від конкретних вимог дослідження. Гнучкість та адаптивність стратегії парсингу можуть допомогти максимізувати швидкість та ефективність процесу збору даних з різних джерел.

2.1.5 Рекомендації

Потреби дослідження. Вибір оптимального методу парсингу повинен враховувати конкретні потреби дослідження. Якщо важлива швидкість отримання даних, *API* може бути перевагою. Якщо важлива гнучкість та доступ до деталей, веб-скрапінг може бути відмінним вибором.

Оптимізація запитів. Для використання *API* важливо оптимізувати запити. Це може включати в себе вибір конкретних параметрів запиту, щоб отримати лише необхідну інформацію, або використання кешування для уникнення повторних запитів.

Управління обмеженнями *API*. Якщо використовується *API*, важливо розуміти та управляти можливими обмеженнями, такими як обмеження кількості запитів за один час чи розмір повернутих даних.

Гнучкість стратегії. Стратегія парсингу повинна бути гнучкою та адаптивною. Варто розглядати зміну методів в залежності від характеру та обсягу даних.

					<i>KPM.KI.1.415-03.1.4</i>	Арк.
						31
Змн.	Арк.	№ докум.	Підпис	Дата		

Аналіз результатів. Після завершення парсингу, важливо аналізувати результати та враховувати їх при прийнятті рішень щодо оптимізації стратегії парсингу для майбутніх завдань.

Ці рекомендації враховують ключові аспекти вибору та оптимізації методу парсингу даних. Гнучкість, аналіз результатів та управління обмеженнями є ключовими компонентами успішної стратегії парсингу.

2.2 Принцип роботи парсерів

У сучасному інформаційному середовищі, де оновлення даних надзвичайно швидкі, процес ручного збору та обробки інформації стає неефективним і часомістким завданням. Крім того, існує ризик пропустити важливі дані в процесі вручного аналізу. З цією метою були розроблені спеціальні програми, відомі як "парсери", які здатні автоматично аналізувати та збирати дані з різних джерел. Вони дозволяють ефективно обробляти великі обсяги інформації, що постійно оновлюються.

Парсинг, в цьому контексті, представляє собою процес автоматизованого збору даних з різних веб-сайтів або сервісів. Результати парсингу можуть бути збережені у файлах або в спеціальному форматі та використані для різних цілей, включаючи наповнення веб-сайтів. Отриманий контент може бути подальше унікалізованим шляхом замовлення рерайту на фріланс-біржах.

Парсери – це програми, створені для автоматичного збору та аналізу інформації, яка відповідає заданим параметрам. Вони здатні відокремлювати та структурувати інформацію відповідно до вимог користувача.

1. Алгоритм дії парсера полягає в наступних кроках:
2. Програма сканує вхідні дані, такі як текст або веб-сторінки, і виділяє певні елементи.
3. Виділені елементи залежать від конкретного завдання та правил, які задає користувач.
4. Правила пошуку зазвичай визначаються за допомогою регулярних виразів, які вказують, що і як шукати.

					<i>KPM.KI.1.415-03.1.4</i>	<i>Арк.</i>
<i>Змн.</i>	<i>Арк.</i>	<i>№ докум.</i>	<i>Підпис</i>	<i>Дата</i>		32

5. На основі зібраної інформації формується звіт або таблиця, яка містить отримані результати.

Робота парсера може бути поділена на три основні етапи: сканування вхідних даних, виділення необхідних даних залежно від заданих правил та створення звіту про знайдені елементи.

В цілому, парсери подібні до того, як людина читає та розуміє текст. Вони використовують синтаксичний аналіз та правила, щоб виділити інформацію з масиву даних. Принцип роботи парсера може відрізнятися в залежності від конкретної реалізації, але основний принцип залишається незмінним.

2.3 Сфери використання парсерів

Парсери знаходять застосування в різних галузях та сферах діяльності, де потрібно здійснювати збір та аналіз даних з веб-сайтів та інших джерел. Давайте розглянемо ці сфери більш детально.

Інтернет-магазини та електронна комерція:

1. Збір товарних позицій: власники інтернет-магазинів використовують парсери для автоматичного збору інформації про товари з інших магазинів або конкурентів. Це допомагає встановити конкурентоспроможні ціни та асортимент товарів.
2. Оновлення інформації: парсери дозволяють автоматично оновлювати інформацію про товари, їх ціни та наявність на сайті.

Ріелтори та автодилери:

1. Моніторинг оголошень: професіонали, які торгують нерухомістю або автотранспортом, використовують парсери для постійного моніторингу оголошень та оновлення інформації про об'єкти.
2. Аналіз цін: парсери допомагають зіставляти ціни на нерухомість або автомобілі в різних джерелах.

Розробка веб-сайтів та контент-менеджмент:

					<i>KPM.KI.1.415-03.1.4</i>	<i>Арк.</i>
<i>Змн.</i>	<i>Арк.</i>	<i>№ докум.</i>	<i>Підпис</i>	<i>Дата</i>		33

1. Збір контенту: при створенні веб-сайтів та блогів, парсери використовуються для автоматичного збору контенту з інших джерел, таких як новинні сайти, форуми або бази даних.

2. Унікальний контент: парсери можуть допомогти створювати унікальний контент шляхом заміни синонімів або автоматичного перекладу тексту.

Маркетинг та пошук клієнтів:

1. Пошук контактів: парсери використовуються для знаходження контактних даних потенційних клієнтів або партнерів.

2. Аналіз соціальних мереж: парсери можуть аналізувати активність користувачів у соціальних мережах, таку як лайки, коментарі та репости.

SEO (пошукова оптимізація):

1. Аналіз посилань: парсери використовуються для аналізу посилань на сайт з пошукових систем та визначення їх ефективності.

2. Моніторинг рейтингів та позицій: парсери можуть відстежувати рейтинги та позиції сайту в пошукових системах.

Оновлення даних в режимі реального часу:

1. Біржі курсів валют: парсери використовуються для постійного оновлення курсів валют та фінансових індексів.

2. Прогнози погоди: парсери допомагають оновлювати інформацію про погодні умови в режимі реального часу.

Сайти-агрегатори:

1. Порівняльний пошук: парсери використовуються на сайтах-агрегаторах для збору та порівняння інформації з різних джерел, таких як сайти з працевлаштуванням, інтернет-магазини, новинні ресурси тощо.

Туристичні фірми:

1. Оновлення інформації про туристичні напрями: туристичні компанії використовують парсери для оновлення інформації про місця відпочинку, ціни, умови та погодні умови.

					<i>KPM.KI.1.415-03.1.4</i>	Арк.
						34
Змн.	Арк.	№ докум.	Підпис	Дата		

Новинні сайти:

2. Збір актуальних новин: новинні сайти використовують парсери для автоматичного збору актуальних новин та інформації з різних джерел.

Аналіз соціальних мереж:

1. Збір даних з соціальних мереж: парсери використовуються для збору даних зі сторінок користувачів у соціальних мережах для подальшого аналізу та використання в рекламних кампаніях.

Ці сфери використання парсерів постійно розширюються, оскільки вони допомагають ефективно обробляти та аналізувати великі обсяги даних з Інтернету, що дозволяє підтримувати актуальну інформацію та конкурентність в різних галузях бізнесу та інших сферах діяльності.

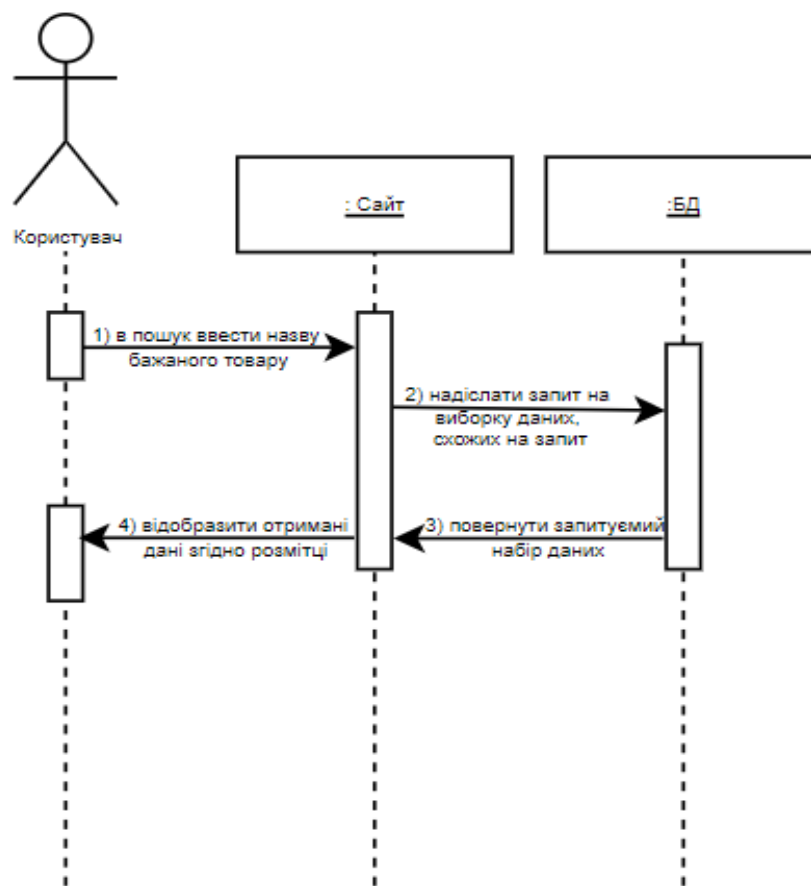


Рис. 2.4 – Діаграма послідовності. Сценарій роботи з сайтами-агрегаторами

Збір та аналіз інформації в Інтернеті часто є часовою та ресурсомісткою задачею. Проте, автоматизовані парсери вирішують цю проблему швидко та

легко. Протягом 24 годин вони здатні просканувати велику частину вмісту в мережі, знаходячи та аналізуючи потрібні дані.

Більшість інформації, яку парсери збирають для загального використання, розміщена на сайтах-агрегаторах. Рисунок 2.4 демонструє найпоширеніший метод отримання інформації з таких сайтів. В подальшому, в залежності від специфікації таких ресурсів, користувачі можуть вибирати різноманітні дії: детально порівнювати інформацію на сайті, переходити до джерела первинної інформації, додавати товари або послуги до кошика безпосередньо на сайті-агрегаторі, і так далі.

Пошукові системи, програми для перевірки на унікальність, інші швидкі інструменти проводять аналіз сотень веб-сторінок, які містять схожий текст. Відповідно, програми-парсери дозволяють знаходити контент для наповнення власних веб-сайтів.

Серед можливих типів даних, які можна отримати за допомогою парсера, включають:

1. Списки товарів, разом із їх характеристиками, фотографіями, описами тощо.
2. Веб-сторінки з помилками, такими як сторінки з кодом помилки 404 або ті, які не мають заголовків (*title*).
3. Інформацію про ціни на товари в конкуруючих магазинах.
4. Рівень активності користувачів на веб-сайтах, включаючи лайки, коментарі та репости.
5. Потенційну цільову аудиторію для рекламних кампаній та продажу товарів чи послуг.

Власники інтернет-магазинів, використовуючи парсери, можуть автоматизувати процес збору даних для карток товарів. Створення описів для товарів часом може бути витратним завданням. Парсери допомагають вирішити такі завдання:

					КРМ.КІ.1.415-03.1.4	Арк.
						36
Змн.	Арк.	№ докум.	Підпис	Дата		

1. Масштабний збір контенту: З конкурентним ринком необхідно збирати та розміщувати величезну кількість інформації на веб-ресурсах, і це вже неможливо зробити вручну.
2. Постійне оновлення контенту: Інформація на веб-сайтах постійно змінюється. Оновлення даних проходить щоденно, і ручне відслідковування їх стає майже неможливим. Автоматизовані парсери є ефективним інструментом для постійного оновлення інформації.

Використання парсерів є сучасним та ефективним способом отримання та оновлення контенту в автоматичному режимі. Переваги використання парсерів включають високу швидкість, точність та ефективність в обробці та перетворенні даних.

2.4 Етапи парсинга

Етапи парсингу даних можуть включати кілька кроків, спрямованих на збір та обробку інформації з веб-сайтів. В рамках цієї дослідницької роботи ми розглядаємо сфери використання парсерів, і приклади застосування парсерів в галузі ціноутворення на різні товари, тому можна виділити наступні основні етапи роботи парсеру:

1. Вибір джерела даних. На першому етапі ви обираєте джерело даних, які ви хочете спарсити. Це може бути веб-сайт, API, соціальна мережа, інтернет-магазин або будь-який інший ресурс.
2. Визначення структури даних. Важливим кроком є визначення того, які саме дані вас цікавлять і яку структуру вони мають. Наприклад, це може бути інформація про товари на інтернет-магазині, статті на новинному сайті, вакансії на сайті з працевлаштуванням тощо.
3. Вибір та налаштування парсеру. Залежно від типу джерела даних і його особливостей, потрібно обрати відповідний парсер або створити свій власний. Парсери можуть бути спеціалізованими для конкретних видів даних або загальні. Важливо враховувати особливості сайту або джерела даних під час налаштування парсера.

					<i>KPM.KI.1.415-03.1.4</i>	<i>Арк.</i>
						37
<i>Змн.</i>	<i>Арк.</i>	<i>№ докум.</i>	<i>Підпис</i>	<i>Дата</i>		

4. Збір даних. На цьому етапі парсер починає збирати інформацію з обраних джерел. Це може включати завантаження *HTML*-сторінок, виконання запитів до *API* або інші методи збору даних. Здебільшого цей процес вимагає обробки багатьох сторінок або запитів.
5. Очищення та обробка даних. Зібрані дані зазвичай потребують очищення та обробки. Це може включати видалення зайвих тегів *HTML*, фільтрацію та перетворення тексту, обрізку зображень або будь-які інші операції, необхідні для покращення якості даних.
6. Збереження даних. Оброблені дані зазвичай зберігаються у відповідному форматі, який може бути базою даних, *CSV*-файлом, *JSON* або іншими форматами. Вибір формату залежить від вашо проекту.
7. Аналіз і використання даних. Зібрані та оброблені дані можуть бути використані для виконання конкретних завдань у вашій роботі. Наприклад, ви можете використовувати дані для аналізу ринку, створення статистичних звітів, автоматизації процесів або інших цілей.
8. Підтримка та оновлення. Після завершення парсингу вам слід регулярно оновлювати дані та підтримувати парсер в робочому стані. Зміни на джерелі даних можуть вимагати адаптації парсера.

Важливо враховувати, що парсинг даних повинен відбуватися відповідно до правил та законів, що регулюють збір та використання інформації в Інтернеті, і з урахуванням етичних аспектів вашого дослідження.

Нижче, на рис. 2.5 зображена взаємодія між об'єктами при зборі даних.

На ньому виділені наступні сутності:

1. «Розробник» - спеціаліст, що відповідає за працездатність скрипту.
2. «Скрипт оновлення БД» - сам скрипт, який після запуску автоматично виконує запрограмовані дії.
3. «Сайт торгової мережі» - об'єкт парсингу, з якого будуть зібрані дані.
4. «БД» - база даних для зберігання отриманої інформації в результаті парсингу.

					<i>KPM.KI.1.415-03.1.4</i>	<i>Арк.</i>
						38
<i>Змн.</i>	<i>Арк.</i>	<i>№ докум.</i>	<i>Підпис</i>	<i>Дата</i>		

3. Використання *CAPTCHA*. Додавання *CAPTCHA* на сторінки сайту може стати перешкодою для автоматизованих парсерів, оскільки вимагає від користувача вирішення завдань, які складні для комп'ютерів, але легкі для людей.

4. Захист *API*. Якщо сайт надає доступ до даних через *API*, використання аутентифікації з використанням токенів або ключів доступу дозволяє контролювати доступ і змінювати обмеження швидкості запитів.

5. Аналітичні інструменти. Сайти можуть використовувати аналітичні інструменти для виявлення підозрілих активностей, таких як несподіваний зріст обсягу запитів з певних *IP*-адрес. Аналіз може допомогти виявити парсери.

6. Юридичний захист. Компанії можуть вживати юридичних заходів для захисту своїх даних від конкурентного парсингу. Це включає укладання угод з конфіденційності, звернення до правопорушень, навіть судові справи у випадках порушення авторських прав.

Захист від конкурентного парсингу стає необхідним, оскільки конфіденційні дані можуть бути використані конкурентами для копіювання бізнес-стратегій, ціноутворення, або для завдання шкоди репутації компанії. Успішний захист включає технічні, юридичні та аналітичні заходи з метою збереження даних і конфіденційності бізнесу.

2.5 Етична сторона парсингу

Етична сторона парсингу виникає з моральних аспектів та питань, пов'язаних із збором та використанням даних з веб-сайтів. Тут розглянемо декілька ключових питань, які виникають при обговоренні етики парсингу:

1. Доступ до публічних даних. Перший аспект етики парсингу стосується доступу до даних, які публічно доступні на веб-сайтах. Багато сайтів розміщують інформацію для загального доступу. Тут питання полягає в тому, чи може парсер збирати та використовувати ці дані без обмежень.

					<i>KPM.KI.1.415-03.1.4</i>	Арк.
						40
Змн.	Арк.	№ докум.	Підпис	Дата		

3. Завдання навмисного завантаження. Важливо пам'ятати, що інтенсивне завантаження сайту через парсер може призвести до перевантаження серверів та негативно вплинути на доступність сайту для інших користувачів. Оцінка впливу парсера на доступність сайту є етичною проблемою.
4. Збір особистої інформації. Парсинг сайтів для збору особистої інформації користувачів без їхньої згоди або обробка цієї інформації в незаконних цілях є явно неетичними діями.
5. Збір інформації для масового надсилання спаму. Використання парсерів для збору електронних адрес інших особистих даних для масового розсилання спаму також є неетичним і часто є незаконним.
6. Доцільність парсингу. Питання ефективності та доцільності парсингу даних також має етичний аспект. Збір даних повинен мати якийсь обґрунтований інтерес або користь для користувачів або суспільства в цілому.
7. Порухення правил сайту. Деякі веб-сайти встановлюють правила щодо доступу та використання свого контенту. Порухення цих правил парсером може бути неетичним.
8. Юридичні обмеження. Кожна країна має свої закони, які регулюють збір та використання даних з Інтернету. Порухення цих законів також може бути неетичним і незаконним.

Для забезпечення етичного парсингу важливо дотримуватися кращих практик, законодавства, моральних норм, звертати увагу на правила сайтів та враховувати вплив на доступність інших користувачів. Також важливо мати чітке розуміння мети та доцільності збору даних для мінімізації можливостей порушення етичних стандартів.

Висновок до другого розділу

Представлені в другому розділі рішення, прийняті на етапі підготовки до дослідження ефективності методів збору даних про ціни на товари та послуги

					<i>KPM.KI.1.415-03.1.4</i>	<i>Арк.</i>
						41
<i>Змн.</i>	<i>Арк.</i>	<i>№ докум.</i>	<i>Підпис</i>	<i>Дата</i>		

стосуються структурних, поведінкових та інших аспектів досліджуваної області. Вони дають змогу перейти до практичної реалізації застосунок для збору інформації і подальших розрахунків з використанням відповідних інструментальних засобів в обраному середовищі розробки. Застосунок спроектований таким чином, що можливий перегляд середнього рівня важкості цих рішень з наступним внесенням змін до реалізації.

					<i>KPM.KI.1.415-03.1.4</i>	<i>Арк.</i>
<i>Змн.</i>	<i>Арк.</i>	<i>№ докум.</i>	<i>Підпис</i>	<i>Дата</i>		42

РОЗДІЛ 3

ПРОЄКТУВАННЯ І ПРОГРАМНА РЕАЛІЗАЦІЯ СИСТЕМИ

3.1 Визначення методів збору даних та джерела інформації

Серед розглянутих вище методів збору інформації, виділимо декілька основних і найпопулярніших:

1. Веб-скрапінг (*Web Scraping*). Веб-скрапінг – це процес автоматичного збору даних з веб-сайтів. Зазвичай використовується для отримання структурованої інформації з *HTML*-коду веб-сторінок. *Python* має багато бібліотек, таких як *Beautiful Soup* та *Scrapy*, які полегшують цей процес.
2. *API* (*Application Programming Interface*). Багато веб-сайтів надають *API* для доступу до своїх даних. Це дозволяє програмістам отримувати доступ до конкретних ресурсів або функціональності веб-сайту. Використання *API* є легальним і зручним методом збору даних.
3. Аналіз *HTTP*-запитів і відповідей. Вивчення мережевого трафіку між вашим браузером і сервером за допомогою інструментів розробника браузера (наприклад, вбудованих в *Chrome* або *Firefox*) дозволяє вам детально вивчати обмін даними.

Для дослідження описаних методів були розглянуті сайти таких торговельних мереж як:

1. «АТБ» – <https://www.atbmarket.com/>.

Переваги:

- великий обсяг даних: АТБ – одна з найбільших мереж, тому може мати великий обсяг даних для парсингу;
- технічна стабільність: сайт добре структурований і технічно стабільний, парсинг може бути ефективним.

Недоліки:

					КРМ.КІ.1.415-03.1.4	Арк.
						43
Змн.	Арк.	№ докум.	Підпис	Дата		

– немає можливості робити *API* – запити та аналіз *HTTP*-запитів і відповідей, доступний лише веб-скрапінг.

2. «Таврія В» – <https://www.tavriav.ua/>.

Переваги:

- регіональна особливість: «Таврія В» має регіональні особливості, тому парсинг може надати унікальну інформацію;
- програми лояльності: «Таврія В» має дані про програми лояльності, які можуть бути корисними для аналізу споживчих звичок.

Недоліки:

- Немає можливості робити *API* – запити та аналіз *HTTP*-запитів і відповідей, доступний лише веб-скрапінг.

3. «Сільпо» – <https://shop.silpo.ua/>.

Переваги:

- розмаїтість інформації: «Сільпо» є великим роздрібним мережевим супермаркетом, і на їхньому веб-сайті доступне різноманіття даних, що включає інформацію про продукти, ціни, акції, графіки роботи магазинів тощо;
- загальна доступність: «Сільпо» є популярним та широкодоступним мережевим магазином, що означає, що багато користувачів відвідують їхній веб-сайт, що забезпечить достатню кількість даних для аналізу та порівняння;
- актуальність даних: у сфері роздрібної торгівлі актуальність даних є важливою, що допомагає дослідженню на сайті Сільпо визначити, наскільки швидко та ефективно можна збирати та оновлювати дані;
- електронні торговельні платформи: оскільки в Сільпо є електронною торговельною платформою, його можна розглядати як організовану та відображену інформацію в онлайн-середовищі;
- доступність для збору: серед конкурентів, таких як «Таврія В», «АТБ», «Копійка» та інших, «Сільпо» надає можливість збору даних за усіма вищезазначеними методами.

					<i>KPM.KI.1.415-03.1.4</i>	<i>Арк.</i>
						44
<i>Змн.</i>	<i>Арк.</i>	<i>№ докум.</i>	<i>Підпис</i>	<i>Дата</i>		

Недоліки:

- авторизація та обмеження: деякі дані можуть бути недоступні без авторизації, і парсинг може порушити умови використання сайту.

Тож, за висновками з фактів, приведених вище, можна дійти висновку, що для дослідження найкраще підходить сайт торгової мережі «Сільпо».

3.2 Збір даних

Розпочавши створення програмного забезпечення для автоматичного збору інформації, був розроблений новий проект у середовищі програмування *pyCharm*. На поточний момент проект містить три скрипти *Python*, які представляють собою алгоритми для виконання пошуку, аналізу, імітації дій звичайного користувача та збору інформації до баз даних. Кожен з цих файлів призначений для роботи з конкретним методом збору: Веб-скрапінг, *API* та Аналіз *HTTP*-запитів і відповідей.

Нижче представлений перелік бібліотек та їхні методи, які використовуються для збору та обробки інформації, а саме:

1. *pandas*: бібліотека для збору, очистки та форматування даних у базах даних;
2. *bs4*: *Python*-бібліотека для вилучення даних з файлів *HTML*, які можна отримати шляхом виконання запитів до сайту;
3. *selenium*: бібліотека, що дозволяє імітувати поведінку звичайного користувача для отримання невидимих частин розмітки та уникнення "зрізання" можливої кількості запитів. Зокрема використовуються підбібліотеки, такі як:
 - *selenium.webdriver.support.ui*;
 - *selenium.webdriver.common.keys*;
 - *selenium.webdriver.common.by*;
 - *time, datetime*: бібліотеки, які дозволяють обробляти всі дані, що стосуються часу, включаючи приведення часу до одного формату, розрахунок часу виконання запиту та інше;

					КРМ.КІ.1.415-03.1.4	Арк.
						45
Змн.	Арк.	№ докум.	Підпис	Дата		

- назва товару;
- ціна;
- вага;
- фото;
- додаткові параметри.

Ці дані збиралися, оброблялися за допомогою бібліотеки *pandas* і додавалися до тимчасової міні-бази даних у форматі *Data Frame*. На закінчення дані з цієї міні-бази можна об'єднати та додати у csv файл або створити базу даних, в якій можна зберігати та оновлювати інформацію.

```
def main():
    start = time.time()
    df = pandas.DataFrame()
    headers = {
        "sec-ch-ua": "Chromium;v=94,\"GoogleChrome\";v=94,\"Not A Brand\";v=99\",",
        "X-sessionid": "86e027e15823ab6e933e7909991ead9e",
        "sec-ch-ua-mobile": "0",
        "User-Agent": "Mozilla / 5.0 (WindowsNT10.0; WOW64) AppleWebKit / 537.36 (KHTML, like Gecko) Chrome / 94.0.4600.61 Safari / 537.36",
        "Content-Type": "application / json; charset = UTF-8",
        "accept": "application / json, text / plain, */*",
        "user-uid": "5c07730b-70e9-9901-a971-c625280c2e1",
        "sec-ch-ua-platform": "Windows",
        "origin": "https://shop.silpo.ua",
        "sec-fetch-site": "same-site",
        "sec-fetch-mode": "cors",
        "sec-fetch-dest": "empty",
        "referrer": "https://shop.silpo.ua/",
        "Accept-Encoding": "gzip, deflate, br",
        "Accept-Language": "en, ru;q = 0.9, en - US;q = 0.8"
    }

    data = {
        "method": "GetSimpleCatalogItems",
        "data": {
            "merchantId": 1,
            "basketId": "8e990bbd-b9cd-4273-bcaf-52d19a418a5",
            "deliveryType": 1,
            "filter": 2000,
            "From": 1,
            "businessId": 1,
            "top": 32,
            "ingredients": False,
            "categoryId": 22,
            "rangeFilters": {},
            "multiFilters": {},
            "universalFilters": {},
            "categoryFilter": {},
            "Promos": {}
        }
    }

    url = "https://api.catalog.ecom.silpo.ua/api/2.0/exec/EcomCatalogGlobal"
```

Рис. 3.2 – Приклад параметрів, що передавалися до API Сільпо

3.3 Опис програмного комплексу для програміста

3.3.1 Збір даних за допомогою API

Метод *data_process* використовується для обробки отриманої відповіді від API. На початку створюється Дата Фрейм, що міститиме в собі оброблений результат (лістинг 3.1).

Лістинг 3.1 Фрагмент коду метода *data_process*

```
def data_process(response):
    df = pandas.DataFrame()
```

Далі відбувається перевірка на коректну відповідь (лістинг 3.2). Про це свідчить спеціальний системний код, що дорівнює 200.


```

"sec-ch-ua-platform": "Windows",
"Origin": "https://shop.silpo.ua",
"Sec-Fetch-Site": "same-site",
"Sec-Fetch-Mode": "cors",
"Sec-Fetch-Dest": "empty",
"Referer": "https://shop.silpo.ua/",
"Accept-Encoding": "gzip, deflate, br",
"Accept-Language": "en, ru;q = 0.9, en - US;q = 0.8"
}

data = {
  "method": "GetSimpleCatalogItems",
  "data": {
    "merchantId": 1,
    "basketGuid": "8e99dbbd-b9cd-4273-bcaf-52d19a4118a5",
    "deliveryType": 1,
    "filialId": 2030,
    "From": 1,
    "businessId": 1,
    "To": 32,
    "ingredients": False,
    "categoryId": 22,
    "RangeFilters": {},
    "MultiFilters": {},
    "UniversalFilters": [],
    "CategoryFilter": [],
    "Promos": []
  }
}

```

Для запитів конкретно до *API*, а не до сайту, використовується нижче наведене посилання. Для проходження по всім можливим категоріям використовується цикл, в рамках якого змінюється номер бажаної категорії, відправляється запит та пересилка результату запитів до методу, що обробляє та зберігає отримані дані (лістинг 3.5).

Лістинг 3.5 Фрагмент коду циклу

```

url = "https://api.catalog.ecom.silpo.ua/api
      /2.0/exec/EcomCatalogGlobal"
for category_id in range(0, 5000):
    data["data"]["categoryId"] = category_id
    response = requests.post(url, headers=headers,
                             data=json.dumps(data))
    df = df.append(data_process(response))
df.to_csv("silpo_result.csv")
print("Time spented on " + str(len(df)) + " elements:
      " + str(time.time() - start))
if __name__ == '__main__':
    main()

```

						<i>Арк.</i>
						49
<i>Змн.</i>	<i>Арк.</i>	<i>№ докум.</i>	<i>Підпис</i>	<i>Дата</i>		

3.3.2 Веб-скрапінг

Одразу бачимо функцію, що, згідно з її назвою, відповідає за отримання та першу обробку коду сайту, з якого збирається інформація (лістинг 3.6).

Лістинг 3.6 Фрагмент коду функції першої обробки сайту

```
def get_site_code(driver): # получить код сайта сейчас
    res = driver.execute_script("return
        document.documentElement.outerHTML")
    soup = BeautifulSoup(res, 'lxml')
    return soup
```

Далі розглянемо метод *find_categories*. В ньому відбувається аналіз та пошук елементів, що відповідають за посилання категорій. Ця інформація допоможе переміщатися безпосередньо по згуртованим за спільною категорією товарам та збирати інформацію одразу з декількох продуктів замість одного (лістинг 3.7).

Лістинг 3.7 Фрагмент коду метода *find_categories*

```
def find_categories(soup):
    code = soup.find_all('li', class_='category-menu__item')
    categories = []
    for item in code:
        categorie = re.search('class="category-menu__link"
            href="(.*)">(.*?)<', str(item))
        subcategories = re.findall('<a class="submenu__link"
            href="(.*?)">(.*?)</a>', str(item))
        categories.append({"categorie": categorie.group(2),
            "url": categorie.group(1),
            "subcategories": subcategories})
    print(categories)
    return categories
```

Для більш ефективного переміщення по категоріям на сайті є пагінація – це розбивка величезного масиву даних (зазвичай однотипного) на невеликі за обсягом сторінки. Причому з відображенням нумерованої навігації по ним. Найчастіше пагінація сторінок застосовується в інтернет-магазинах – в категоріях та розділах. Також вона зустрічається в блогах, де публікується дуже багато статей.

Для того, аби одразу дізнатися кількість і номера сторінок з коду за спеціальним шаблоном було виділено конкретний елемент, в якому міститься максимальна кількість сторінок (лістинг 3.8).

Лістинг 3.8 Фрагмент коду пагінації

```
def get_pagination(soup):
    page = 1
    max_page_index = soup.find_all('a', class_ = 'product-
        pagination__link')
    if max_page_index != []:
        page = re.search('a class="product-pagination__link"
            data-page="[0-9]" href="(.*)">([0-9])',
            str(max_page_index)).group(2)
    return int(page)
```

Перейшовши на конкретну сторінку та отримавши її код можна приступати до пошуку потрібної інформації (лістинг 3.9). Так, як на сторінках – категоріях знаходиться від одного до декількох десятків товарів, використовується цикл, в рамках якого для кожного з елементів знаходяться потрібні категорії інформації. А результат такого пошуку записується у заздалегідь створений Дата Фрейм.

Лістинг 3.9 Фрагмент коду пошуку

```
def find_element(soup):
    items = soup.find_all('article', class_='catalog-item js-
        product-container')
    result = pd.DataFrame()
    for item in items:
        title = re.search('href="\~/product\/(.*?)\/(.*?)">(.*?)<',
            str(item))
        price = re.search('data class="product-price__top"
            value="(.*?)"', str(item)).group(1)
        currency = re.search('abbr class="product-price__currency-
            abbr" title="(.*?)"', str(item)).group(1)
        unit = re.search('span class="product-price__unit"
            \n(\s*)(.*)', str(item)).group(2).split(' ')[0]
        row = {'city_id': title.group(1),
            'product_id': title.group(2),
            'product_name': title.group(3)[1:],
            'price': price,
            'currency': currency,
            'unit': unit,
            'shop' : "silpo_scrap"
        }
        result = result.append(row, ignore_index= True)
    print(result)
    return result
```

					<i>KPM.KI.1.415-03.1.4</i>	<i>Арк.</i>
<i>Змн.</i>	<i>Арк.</i>	<i>№ докум.</i>	<i>Підпис</i>	<i>Дата</i>		51

Одним з основних методів являється *collect_info* (лістинг 3.9). З його допомогою можна пройтись по всім сторінкам обраної категорії та зібрати їх код для подальшої обробки. Через наявність на сайті перевірки на вік, була створена спеціальна зміна, що перевіряє сторінку на наявність спеціального поп-апу і, якщо він є, то передає керування функції *alcohol_check*, в рамках якої імітується поведінка звичайного користувача і відбувається клік на кнопку погодження з твердженням у поп-апі.

Лістинг 3.10 Фрагмент коду метода *collect_info*

```
def collect_info(driver, url):
    result = pd.DataFrame()
    driver.get(url)
    soup = get_site_code(driver)
    alco = alcohol_check(driver, soup)
    if alco == 1:
        soup = get_site_code(driver)
    else:
        pass
    pagination = get_pagination(soup)
    for i in range (1, pagination + 1):
        driver.get(url + '?page=' + str(i))
        soup = get_site_code(driver)
        result = result.append(find_element(soup))
    return result

def alcohol_check(driver, soup):
    is_alco = 0
    button = soup.find_all('button', class_ = 'custom-blue-btn alcohol-
                                modal__submit')
    fisical_button = driver.find_elements(By.XPATH, "//*[contains(text(),
                                'Так, мені вже є 18')]")
    if fisical_button != '':
        for but in fisical_button:
            but.click()
            is_alco = 1
    return is_alco
```

Ну і частина коду, що запускає, збирає та до оброблює кінцеві файли в одному великому для подальшого його запису в БД (лістинг 3.11).

Лістинг 3.11 Фрагмент коду обробки, запуску та збіру кінцевого файлу

```
if __name__ == '__main__':
    start = time.time()
    driver = webdriver.Chrome(r'C:\Users\User
                                \Downloads\chromedriver.exe')
    url = 'https://shop.silpo.ua/'
    df = pd.DataFrame()
```

					КРМ.КІ.1.415-03.1.4	Арк.
						52
Змн.	Арк.	№ докум.	Підпис	Дата		

```

driver.get(url)
soup = get_site_code(driver)
categories = find_categories(soup)
for categorie in categories:
    print(categorie)
    if categorie['subcategories'] == []:
        page_url = url + categorie['url']
        result = collect_info(driver, page_url)
        if result.empty != True:
            result['category'] = categorie['categorie']
            result['subcategory'] = 'empty'
            result['date'] = date.today()
            df = pd.concat([df, result])
    for elem in categorie['subcategories']:
        page_url = url + elem[0]
        result = collect_info(driver, page_url)
        if result.empty != True:
            result['category'] = categorie['categorie']
            result['subcategory'] = elem[-1]
            result['date'] = date.today()
            df = pd.concat([df, result])
df.to_csv('silpo_scrap_result.csv')
print("Time spendened on " + str(len(df)) + " elements:
      " + str(time.time() - start))
print(df)

```

3.2.3 Аналіз *HTTP*-запитів і відповідей

Для цього нам знадобиться схожий скрипт на той, що ми використовували при зборі даних за допомогою *API* запитів, проте буде декілька ключових відмінностей.

По-перше – посилання. В цьому скрипті ми будемо звертатися не до спеціалізованого *API*-посилання, а до звичайного, яке можна побачити просто відвідуючи сайт – <https://shop.silpo.ua/category/>.

```
url = "https://shop.silpo.ua/category/"
```

По-друге – методи, які використовувались для обробки відповідей (лістинг 3.12). Якщо при запиті *API* повертало *json* відповідь, в якій вже були розмічені поля і їх можна було отримати за допомогою обробки *json* бібліотекою, то зараз потрібно скористатися бібліотекою *bs4*, яка допоможе дістати данні з *html* розмітки сторінки наступним чином:

Лістинг 3.12 Фрагмент коду метода обробки відповідей

```

product_id = items['data-product-id']
name = items.find('span', class_='product-card__name').text.strip()
unit = items.find('span', class_='product-card__unit').text.strip()

```

					КРМ.КІ.1.415-03.1.4	Арк.
						53
Змн.	Арк.	№ докум.	Підпис	Дата		

```

price = items.find('span', class_='product-card__price').text.strip()
main_image = items.find('img', class_='product-card__img')['src']
parameters = []
parameters_element = items.find('div', class_='product-card__parameters')
if parameters_element:
    parameters = [param.text.strip() for param in parameters_element.
                  find_all('span', class_='parameter')]

```

Висновки до третього розділу

У даному розділі вирішено завдання розробки системи збору даних, спрямованої на дослідження методів збору даних про ринкові ціни. Проектування системи передбачало розробку скриптів, за допомогою яких можна проводити дослідження методів збору даних. Програмна реалізація системи відповідає всім вимогам та цілям дослідження. Під час розробки враховано висновки з попередніх розділів диплому, що дозволило вдосконалити функціонал системи. Застосування сучасних технологій програмування забезпечило зручний доступ до інформації та високу продуктивність системи.

					<i>КРМ.КІ.1.415-03.1.4</i>	<i>Арк.</i>
<i>Змн.</i>	<i>Арк.</i>	<i>№ докум.</i>	<i>Підпис</i>	<i>Дата</i>		54

РОЗДІЛ 4 ЕКОНОМІЧНІ РОЗРАХУНКИ

4.1 Організаційно-економічне та маркетингове обґрунтування проекту

Перед початком цього розділу було проведено дослідження більшості існуючих методів парсингу та інструментів, що можуть для цього використовуватись. Нижче, на таблиці 4.1, можна детально ознайомитись з прикладами.

Таблиця 4.1

Методи збору даних і їх порівняння

Метод збору даних	Переваги	Недоліки
<i>HTML Parsing</i> з використанням <i>BeautifulSoup</i> (<i>Python</i>)	<ol style="list-style-type: none"> Простий у використанні. Добре впорядковує та розуміє <i>HTML</i>-структуру. 	<ol style="list-style-type: none"> Потребує додаткового коду для навігації між сторінками. Може не коректно обробляти <i>HTML</i>.
<i>XPath</i> (за допомогою мови програмування, такої як <i>Python</i>)	<ol style="list-style-type: none"> Дозволяє точно вказати шлях до елементів. Ефективний для складних структур сторінок. 	<ol style="list-style-type: none"> Вимагає знань <i>XPath</i>-виразів, що може бути складним для новачків. Вразливий до змін у структурі <i>HTML</i>.
<i>Regular Expressions</i> (<i>Regex</i>)	<ol style="list-style-type: none"> Дуже гнучкий та потужний для пошуку та вилучення даних. 	<ol style="list-style-type: none"> Складний для використання при роботі з великими обсягами <i>HTML</i>. Може викликати помилки при зміні структури.

Аналог	Переваги	Недоліки
<i>ParseHub</i>	Має інтуїтивний інтерфейс для конфігурації правил парсингу.	Безкоштовна версія обмежена обсягами даних та швидкістю парсингу.
<i>DiffPlex</i> (<i>Luminati</i>)	Використовує мережу проксі для анонімного та розподіленого парсингу.	Потребує платформу <i>Luminati Proxy</i> для розширеного парсингу, що може бути дорогим для користувачів.
<i>Common Crawl</i>	Надає доступ до великого обсягу даних, які збираються автоматично.	Деякі дані можуть бути застарілими, оскільки <i>Common Crawl</i> робить обмежену кількість веб-краулінгу щомісяця.
<i>WebHarvy</i>	Має інтерфейс, який дозволяє користувачам визначати правила екстракції без програмування.	Обмежений у порівнянні з програмними інструментами щодо розширеної автоматизації та складних завдань.

Зважаючи на розглянуті інструменти у даній магістерській роботі проведено комплексне дослідження методів збору даних про ціни на товари та послуги для підвищення їх ефективності та аналізу даних з різноманітних джерел в Інтернеті.

Для дослідження методів було проаналізовано доступні алгоритми та інструменти збору даних про ціни на товари та послуги.

Мета проекту: полягає в оцінці та порівнянні різних методів, використовуваних для збору інформації про ціни на товари та послуги в сучасному бізнес-середовищі.

Очікувані результати впровадження проекту:

					KPM.KI.1.415-03.1.4	Арк.
						58
Змн.	Арк.	№ докум.	Підпис	Дата		

- пошук оптимального методу парсингу для конкретних завдань;
- зменшення витрат часу та ресурсів на збір та обробку даних;
- підвищення точності та надійності отриманих даних.

Класифікаційна оцінка різновиду проекту:

- клас – монопроект, тому що проект орієнтований на певне середовище застосунку й він невеликий;
- тип – технічний (проект направлено на отримання технічних знань та навиків);
- вид – інноваційний;
- тривалість – коротко-терміновий, тому що створюється за порівняно малі строки;
- складність – середній;
- рівень – галузевий.

Етапи виконання проекту:

- підготовчий етап (серпень-вересень) – аналіз предметної галузі, аналіз аналогічних досліджень і систем, вибір засобів реалізації, збір технічних і економічних даних необхідних для проекту;
- основний (вересень-листопад) – проведення експерименту, опис методики дослідження, розробка структурних схем, створення концепту потрібної системи;
- заключний (листопад-грудень) – техніко-економічне обґрунтування проекту, аналіз отриманих результатів, тестування, виявлення та виправлення недоліків.

4.2 Економічні розрахунки проекту

4.2.1 Розрахунок трудомісткості розробки методом сітьового планування

Сітьове планування полягає у створенні логічних діаграм послідовності виконання проектних робіт.

					<i>KPM.KI.1.415-03.1.4</i>	Арк.
						59
Змн.	Арк.	№ докум.	Підпис	Дата		

Основні цілі його проведення:

1. Планування часових та ресурсних витрат.

- визначення, скільки часу буде потрібно для завершення кожного завдання та проекту в цілому;
- встановлення, які ресурси (людські, фінансові, матеріальні) будуть задіяні під час виконання проекту.

2. Оцінка реалістичних термінів виконання:

- уникнення неправдоподібних обіцянок щодо термінів виконання проекту;
- виявлення можливих затримок та факторів, які можуть вплинути на графік.

3. Визначення критичних завдань та шляхів:

- визначення завдань, які мають найбільший вплив на завершення проекту;
- виявлення критичного шляху, на якому будь-яке затримання призведе до затримки всього проекту.

4. Прийняття рішень щодо ресурсів:

- визначення потреб у ресурсах для ефективного планування і використання команди;
- оптимізація розподілу ресурсів для досягнення оптимальної продуктивності.

5. Оцінка можливостей для управління ризиками:

- визначення можливих ризиків та створення буферного часу для компенсації можливих затримок.

6. Моніторинг та контроль:

- забезпечення ефективного моніторингу та контролю за ходом виконання проекту;
- виявлення відхилень від плану та прийняття вчасних корективних заходів.

Сітьовий графік будується за даними таблиці 4.3 і представлений на

					<i>KPM.KI.1.415-03.1.4</i>	<i>Арк.</i>
						60
<i>Змн.</i>	<i>Арк.</i>	<i>№ докум.</i>	<i>Підпис</i>	<i>Дата</i>		

рисунку 4.1.

Розрахунок параметрів сітьового графіку:

- ранній строк здійснення роботи, $tro(i,j)$;
- тривалість роботи, $tTF(i,j)$;
- ранній строк здійснення події, $tr(i)$;
- пізній строк здійснення події, $tTF(i)$;
- повний резерв часу роботи, $Rп(i,j)$.

Таблиця 4.3

Склад робіт проекту і їх тривалість

№ коду роботи	Назва роботи	T (дні)	№ коду попередньої роботи
0-1	Постановка завдання	7	----
1-2	Вивчення предметної області	13	0-1
1-3	Збір та аналіз даних	4	0-1
2-4	Створення ТЗ	10	1-2
3-4	Фіктивна робота	0	2-4
4-5	Розробка моделі дослідження	30	3-5, 4-5
5-6	Проектування структури і алгоритмів програмної системи	29	5-6
6-7	Організація дослідження ефективності методів збору даних	27	5-6
6-8	Проведення імітаційних експериментів	15	5-6
7-9	Обробка результатів дослідження	12	6-7
8-9	Створення супроводжуючої документації	6	6-8
9-10	Оцінка результатів проекту й підведення підсумків	13	8-11, 9-11
5-6	Проектування структури і алгоритмів програмної системи	29	5-6

Повний резерв

Робота (i,j)	Кількість попередніх робіт	Трива- лість t_{ij}	Ранні терміни		Пізні терміни		Резерви часу:
			Поча-ток t_{ij} Р.Н.	Закін- чення t_{ij} Р.О.	Початок t_{ij} П.Н.	Закін- чення t_{ij} П.О.	Повний R_{ij} П
(0,1)	0	7	0	7	0	7	0
(1,2)	1	13	7	20	18	31	11
(1,3)	1	4	7	11	7	11	0
(2,4)	1	10	21	31	22	32	1
(3,4)	1	0	11	11	11	11	0
(4,5)	1	30	1	31	12	42	11
(5,6)	2	29	42	71	42	71	0
(6,7)	1	27	71	98	71	98	0
(6,8)	1	15	71	86	71	86	0
(7,9)	1	12	98	110	98	110	0
(8,9)	1	6	86	92	104	110	18
(9,10)	1	13	110	123	110	123	0

4.2.2 Визначення трудомісткості розробки програмного продукту

Розрахунок трудомісткості програмного продукту (ПП), що розробляється, проводиться за формулою (4.1):

$$T_{ПП} = T_{ТЗ} + T_{ТП} + T_{РП} + T_{ВН}, \quad (4.1)$$

де

$T_{ТЗ}$ – трудомісткість розробки технічного завдання створення ПП;

$T_{ТП}$ – трудомісткість розробки технічного проекту ПП;

					КРМ.КІ.1.415-03.1.4	Арк.
						63
Змн.	Арк.	№ докум.	Підпис	Дата		

T_{PP} – трудомісткість розробки робочого проекту ПП;

T_{BH} – трудомісткість впровадження розробленого ПП.

Розраховуючи трудомісткість, необхідно враховувати, що програмний засіб (ПЗ), ПЗ введення інформації з $V_0 = 567$ умовних машинних команд із трудомісткістю приблизно $T_p = 285$ чол/год.

Трудомісткість розроблювального ПП визначається на кожному етапі окремо на підставі трудомісткості аналога з урахуванням складності розробки, ступеня новизни й ступеня використання в розробці стандартних модулів на підставі формул:

$$T_{m3} = T_p \cdot L_i \cdot K_n; \quad (4.2)$$

$$T_{mn} = T_p \cdot L_2 \cdot K_n; \quad (4.3)$$

$$T_{pn} = T_p \cdot L_3 \cdot K_n \cdot K_m; \quad (4.4)$$

$$T_{bn} = T_p \cdot L_4 \cdot K_n, \quad (4.5)$$

де:

T_p – укрупнена норма часу на розробку аналога ПЗ, чол/год (табл. 4.3), що корегується поправочним коефіцієнтом, що враховує умови розробки ПЗ, тобто в умовах комп'ютера: $K_k = (0,7/0,8)$;

L_j – питома вага даного етапу розробки з урахуванням ступеню новизни. Код ступеню новизни даного ПП – В, отже $L_1 = 0,12$; $L_2 = 0,11$; $L_3 = 0,61$; $L_4 = 0,16$.

K_n – поправочний коефіцієнт, що враховує ступінь новизни ($K_n = 0,7$);

K_m – поправочний коефіцієнт, що враховує ступінь використання в розробці типових програм ($K_m = 0,8$).

Розрахунок трудомісткості розробки технічного завдання розраховується за формулою (6.2):

$$T_{m3} = 285 \cdot 0,12 \cdot 0,7 = 24 \text{ чол/год.}$$

					<i>KPM.KI.1.415-03.1.4</i>	Арк.
						64
Змн.	Арк.	№ докум.	Підпис	Дата		

Розрахунок трудомісткості розробки технічного проекту розраховується за формулою (4.3):

$$T_{mn} = 285 \cdot 0,11 \cdot 0,7 = 22 \text{ чол/год.}$$

Розрахунок трудомісткості розробки робочого проекту проводиться за формулою (4.4):

$$T_{pn} = 285 \cdot 0,61 \cdot 0,7 \cdot 0,8 = 98.$$

Розрахунок трудомісткості впровадження проводиться за формулою (4.5):

$$T_{vn} = 285 \cdot 0,16 \cdot 0,7 = 32.$$

В таблиці 4.3 приведено розрахунок трудомісткості розробки проекту.

N_{mz} – 2 стор., N_{mn} – 8 стор., N_{pn} – 20 стор., N_{vn} – 30 стор., N_{pz} – 128 стор..

Таблиця 4.5

Трудомісткість розробки проекту

Найменування етапів	Розрахунок, чол./год.
1. Технічне завдання	$T_{ТЗ} = 24;$ $T_{КК} = 0,7 \cdot N_{ТЗ} = 0,7 \cdot 2 = 1,4;$ $T_{НК} = 0,15 \cdot N_{ТЗ} = 0,15 \cdot 2 = 0,3;$
2. Розробка ТП (алгоритму й блок-схеми)	$T_{ТП} = 22$ $T_{КК} = 0,7 \cdot N_{ТП} = 0,7 \cdot 8 = 5,6;$ $T_{НК} = 0,15 \cdot N_{ТП} = 0,15 \cdot 8 = 1,2;$
3. Розробка робочого проекту (складання програм)	$T_{рп} = 98;$ $T_{КК} = 0,7 \cdot N_{рп} = 0,7 \cdot 20 = 14;$ $T_{НК} = 0,15 \cdot N_{рп} = 0,15 \cdot 20 = 3;$

розробку. Розрахунок проводиться за формулою (4.7):

$$Ц = K \cdot C + Пр, \quad (4.7)$$

де

$Ц$ – коефіцієнт обліку витрат на виготовлення дослідного зразка як продукції виробничо-технічного призначення;

C – витрати на розробку програмної продукції (кошторисна собівартість);

$Пр$ – нормативний прибуток.

Розрахунок нормативного прибутку проводиться за формулою (4.8):

$$Пр = (C - C_m) \cdot P_n / 100, \quad (4.8)$$

де

P_n – норматив рентабельності, (рекомендується 25 %);

C_m – матеріальні витрати, грн.

Форма для розрахунку матеріальних витрат наведена в табл. 4.4.

Таблиця 4.6

Розрахунок матеріальних витрат

Назва матеріальних витрат	Тип, модель	Кількість	Ціна за одинцю, грн	Сума, грн.
Папір	A4	1	150,00	150,00
Флеш-накопичувач	USB Флешка 2.0 16Gb Xiaomi	1	200,00	200,00
Лазерний диск	Samsung	2	8,00	16,00
Роздруківка	-	114	2,50	285,00
Всього:				651,00
$K_{TP} = 0,1$				65,1
Разом:				716,1

					KPM.KI.1.415-03.1.4	Арк.
Змн.	Арк.	№ докум.	Підпис	Дата		67

Витрати, зв'язані з використанням Інтернету, визначають по формулі (4.9):

$$C_{\text{інт}} = n \cdot p, \quad (4.9)$$

Де

n – час використання Інтернету (30 днів = 1 місяць);

p – вартість підключення Інтернету в день.

Ціна Інтернету в місяць 225 грн.

$$C_{\text{інт}} = 120 \cdot 7,5 = 900 \text{ грн.}$$

Основна заробітна плата: у статтю входить основна заробітна плата виконавця, безпосередньо зайнятого розробкою даного ПП, з урахуванням його посадового окладу і часу участі в розробці. Розрахунок ведеться по формулі:

$$C_{30} = \sum Z_i \cdot \tau_i / \alpha, \quad (4.10)$$

де

Z_i – середньомісячний оклад i -того виконавця, грн;

α – середня кількість робочих днів в місяці (рекомендовано 21-22);

τ_i – трудомісткість робіт, що виконуються i -тим виконавцем, чол/дні

У розробці задіяні постановник задачі, розробник, середньомісячний оклад яких складає 10000 грн. та 12000 грн. відповідно.

$$C_{301} = 10000 \cdot 123 \cdot 0,1 / 22 = 5590,9 \text{ грн.}$$

$$C_{302} = 12000 \cdot 123 \cdot 0,9 / 22 = 60381,8 \text{ грн.}$$

$$C_{30} = 5590,9 + 60381,8 = 65972,7 \text{ грн.}$$

Додаткова заробітна плата: у статті враховуються всі виплати безпосереднім виконавцям за час, що не пропрацювало на виробництві, зокрема: оплата чергових відпусток, компенсації за невикористану відпустку,

					<i>KPM.KI.1.415-03.1.4</i>	Арк.
						68
Змн.	Арк.	№ докум.	Підпис	Дата		

оплата пільгового годинника підліткам та ін.

Розрахунок ведеться по формулі:

$$C_{зд} = C_{зо} \cdot \alpha_{д}, \quad (4.11)$$

де

$\alpha_{д}$ – коефіцієнт відрахувань та заробітну плату $\alpha = 0,1$

$$C_{зд} = 65972,7 \cdot 0,1 = 6597,27 \text{ грн.}$$

Відрахування на соціальне страхування: у статті враховуються відрахування до бюджету соціального страхування за встановленим законодавством тарифом від суми основної і додаткової заробітної плати, тобто:

$$C_{сс} = \alpha_{сс} (C_{зо} + C_{зд}), \quad (4.12)$$

де

$\alpha_{сс}$ – коефіцієнт відрахувань на соціальне страхування (рекомендується 22 %).

$$C_{сс} = 0,22 \cdot (65972,7 + 6597,27) = 15965,4 \text{ грн.}$$

Накладні витрати: у статті враховуються витрати на загальногосподарські витрати, позавиробничі (комерційні) витрати і витрати на управління. Накладні витрати визначають в процентному відношенні до основної заробітної плати, тобто:

$$C_{н} = K_{н} \cdot C_{зо}, \quad (4.13)$$

де

$K_{н}$ – коефіцієнт накладних витрат (0,5).

$$C_{н} = 0,5 \cdot 65972,7 = 32986,35 \text{ грн.}$$

					КРМ.КІ.1.415-03.1.4	Арк.
						69
Змн.	Арк.	№ докум.	Підпис	Дата		

Результати розрахунку кошторисної вартості

Найменування статті	Кошторисна собівартість, грн.	Питома вага, %
Матеріали	716,1	0,58
Internet	900,0	0,73
Основна заробітна плата	65972,7	54
Додаткова заробітна плата	6597,27	5,4
Відрахування на соціальне страхування	15965,4	13,06
Накладні витрати	32986,35	27
Разом	123 137,82	100

Визначимо нормативний прибуток по формулі:

$$P_p = (C - C_m) \cdot P_n / 100, \quad (4.14)$$

де

P_n – норматив рентабельності (рекомендується 25 %);

C_m – матеріальні витрати, грн.;

C – витрати на розробку програмної продукції (кошторисна собівартість);

$$P_p = ((123137,82 - 716,1) \cdot 25) / 100 = 30605,43 \text{ грн.}$$

Визначимо ціну програмного продукту за формулою:

$$Ц = K \cdot C + P_p, \quad (4.15)$$

де

K – коефіцієнт обліку витрат на виготовлення дослідного зразка ПП як продукції виробничо-технічного призначення ($K = 1,1$);

					<i>KPM.KI.1.415-03.1.4</i>	Арк.
						70
Змн.	Арк.	№ докум.	Підпис	Дата		

$$Ц = 1,1 \cdot 123137,82 + 30605,43 = 166057,03 \text{ грн.}$$

Розрахунок капітальних витрат, пов'язаних з впровадженням ПП, здійснюється по формулі:

$$K = K_{II} + K_{KO} + K_{BO} + K_C, \quad (4.16)$$

де

K_{II} – довиробничі витрати;

K_{KO} – вартість комп'ютерного устаткування;

K_{BO} – вартість допоміжного устаткування, необхідного для надійної роботи ПП;

K_C – вартість будівництва у зв'язку з впровадженням ПП.

Довиробничі витрати K_{II} включають всі витрати, пов'язані з проектуванням, розробкою, відладкою і впровадженням ПП – допроектні і проектні дослідження, постановка завдань і їх алгоритмізація, розробка, відладка і впровадження програмного забезпечення, навчання обслуговуючого персоналу, перепідготовка існуючого персоналу і так далі. Приймаємо K_{II} за 100 % загальної вартості розробленого ПП.

$$K_{II} = 166057,03 \text{ грн.}$$

Комп'ютерне устаткування $K_{КУ}$ – ноутбук, що коштує в середньому 20000 грн.

Вартість допоміжного устаткування – 10 % від вартості $K_{КУ}$.

$$K_{доп} = 2000 \text{ грн.}$$

Вартість будівництва (реконструкції) у зв'язку із впровадженням ПП $K_C = 0$.

Таким чином,

$$K = 166057,03 + 20000 + 2000 + 0 = 188057,03 \text{ грн.}$$

					КРМ.КІ.1.415-03.1.4	Арк.
						71
Змн.	Арк.	№ докум.	Підпис	Дата		

Розрахунок поточних (експлуатаційних) витрат за рік, пов'язаних з впровадженням ІС, здійснюється по формулі:

$$C = C_{ОПЛ} + C_A + C_{ЭЛ} + C_P + C_{ВСП} + C_{П}, \quad (4.17)$$

де

$C_{ОПЛ}$ – річний фонд основної і додаткової оплати праці персоналу, обслуговуючого ПП з нарахуваннями;

C_A – сума річних амортизаційних відрахувань від вартості основного і допоміжного устаткування;

$C_{ЭЛ}$ – вартість витрат на електроенергію за рік; $C_{П}$ – вартість річного утримування приміщень.

Річний фонд основної заробітної плати персоналу, що обслуговує ПП розраховуємо за формулою:

$$З_{ОСН} = Ч_{П} \cdot ЗП_{М} \cdot \tau_{P}, \quad (4.18)$$

де

$Ч_{П}$ – чисельність фахівців, обслуговуючих ПП;

$ЗП_{М}$ – посадовий оклад працівника, грн/місяць; річний фонд робочого часу працівника, місяць.

Необхідно підтримувати та слідкувати за використанням ПП. Для цього достатньо двох робітників, що працює над цим 6 години у день з місячним окладом 10000 грн., тому необхідно знайти місячну заробітну плату – $З_M$. Приблизно в місяць робітник працює 20 днів по 8 годин, отже 160 годин в місяць, звідси:

$$З_M = (10000 / (8 \cdot 20)) \cdot 20 \cdot 6 = 7500 \text{ грн/міс для одного робітника.}$$

$$З_{ОСН} 1 = 7500 \cdot 12 \cdot 2 = 180000 \text{ грн/рік.}$$

Після впровадження проекту для підтримки ПП буде достатньо одного працівника, що працює над цим 2 години у день з місячним окладом 8000 грн.

					<i>KPM.KI.1.415-03.1.4</i>	Арк.
						72
Змн.	Арк.	№ докум.	Підпис	Дата		

$$Z_{ОСН 2} = 8000 \cdot 12 \cdot 1 = 96000 \text{ грн/рік.}$$

За формулою 6.18 розраховується фонд додаткової заробітної платні:

$$Z_{ДОП1} = 180000 \cdot 0,1 = 18000 \text{ грн.}$$

$$Z_{ДОП2} = 96000 \cdot 0,1 = 9600 \text{ грн.}$$

Розрахунок нарахувань на заробітну платню проводиться за формулою 4.19:

$$Z_{НАР1} = (180000 + 18000) \cdot 0,22 = 43560 \text{ грн.}$$

$$Z_{НАР2} = (96000 + 9600) \cdot 0,22 = 23232 \text{ грн.}$$

Таким чином, загальні витрати на оплату праці розраховуються за формулою 4.20:

$$C_{ОПЛ1} = 180000 + 18000 + 43560 = 241560 \text{ грн.}$$

$$C_{ОПЛ2} = 96000 + 9600 + 23232 = 128832 \text{ грн.}$$

Розрахунок амортизаційних відрахувань визначається по формулі (4.21):

$$C_a = K_{ко} \cdot Ha / 100, \quad (4.21)$$

де

Ha – норма амортизаційних відрахувань (для комп'ютерного устаткування $Ha = 60 \%$).

$$C_a = 20000 \cdot 60 / 100 = 12000 \text{ грн.}$$

Річна вартість споживаної електроенергії визначається за формулою (4.22):

$$C_{ЕЛ} = M_y \cdot T_{КО} \cdot Ц_E \cdot K_{И}, \quad (4.22)$$

де

M_y – встановлена сумарна потужність комп'ютерного устаткування,

					<i>KPM.KI.1.415-03.1.4</i>	Арк.
						73
Змн.	Арк.	№ докум.	Підпис	Дата		

приймається рівною 0,25 кВт;

T_{KO} – річний фонд роботи ЕОМ з урахуванням часу на профілактичні огляди складає 240;

$K_{и}$ – коефіцієнт інтенсивного використання потужності, приймається рівним 0,9.

Вартість 1кВт*год електроенергії на даний момент складає 4 грн/ кВт · год.

Вартість річного ремонту основного і допоміжного устаткування становить 6 % від вартості комп'ютерного і допоміжного устаткування і складає:

$$C_p = 20000 \cdot 0,06 = 1200 \text{ грн.}$$

Річна вартість допоміжних матеріалів, пов'язаних з експлуатацією ПП становить 2 % від вартості комп'ютерного устаткування і складає:

$$C_{всп} = 20000 \cdot 0,02 = 400 \text{ грн.}$$

Таким чином, за формулою (4.23) поточні витрати до і після впровадження даного проекту складають:

$$C_1 = 241560;$$

$$C_2 = 128832 + 12000 + 2300 + 400 + 0 = 143738.$$

Очікуваний економічний ефект розраховується за формулою

$$E_0 = E_p - E_n \cdot K_n, \quad (4.22)$$

де

E_p – річна економія на поточних витратах, грн;

E_n – нормативний коефіцієнт ефективності одноразових витрат (приймаємо 0,25);

K_n – одноразові витрати на проект, грн.

					<i>KPM.KI.1.415-03.1.4</i>	Арк.
						74
Змн.	Арк.	№ докум.	Підпис	Дата		

Річна економія складається з поточних витрат і приросту прибутку у зв'язку із впровадженням проекту, у такий спосіб:

$$E_p = (C_1 - C_2), \quad (4.23)$$

де

C_1, C_2 – поточні витрати до й після впровадження проекту (грн).

$$E_p = (241560 - 143738) = 97822;$$

$$E_0 = 97822 - 0,25 \cdot 113930 = 69339,5.$$

Розраховуємо коефіцієнт ефективності одноразових витрат за формулою:

$$E = E_p / K_{\Pi} \quad (4.24)$$

$$E = 97822 / 113930 = 0,86.$$

$E > E_n$, тому проект економічно ефективний.

Термін окупності одноразових витрат обчислюється за формулою:

$$T = 1 / E \quad (4.25)$$

$$T = 1 / 0,86 = 1,27 \text{ року.}$$

Таблиця 4.8

Техніко-економічні показники проекту

Найменування показника	Одиниця виміру	Значення показника
Час роботи над проектом	дн	123
Ціна ПП	грн	166057,03
Капітальні витрати	грн	188057,03
Річна економія на поточних витратах	грн	97822
Економічний ефект від реалізації проекту	грн/рік	69339,5

					<i>KPM.KI.1.415-03.1.4</i>	Арк.
						75
Змн.	Арк.	№ докум.	Підпис	Дата		

Найменування показника	Одиниця виміру	Значення показника
Термін окупності	рік	1,27
Економічна ефективність (рентабельність) проекту		0,786

Висновок до четвертого розділу

Після ряду розрахунків було отримано тривалість та вартість розробки проекту, що сягають собою приблизно 166 057 грн. Сам термін окупності становить трохи більше 15 місяців без урахування можливих змін на ринку, та можливого сплеску на товар, що є прийнятним результатом і доводить економічну доцільність впровадження даного проекту.

					<i>KPM.KI.1.415-03.1.4</i>	Арк.
						76
<i>Змн.</i>	<i>Арк.</i>	<i>№ докум.</i>	<i>Підпис</i>	<i>Дата</i>		

РОЗДІЛ 5 ОХОРОНА ПРАЦІ

Охорона праці – це система законодавчих актів, соціально-економічних, організаційних, технічних, гігієнічних і лікувально-профілактичних заходів і засобів, що забезпечують безпеку, збереження здоров'я й працездатності людини в процесі праці.

5.1 Вимоги до особистого робочого місця працівника

Згідно з нормативами розмір одного робочого міста має бути не менш ніж 6 квадратних метрів, але якщо робочі місця є спареними, то між столами обов'язково повинна бути прозора перегородка та висота її повинна бути 2 та більше метрів, але це все застарілі норми та сьогодні так вже не роблять, а просто ставлять робочі столи на достатній відстані задля комфорту. У кожного працівника є навушники, якщо йому треба поговорити на «мітінгу» чи прослухати якусь інформацію стосовно роботи, щоб не заважати іншим. Однак відсутність перегородки, дає більше свіжого повітря та покращує освітлення;

У кожного працівника має бути достатньо місця на робочій поверхні для всієї необхідної техніки. На просторому столі має бути місце для: двох моніторів, принтера, планшету, колонок, тощо. Також біля робочого місця мають бути різні ящики та тумбочки та їх теж повинно теж бути достатньо;

Робочий стілець повинен бути зручним, легко регулюватися за висотою, щоб його було можливо виставити для ніг, та має бути спинка теж регульованою, щоб виставити зручне положення для спини працівника. Спина та хребет повинен мати належну та зручну підтримку;

Монітор для роботи повинен бути досить великим з добрим дозволом, щоб на ньому не було нечітких зображень та щоб всі зображення не розпливалися. Розмір монітору Хоча сьогодні багато роботодавців ставлять два монітори, але це не являється обов'язковою умовою. Кількість моніторів,

					<i>КРМ.КІ.1.415-03.1.4</i>	Арк.
						77
Змн.	Арк.	№ докум.	Підпис	Дата		

рекомендована для програміста, варіюється в залежності від специфіки виду діяльності та індивідуальних переваг співробітника.

На робочому столі повинні бути антивібраційні килимки. Працівник повинен слідкувати за своїм робочим місцем вчасно прибирати на робочому місці всі зайві документи та непотрібні для роботи предмети.

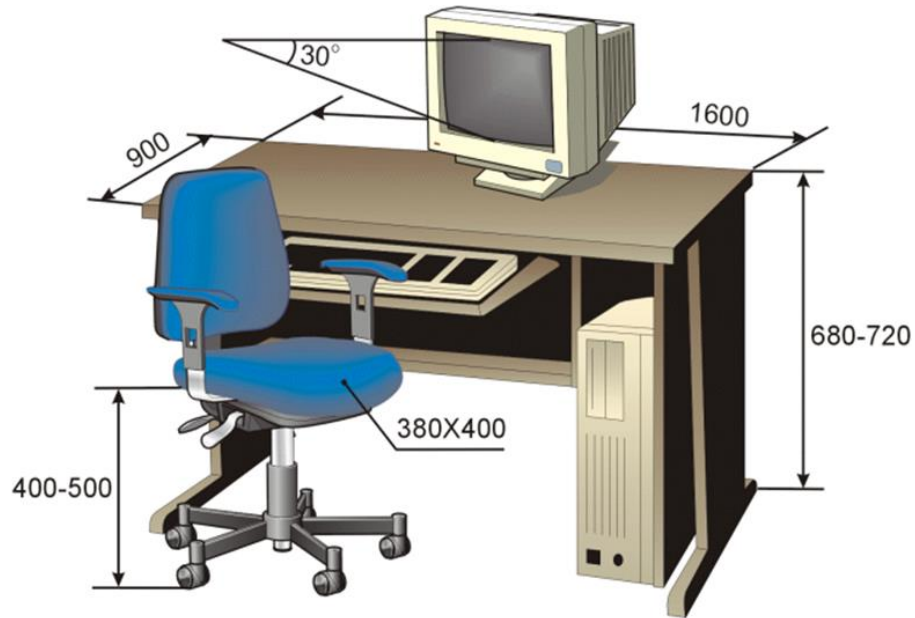


Рис. 5.1 – Розмірна схема для особистого робочого місця працівника

5.2 Рішення з розміщенням проектованого обладнання

Розміри заданого приміщення наступні: довжина – $A = 5$ м; ширина – $B = 6$ м; висота – $H = 3$ м.

Площа приміщень, у яких розташовують відео-термінали, визначають відповідно діючим нормативним документам з розрахунку на одне робоче місце: площа – не менше 6,0 кв. м, обсяг – не менше 20,0 куб. м, з урахуванням максимальної кількості осіб, що одночасно працюють в приміщенні.

У нашому випадку площа – 30 кв.м, а об’єм – 90 куб.м. Отже найбільша кількість людей, що одночасно працюють у приміщенні, дорівнює 4 людини.

5.3 Електробезпека

Електробезпека – це система організаційних, технічних заходів і засобів, що забезпечують захист людей від шкідливого і небезпечного впливу

					<i>KPM.KI.1.415-03.1.4</i>	Арк.
						78
Змн.	Арк.	№ докум.	Підпис	Дата		

електричного струму, електричної дуги, електромагнітного поля і статичної електрики. Електробезпека на підприємстві забезпечується завдяки дотриманню вимог, викладених у таких актах законодавства: Правила безпечної експлуатації електроустановок споживачів (наказ Держнаглядодохоронпраці від 09.01.1998 № 4); Правила безпечної експлуатації електроустановок (наказ Держнаглядодохоронпраці України від 06.10.1997 № 257), дія яких поширюються на працівників, що виконують роботи в електроустановках Міністерства енергетики України; Правила технічної експлуатації електроустановок споживачів, затверджені наказом Мінпаливенерго України від 25.07.2006 № 258 (у редакції наказу від 13.02.2012 №91), якими унормовано організаційні й технічні вимоги щодо експлуатації електроустановок споживачів;

Правила експлуатації електро-захисних засобів, затверджені наказом Міністерства праці та соціальної політики України від 05.06.2001 № 253, в яких визначено перелік засобів захисту, вимоги до них, норми випробувань, порядок їх застосування, зберігання, а також норми комплектування електроустановок;

Правила улаштування електроустановок визначають вимоги до електроустановок, принципи будови електроустановок, особливі вимоги до окремих вузлів і комунікацій (наказ Міністерства енергетики та вугільної промисловості України від 24.07.2017 № 476); ДСТУ 2843-94 «Електротехніка. Основні поняття. Терміни та визначення», яким установлені терміни і визначення основних понять з електробезпеки.

Одним з основних засобів захисту від напруги є правильно виконане заземлення. Далі проводяться його розрахунки.

Визначення розрахункового значення питомого опору ґрунту:

$$\rho_{\text{розр}} = \rho_{\text{гр}}\psi, (\text{Ом}) \quad (5.1)$$

де

$\rho_{\text{гр}} = 30 \text{ Ом}$ – питомий опір ґрунту (чорнозем);

					<i>KPM.KI.1.415-03.1.4</i>	Арк.
						79
Змн.	Арк.	№ докум.	Підпис	Дата		

ψ – кліматичний коефіцієнт, що враховує сезонні коливання вологості ґрунту, приймають від 1.2 до 2.0. Для даного випадку приймається $\psi = 1.8$.

$$P_{\text{розр}} = 30 * 1,8 = 54(\text{Ом}).$$

Довжина вертикального заземлювача приймається з умови $\frac{l'}{l} = 1; 2; 3(3)$,

де

l' – відстань між заземлювачами, м;

l – довжина заземлювачів, м.

$$t = \frac{1}{2} + t_0, (\text{м}) \quad (5.2)$$

де $t_0 \geq 0,5$ (м)

t_0 – глибина, на яку закопали заземлювач, $t_0 = 1,5$.

Довжина заземлювачів приймається рівною 9 м, а відстань між заземлювачами – 3 м. Тоді $\frac{l'}{l} = 2.1$. Розраховується t :

$$t = \frac{9}{3} + 1,5 = 4,5 (\text{м})$$

Вибирається система розподілу заземлювачів по контуру.

Далі відбувається розгляд опору одного вертикального заземлювача:

$$R_0 = \frac{54}{2 * 3,14 * 9} \left[\ln \frac{2 * 9}{0,00175} + \frac{1}{2} \ln \frac{4 * 4,5 + 9}{4 * 4,5 - 9} \right] \quad (5.3)$$

Для штучного заземлення в якості електродів приймають сталеві труби $d = 0.035 \times 0.05$ (м);

Визначається кількість вертикальних заземлювачів. При $U = 1000\text{В}$ необхідний опір $R_m \leq 4$ (Ом) для мережі змінного току $R_m = 2$.

$$N = \frac{R_0}{R_m} \quad (5.4)$$

					КРМ.КІ.1.415-03.1.4	Арк.
						80
Змн.	Арк.	№ докум.	Підпис	Дата		

де

$n = \frac{2,77}{2} = 1,375$, округляється до найближчого стандарту й стає рівним.

Далі визначається опір системи вертикальних заземлювачів:

$$R_{cv} = \frac{R_0}{n \cdot \eta_b}, \text{ (Ом)} \quad (5.5)$$

де

n – округлене число вертикальних заземлювачів до найближчого стандарту,

$n = n$ (шт).

η_b – коефіцієнт використання вертикальних заземлювачів (0.78).

$$R_{cv} = \frac{2,77}{2 \cdot 0,78} = 1,78 \text{ (Ом)}.$$

Потім розраховується опір сполучної смуги при розміщенні по контуру:

$$L = n \cdot \Gamma, \text{ (м)} \quad (5.6)$$

$$L = 2 \cdot 3 = 6 \text{ (м)}$$

$$R_n = \frac{\rho_{рвсч}}{2\pi L \eta_\Gamma} * \ln \frac{L^2}{dt_0}, \text{ (Ом)} \quad (5.7)$$

де

η_Γ – коефіцієнт використання горизонтальних заземлювачів (0.55).

$$R_n = \frac{54}{2 \cdot 3,14 \cdot 6 \cdot 0,55} * \ln \frac{36}{0,00175 \cdot 1,5} = 24,83 \text{ (Ом)}$$

Визначається загальний опір системи:

$$R_c = \frac{R_n \cdot R_{cv}}{R_n + R_{cv}}, \text{ (Ом)} \quad (5.8)$$

Якщо $R_{cv} > R_n$, то треба збільшити n і провести повторний розрахунок.

					<i>KPM.KI.1.415-03.1.4</i>	Арк.
						81
Змн.	Арк.	№ докум.	Підпис	Дата		

$$R_c = \frac{5,784 * 4,96}{5,784 + 4,96} = 2,67 \text{ (Ом)};$$

Отримали заземлення менше чотирьох Ом, що задовольняє умові.

Лінія електромережі для живлення ПК виконується шляхом прокладання фазового, нульового робочого й нульового захисного провідників, як окрема групова трьох-провідна мережа. ПК підключається до електричної мережі за допомогою справних штепсельних з'єднань і електророзеток заводського виготовлення. Тому що ПК буде розташований біля стіни, електромережа штепсельних розеток для живлення ПК прокладається по підлозі поруч зі стінами приміщення в заземлених металевих трубах і гнучких металевих рукавах.

Якщо в приміщенні обслуговуються більше 5 комп'ютерів, то необхідно на видному і доступному місці встановити аварійний резервний вимикач, який може відключити електричне живлення приміщення крім освітлення.

Підлога всієї зони обслуговування, ремонту й налагодження ПК повинна бути покрита діелектричними килимками, або викладена ізолюючими підстилками (шириною не менше 0.75-0.8 м) для ніг.

Заземлені конструкції, що знаходяться у приміщенні повинні бути надійно захищені діелектричними щитками або сіткам.

5.4 Пожежна профілактика

При визначенні видів та кількості первинних засобів пожежогасіння слід враховувати фізико-хімічні та пожежонебезпечні властивості горючих речовин, їх ставлення до вогнегасною речовин, а також площа виробничих приміщень, відкритих майданчиків і установок.

Азбестові полотна, грубошерсті тканини і повсть розміром не менше 1 x 1 м призначені для гасіння невеликих вогнищ пожеж при займанні речовин, горіння яких не може відбутися без доступу повітря.

					<i>KPM.KI.1.415-03.1.4</i>	Арк.
						82
Змн.	Арк.	№ докум.	Підпис	Дата		

Ємності для піску, що входять в конструкцію пожежного стенду, повинні бути місткістю не менше 0,1 куб. м. Конструкція ящика повинна забезпечувати зручність діставання піску та виключати попадання опадів.

Комплектування технологічного обладнання вогнегасниками здійснюється згідно з вимогами технічних умов (паспортів) на це обладнання або відповідними «Правилами пожежної безпеки».

Вибір типу і розрахунок необхідної кількості вогнегасників слід проводити в залежності від їх вогнегасної здатності, граничної площі, класу пожежі горючих речовин і матеріалів в закритому приміщенні або на об'єкті згідно *ISO N 3941 – 77*.

В замкнутих приміщеннях об'ємом не більше 50 куб. м для гасіння пожеж замість переносних вогнегасників, або додатково до них можуть бути використані вогнегасники порошкові.

Вибираючи вогнегасник з відповідним температурним межею використання, необхідно враховувати кліматичні умови експлуатації будівель та споруд.

Вогнегасники, відправлені з підприємства на перезарядку, повинні замінюватися відповідною кількістю заряджених вогнегасників.

При захисті приміщень ПК слід враховувати специфіку взаємодії вогнегасних речовин з захищеними обладнанням, виробами, матеріалами і т. п. Дані приміщення слід обладнати вуглекислотними вогнегасниками з урахуванням гранично допустимої концентрації вогнегасної речовини.

Приміщення, обладнані автоматичними стаціонарними установками пожежогасіння, забезпечуються вогнегасниками на 50%, виходячи з їх розрахункової кількості. Відстань від можливого вогнища пожежі до місця розміщення вогнегасник не повинна перевищувати 20 м для громадських будівель і споруд; 30 м для приміщень категорій А, Б і в; 40 м для приміщень категорій Г; 70 м для приміщень категорії Д.

Тип даної будівлі – громадське приміщення, а можливий клас пожежі – (Е) (оскільки у приміщенні багато комп'ютерів). Використовуючи ці дані,

					<i>KPM.KI.1.415-03.1.4</i>	Арк.
						83
Змн.	Арк.	№ докум.	Підпис	Дата		

визначається необхідна кількість вогнегасників – один порошковий вогнегасник місткістю 5 літрів.

5.5 Виробнича санітарія

Умови праці визначаються характером і важкістю роботи, яка виконується, а також параметрами навколишнього виробничого середовища.

Плануючи робоче місце, виходять з того, щоб робітник протягом зміни не знижував встановленої продуктивності праці, витрачаючи мінімум фізичних зусиль, був захищений від впливу небезпечних і шкідливих виробничих факторів. Важливе значення мають також робоча поза, темп і ритм виконуваної роботи.

Виключаються статичні м'язові напруги і, як наслідок, профзахворювання (викривлення хребта, розширення вен, плоскостопість) при робочій позі, яка забезпечує розташування центру ваги тіла в межах площі опори.

Загальні ергономічні вимоги до робочих місць викладено в міждержавних стандартах ГОСТ 12.2.032-78 (під час виконання робіт сидячи) та ГОСТ 12.2.033-78 (під час виконання робіт стоячи).

Робочим місцем вважають місце постійного або тимчасового перебування працівника в процесі трудової діяльності.

У разі вільного ритму роботи темп її визначається самим працівником відповідно до його індивідуальних можливостей.

В процесі одноманітних, короткочасних робочих операцій (наприклад, розфасовці та упаковці товарів) можливе відчуття монотонності праці. Чітко виражена монотонність виконуваної роботи призводить не тільки до передчасної втоми, але й до швидкого нервового виснаження.

Втомлюваність можна знизити періодичними змінами робочої пози і ритму виконуваних операцій, зміною робочих місць, естетизацією праці, використанням раціональної організації режиму праці та відпочинку.

					<i>КРМ.КІ.1.415-03.1.4</i>	<i>Арк.</i>
						84
<i>Змн.</i>	<i>Арк.</i>	<i>№ докум.</i>	<i>Підпис</i>	<i>Дата</i>		

У роботі, пов'язаній з великим м'язовим навантаженням, перерви повинні бути тривалими (до 10 хв.), але не частими. Причому, їх тривалість і частота повинні змінюватися протягом робочої зміни.

На основі загальних енерговитрат організму людини розрізняють (згідно ГОСТ 1.005-88) фізичні роботи :

- легкі (категорія I);
- середньої важкості (категорія II);
- важкі (категорія III).

Легкі фізичні роботи :

– категорії I-а з енерговитратами до 139 Вт – роботи, які виконуються сидячи і супроводжуються незначною фізичною напругою;

– I-б – з енерговитратами до 140 ... 174 Вт – роботи, які виконуються сидячи, стоячи або пов'язані з ходінням та супроводжуються деяким фізичним напруженням;

Фізичні роботи середньої важкості охоплюють види діяльності із затратами енергії 175 ... 232 Вт – категорія II-а і 233 ... 290 Вт – категорія II-б.

– до категорії II-а відносяться роботи, пов'язані з постійним ходінням і переміщенням дрібних (до 1 кг) виробів чи предметів в положенні сидячи та вимагають деякої фізичної напруги;

– до категорії II-б – роботи, пов'язані з ходінням, переміщенням і перенесенням вантажів до 10 кг і супроводжуються помірним фізичним напруженням.

Важкі фізичні роботи (категорії III) з енерговитратами понад 290 Вт пов'язані з постійним переміщенням і перенесенням значних (більше 10 кг) вантажів та вимагають великих фізичних зусиль.

На ефективність виконання робіт, самопочуття і стан здоров'я працівника значно впливає навколишнє виробниче середовище.

Виробниче середовище характеризують мікрокліматичні умови, склад і ступінь запиленості повітря, рівні шуму та вібрації, вид і якість освітлення, наявність та інтенсивність теплових, електромагнітних випромінювань і деяких

					<i>KPM.KI.1.415-03.1.4</i>	<i>Арк.</i>
						85
<i>Змн.</i>	<i>Арк.</i>	<i>№ докум.</i>	<i>Підпис</i>	<i>Дата</i>		

інших чинників. Його вважають оптимальним, якщо несприятливі впливи на робітника відсутні, і нормальним, якщо вони знаходяться в межах допустимих рівнів.

Важливе значення у формуванні виробничого середовища відіграє технічна естетика, яка дозволяє попередити нервово-психічні перевантаження робітників. Це досягається використанням раціональних архітектурно-планувальних рішень, високохудожніх інтер'єрів, а також науково обґрунтованого кольорового фарбування стін, стель та устаткування. Кольори для фарбування та устаткування вибирають залежно від їх функціонального призначення, умов зорової роботи, особливостей клімату, орієнтації вікон відносно сторін світу. Так, в приміщеннях з великим тепловиділенням огорожувальні конструкції та поверхні устаткування фарбують в кольори холодних тонів; в приміщеннях з природним освітленням – в кольори теплих тонів.

Використання з метою охорони праці різноманітних кольорів та знаків безпеки регламентовано ГОСТ 12.4.026-76.

Для забезпечення нормальних умов праці санітарні норми встановлюють на одного працівника: об'єм виробничих приміщень не менше 15 м³; площа приміщення, огороженого стінами чи глухими перегородками – не менше 4,5 м²; висота – 3,2 м. На підприємствах передбачаються кабінети психофізичного розвантаження (емоційно-вольового тренування), приміщення для виконання фізичних вправ, санітарно-побутові приміщення.

Санітарно-гігієнічні умови праці на підприємствах визначаються в основному рівнем наявних на робочих місцях шкідливих виробничих чинників. Перевірка відповідності умов праці вимогам санітарного законодавства здійснюється шляхом паспортизації цехів, дільниць, робочих місць (карти умов праці).

Вимірювання параметрів виробничого середовища виконують за договором працівники санепідемстанцій. За результатами цих вимірювань заповнюють відповідні паспорти. У випадку невідповідності значень виміряних

					<i>КРМ.КІ.1.415-03.1.4</i>	<i>Арк.</i>
<i>Змн.</i>	<i>Арк.</i>	<i>№ докум.</i>	<i>Підпис</i>	<i>Дата</i>		86

параметрів чинним нормам на підприємстві розробляються і здійснюються комплексні заходи щодо досягнення встановлених нормативів безпеки, гігієни праці та виробничої санітарії.

Завдання попередження гострих і хронічних професійних захворювань та отруєнь вирішує виробнича санітарія – система організаційних заходів і технічних засобів, які запобігають впливу на працівників шкідливих виробничих чинників чи зменшують його. Проводиться розрахунок світлового потоку.

$$\Phi = \frac{K * E_n * Z * S * 100}{n * \eta}, \text{ (лк)} \quad (5.9)$$

де

S – площа приміщення.

K – коефіцієнт запасу. Тому що в цьому випадку передбачається робота із ПК, то приймається розмір, як для диспетчерських, пультів операторів, контрольно-вимірювальних приладів.

$E_n = 200$ лк. – нормована мінімальна освітленість.

Тип світильника – ЛСП, отже коефіцієнт нерівномірності освітлення $Z = 1.1$.

Для того щоб порахувати світловий потік необхідно обчислити висоту над робочою поверхнею, відстань між центрами світильників (L), кількість світильників (n), індекс приміщення (i) й коефіцієнт використання світлового потоку (η).

$$H_p = H - h_p - 1, \text{ (м)} \quad (5.10)$$

де

$$H_p = 2$$

$$\frac{L}{H_p} = x, \text{ (м)} \quad (5.11)$$

					<i>KPM.KI.1.415-03.1.4</i>	Арк.
						87
Змн.	Арк.	№ докум.	Підпис	Дата		

$$n = \frac{S}{L^2}, \text{ (шт)} \quad (5.12)$$

$$i = \frac{S}{H*(A+B)} \quad (5.13)$$

де

$x = 1.4$ для світильників денного світла ПВЛП.

$A = 5$ м і $B = 6$ м, довжина і ширина приміщення.

$H = 3$ м – висота приміщення.

$H_p = 0$ м – відстань від світильника до стелі

$$H_p = 3 - 0 - 1 = 2 \text{ (м)}$$

$$L = H_p * 1.4 = 2 * 1.4 = 2.8 \text{ (м)}$$

$$n = \frac{30}{7,84} = 3.83,$$

Проводиться округлення до більшого найближчого цілого. Результат – 4 світильника.

$$I = \frac{30}{2*(5+6)} = 1.36.$$

Визначаємо коефіцієнт використання світлового потоку, при $P_n = 50\%$ (коефіцієнт відображення стіни), $P_c = 30\%$ (коефіцієнт відображення), $\eta = 36\%$.

$$\Phi = \frac{1,5*200*1,1*30*100}{3,83*0,42} = 6,155 \text{ (лм)}.$$

$$\Phi = 6,155 \text{ (лм)}$$

Висновки до п'ятого розділу

Було розглянуто потрібний тип приміщення і виробництва та визначені їх належні умови безпеки. Представлені необхідні норми пожежної профілактики та виробничої санітарії. Був здійснений розрахунок електричного опору

					<i>KPM.KI.1.415-03.1.4</i>	Арк.
						88
Змн.	Арк.	№ докум.	Підпис	Дата		

системи та визначення потрібної кількості освітлювачів, потрібних для безпечної роботи.

					<i>КРМ.КІ.1.415-03.1.4</i>	<i>Арк.</i>
<i>Змн.</i>	<i>Арк.</i>	<i>№ докум.</i>	<i>Підпис</i>	<i>Дата</i>		89

ЗАГАЛЬНІ ВИСНОВКИ

1. В ході аналізу сфери збору цін на товари та послуги були визначені ключові характеристики та методи цього процесу.
2. Вивчено сучасні системи та інструменти для збору цін, виявлено їх переваги та недоліки, а також потенціал для подальшого розвитку.
3. Визначено основні цілі та очікувані результати дослідження.
4. Розроблено теоретичну базу для системи збору цін, яка включає основні принципи та критерії ефективності.
5. Створено детальний план розробки, в якому описано ключові компоненти та функції системи.
6. Для збору та аналізу даних обрано відповідні інструменти і технології, зокрема *PyCharm 2023* – як *IDE* та *Python 3.8* – як мову розробки.
7. Детально описано процес збору даних, включаючи вибір джерел інформації та методи валідації зібраних даних.
8. Продемонстровано розробку методології аналізу зібраних даних, включаючи використання інструментів візуалізації.
9. Після завершення проекту було надано рекомендації для спрощення вибору методу збору даних та легшої їх інтеграції.

					<i>KPM.KI.1.415-03.1.4</i>	Арк.
						90
Змн.	Арк.	№ докум.	Підпис	Дата		

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. *Learn PyCharm course* : веб-сайт. URL: <https://www.jetbrains.com/help/pycharm/quick-start-guide.html> (дата звернення: 02.09.2023).
2. *Laragon Documentation* : веб-сайт. URL: <https://laragon.org/docs/index.html> (дата звернення: 03.10.2023).
3. *SQLite Documentation* : веб-сайт. URL: <https://www.sqlite.org/docs.html> (дата звернення: 12.11.2023).
4. *SQLite Architecture – SQLite – Вікіпедія* : веб-сайт. URL: https://ru.wikipedia.org/wiki/SQLite#/media/%D0%A4%D0%B0%D0%B9%D0%BB:SQLite_Architecture.svg (дата звернення: 12.09.2023).
5. *Laravel – Вікіпедія* : веб-сайт. URL: <https://uk.wikipedia.org/wiki/Laravel> (дата звернення: 12.09.2023).
6. Парсер: що це таке простими словами, приклади реклами : веб-сайт. URL: <https://www.activetraffic.ru/wiki/parser/> (дата звернення: 12.11.2023).
7. Фактори ціноутворення : веб-сайт. URL: <https://sendpulse.ua/ru/support-/glossary/pricing> (дата звернення: 01.10.2023).
8. Ціноутворення Вікіпедія : веб-сайт. URL: <https://ru.wikipedia.org/wiki/price/> (дата звернення: 01.10.2023).
9. *Microsoft Windows* : веб-сайт. URL: https://en.wikipedia.org//Microsoft_Windows. (дата звернення: 12.11.2023).
10. *macOS* : веб-сайт. URL: <https://en.wikipedia.org/wiki/MacOS/> (дата звернення: 12.10.2023).
11. *What is an API?* : веб-сайт. URL: <https://www.mulesoft.com/resources-/api/what-is-an-api> (дата звернення: 17.08.2023)
12. *HTTP Communications: веб-сайт. URL: Requests, Responses, and Tools for Web Development and Security Analysis:* <https://medium.com/@gwenilorac/http-requests-and-responses-cc989147fd8c> (дата звернення: 17.08.2023)

					КРМ.КІ.1.415-03.1.4	Арк.
Змн.	Арк.	№ докум.	Підпис	Дата		91

13. АТБ : веб-сайт. URL: <https://www.atbmarket.com/ru> (дата звернення: 29.10.2023)
14. *What is web scraping* : веб-сайт. URL: <https://www.imperva.com/learn/application-security/web-scraping-attack/> (дата звернення: 17.08.2023)
15. ТавріяВ : веб-сайт. URL: <https://tavriav.ua/> (дата звернення: 29.10.2023)
16. Сільпо : веб-сайт. URL: <https://shop.silpo.ua/> (дата звернення: 27.10.2023)
17. *Silpo API* : веб-сайт. URL: <https://gebeto.github.io/info/6-apis/silpo-shop.html> (дата звернення: 15.08.2023)
18. *Selenium – documentation* : веб-сайт. URL: <https://selenium-python.readthedocs.io/> (дата звернення: 16.08.2023)
19. *Requests – documentation* : веб-сайт. URL: <https://requests.readthedocs.io/en/latest/> (дата звернення: 17.08.2023)
20. Державна служба України з питань праці - Охорона праці на підприємстві: що потрібно знати? : веб-сайт. URL: <https://te.dsp.gov.ua/ohorona-pratsi-na-pidpryyemstvi-shho-potribno-znaty/> (дата звернення: 19.11.2023)
21. Верховна Рада України - Про охорону праці - Законодавство - Верховна Рада України : веб-сайт. URL: <https://zakon.rada.gov.ua/-laws/show/2694-12#Text> (дата звернення: 18.11.2023)
22. *Costless – порівняння цін на продукти* : веб-сайт. URL: <https://costless.com.ua/ru> (дата звернення: 06.03.2023)
23. *e-katalog – порівняння цін на товари* : веб-сайт. URL: <https://ek.ua/ua/> (дата звернення: 06.03.2023)
24. *Hotline – порівняння цін на товари* : веб-сайт. URL: <https://hotline.ua/> (дата звернення: 06.03.2023)
25. *Google shopping – Shop online* : веб-сайт. URL: <https://shopping.google.com/?pli=1> (дата звернення: 06.03.2023)
26. Так Аудит – Метод «витрати +» : веб-сайт. URL: <http://tak-audit.com.ua/material/89787-3-metod-vitrati-zatratnyj-metod-cost-plus-method> (дата звернення: 06.03.2023)

					КРМ.КІ.1.415-03.1.4	Арк.
						92
Змн.	Арк.	№ докум.	Підпис	Дата		

27. *Pricecop* – Моніторинг цін : веб-сайт. URL: <https://pricecop.net-uk/blog/monitoring-tsen-v-internetel/> (дата звернення: 03.03.2023)
28. *Hostiq* – А/В-тестування : веб-сайт. URL: <https://hostiq.ua/blog/-ukr/ab-testing/> (дата звернення: 12.03.2023)
29. *Fractus* – Фріміум: веб-сайт. URL: як заробляють на безплатних версіях: <https://fractus.com.ua/uk/blog/frimium-yak-zaroblyajut-na-bez-platnih-versiyah/> (дата звернення: 17.11.2023)
30. *Azure - What is SaaS?* : веб-сайт. URL: <https://azure.microsoft.com/-en-us/resources/cloud-computing-dictionary/what-is-saas> (дата звернення: 16.11.2023)
31. Дипломне проектування: Методичні вказівки до виконання кваліфікаційних робіт для здобувачів освіти СВО «Магістр» спеціальності 123 «Комп’ютерна інженерія» / С.В. Артеменко та ін. Одеса : ОНАХТ «Технолог», 2023. 52 с. (дата звернення: 15.11.2023)
32. Вігуржинська С.Ю., Колесник В.І. Дипломне проектування економічної частини проекту: Методичні вказівки для студентів, що навчаються за спеціальностями 05010101- “Інформаційні управляючі системи та технології”; 05010102 - “Інформаційні технології проектування”; 05010201 – “Комп’ютерні системи та мережі”; 05010203 – “Спеціалізовані комп’ютерні системи”. Одеса : ОНАХТ «Технолог», 2016. 23 с. (дата звернення: 17.11.2023)

					<i>KPM.KI.1.415-03.1.4</i>	Арк.
						93
Змн.	Арк.	№ докум.	Підпис	Дата		


```

        "Sec-Fetch-Mode": "cors",
        "Sec-Fetch-Dest": "empty",
        "Referer": "https://shop.silpo.ua/",
        "Accept-Encoding": "gzip, deflate, br",
        "Accept-Language": "en, ru;q = 0.9, en - US;q = 0.8"
    }
    data = {
        "method": "GetSimpleCatalogItems",
        "data": {
            "merchantId": 1,
            "basketGuid": "8e99dbbd-b9cd-4273-bcaf-52d19a4118a5",
            "deliveryType": 1,
            "filialId": 2043,
            "From": 1,
            "businessId": 1,
            "To": 32,
            "ingredients": False,
            "categoryId": 22,
            "RangeFilters": {},
            "MultiFilters": {},
            "UniversalFilters": [],
            "CategoryFilter": [],
            "Promos": []
        }
    }
    url = https://shop.silpo.ua/category/

    for category_id in range(0, 5000):
        data["data"]["categoryId"] = category_id
        response = requests.post(url, headers=headers,
                                data=json.dumps(data))
        df = df.append(pd.DataFrame([data_process(items)
                                    for items in response.json().get("items", [])]))
        df['date'] = date.today()
    df = df.astype({'parameters': 'string'})
    df.drop(columns=df.columns[0], axis=1, inplace=True)
    df.to_sql('silpo', con=engine, if_exists='append')
    df.to_csv("silpo_result.csv")
    print("Time spent on " + str(len(df)) + " elements: " +
          str(time.time() - start))
if __name__ == '__main__':
    main()

import json
import requests
import pandas
import time
from datetime import date, timedelta
from sqlalchemy import create_engine
def data_process(response):
    df = pandas.DataFrame()
    if response.status_code == 200:
        if response.json()["items"] != []:
            for items in response.json()["items"]:
                try:
                    print(items)
                    result = {"id": items["id"],
                              "name": items["name"],

```

						Арк.
						95
Змн.	Арк.	№ докум.	Підпис	Дата		

```

        "unit": items["unit"],
        "price": items["price"],
        "mainImage": items["mainImage"],
        "parameters": items["parameters"],
        "shop": 'silpo'}
    except IndexError:
        print(items)
        exit()
    print(result)
    df = df.append(result, ignore_index=True)
    df['date'] = date.today()

    else:
        pass
    return df

def main():
    engine =
    create_engine(r'sqlite:///C:\Users\User\Downloads\TelegramDesktop
        \shops.db', echo=False)
    start = time.time()
    df = pandas.DataFrame()
    headers = {
        "sec-ch-ua": '"Chromium";v="94", "GoogleChrome";v="94",';
        Not A Brand";v = "99"',
        "X-SessionId": "8be027e15523aba4953479b9951e4d58",
        "sec-ch-ua-mobile": "?0",
        "User-Agent": "Mozilla / 5.0 (WindowsNT10.0; WOW64) AppleWebKit /
        537.36 (KHTML, likeGecko) Chrome / 94.0.4606.61 Safari / 537.36",
        "Content-Type": "application / json; charset = UTF-8",
        "Accept": "application / json,
        text / plain, */*",
        "user-uid": "5c0773b8-7be9-9961-a971-cf625280c2a1",
        "sec-ch-ua-platform": "Windows",
        "Origin": "https://shop.silpo.ua",
        "Sec-Fetch-Site": "same-site",
        "Sec-Fetch-Mode": "cors",
        "Sec-Fetch-Dest": "empty",
        "Referer": "https://shop.silpo.ua/",
        "Accept-Encoding": "gzip, deflate, br",
        "Accept-Language": "en, ru;q = 0.9, en - US;q = 0.8"
    }
    data = {
        "method": "GetSimpleCatalogItems",
        "data": {
            "merchantId": 1,
            "basketGuid": "8e99dbbd-b9cd-4273-bcaf-52d19a4118a5",
            "deliveryType": 1,
            "filialId": 2043,
            "From": 1,
            "businessId": 1,
            "To": 32,
            "ingredients": False,
            "categoryId": 22,
            "RangeFilters": {},
            "MultiFilters": {},
            "UniversalFilters": [],
            "CategoryFilter": [],
            "Promos": []
        }
    }

```

					Арк.
					96
Змн.	Арк.	№ докум.	Підпис	Дата	

KPM.KI.2. 415-03.1.4

```

    }
}
url = https://api.catalog.ecom.silpo.ua/api/2.0/exec/
    EcomCatalogGlobal
for category_id in range(0, 5000):
    data["data"]["categoryId"] = category_id
    response = requests.post(url, headers=headers,
                             data=json.dumps(data))
    df = df.append(data_process(response))
df = df.astype({'parameters': 'string'})
df.drop(columns=df.columns[0], axis=1, inplace=True)
df.to_sql('silpo', con=engine, if_exists='append')
df.to_csv("silpo_result.csv")
print("Time spend on " + str(len(df)) + " elements: " +
      str(time.time() - start))
if __name__ == '__main__':
    main()

import time
from selenium import webdriver
import pandas as pd
from selenium.webdriver.common.keys import Keys
from selenium.webdriver.common.by import By
start = time.time()
driver = webdriver.Chrome()
titles_list = []
weights_list = []
prices_list = []

def collect_info(pagination):
    url = 'https://shop.silpo.ua/sets/solodkipodarunky'
    additional_url = '?to={}&from={}'.format(pagination, pagination)
    link = url + additional_url
    driver.get(link)
    try:
        time.sleep(2)
        for _ in range(11):
            driver.execute_script("window.scrollTo(0, 550);")
            time.sleep(1)
        time.sleep(2)
        titles = driver.find_elements(By.CLASS_NAME, "product-title")
        prices = driver.find_elements(By.CLASS_NAME, "current-integer")
        weights = driver.find_elements(By.CLASS_NAME, 'product-weight')
        for title, weight, price in zip(titles, weights, prices):
            title_text = title.text
            weight_text = weight.text
            price_text = price.text
            print(title_text)
            titles_list.append(title_text)
            weights_list.append(weight_text)
            prices_list.append(price_text)
    except Exception as e:
        print(f"An error occurred: {e}")

def main():
    for i in range(1, 5):
        collect_info(i)
    driver.quit()

```

						Арк.
						97
Змн.	Арк.	№ докум.	Підпис	Дата		

```

# Create a Pandas DataFrame from the lists
data = {"Title": titles_list, "Weight": weights_list, "Price":
        prices_list}
df = pd.DataFrame(data)
end = time.time()
parse_time = str(end - start)
print("Category parsed for: " + parse_time + ' seconds')
df.to_csv('test.csv')
if __name__ == '__main__':
    main()

```

					<i>KPM.KI.2. 415-03.1.4</i>	<i>Арк.</i>
						98
<i>Змн.</i>	<i>Арк.</i>	<i>№ докум.</i>	<i>Підпис</i>	<i>Дата</i>		

Огляд існуючих рішень

- **Hotline.ua** – український онлайн-сервіс для вибору товарів і порівняння цін. Він дозволяє знайти **нон-фуд** (**Нон-фуд** - промислові товари, які є невід'ємною частиною нашого повсякденного життя. У категорію товарів нон-фуд входить ряд товарних груп, до яких можна віднести: іграшки, товари для інтер'єру, спортивні товари, канцелярське приладдя, косметичні засоби і товари побутової хімії) товари, порівняти їх ціни у різних магазинах та характеристики.
- **E-Katalog** – багатофункціональний сервіс пошуку товарів в інтернет-магазинах і порівняння цін. Він охоплює найрізноманітніші категорії товарів: електроніка, комп'ютери, побутова техніка, автотовари, обладнання для ремонту та будівництва, туристичне спорядження, дитячі товари і багато іншого (нон-фуд).
- **Costless** – застосунок, що дозволяє порівнювати ціни на окремі продукти або цілі продукти/кошки (переважно фуд-онлі), визначаючи оптимальний для вас магазин поблизу.
- **Google Shopping** – сервіс, що був створений для пошуку товарів в інтернет-магазинах та порівняння цін між різними пропозиціями. Спочатку був списком з цінами. На даний момент Google Shopping стає більш привабливим для розміщення своїх товарів для багатьох гравців ринку eCommerce. Тут можна знайти товари як самостійних інтернет-магазинів, так і пропозиції маркетплейсів.

hotline

e.katalog



Google Shopping

Рис. Б.1 – Слайд №3

Постановка задачі

1. Аналіз існуючих методів:

- Ретельно проаналізувати різні методи збору даних про ціни на товари та послуги.
- Вивчити переваги та недоліки кожного методу.

2. Оцінка точності та достовірності даних:

- Визначити метрики точності та надійності для оцінки якості даних, зібраних за допомогою кожного методу.

3. Визначення ефективності в залежності від галузі та масштабу бізнесу:

- Вивчити, як різні методи впливають на ефективність збору даних в різних галузях.

4. Аналіз вартості та часу:

- Визначити витрати, пов'язані з впровадженням та підтриманням кожного методу.

5. Розробка рекомендацій:

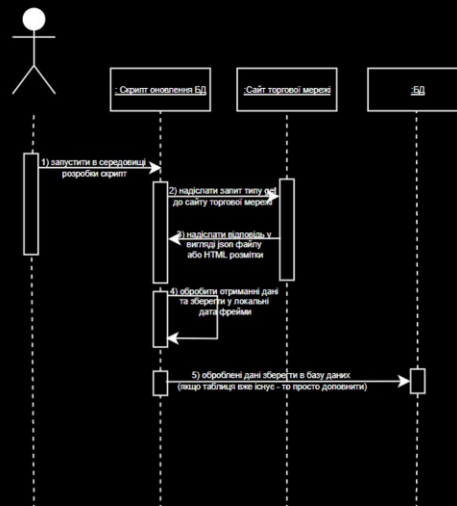
- На основі отриманих результатів розробити рекомендації для бізнес-середовища щодо вибору оптимального методу чи їх комбінації.

Рис. Б.1 – Слайд №4

					KPM.KI.0. 415-03.1.4	Арк.
						100
Змн.	Арк.	№ докум.	Підпис	Дата		

Основні етапи роботи парсера

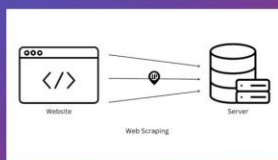
- Програма сканує дані, що надходять на вхід, будь то текст, веб-сторінка або інший набір інформації, і відокремлює деякі елементи.
- Що саме виділятиме парсер із масиву даних – залежить від конкретного завдання. Зазвичай програми можна налаштувати таким чином, щоб отримати потрібні результати.
- Правила пошуку найчастіше задаються регулярними висловлюваннями – рядками, складеними за певними правилами що дають програмі пояснення, що і як шукати.
- На основі зібраної інформації формується звіт або таблиця, в якій відображено всі отримані результати.



Діаграма послідовності. Робота скрипта збору даних

Рис. Б.1 – Слайд №5


Визначення методів збору даних



- Веб-скрапінг (Web Scraping):**
 - Веб-скрапінг - це процес автоматичного збору даних з веб-сайтів. Зазвичай використовується для отримання структурованої інформації з HTML-коду веб-сторінок. Python має багато бібліотек, таких як BeautifulSoup та Scrapy, які полегшують цей процес.
- API (Application Programming Interface):**
 - Багато веб-сайтів надають API для доступу до своїх даних. Це дозволяє програмам отримувати доступ до конкретних ресурсів або функціональності веб-сайту. Використання API є легальним і зручним методом збору даних.
- Аналіз HTTP-запитів і відповідей:**
 - Вивчення мережевого трафіку між вашим браузером і сервером за допомогою інструментів розробника браузера (наприклад, вбудованих в Chrome або Firefox) дозволяє вам детально вивчати обмін даними.

Рис. Б.1 – Слайд №6

Змн.	Арк.	№ докум.	Підпис	Дата




Визначення джерел інформації для збору даних

Для дослідження описаних методів були розглянуті сайти таких торговельних мереж як:

- «АТБ»
 - Переваги:
 - Великий обсяг даних
 - Технічна стабільність
 - Недоліки:
 - Доступний лише веб-скрапінг.
- «Таврія В»
 - Переваги:
 - Регіональна особливість
 - Програми лояльності
 - Недоліки:
 - Доступний лише веб-скрапінг
- «Сільпо»
 - Переваги:
 - Розмаїтість інформації
 - Загальна доступність
 - Актуальність даних.
 - Доступність для збору.
 - Недоліки:
 - Авторизація та обмеження.

Рис. Б.1 – Слайд №7



Загальний час, витрачений на всі категорії

Середній час, витрачений на категорію

Швидкість методів

- Веб-скрапінг часто є найчасоємкішим методом через потребу в обробці HTML-сторінок та можливість виникнення проблем при аналізі складних структур. API надає швидший та більш ефективний спосіб отримання структурованих даних. HTTP-запити можуть поєднувати обидва методи, але його швидкість також залежить від різних факторів, таких як швидкість сервера та обмежень API.

Рис. Б.1 – Слайд №8

Розробка та підтримка

- API-запити:
 - Використання API є найшвидшим способом отримання даних, оскільки взаємодія відбувається безпосередньо з сервером, використовуючи визначений інтерфейс. На розробку такого витрачається менше часу в порівнянні з іншими методами.
 - Зазвичай API забезпечують стабільний та документований спосіб отримання даних, тому підтримка не повинна викликати проблем.
- HTTP-запити:
 - Використання HTTP-запитів може бути ефективним для швидкого отримання даних, особливо при роботі з простими операціями, такими як витягування сторінок HTML або обмін даними у форматі JSON. Проте, варто враховувати, що цей метод може бути складнішим порівняно з використанням API.
 - Залежно від конкретного використання, може бути менш стабільним, порівняно з API. Однак це може бути ефективним методом для взаємодії з веб-сайтами, якщо API недоступне.
- Веб-скрапінг:
 - Скрапінг веб-сайтів може бути більш витратним з точки зору часу, оскільки потрібно аналізувати HTML-код та витягувати дані з нього.
 - Підтримка для веб-скрапінгу може бути менш стабільною, оскільки власник веб-сайту може вносити зміни в розміщення або структуру сторінок. Також, його використання може порушувати правила використання веб-сайтів і призвести до юридичних питань.

Рис. Б.1 – Слайд №9

Специфіка категорій

- Різна кількість даних: Різні категорії товарів можуть мати різну кількість доступної інформації. Наприклад, категорія "Кулінарія" може містити багато різних товарів з додатковою інформацією, що може впливати на час парсингу.
- Складні структури сторінок: Різні категорії можуть використовувати різні структури сторінок та розміщення даних. Структура сторінок може варіюватися від простої до складної, що впливає на час, необхідний для аналізу.
- Величина фотографій та інших мультимедійних даних: Деякі категорії можуть включати багато зображень, відео або інших мультимедійних даних, що також може вплинути на час парсингу, особливо при використанні веб-скрапінгу.

Категорія	API (секунди)	HTTP-запити (секунди)	Веб-скрапінг (секунди)
18+	29	31	59
Алкоголь	118	179	241
Бакалія	63	121	183
Гігієна	58	120	182
Дитячі товари	116	181	239
Для дому	119	177	244
Для тварин	57	122	177
Заморожена продукція	62	118	243
Кава, чай	59	121	181
Квіти, товари для саду та городу	62	57	121
Консерви, соуси, приправи	118	179	245
Кулінарія	121	239	365
Молочні продукти та яйця	61	59	121
М'ясо	178	122	235
М'ясо-ковбасні вироби	62	118	185
Напції	61	119	178
Риба	122	238	355
Сири	59	122	175
Снеки	63	58	123
Солодощі	57	123	179
Фрукти, овочі, соління	128	175	310
Хліб та хлібобулочні вироби	58	61	125
Total	1,831	2,840	4,466

Рис. Б.1 – Слайд №10

Змн.	Арк.	№ докум.	Підпис	Дата

KPM.KI.O. 415-03.1.4

Оптимізація і вибір методу

- Комбінація методів: Одним з підходів до оптимізації може бути використання комбінації різних методів парсингу для різних завдань. Наприклад, використання API для отримання загальної інформації, а веб-скрапінгу для детального аналізу.
- Оптимізація запитів: Для використання API може бути важливим оптимізувати запити, вибираючи лише необхідні дані та обмежуючи кількість запитів для зменшення часу відповіді.
- Кешування даних: Збереження результатів попередніх запитів у кеші може допомогти уникнути повторних запитів на одні й ті самі дані, що прискорює процес.
- Асинхронний парсинг: Використання асинхронних запитів або паралельного парсингу може покращити ефективність, дозволяючи виконувати декілька операцій одночасно.
- Оптимізація методу парсингу важлива для максимальної продуктивності та швидкості збору даних. Вибір методу повинен враховувати особливості конкретного дослідження, вимоги до швидкості та обсягу даних, які необхідно зібрати. Комбінування різних методів та використання оптимізаційних стратегій може сприяти ефективному використанню ресурсів при парсингу великих обсягів даних.

Рис. Б.1 – Слайд №11

Основні економічні показники

Показники	Одиниця виміру	Значення показника	
		до впровадження	після Впровадження
Трудомісткість розробки проекту	дні	-	123
Ціна ПП	грн.	-	166 057,03
Капітальні витрати	грн.	-	188 057,03
Поточні витрати	грн./рік	241 560	123 137,82
Економічний ефект від реалізації	грн./рік	-	69 339,5
Строк окупності	рік	-	1,27
Економічна ефективність		-	0,786

Рис. Б.1 – Слайд №12

ЗАГАЛЬНІ ВИСНОВКИ

- В ході аналізу сфери збору цін на товари та послуги були визначені ключові характеристики та методи цього процесу.
- Вивчено сучасні системи та інструменти для збору цін, виявлено їх переваги та недоліки, а також потенціал для подальшого розвитку.
- Визначено основні цілі та очікувані результати дослідження.
- Розроблено теоретичну базу для системи збору цін, яка включає основні принципи та критерії ефективності.
- Створено детальний план розробки, в якому описано ключові компоненти та функції системи.
- Для збору та аналізу даних обрано відповідні інструменти і технології, зокрема *PyChart 2023* – як *IDE* та *Python 3.8* – як мову розробки.
- Детально описано процес збору даних, включаючи вибір джерел інформації та методи валідації зібраних даних.
- Продемонстровано розробку методології аналізу зібраних даних, включаючи використання інструментів візуалізації.
- Після завершення проекту було надано рекомендації для спрощення вибору методу збору даних та легшої їх інтеграції.

Conclusion

Рис. Б.1 – Слайд №13

					КРМ.КІ.0. 415-03.1.4	Арк.
						105
Змн.	Арк.	№ докум.	Підпис	Дата		