

Київський національний торговельно-економічний університет

Кафедра цифрової економіки та системного аналізу

ВИПУСКНА КВАЛІФІКАЦІЙНА РОБОТА

на тему:

«Технології та інструменти обробки великих даних в екосистемі мобільного меседжингу»

Студентки 2 курсу, 1м групи,

спеціальності
051 «Економіка»

спеціалізації
«Цифрова економіка»

Науковий керівник
кандидат технічних наук,
доцент

Гарант освітньої програми
доктор фізико-математичних
наук, професор

Ярославцева Анастасія
Олександрівна

підпис студента

Геселева Наталія
Валеріївна

підпис керівника

Гамалій
Володимир Федорович

підпис гаранта

Київ 2021

Київський національний торговельно-економічний університет

Факультет інформаційних технологій
Кафедра цифрової економіки та системного аналізу
Освітній ступінь магістр
Спеціальність 051 «Економіка»
Спеціалізація «Цифрова економіка»

Затверджую

Зав. кафедри _____ Роскладка А.А.
«15» листопада 2020 р.

**Завдання
на випускню кваліфікаційну роботу (проект) студентці**

Ярославцевій Анастасії Олександрівні

(прізвище, ім'я, по батькові)

1. Тема випускної кваліфікаційної роботи (проекту)
«Технології та інструменти обробки великих даних в екосистемі мобільного меседжінгу»
Затверджена наказом КНТЕУ від «22» жовтня 2020 р. № 3066
2. Строк здачі студентом закінченої роботи «05» листопада 2021 року
3. Цільова установка та вихідні дані до роботи
Мета роботи полягає у дослідженні інтеграції технологій обробки великих даних в екосистему мобільного меседжінгу.
Об'єкт дослідження процес обробки великих даних у сфері телекомунікацій.
Предметом дослідження є теоретичні, методичні та практичні підходи до організації обробки великих даних у екосистемі мобільного меседжінгу.

4. Консультанти по роботі (проекту) із зазначенням розділів, за якими здійснюється консультування:

Розділ	Консультант (прізвище, ініціали)	Підпис, дата	
		Завдання видано	Завдання прийнято
1	Геселева Н. В.	15.11.2020 р.	15.11.2020 р.
2	Геселева Н. В.	15.11.2020 р.	15.11.2020 р.
3	Геселева Н. В.	15.11.2020 р.	15.11.2020 р.

5. Зміст випускної кваліфікаційної роботи (проекту) (перелік питань за кожним розділом)

ВСТУП

РОЗДІЛ 1. BIG DATA, МЕТОДИ І ТЕХНІКИ АНАЛІЗУ ВЕЛИКИХ ДАНИХ.

1.1. Ключові характеристики, роль та функції великих даних в інформаційних технологіях

1.2. Методи та техніки аналізу великих даних

1.3. Досвід використання великих даних в Україні та світі.

Висновки до розділу 1

РОЗДІЛ 2. СТАН ЕКОСИСТЕМИ МОБІЛЬНОГО МЕСЕДЖИНГУ .

2.1. Мобільний меседжінг як окрема бізнес-екосистема

2.2. Закономірності функціонування ринку телекомунікаційних послуг.

2.3. Основні елементи екосистеми мобільного меседжінгу

Висновки до розділу 2

РОЗДІЛ 3. BIG DATA РІШЕННЯ ОПЕРАТОРА МОБІЛЬНОГО МЕСЕДЖИНГУ НА ПРИКЛАДІ БІЗНЕС-КЕЙСУ ДЛЯ АВТОМОБІЛЬНОГО ЦЕНТРУ ТОВ «ГРАНД АУТОС»

3.1. Формування основних функціональних вимог та моделювання бізнес-процесів задачі просування продукції ТОВ «Гранд Аутос».

3.2. Веб-орієнтований додаток особистого кабінету для оптимізації процесу проведення таргетованих розсилок.

Висновки до розділу 3

ВИСНОВКИ ТА ПРОПОЗИЦІЇ

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

ДОДАТКИ

6. Календарний план виконання роботи (проекту)

№ з/п	Назва етапів випускної кваліфікаційної роботи	Строк виконання етапів роботи	
		За планом	фактично
1	2	3	4
1	<i>Вибір теми випускної кваліфікаційної роботи</i>	20.10.2020	20.10.2020
2	<i>Розробка та затвердження завдання на випускну кваліфікаційну роботу</i>	15.11.2020	15.11.2020
3	<i>Вступ</i>	01.03.2021	
4	<i>Розділ 1. Big Data, методи і техніки аналізу великих даних</i>	25.06.2021	
5	<i>Розділ 2. Стан екосистеми мобільного меседжингу</i>	01.09.2021	
6	<i>Підготовка статті у збірник наукових статей магістрів</i>	15.09.2021	
7	<i>Розділ 3. Big Data рішення оператора мобільного меседжингу на прикладі бізнес-кейсу для автомобільного центру ТОВ «Гранд Аутос».</i>	18.10.2021	
8	<i>Висновки</i>	01.11.2021	
9	<i>Здача випускної кваліфікаційної роботи на кафедрі науковому керівнику</i>	05.11.2021	
10	<i>Попередній захист випускної кваліфікаційної роботи</i>	25.11.2021	
11	<i>Виправлення зауважень, зовнішнє рецензування випускної кваліфікаційної роботи</i>	28.11.2021	
12	<i>Представлення готової зшитої випускної кваліфікаційної роботи на кафедрі</i>	30.11.2021	
13	<i>Публічний захист випускної кваліфікаційної роботи</i>	За розкладом роботи ЕК	

7. Дата видачі завдання «15» листопада 2020 р.

8. Науковий керівник випускної кваліфікаційної роботи (проекту)

Геселева Н. В.
(підпис, прізвище, ініціали)

9. Гарант освітньої програми

Гамалій В. Ф.
(підпис, прізвище, ініціали)

10. Завдання прийняв до виконання студент

Ярославцева А. О.
(підпис, прізвище, ініціали)

Анотація

В даній роботі досліджено сферу телекомунікацій та мобільного меседжингу. Розглянуто теоретичні положення великих даних, ключових характеристик, роль та функції великих даних в інформаційних технологіях. Проаналізовано характеристики даних, що збираються, зберігаються, обробляються та аналізуються за допомогою методу Big Data. Наведено характеристику методів та техніки аналізу великих даних. Досліджено основні технології управління великими даними, принципи роботи та безпосередньо вплив великих даних на різні сфери життя. Розглянуто Big Data як сучасний метод дослідження цільової аудиторії по запити підприємств.

Проаналізовано та охарактеризовано мобільний меседжинг як окрему бізнес-екосистему. Розглянуто роботу захищених та класичних месенджерів, їх рейтинги. Досліджено ринок мобільного зв'язку України, розглянуто його значення для національної економіки. Проаналізовано структуру ринку, стан конкурентного середовища та виявлено частку основних гравців на ньому.

Надані рекомендації щодо практичної реалізації запропонованих рішень для оптимізації процесу надсилення запиту на обробку великих даних та проведення таргетованих розсилок. Було розроблено та запропоновано рішення у формі веб-орієнтованого додатку особистого кабінету.

Ключові слова: телекомунікації, ринок мобільного зв'язку, BigData, Data Mining, таргетинг.

SUMMARY

This research studies the field of telecommunications and mobile messaging. Theoretical provisions of big data, key characteristics, role and functions of big data in information technologies are considered. The characteristics of the data collected, stored, processed and analyzed using the Big Data method are analyzed. The characteristics of methods and techniques of big data analysis are given. The basic technologies of big data management, principles of work and direct influence of big data on various spheres of life are investigated. Big Data is considered as a modern

method of researching the target audience at the request of enterprises.

Mobile messaging as a separate business ecosystem is analyzed and characterized. The work of protected and classic messengers, their ratings are considered. The market of mobile communication of Ukraine is investigated, its value for national economy is considered. The structure of the market, the state of the competitive environment are analyzed and the share of the main players in it is revealed.

Recommendations for the practical implementation of the proposed solutions to optimize the process of meeting the request for large data processing and targeted mailings. A solution in the form of a web-based personal account application was developed and proposed.

Keywords: telecommunications, mobile market, mobile operators, Big Data, Data Mining, targeting.

Зміст

ВСТУП.....	3
РОЗДІЛ 1. BIG DATA, МЕТОДИ ТА ТЕХНІКИ АНАЛІЗУ ВЕЛИКИХ ДАНИХ... 6	6
1.1 Ключові характеристики, роль та функції великих даних в інформаційних технологіях.....	6
1.2 Методи та техніки аналізу великих даних.....	13
1.3 Досвід використання великих даних в Україні та світі.....	19
Висновки до розділу 1	25
РОЗДІЛ 2. СТАН ЕКОСИСТЕМИ МОБІЛЬНОГО МЕСЕДЖИНГУ	26
2.1 Мобільний меседжинг як окрема бізнес-екосистема	26
2.2 Закономірності функціонування ринку телекомунікаційних послуг	32
2.3 Основні елементи екосистеми мобільного меседжингу	39
Висновки до розділу 2	44
РОЗДІЛ 3. BIG DATA РІШЕННЯ ОПЕРАТОРА МОБІЛЬНОГО МЕСЕДЖИНГУ НА ПРИКЛАДІ БІЗНЕС-КЕЙСУ ДЛЯ АВТОМОБІЛЬНОГО ЦЕНТРУ ТОВ «ГРАНД АУТОС»	46
3.1 Формування основних функціональних вимог та моделювання бізнес-процесів задачі просування продукції ТОВ «Гранд Аутос».	46
3.2. Веб-орієнтований додаток особистого кабінету для оптимізації процесу проведення таргетованих розсилок.	53
Висновки до розділу 3	61
ВИСНОВКИ ТА ПРОПОЗИЦІЇ:.....	63
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ:.....	65

ВСТУП

Актуальність теми дослідження обумовлена тим, що ринок телекомунікаційних послуг один із найбільш прогресивних та досить специфічні з тієї причини, що він тією чи іншою мірою охоплює майже все населення нашої держави та світу. Інформація в умовах активного розвитку продуктивних сил стає вагомим фактором виробництва та національного благополуччя, як наслідок, глобальним трендом стає інформатизація суспільства, яка сприяє залученню все більшого кола людей до обміну даними. За цих обставин галузь телекомунікацій вимагає детального аналізу ринкової кон'юнктури, особливо за умов олігополізації даної галузі, в тому числі і в Україні.

Використовуючи новітні досягнення у галузі інформаційних технологій задля результативного управління бізнесом допомагає асекурації своєчасності та деталізації інформації стосовно управлінських процесів, що надає спроможність здійснення поглибленого аналізу, моделювання процесів та прогнозування їхнього перспективного розвитку. Розробка автоматизованих рішень в умовах розвитку сучасних інформаційних технологій є важливим завданням. Автоматизація стосується нині практично усіх без винятку сфер господарської діяльності. Не виключення є і сфера торгівлі, зокрема продаж автозапчастин. Важливість впровадження інформаційних систем в автоматизацію бізнес-процесів є досить високою з огляду на доволі значні витрати часу працівників в умовах відсутності автоматизації. Наявність інформаційних систем дозволяє знизити витрати підприємств, посилити рівень контролю за організацією бізнес-процесів.

Дослідженням ринку мобільного зв'язку та наявної конкуренції присвячено дослідження таких зарубіжних і вітчизняних вчених та науковців як: В.М. Гранатурова, С.П. Воробієнко, Є.Г. Іванова, А. Пономарева, С. Таракановський, І. Балабанов та ін. Окремо олігополію та її вплив розглядали багато економістів. Зокрема, Й. Шумпетер навіть стверджував, що лише олігополія сприяє технічному прогресу. Технологія Big Data в свою чергу є також цікавою темою для досліджень, тому її вплив на організацію бізнесу став темою досліджень

таких вчених, як Н. Бессіс, В. Майєр-Шенбергер, К. Кукер, К. Девіс та інші. Оскільки ринок мобільного зв'язку достатньо швидко розвивається, як і інноваційні технології – це впливає на те, що недостатньо актуальної інформації у науковій літературі з приводу сучасного стану даного ринку, нових сервісів та послуг, тенденції розвитку та інше.

Мета роботи полягає у дослідженні ключових аспектів розвитку великих даних та їх обробки, інтеграції в роботу екосистеми мобільного меседжингу та вплив на вирішення задач відповідно до запитів ринку.

Завданнями даної роботи є наступні:

- визначити основні принципи роботи з великими даними, дослідити методи та техніки аналізу великих даних з використанням українського та світового досвіду;
- охарактеризувати мобільний меседжинг як окрему бізнес-екосистему та дослідити його складові, а саме захищені та класичні месенджери, їх особливості;
- дослідити основні бізнес-процеси підприємства ТОВ «Гранд Аутос»;
- запропонувати шляхи оптимізації комунікації, між операторами мобільного меседжингу та замовниками Big Data рішень;
- здійснити розробку програмного рішення для проведення таргетованих розсилок для підприємства ТОВ «Гранд Аутос».

Об'єктом дослідження процес обробки великих даних у сфері телекомунікацій.

Предметом дослідження є теоретичні, методичні та практичні підходи до організації обробки великих даних у екосистемі мобільного меседжингу.

Основними методами дослідження в магістерській роботі є як загальнонаукові методи (порівняння, дедукції та індукції, історико-логічний підхід, синтезу наукових категорій, тощо), так і спеціальні методи економічних досліджень (статистичний аналіз, графічний аналіз, економіко-математичні методи, методи статистичного моделювання, тощо).

Інформаційною базою виконання випускної кваліфікаційної роботи є: підручники та посібники з інформаційних технологій, періодична література, ресурси Інтернет.

Апробація результатів дослідження. За результатами проведеного дослідження підготовлено та опубліковано наукову статтю на тему: «Технології та інструменти обробки великих даних в екосистемі мобільного меседжингу» у збірнику наукових праць студентів КНТЕУ.

Практичне значення випускної кваліфікаційної роботи полягає у дослідженні особливостей та тенденцій роботи з великими масивами даних, проблем та перспектив подальшого функціонування в умовах конкуренції на ринку.

Структура роботи. Випускна кваліфікаційна робота структурно складається зі вступу, 3-х розділів, висновків та пропозицій, списку використаних джерел зі 32 найменувань та додатків. Загальний обсяг роботи становить 69 сторінок друкованого тексту.

РОЗДІЛ 1. BIG DATA, МЕТОДИ ТА ТЕХНІКИ АНАЛІЗУ ВЕЛИКИХ ДАНИХ

1.1 Ключові характеристики, роль та функції великих даних в інформаційних технологіях

У сучасному світі Big Data — соціально-економічний феномен, який пов'язаний з тим, що з'явилися нові можливості для аналізу величезної кількості даних. Величезні обсяги даних обробляються для того, щоб людина могла отримати конкретні і потрібні їй результати для їх подальшого ефективного застосування. Тому важливість великих даних досить складно переоцінити в сучасному житті. Основні підходи до визначення сутності великих даних (Big Data) наведено у табл. 1.1.

Таблиця 1.1

Основні підходи до визначення сутності великих даних (BigData)

№	Автор	Визначення
1	Барсегян А.А. [1]	різні інструменти, підходи та методи обробки як структурованих, так і неструктурованих даних для того, щоб їх використовувати для конкретних завдань та цілей.
2	Майер-Шенбергер [8]	позначення структурованих і неструктурованих даних величезних обсягів і значного різноманіття, що ефективно обробляються горизонтально масштабованими програмними інструментами
3	Резниченко В.А. [10]	величезні обсяги неоднорідної, неструктурованої або слабо структурованої, істотно розподіленої і інтенсивно зростаючої, змінної і використовуваної цифрової інформації, яку неможливо обробити традиційними засобами
4	Силен Д, Мейсман А., Али М. [11]	методи, технології та засоби їх збору, зберігання та обробки та аналізу з метою отримання сприйманих людиною результатів.
5	Беркана А. [2]	великі обсяги інформації, яка не має заздалегідь певної структури або не організована в певному порядку.

*складено автором на основі [1-11]

Термін «великі дані» запровадив редактор журналу Nature Кліффорд Лінч ще у 2008 році у спецвипуску, присвяченому вибуховому зростанню світових

обсягів інформації. Хоча, звичайно, найбільші дані існували і раніше. За словами фахівців, до категорії Big Data належить більшість потоків даних понад 100 Гб на день [15].

Як бачимо, у цьому визначенні є такі невизначені терміни, як «величезних», «значного», «ефективно» та «альтернативних». Навіть сама назва дуже суб'єктивна. Наприклад, 4 Терабайти (ємність сучасного зовнішнього жорсткого диска для ноутбука) – це вже великі дані чи ще ні? До цього визначення Вікіпедія додає таке: «у широкому розумінні про "великі дані" говорять як про соціально-економічний феномен, пов'язаний з появою технологічних можливостей аналізувати величезні масиви даних, в деяких проблемних областях — весь світовий обсяг даних, і з цього трансформаційних наслідків» [1].

Фактично, Big Data - це вирішення проблем та альтернатива традиційним системам управління даними. Обсяги неоднорідної цифрової інформації, що швидко надходить, обробити традиційними інструментами неможливо. Сам аналіз даних дозволяє побачити певні та непомітні закономірності, які не може побачити людина. Це дозволяє оптимізувати всі сфери нашого життя – від державного управління до виробництва та телекомунікацій [2].

Наприклад, деякі компанії ще кілька років тому захищали своїх клієнтів від шахрайства, а турбота про гроші – турбота про свої власні гроші. Перелічені особливості великих даних підштовхують до перегляду використання традиційних методів інформаційного взаємодії всіх рівнях організації, що у свою чергу, супроводжується необхідністю перебудови IT-інфраструктури організації, зокрема впровадження інформаційних систем, основу яких закладено технології Big Data – Big Data-орієнтованих інформаційних систем (IB) [2].

Big Data-орієнтована інформаційна системи – сукупність апаратно-програмних засобів, що дозволяє, накопичувати, обробляти та зберігати великий масив даних, тим самим надаючи можливість використання технології Big Data для аналізу цих даних. Концептуальну схему використання Big Data-орієнтованої інформаційної системи наведено на рис. 1.1.



Рис. 1.1 Концептуальна схема використання BigData-орієнтованої інформаційної системи

*складено автором на основі [2, 10]

Представлена концептуальна схема ілюструє місце застосування технологій Big Data за допомогою кількох рівнів та взаємодії між ними. Як видно, в основі лежить рівень, що відображає події та процеси у відповідному домені, що є джерелом Big Data (1-й рівень). На наступному рівні (2-й рівень) розгортаються бізнес-процеси та техніки первинного збору інформації. Третій рівень втілює Big Data-орієнтовану інформаційну систему. Як бачимо, ключовими компонентами тут є сховище даних. У ньому акумулюються дані, що надходять із 2-го рівня, на який, з іншого боку, повертається інформація у вигляді відповіді на запити користувачів даних. Важливим завданням, яке вирішується на даному рівні, є вибір оптимальної моделі сховища даних. Необхідність обробки та зберігання великого масиву інформації підштовхнула проектувальників баз даних до відмови від класичних реляційних СУБД на користь розподілених технологій обробки даних та парадигми NoSQL, або комбінування нових та класичних підходів [2].

Важливим елементом інформаційної Big Data-орієнтованої системи є аналітичний блок. На цьому рівні здійснюється аналіз великого масиву даних із репозиторію з використанням технологій Big Data. Дані технології активно застосовуються для вирішення багатьох прикладних бізнес-завдань і все активніше використовуються у сфері державного управління [11].

Для Big Data було сформульовано визначальні характеристики. Вперше у 2001 р. ознаки «Три V» виділив провідний аналітик Gartner Дуг Лані [2], а саме, обсяг, швидкість, різноманітність.

Volume (обсяг). Вважається, що Big Data починаються з обсягів у петабайти (10¹⁵ байт). Щоб уявити, що це за обсяг, наведемо приклад. У Національній бібліотеці України ім. В.І. Вернадського функціонує портал Наукової періодики України. Редакції понад 2700 наукових періодичних видань України надають усі свої статті упродовж 10 років. За цей час обсяг порталу становив близько 1 мільйона статей. Якщо припустити, що розмір статті в середньому становить 1 МБ, обсяг ресурсів цього порталу становить 1 ТБ. Це на три порядки нижче за мінімальний обсяг для Big Data, тобто через десять тисяч років успішного функціонування цього порталу він накопичить обсяги, характерні для Big Data [10].

Big Data з'являються тоді, коли сотні мільйонів людей об'єднуються в спільноти та викладають свої інформаційні ресурси, або об'єднані центри наукових досліджень надають дані результатів своїх досліджень, наприклад, у 2017 році дата-центр CERN перевищив розмір 200 петабайт і щорічно цей обсяг збільшується на 15 петабайт. Якщо помістити в DVD всі породжені в світі за день дані і покласти ці диски один на одного, то вийде чарка, що двічі перевищує відстань до Місяця.

Velocity (швидкість). Є однією з найважливіших характеристик Big Data з точки зору їх практичного використання. Під швидкістю мається на увазі як швидкість приросту (надходження, накопичення) даних, так і швидкість їх обробки з метою отримання кінцевих результатів. Крім того, до цієї категорії включаються характеристики інтенсивності та обсягів інформаційних потоків.

Для цього технологія обробки таких даних повинна допускати можливість їх аналізу вже в момент їхнього породження (іноді званою «оперативною аналітикою» – in-memory analytics), тобто до того, як вони потраплять до сховища даних [24].

Variety (різноманітність). Можливість сприймати, зберігати та обробляти різні дані. Говорячи про різноманіття, маються на увазі різні джерела отримання даних. Наведемо приклади джерел виникнення великих даних:

- дані, що безперервно надходять з вимірювальних пристроїв,
- події від радіочастотних ідентифікаторів,
- потоки повідомлень із соціальних мереж,
- метеорологічні дані,
- дані дистанційного зондування Землі,
- потоки даних про місцезнаходження абонентів мереж стільникового зв'язку, пристроїв аудіо- та відеореєстрації [24].

Можна навести різні способи представлення даних, наприклад сигнали, що надходять від датчиків, відрізняються від текстів наукових статей. Також існують різні формати зберігання (надходження) даних. Це можуть бути тексти, аудіо- та відео дані, зображення. Більше того, ті самі дані можуть бути представлені в різних форматах. Промова, що вимовляється людиною, може бути представлена в аудіо-форматі і у вигляді текстового файлу. Семантика тих самих даних може бути представлена по-різному, наприклад, вік людини може бути вказано кількісно або у вигляді таких термінів, як дитина, юнак, доросла людина.

Традиційні бази даних дозволяють зберігати структуровані дані, але фактично в даний час дані на 80%, що породжуються, є слабо структурованими або навіть неструктурованими. Технологія Big Data дозволяє об'єднувати та обробляти дані, що володіють наведеним вище різноманіттям. Зікопоулоус [8] запропонував додати ще 2 ознаки – достовірність та цінність (значимість), таким чином отримавши «5V»:

Veracity (достовірність). Властивість, що характеризує надійність даних. Технологія створення та використання традиційних БД передбачає, що до БД

надходять ретельно відібрані та перевірені дані. У Big Data справа інакша. Вихідні дані можуть бути «сирими» (неповними, неточними, нечіткими, розпливчастими, спотвореними), тобто надходять без будь-якої попередньої обробки, вони можуть бути суб'єктивними, випадковими та містити багато «шуму». Ще один критерій цієї характеристики - ступінь довіри до даних, що надходять. Хоча Big Data надають чудові можливості для аналізу та прийняття рішень, проте їхня цінність багато в чому залежить від якості вихідних даних. Технологія Big Data враховує цю характеристику та дозволяє надійно працювати з такими даними [11].

Value (цінність). Коли ми говоримо про цінність даних, то маємо на увазі їх значущість з погляду прикладних завдань. За розрахунками IBS, лише 1,5 % накопичених масивів даних має інформаційну значимість. Велика кількість даних - це добре, але якщо вони не становлять жодного інтересу, то вони марні.

Згодом стали пропонувати додаткові визначальні характеристики BigData [10], які отримали назву "7V" та "10V". Наведемо цей додатковий перелік.

Variability (мінливість). Під мінливістю в Big Data мається на увазі ситуація, коли постійно змінюється зміст даних. Наприклад, це має місце, коли збирання та обробка даних відбувається в процесі аналізу природничо-мовних тестів і особливо при перекладі з однієї мови на іншу.

Volatility (волатильність, актуальність). Характеристика, яка визначає, який період старіння даних, коли вони стають нерелевантними або марними. Як довго їх треба зберігати? До ери Big Data дані могли зберігатися невизначено довго, використання цих цілей кілька десятків терабайт був обтяжливим. Більше того, їх можна було зберігати в базі даних, що не діє, не викликаючи при цьому проблем з продуктивністю. Однак за наявності Big Data, враховуючи характеристики об'єму та швидкості, слід ретельно стежити за волатильністю даних. Необхідно встановити правила управління зберіганням даних для того, щоб забезпечити ефективне їх використання.

Vulnerability (уразливість). Великі дані породжують нові проблеми їхньої безпеки. Зламування великих даних призводить до великого злону. Прикладом

може бути зламування бази даних соціальної мережі LinkedIn, в результаті якого було викрадено 167 млн облікових записів і 360 мільйонів відомостей про e-mail.

Validity (придатність, обґрунтованість). Ця характеристика тісно пов'язана з достовірністю і характеризує, якою мірою наявні дані є точними і правильними з точки зору їх передбачуваного використання. За оцінкою Forbes [10] вчені так витрачають свій час для роботи з великими даними (рис. 1.2):

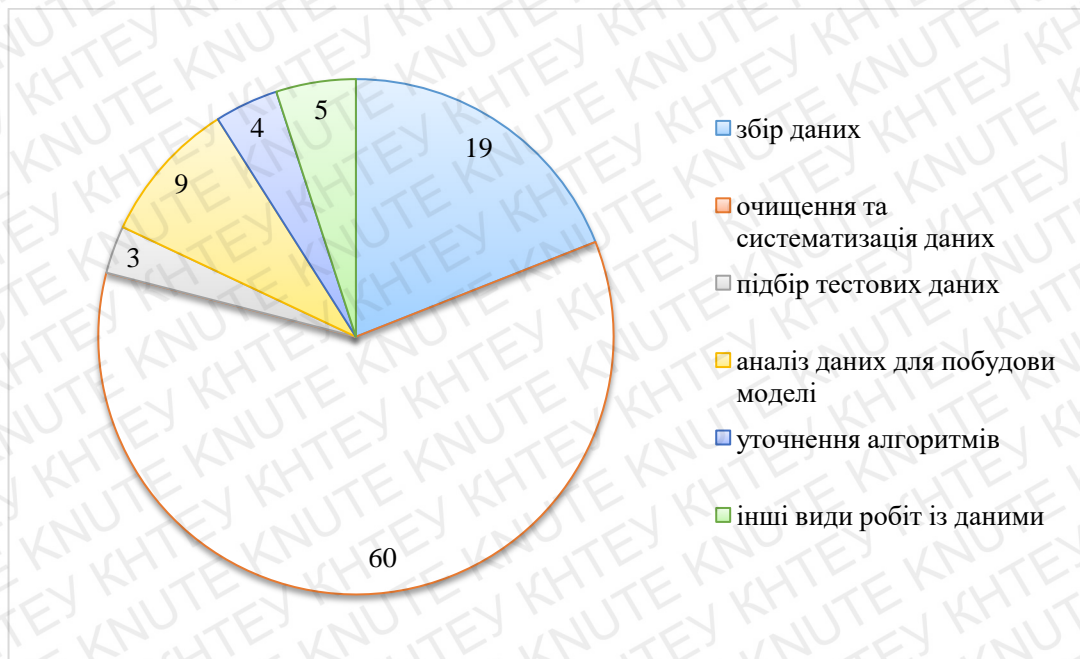


Рис. 1.2. Структура витрат часу на роботу з великими даними

*складено автором на основі [2]

Таким чином, учений витрачає 80% свого часу на підбір і підготовку даних, перш ніж приступити до їх аналізу. Перевагою використання великих даних для проведення аналітичних досліджень можна повною мірою скористатися тільки тоді, коли дані ретельно відібрані, є релевантними та достовірними.

Visualization (Візуалізація). Після отримання та обробки даних їх треба подати таким чином, щоб вони були читабельними та доступними. Саме це має на увазі візуалізація.

Отже, Big Data можна визначити як величезні обсяги неоднорідної, неструктурованої або слабо структурованої, істотно розподіленої і інтенсивно зростаючої, змінної і використовуваної цифрової інформації, яку неможливо

обробити традиційними засобами. Використання Big Data дозволяє побачити певні та непомітні закономірності, які не може побачити людина. Це дозволяє оптимізувати всі сфери нашого життя – від державного управління до виробництва та телекомунікацій. Основними характеристиками Big Data є volume (обсяг), velocity (швидкість), variety (різноманітність), veracity (достовірність), value (цінність), variability (мінливість), volatility (волатильність, актуальність), vulnerability (уразливість), validity (придатність, обґрунтованість), visualization (візуалізація).

1.2 Методи та техніки аналізу великих даних

На сьогодні створено та адаптовано безліч методів та технологій для збору, агрегування, маніпулювання, аналізу та візуалізації великих даних. Ці методи та технології запозичені з різних галузей, включаючи статистику, інформатику, прикладну математику та економіку. Це означає, що для отримання вигоди з великих даних слід використовувати гнучкий міждисциплінарний підхід. Деякі методи та технології були розроблені для оперування значно меншими обсягами та різноманітністю даних, але були успішно адаптовані для Big Data. Інші були розроблені останнім часом, зокрема, для збирання та аналізу великих даних. Далі наводиться перелік і короткий опис методів і технологій аналізу та візуалізації, що застосовуються до Big Data, які взяті зі звіту McKinsey [10].

1. Методи аналізу Big Data Методи класу Data Mining:

- навчання асоціативним правилам (association rule learning) - це метод, що базується на правилах, використовується для навчання машин способом виявлення залежностей між даними у великих базах даних;
- класифікація - методи категоризації нових даних на основі принципів, раніше застосованих до даних, що вже є;
- кластерний аналіз - статистичний метод класифікації об'єктів, що призводить до поділу різноманітних груп більш дрібні групи подібних (подібних) об'єктів, котрим критерій подоби заздалегідь не відомий;

- Регресійний аналіз.

2. Краудсорсинг (crowdsourcing) - спосіб збору, категоризація та збагачення даних силами широкого кола осіб, залучених на підставі публічної оферти, без вступу до трудових відносин, зазвичай за допомогою використання мережевих медіа.

4. Змішання та інтеграція даних (data fusion and integration) - набір методів, що дозволяють інтегрувати та аналізувати різноманітні дані з різноманітних джерел для глибокого аналізу більш точно та ефективно, ніж з єдиного джерела даних. Як приклади методів цього класу є цифрова обробка сигналів та обробка природної мови.

5. Навчання асоціативним правилам (association rule learning). Сукупність методів для аналізу необхідних взаємозв'язків, тобто «асоціативних правил» серед змінних у великих базах даних [3].

6. Машинне навчання (machine learning). Клас методів штучного інтелекту, характерною рисою яких є непряме розв'язання задачі, а навчання в процесі застосування рішень безлічі подібних задач. Включає навчання з вчителем (supervised learning) та без вчителя (unsupervised learning), а також Ensemble learning – використання моделей, побудованих на базі статистичного аналізу чи машинного навчання для отримання комплексних прогнозів на основі базових моделей (constituent models) [10].

7. Обробка природної мови (Natural language processing - NLP). Загальний напрямок штучного інтелекту та математичної лінгвістики. Воно вивчає проблеми комп'ютерного аналізу та синтезу природних мов. Стосовно штучного інтелекту аналіз означає розуміння мови, а синтез - генерацію грамотного тексту. Багато методів NLP є методами машинного навчання [2].

8. Штучні нейронні мережі (Artificial neural networks). Математична модель, побудована за принципом організації та функціонування біологічних нейронних мереж – мереж нервових клітин живого організму.

9. Мережевий аналіз (Network Analysis). Набір методів, що використовуються для опису та аналізу відносин між дискретними вузлами у

графі чи мережі. В аналізі соціальної мережі аналізуються зв'язки між людьми у спільноті чи організації, наприклад, як переміщується інформація чи хто має найбільший вплив на когось.

10. Розпізнавання образів (pattern recognition). Набір методів машинного навчання, що розвивають основи та методи класифікації та ідентифікації предметів, явищ, процесів, сигналів, ситуацій тощо об'єктів, які характеризуються кінцевим набором деяких властивостей та ознак [10].

11. Прогнозна аналітика (Predictive Analytics). Клас методів аналізу даних, що концентрується на прогнозуванні майбутньої поведінки об'єктів та суб'єктів з метою прийняття оптимальних рішень [27].

12. Аналіз тональності тексту (Sentiment analysis). Клас методів контент-аналізу в комп'ютерній лінгвістиці, призначений для автоматизованого виявлення в текстах емоційно забарвленої лексики та емоційної оцінки авторів (думок) стосовно об'єктів, про які йдеться в тексті.

13. Імітаційне моделювання (simulation modeling) - метод дослідження, у якому досліджувана система замінюється моделлю, з достатньою точністю описує реальну систему (побудована модель описує процеси оскільки вони проходили насправді), з якою проводяться експерименти, з одержання інформації про цю систему.

14. Просторовий аналіз (Spatial analysis) – набір методів, які аналізують топологічні, геометричні чи географічні властивості, представлені в наборі даних. Часто дані для просторового аналізу надходять із географічних інформаційних систем (ГІС) [19].

15. Статистичний аналіз, приклади: А/В-тестування (контрольна група елементів порівнюється з набором тестових груп, в яких один або кілька показників були змінені, для того, щоб з'ясувати, які зміни покращують цільовий показник) та аналіз часових рядів.

16. Аналіз часових рядів (time series analysis) - сукупність математико-статистичних методів аналізу, призначених виявлення структури часових рядів та їх прогнозування. Сюди належать, зокрема, методи регресійного аналізу.

Виявлення структури часового ряду необхідно для того, щоб побудувати математичну модель того явища, яке є джерелом аналізованого часового ряду.

Існує безліч технологій для агрегації, маніпулювання, управління та аналізу великих даних. У табл. 1.2 наводиться список найбільш відомих технологій і засобів, що використовуються.

Таблиця 1.2

Технології управління великими даними (Big Data)

<i>Назва технології</i>	<i>Характеристика технології</i>
Big Table	Запатентована розподілена система баз даних побудована на основі Google File System.
Business intelligence (BI) (бізнес-аналітика)	Сукупність методологій, процесів, архітектур та технологій, які перетворюють великі обсяги «сирих» даних в осмислену та корисну інформацію, придатну для бізнес-аналізу та для підтримки прийняття оптимальних тактичних та стратегічних рішень.
Cassandra	Система управління базами даних, що вільно розповсюджується, призначена для маніпулювання даними величезного обсягу в розподілених системах.
Cloud computing (хмарні обчислення)	Обчислювальна парадигма, в якій обчислювальні ресурси, що високомасштабуються, зазвичай сконфігуровані у вигляді розподілених систем, надаються в мережах як сервіси.
Data Warehouse (сховище даних)	Предметно-орієнтована інформаційна база даних, спеціально розроблена та призначена для підготовки звітів та аналізу даних з метою підтримки прийняття рішень в організації та є однією з основних компонентів бізнес-аналізу. Виступає центральним репозиторієм даних, що надходять із різних джерел. Зберігає поточні та історичні дані. Будується на базі систем управління базами даних та систем підтримки прийняття рішень.
Distributed system (розподілена система)	Безліч комп'ютерів, що взаємодіють через мережу та об'єднані для вирішення загального обчислювального завдання.
Dynamo	Запатентована розподілена система зберігання даних розроблена в Amazon.
Google File System	Запатентована розподілена файлова система. На її основі збудовано Hadoop
Hadoop	Проект фонду Apache Software Foundation, набір утиліт, що вільно розповсюджується, бібліотек і фрейм-ворк для розробки та виконання розподілених програм, що працюють на кластерах із сотень і тисяч вузлів. Використовується для реалізації пошукових та контекстних механізмів багатьох високонавантажених веб-сайтів, у тому числі для Yahoo! та Facebook. Базується на MapReduce та Google File System.
Extract, transform, and load (ETL)	ППР, що використовується для вилучення даних із зовнішніх джерел, перетворення їх для задоволення операційних потреб, і завантаження їх у базу даних або сховище даних.

<i>Назва технології</i>	<i>Характеристика технології</i>
HBase	розподілена нереляційна база даних, що вільно розповсюджується, створені на основі Big Table Google.
MapReduce	Модель розподілених обчислень, представлена Google, що використовується для паралельних обчислень над дуже великими, аж до декількох петабайт, наборами даних у комп'ютерних кластерах. Ця модель реалізована в Hadoop.
Mashup	Веб-додаток, що об'єднує дані з кількох джерел в один інтегрований, наприклад, при об'єднанні картографічних даних Google Maps з даними про нерухомість з Craigslist виходить новий унікальний веб-сервіс, який спочатку не пропонує жодним джерелом даних.
R	Розповсюджувана мова програмування середовище програмування для статистичних і графічних обчислень.
Stream processing	Технологія призначена для обробки великих потоків даних у реальному масштабі часу.

*складено автором на основі [2, 10, 11]

Виходячи з наведених вище визначень, основні принципи роботи з великими даними такі [2, 10, 11]:

Горизонтальна масштабованість. Це базовий принцип обробки великих даних. Як вже попередньо було зазначено, великих даних з кожним днем стає дедалі більше. Відповідно, необхідно збільшувати кількість обчислювальних вузлів, якими розподіляються ці дані, причому обробка повинна відбуватися без погіршення продуктивності.

Відмовостійкість. Цей принцип впливає із попереднього. Оскільки обчислювальних вузлів у кластері може бути багато (іноді десятки тисяч) та їх кількість, не виключено, збільшуватиметься, зростає і ймовірність виходу машин з ладу. Методи роботи з великими даними мають враховувати можливість таких ситуацій та передбачати превентивні заходи.

Локальність даних. Так як дані розподілені за великою кількістю обчислювальних вузлів, то якщо вони фізично знаходяться на одному сервері, а обробляються на іншому, витрати на передачу даних можуть стати невиправдано більшими. Тому обробку даних бажано проводити на тій же машині, де вони зберігаються.

Ці принципи відрізняються від тих, що характерні для традиційних, централізованих, вертикальних моделей зберігання добре структурованих даних. Відповідно, для роботи з великими даними розробляють нові підходи та технології.

На рис. 1.3 представлено порівняння особливостей аналізу з використанням інструментів класичної статистичної та інструментів BigData.

	Традиційна статистика	Big Data
Орієнтація	Ретроспектива	Прогнозування
Питання	Що трапилось? Коли? Як?	Що відбудеться? Які можуть бути наслідки? Як зміняться параметри?
Генерування знань	Мануальні способи	Автоматично
Типи даних	Структуровані дані, висока якість вибірки, цільові дані (1 задача – 1 набір даних)	Структуровані та неструктуровані дані, великі масиви, використання суміжних них
Методи	Звіти, моніторинг, OLAP, графіки	Data Mining, Text Mining, Business Intelligence, simulation modeling ...

Рис. 1.3. Порівняння особливостей аналізу з використанням традиційної статистики та BigData

*складено автором на основі [2, 10, 11]

Виходячи з представленої схеми можна зробити висновок, що аналіз даних за допомогою статистичних методів більше орієнтований на пояснення того, «що було?», тобто минулої динаміки, тоді як Big Data більше орієнтована на майбутнє і відповідає питанням «що буде?». Методи статистики більш залежні від якості даних та працюють переважно з вибірками щодо невеликого обсягу. Big Data в свою чергу менш вимоглива до якості даних, що успішно справляється з неструктурованими різномірними даними, а також з великим їх обсягом. Таким

чином, використання технологій Big Data, з урахуванням сучасних тенденцій накопичення великих даних, є більш доцільним та ефективним.

Отже, основними техніками та методами аналізу, застосовні до Big Data по McKinsey є такі: Data Mining; краудсорсінгзмішування та інтеграція даних; машинне навчання; штучні нейронні мережі; розпізнавання образів; прогнозна аналітика; імітаційне моделювання; просторовий аналіз; статистичний аналіз; візуалізація аналітичних даних. Технологіями, що викистовуються для управління великими даними, є NoSQL, MapReduce, Hadoop, R та різні апаратні рішення.

1.3 Досвід використання великих даних в Україні та світі

Аналітики дослідження вважають, що дані стануть життєво важливим активом, а безпека – критично важливим фундаментом у житті. Також автори роботи впевнені, що технологія змінить економічний ландшафт, а звичайний користувач комунікуватиме із підключеними пристроями близько 4800 разів на день.

За даними компанії IBS, до 2003 року світ накопичив 5 ексабайтів даних (1 ЕБ = 1 млрд гігабайтів). До 2008 року цей обсяг зріс до 0,18 зеттабайта (1 ЗБ = 1024 ексабайти), до 2011 року - до 1,76 зеттабайта, до 2013 року - до 4,4 зеттабайта. У травні 2015 року глобальна кількість даних перевищила 6,5 зеттабайт (докладніше).

У 2020 році людство сформувало 40-44 зеттабайти інформації. За прогнозами, до 2025 року обсяг даних зросте в 10 разів, йдеться у доповіді The Data Age 2025, яку підготували аналітики компанії IDC. У доповіді наголошується, що більшу частину даних генеруватимуть самі підприємства, а не звичайні споживачі (рис. 2.4).

Також, варто додати, що досить швидке зростання обсягу великих даних прогнозується завдяки розитку IoT (інтернету речей), коли інформація буде накопичуватися різними пристроями. За словами фахівців, до категорії Big Data належить більшість потоків даних понад 100 Гб на день.

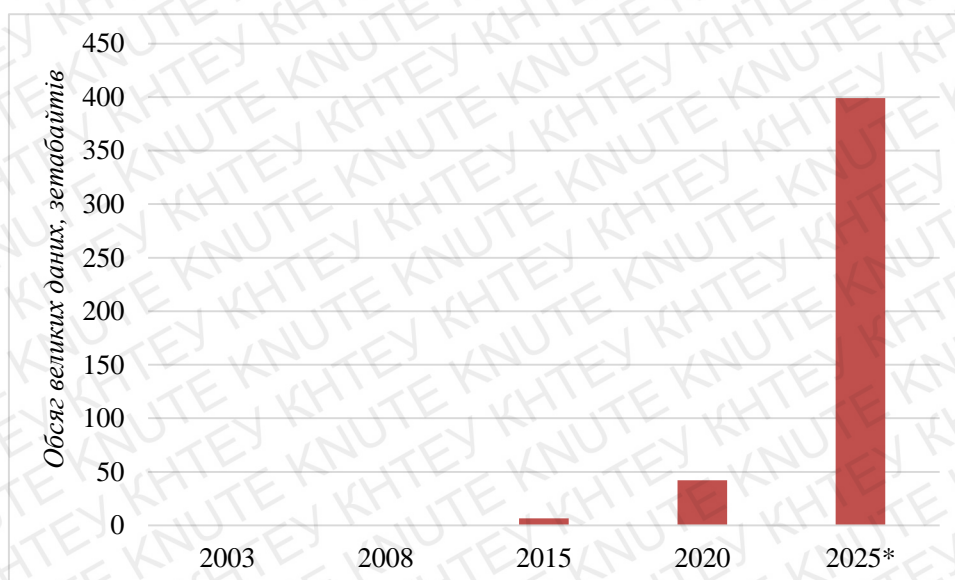


Рис. 1.4. Динаміка обсягу великих даних у світі, зетабайтів

*складено автором на основі

Можна навести декілька фактів, що характеризують темпи та специфіку формування великих даних, які відносяться до соціальних мереж (табл. 1.3).

Таблиця 1.3

Темпи та специфіку формування великих даних, які відносяться до соціальних мереж

Соціальна мережа	Кількість користувачів	Характеристика формування великих даних
YouTube	1,9 мільярди	Щохвилини закачується нових фільмів на 100 годин і завантажується фільмів на 700 тисяч годин.
Facebook	1,4 мільярди	Щодня на сайт вивантажується 100 терабайт даних і щохвилини ставляться понад 34 тисячі лайків. Щохвилини завантажується 200 000 фотографій. Щомісяця викладається у відкритий доступ 30 млрд нових джерел інформації
Twitter	645 мільйонів	Щодня генерується 175 мільйонів твітів.
Google	2937,6 млрд.записів	Щохвилини відпрацьовується 2,4 мільйона пошукових запитів (40 000 запитів на секунду). Щодня обробляється 25 петабайт даних. Щохвилини у світі посилається 204 мільйони е-листів

*складено автором на основі [10]

Американська компанія SAS на своєму сайті поділилася успішним досвідом виявлення шахрайських схем із ПДВ у Бельгії. Застосування розробленої

компанією "гібридної моделі" дозволило Спеціальній податковій інспекції скоротити фінансові втрати Бельгії на 98% [4].

У той же час, згідно з підрахунками Британського експертного центру Policy Exchange, аналіз даних дозволяє виявляти акти шахрайства та підвищувати збирання податків, що дозволить країні заощаджувати до 33 млрд фунтів щорічно [26].

Крім виявлення шахрайства, аналіз «Великих даних» може підвищити ефективність державного управління в цілому. Наприклад, на основі аналізу інформації з баз даних Керівництву служби зайнятості Німеччини вдалося виявити, які заходи безробітним є дієвими, а які на практиці виявилися марними. Це дозволило не лише заощадити бюджет, а й підвищити загальну ефективність роботи служби зайнятості.

Як видно з прикладів, застосування технологій аналізу великих даних має величезний потенціал, при цьому величезні інформаційні масиви дозволяють отримувати найнесподіваніші результати. Варто також звернути увагу на наступне питання: чому використання технологій Big Data є більш доцільним, ніж використання звичайної статистики та бізнес-аналітики, які використовують стандартні статистичні методи аналізу та подання даних [10].

Зі зростанням кількості гаджетів, розвитком пропускної можливості стільникових та провідних мереж та появою нових інформаційних технологій, обсяг економіки даних зростає по експоненті. А разом з нею і можливості, які робота з великими даними відкриває. Робота з Big Data змінює на очах абсолютно кожен сферу людської активності — від розваг до охорони здоров'я, безпеки до продуктів харчування.

Зараз, коли значна частина людства в тій чи іншій формі на карантині, самоізоляції, а то й у ситуації економічної невизначеності, саме час поглянути на те, яку роль великі дані відіграють у нашому житті, а головне — яку роль вони відіграватимуть у найближчому майбутньому [10].

Великі дані визначають не тільки який фільм чи серіал вам запропонувати подивитися наступним, але й які фільми вам знімати. Голлівуд вже впритул

наблизився до застосування Big Data при написанні сценаріїв та кастингу акторів, але поки що знаходиться лише на початку шляху. І зовсім необов'язково обмежуватися лише сегментуванням та аналізом аудиторії. Big Data дозволяє працювати у такий спосіб і з самим контентом. Наприклад, Netflix аналізує та сортує свою величезну бібліотеку контенту за більш ніж 70 тисячами характеристик. Після кіно і телеіндустрією великі дані змінюють і світ музики. Хмарні сервіси, від еппловського iTunes до YouTube Music оперують величезними масивами даних про користувацькі уподобання і звички.

Великі дані змінюють усі напрямки креативної промисловості. Остання гучна подія у світі подкастів — анонсований «переїзд» подкасту Джо Рогана, Joe Rogan Experience, на платформу Spotify, оцінюється у суму щонайменше \$100 мільйонів. При цьому сам контент шоу Рогана - як весь архів, так і нові випуски, будуть повністю безкоштовні для користувачів платформи і програми Spotify. На думку багатьох аналітиків, мета дорогої «купівлі» якраз доступ до даних багатомільйонної аудиторії слухачів шоу: їх смакам, інтересам, перевагам і звичкам.

Спорт — ще один напрямок у світі розваг, що знаходиться на порозі інформаційної революції. Футбольні команди у Великій Британії «виявилися помічені» у використанні розумних датчиків для виміру становища їх гравців на полі та їх пульсу. Але це тільки початок: багато баскетбольних і бейсбольних команд аналізують дані відеозаписів і датчиків як своїх гравців, так і команд-суперників, перетворюючи поле битви м'язів і духу ще й у поле бою даних.

Сенсори і датчики, відеоспостереження і пристрої, що носяться, стають реальністю не тільки людського життя, а й сільського господарства. The Climate Corporation, підрозділ агрогіганта Monsanto веде роботу з агрохолдингами, фермерськими господарствами та розробниками софту для ретейлу, об'єднуючи їх у єдину потужну мережу Climate FieldView, що базується на великих даних. Climate FieldView дає фермерам детальну візуалізацію процесів сільськогосподарського виробництва, об'єднану з даними про потреби ринку продовольства, які дозволяють їм ухвалювати ефективні інформовані рішення.

Фермерам доступні дані про погоду, стан ґрунту, вологість, дозрівання плодів, хід зростання та умови для худоби. Ця інформація дозволяє максимізувати та оптимізувати виробництво продукції під потреби ринку у реальному часі. У Велику депресію 1930-х років фермерським господарствам часто доводилося знищувати зібраний хліб та виливати надоєне молоко. За допомогою технологій великих даних ХХІ столітті подібних проблем можна буде уникнути і потреби таких драматичних заходів вже не залишиться. Завдяки великим даним, інтернету речей і хмарним технологіям якісніші продукти стають доступні дедалі більшій кількості людей.

Онлайн-шопінг можна вважати вже традиційним та передбачуваним способом застосування великих даних. Компанія Amazon була одним із перших, що використали великі дані в інтернет-продажах, порівнюючи покупки зроблені їх користувачами, порівнюючи їх з тим, що купували інші і намагаючись передбачити, що їх клієнту може знадобитися або захочеться придбати слідом. Зараз ці технології застосовуються практично у будь-якому великому інтернет-магазині.

Онлайн-реклама – інше важливе застосування великих даних. Можливо, великі дані зможуть «вбити» рекламу у звичному розумінні. Адаже що зазвичай називають рекламою? Нав'язливе інформування про щось непотрібне. Але при цьому у всіх людей є потреби та проблеми, на пошуки вирішення для яких ми витрачаємо величезну кількість часу та сил. Реклама, що відповідає існуючим потребам користувача — це вже не реклама, а корисна порада [3].

Як показав аналіз поведінки 20 мільйонів власників карток, у разі двох подібних інцидентів протягом півроку, коли транзакція, що здійснюється користувачем, приймалася за підозрілу, витрати з «підозрілої» картки в наступні шість місяців знижувалися в середньому на 15%, а кожен п'ятий власник взагалі переставав їй користуватися. Для бізнесів важливо не лише боротися з фродом, але й зробити цю боротьбу якнайменш помітною для користувачів. Аналіз великих даних за допомогою машинного навчання дозволяє створювати моделі

добропорядної поведінки користувачів, розповідає Курт Лонг, засновник компанії захисту даних FairWarning.

Ще важливіша роль Big Data у контролі поширення коронавірусу. Як показує вже накопичений досвід, багато випадків спалахів пов'язані з супер-івентами, коли масове зараження почалося в певному місці або під час великої події. Наприклад, останній помітний спалах захворювань у Південній Кореї був викликаний лише однією людиною, яка, будучи безсимптомним носієм, за одні вихідні тусовок у клубному кварталі Сеула заразила сотні людей [2].

Тому багато країн світу, від Німеччини до Індії вже випускають програми, які дозволяють у тому чи іншому вигляді відстежувати переміщення та соціальні контакти користувачів, щоб у разі виявлення захворювання в одного з них вчасно попередити необхідність самоізоляції інших людей.

Фактично робота з великими даними дозволяє зробити карантин розумним, мінімізуючи наслідки пандемії як для здоров'я людей, так і для економік. Складність, масштабність і тривалість розробки вакцини, її виведення ринку і вакцинування мільярдів людей у світі означає, що загроза коронавірусу мине швидко. За прогнозами вчених, до масового вакцинування ще менше року. Але посадити економіку всього світу на самоізоляцію на рік не вдасться. Розумний карантин із контролем поширення вірусу через роботу з великими даними — головне, що допоможе людству пережити час, що залишився до появи ефективного засобу боротьби з самим коронавірусом.

Отже, великі дані набули свого поширення практично у всіх сферах людського життя, які мають відношення до автоматизації та наявності певних систем, які здатні фіксувати інформацію. Кожні 5 років обсяг великих даних мав тенденцію до значного розширення, починаючи з 2000-х років. Досить швидке зростання обсягу великих даних прогнозується завдяки розитку IoT (інтернету речей), коли інформація буде накопичуватися різними пристроями. За словами фахівців, до категорії BigData належить більшість потоків даних понад 100 Гб на день.

Висновки до розділу 1

Визначення поняття Big Data характеризується як великі обсяги неоднорідної, неструктурованої або слабо структурованої, істотно розподіленої і інтенсивно зростаючої, змінної і використовуваної цифрової інформації, яку неможливо обробити традиційними засобами. Використання Big Data дозволяє побачити певні та непомітні закономірності, які не може побачити людина.

Було розглянуто техніки та методи аналізу, застосовні до BigData по McKinsey є такі: Data Mining; краудсорсінгзмішування та інтеграція даних; машинне навчання; штучні нейронні мережі; розпізнавання образів; прогнозна аналітика; імітаційне моделювання; просторовий аналіз; статистичний аналіз; візуалізація аналітичних даних.

Доведено, що великі дані набули свого поширення практично у всіх сферах людського життя, які мають відношення до автоматизації та наявності певних систем, які здатні фіксувати інформацію. Кожні 5 років обсяг великих даних мав тенденцію до значного розширення, починаючи з 2000-х років. Досить швидке зростання обсягу великих даних прогнозується завдяки розвитку IoT (інтернету речей), коли інформація буде накопичуватися різними пристроями.

РОЗДІЛ 2. СТАН ЕКОСИСТЕМИ МОБІЛЬНОГО МЕСЕДЖИНГУ

2.1 Мобільний меседжинг як окрема бізнес-екосистема

Мобільні програми стали невід'ємною частиною бізнесу. З ними відкрилися і нові можливості — використання додатків як каналів просування та продажу, і нові загрози — зламування додатків, витік інформації, втрата ділової репутації. За 2019 рік світова економіка втратила 2,5 трильйона доларів через кіберзлочини, переважно хакерські атаки і зломи. У цю суму входять збитки, спричинені втратою конфіденційних даних, крадіжкою інтелектуальної власності, порушенням бізнес-операцій [27].

Конфіденційність є найчутливішою темою доби мобільного Інтернету. Регулярно з'являються новини про порушення безпеки через недоліки додатків. Ніщо не безпечне, навіть месенджери техногігантів. Щодня внутрішні повідомлення просочуються до мережі, а облікові записи викрадаються. Причин цьому маса - від недбалого обслуговування бази даних до використання найпростіших паролів. Такий стан справ підвищив попит на створення месенджерів із шифруванням, їх ще називають захищеними чи безпечними [29].

Розглянемо принцип роботи захищених та класичних месенджерів. Додатками для обміну повідомленнями користується щонайменше половина населення планети. Одні тільки WhatsApp та Facebook Messenger нараховують 2 мільярди та 1,3 мільярда активних користувачів відповідно x32i.

Більшість популярних месенджерів будуються на двох типах безпечних протоколів обміну повідомленнями [28]:

1. HTTP + push-сповіщення. Це коли користувач отримує повідомлення про те, що надійшло повідомлення, а сервер відповідає вже тільки після того, як відкрита програма.
2. Розширюваний протокол обміну повідомленнями (XMPP) на основі сокетів. Цей протокол використовується частіше, тому що тут завжди

залишаються підключення до сервера, а якщо раптом з'єднання зникне, то месенджер перейдете в автономний режим.

Кожен месенджер має свою мету і завдання, як і слабкі місця, що загрожують кібербезпеці. А зловмисники саме вибирають програми, які обробляють конфіденційні дані. Вони потім можуть використовуватися проти користувачів компанії, які ведуть приватне чи корпоративне листування. Інформація, що передається в їхніх повідомленнях, будь то текст, зображення або відео, повинна бути надійно захищена [27].

Кожен додаток для обміну повідомленнями використовує власні методи забезпечення безпеки, але головна різниця між захищеними та класичними месенджерами полягає в галузі реалізації та підходу до даних користувача. Захищені месенджери стали окремою категорією через усвідомлення, що спілкування в Інтернеті небезпечно — там третя сторона може отримати доступ до листування. Відомі випадки, коли великі компанії використовували дані користувачів, включаючи особисті повідомлення для таргетованої реклами [27].

Основним принципом безпечного обміну повідомленнями є наскрізне шифрування. Наприклад, програма Signal працює на однойменному криптографічному протоколі, розробленому Open Whisper Systems ще для першої програми (TextSecure). Такий тип шифрування використовує багаторівневий підхід, який ускладнює грубе проникнення листування з доступом до даних. Signal Protocol також застосовується за умовчанням у WhatsApp, і як додаткова функція у Facebook Messenger (Secret conversations) та Skype (Private conversations) [27].

Будь-який захищений месенджер будується на принципі наскрізного шифрування, який має такий вигляд [27]:

Два користувача розпочинають діалог. Ця подія генерує два набори ключів:

- закритий ключ (залишається на пристрої користувача);
- відкритий ключ (зберігається на сервері постачальника послуг).

Коли перший користувач пише другому, відкритий ключ виймається та використовується для шифрування повідомлення, щоб воно було доступне лише

через закритий ключ. Потім повідомлення надсилається другому користувачу через сервер і розшифровується за допомогою закритого ключа [27].

У такій схемі дані, що зберігаються на сервері, є в зашифрованому вигляді. Вони виглядають як набір літер із цифрами, які ніхто не може прочитати, не маючи закритого ключа. А ключ настільки складний, що методи дешифрування йому не підходять — його випадково згенеровані символи не можна порівняти з символами відкритого ключа. Якщо зловмисник не може знайти спосіб отримати закритий ключ із пристрою користувача, то ймовірність прочитання листування та отримання інформації зводиться до нуля.

Можливість видалення повідомлень – це ще одна важлива частина роботи месенджингу. Хоча функція видалення повідомлень є в багатьох месенджерах, не можна бути впевненим на 100%, що повідомлення дійсно видаляються з серверів додатків та баз даних. Тільки постачальники послуг знають правду, але не розкривають її. Достовірно відомо, що Facebook та Google зберігають повідомлення з листування. Але так не у всіх месенджерах, принаймні можна лише сподіватися на це. У будь-якому випадку, можливість видалення повідомлень — великий крок до побудови довірчих відносин з користувачами [29].

Метадані також вважаються проблемним елементом за дотримання конфіденційності. Вони використовуються для ідентифікації користувачів та їх облікових даних. Більшість месенджерів за замовчуванням зберігають метадані повідомлень, такі як час, відправник та одержувач, список контактів, ідентифікатор пристрою. Ця інформація може бути використана хакерами для ідентифікації користувача та застосування соціальної інженерії, щоб отримати ключ дешифрування. Як приклад програми, що зберігає метадані, можна навести той самий WhatsApp [29].

Прозорість для захищених месенджерів звучить подвійно. З одного боку, у користувача мають бути прописані певні умови обслуговування, в яких йдеться про намір надати користувачеві безпечну платформу із захистом конфіденційності. Отже, листування та дані про користувачів мають бути

приватними, а не відкритими. З іншого боку, справжньою ознакою безпеки програми є відкритий код, який кожен можна перевірити на надійність. Водночас це легкий спосіб підвищити якість силами програмістів-ентузіастів, які готові проводити безкоштовне тестування цікавих новинок.

Разом виходить, що цінність захищених месенджерів ґрунтується на чотирьох принципах: 1. Наскрізне шифрування; 2. Можливість видалення повідомлень; 3. Обмежене використання метаданих; 4. Прозорість сервісу та продукту. Розглянемо основні можливості, які надає користувачам захищений месенджер (табл. 2.1).

Таблиця 2.1

Можливості, які надає користувачам захищений месенджер

<i>Назва можливості</i>	<i>Характеристика</i>
Реєстрація за номером телефону	Зручність має стояти на першому місці. Враховуючи, що користування починається з реєстрації, воно стане простішим, якщо людям не доведеться запам'ятовувати паролі та іншу інформацію, необхідну для входу в додаток. До того ж зловмисники можуть отримати доступ до електронної пошти, а до сім-картки навряд чи.
Видалення повідомлень та облікового запису	Можливість видалення повідомлень — невід'ємна функція, яка має бути у захищеному месенджері. У деяких месенджерах можна увімкнути видалення повідомлень через проміжок години. Також і з видаленням облікового запису: якщо користувач не заходив у програму протягом установленого терміну, обліковий запис автоматично вилучається разом зі всіма діалогами та іншими даними.
Голосові та відеодзвінки	Спілкування не лише за допомогою текстового листування, а й у форматі аудіо- та відеодзвінка — корисна властивість месенджерів, про яку немає рації багато розповідати. Назначення та необхідність такого способу зв'язку очевидні.
Групові чати	Створюючи безпечні приватні чати, користувачі можуть вести зашифровані розмови: з родичами та друзями про особисте, з колегами та партнерами про бізнес. При цьому програма стане ідеальною, якщо сервер не матиме доступу до жодних групових метаданих, включаючи значки, заголовки, списки учасників.
Обмін контентом	Передача зображень, аудіо- та відеофайлів, документів — обов'язкова функція будь-якого месенджера. У захищених вона також має бути присутньою. Тільки бажано додати опції в налаштування не зберігати або автоматично удаляти отримані файли з папки за замовчуванням.

*Джерело: складено на основі [9]

Потоки інформації в Інтернеті завжди контролюються кимось. Дані передаються через сервери з точки А до точки Б і далі. Листування через месенджер відбувається через стороннього постачальника послуг. Хоча умови сервісу припускають, що всі дані користувача вважаються конфіденційними і тому недоторканими, вони доступні постачальнику послуг. Це означає, що теоретично вони можуть використовуватись і в інших цілях, крім зберігання інформації, оскільки користувачі ухвалюють умови угоди. А люди зазвичай не читають умови, або не помічають хитрих формулювань, де представник сервісу знімає з себе відповідальність за конфіденційність даних [9].

Виходить, що користуватися деякими програмами обміну повідомленнями небезпечно: постачальник послуг може виявитися ненадійним. Одна річ, коли користувач йде на компроміс, і зовсім інша, коли сервіс скомпрометований. Бізнесмени не повинні ризикувати своєю конфіденційною інформацією.

У такому разі вирішити проблему із шифруванням та захисту даних під час спілкування може власний месенджер. Бізнес може зробити програму закритою: реєстрація буде доступна лише на запрошення існуючим користувачем або конкретною людиною. Зберігати дані на своєму сервері, завжди мати можливість повністю видалити їх - користувач сам контролюватимете інформацію, яка проходить через месенджер [9].

Розробка захищеного месенджера – це складний та трудомісткий процес. Не можна зібрати натовп із недорогих фрілансерів та виконати завдання за кілька днів. Бекенд для такого додатка має особливі нюанси, і лише з грамотним підходом можна спростити процес, зменшити термін розробки та її вартість. Все залежить від стека технологій і реалізованих функцій.

Для такої роботи потрібна кваліфікована команда розробників із тестувальниками. Професіонали зроблять так, щоб безпека спілкування та вирішення бізнес-питань через месенджер була на 100% конфіденційною, а інформація не втекла у сторонні руки. Якщо програма розробляється не для вузького кола користувачів, а для всього світу, то треба планувати ширше і далі. Щоб новий месенджер помітили, знадобиться унікальна бізнес-модель, зручний і

доброзичливий інтерфейс користувача, а також опрацьований маркетинговий план.

Саме створення захищеного месенджера складається з таких етапів:

1. Написання технічного завдання – від 50 годин.
2. Бізнес-аналіз та специфікація - від 40 годин.
3. UI та UX дизайн – близько 150 годин.
4. Розробка бекенда – більше 500 годин, залежно від необхідних функціональних можливостей програми та кількості платформ (ОС).
5. Розробка мобільного додатка – понад 320 годин.
6. Тестування – близько 250 годин, залежно від кількості платформ та функціональності.

Виходячи з цих даних, приблизна вартість MVP програми починається від 50 000 доларів і далі залежить від факторів, що ускладнюють роботу. Також доведеться заздалегідь врахувати, що після завершення розробки месенджер ще потрібно підтримувати: виправляти можливі помилки, доопрацьовувати та реалізовувати нові функції. Таке обслуговування після релізу оплачується окремо.

Отже, було розглянуто особливості мобільного меседжингу як бізнес-екосистеми. У бізнес-операціях спілкування стало важливим елементом підтримки ефективного та динамічного робочого процесу. Оскільки бізнесмени та їхні працівники завантажені безліччю завдань, месенджери перетворилися на їхніх головних помічників. Однак так само, як допомогти, вони можуть нашкодити. Основним принципом безпечного обміну повідомленнями є наскрізне шифрування. Наприклад, програма Signal працює на однойменному криптографічному протоколі, розробленому Open Whisper Systems ще для першої програми (TextSecure). Такий тип шифрування використовує багаторівневий підхід, який ускладнює грубе проникнення листування з доступом до даних. Signal Protocol також застосовується за умовчанням у WhatsApp, і як додаткова функція у Facebook Messenger (Secret conversations) та Skype (Private conversations).

2.2 Закономірності функціонування ринку телекомунікаційних послуг

Історія виникнення мереж мобільного зв'язку почалася у 1980-х роках з появою кількох інноваційних мережевих технологій AMPS у США та TACS, NMT у Європі. У цей же час у Європі розробили стандарт GSM (Group Special Mobile, пізніше перейменований на Global System for Mobile Communications), що також відноситься до систем другого покоління (2G) і підтримує чотири різні діапазони: 850,900,1800 та 1900 МГц [30].

Наступним поколінням мереж мобільного зв'язку 3G було обрано зовсім інший підхід доступу до інформації, що базується на основі пакетної передачі інформації. На сьогоднішній день існує два види 3G стандарту. Наступними поколіннями мереж мобільного зв'язку стали 3.5G та 4G технології. Технологія HSDPA (High - Speed Downlink Packet Access) розглядається як один із перехідних етапів до четвертого покоління мереж. Стандарт базується на високошвидкісній пакетній передачі даних від базової станції до мобільного телефону. 4G покоління мереж мобільного зв'язку представлені такими технологіями: LTE, mobile WiMAX, UMB [30].

У 2020 році витрати компаній та споживачів на телекомунікаційні послуги, включаючи платне телебачення, сягнули \$1,53 трлн. Такий самий обсяг був зареєстрований і роком раніше, повідомляють аналітики IDC [31]. За їхніми даними, у країнах Північної та Південної Америки ринок телеком-послуг за підсумками 2020 року становив \$583 млрд, що на 0,7% більше у порівнянні з 2019-м. В Азіатсько-Тихоокеанському регіоні динаміка була нульовою., а обсяг ринку сягнув \$482 млрд. На 0,8% зменшилися витрати фізичних та юридичних осіб на телекомунікаційні послуги в регіоні EMEA — до \$467 млрд (рис. 2.1).

У 2020 році глобальний телекомунікаційний ринок пройшов серйозне випробування, і воно успішно впоралося з ним. IDC вважає, що уроки, що були зроблені в 2020 році, допоможуть галузі забезпечити стабільне зростання в майбутньому, — йдеться у дослідженні [31].

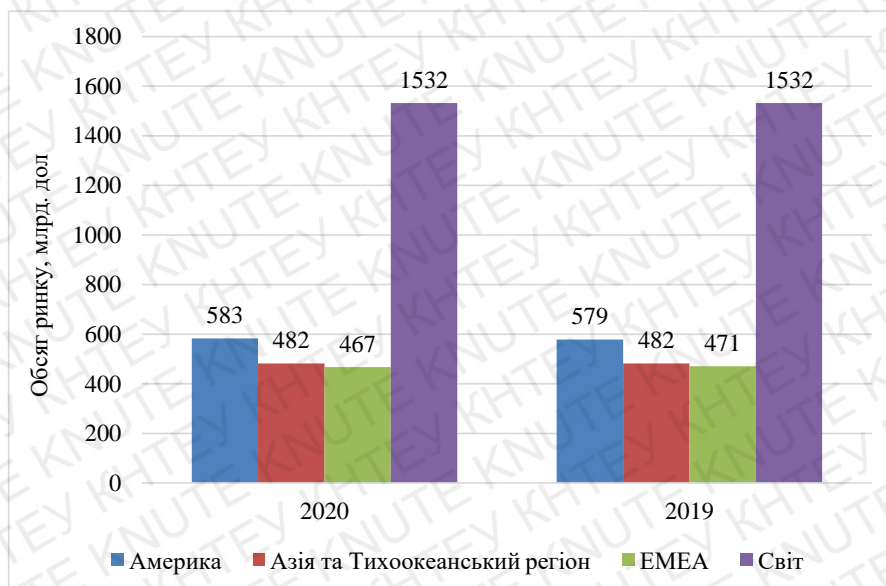


Рис. 2.1. Динаміка обсягу світового ринку телекомунікаційних послуг у 2019-2020 рр, млрд.дол

*Джерело: складено на основі [30]

На думку аналітиків, перехід на мережі FTTP (Fibre-To-The-Premises) пришвидшуватиметься у багатьох країнах, а B2B-сегмент ринку послуг фіксованого зв'язку відновлюватиметься у довгостроковій перспективі, оскільки відновлення економіки підштовхує інвестиції у хмарні інструменти корпоративного рівня. Зростання доходів у сегменті мобільних послуг певною мірою підтримуватиметься впровадженням 5G-мереж, які запропонують користувачам більш просунуті можливості обробки даних та використовувати контент та послуги, що залежать від високошвидкісних каналів зв'язку.

Віце-президент з дослідження європейського ринку телекомунікаційних послуг та відповідної інфраструктури Кріс Барнард (Chris Barnard) зазначив, що пандемія коронавірусу COVID-19 показала стійкість та значущість телекомунікаційної галузі. Нові методи роботи зберуться і після пандемії, формуючи майбутні можливості для отримання доходів, у той час як увага споживачів до мереж визначатиме вимоги до пропускної спроможності на ринку [31].

Динаміка на телекомунікаційному ринку у 2020 році була нерівномірною в умовах пандемії. Через неї у першому півріччі суттєво скоротилися витрати

користувачів та абонентські бази операторів. Зокрема, спад стався через обмеження, впроваджені владою країн для боротьби з поширенням інфекції, а також через загальний песимізм і занепокоєння, які змушували людей витратити менше грошей на несуттєві продукти та послуги [31].

У другій половині 2020 року попит у телекомі відновився, чому сприяли заходи підтримки економіки та вакцинація населення. Оптимізм, що відновився, допоміг глобальним і регіональним ринкам компенсувати збитки, зафіксовані в першому півріччі, і наблизитися до результатів 2019 року [31].

Як зазначається у дослідженні, пандемія різко змінила тенденції, які тривалий час визначали ситуацію на телекомунікаційному ринку. Послуги фіксованого зв'язку раптово стали для споживачів найважливішим засобом підключення до Інтернету, оскільки вони допомагали їм працювати, навчатися та розважатися в домашніх умовах [31].

Динаміка зростання витрат на послуги фіксованого зв'язку у B2B-сегменті, навпаки, суттєво погіршилася, але операторів зв'язку врятували довгострокові контракти з компаніями. Ще один тренд в умовах пандемії – бізнес став частіше відмовлятися від традиційної телефонії. Багато компаній, що представляють малий бізнес, закрилися, а більшість підприємств, що залишилися, урізали витрати на такі послуги [31].

На ринку мобільного зв'язку у 2020 році також спостерігалось зниження витрат багато в чому через уповільнене продовження контрактів, скорочення витрат поза пакетними тарифами та різке падіння доходів операторів від роумінгу [31].

У світі 4G становить половину всіх бездротових підключень; 5G має тенденцію стрімкого зростання у 2020 році (рис. 2.2) [27, 30]. У різних країнах сформувалися свої рейтинги месенджерів на основі статистики активних користувачів протягом місяця. У 169 країнах перше місце посідає WhatsApp, але ще в 25 він поступився першістю. З ряді країн у лідирує Facebook Messenger, і лише у 10 країнах переважають месенджери, які не належать корпорації Фейсбук.

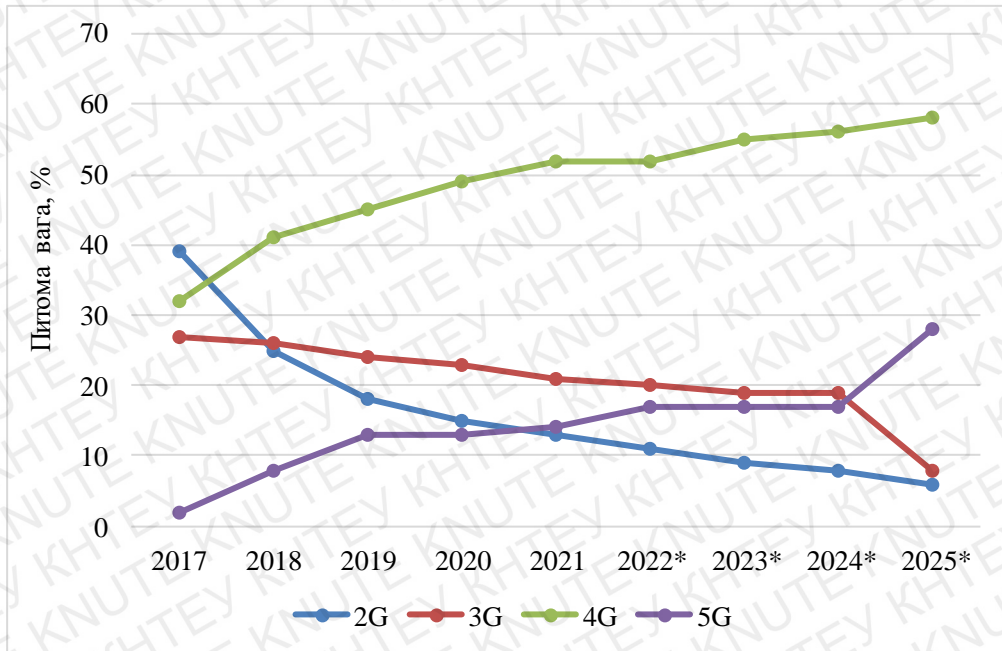


Рис. 2.2. Розподіл технологій світової мобільної бази користувачів

*Джерело: складено на основі [32]

Якщо ринок настільки глобалізований, підемо від загального до приватного. Ось так зараз виглядає рейтинг додатків для обміну повідомленнями у глобальному масштабі (рис. 2.2):

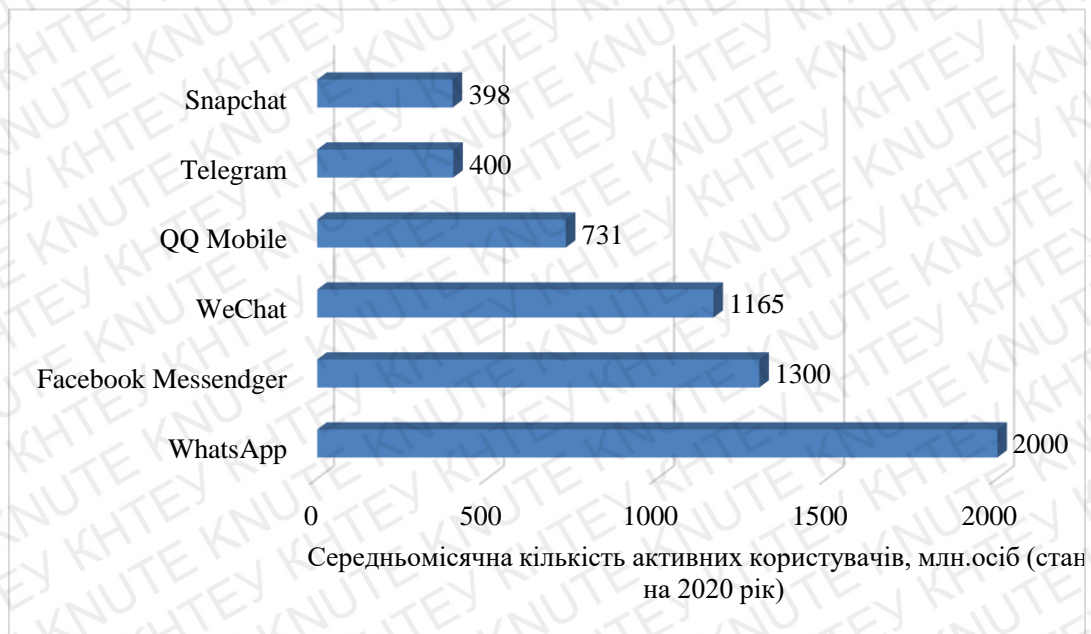


Рис. 2.3. Середньомісячна кількість активних користувачів, млн.осіб (станом на 2020 рік)

*Джерело: складено на основі [32]

Дані про щомісячну кількість активних користувачів складено наприкінці 2020 року, а кількість вказана у мільйонах. Дивний факт: світовий лідер WhatsApp на внутрішньому ринку поступається і іншому месенджеру Фейсбука, і конкурентному додатку, хоча вони з'явилися на 2,5 роки пізніше.

А ось топ-3 програм для обміну повідомленнями в окремо взятих у топ-країнах, які є лідерами у використанні мобільного меседжингу у світі:

Таблиця 2.2

**Рейтинг мобільних месенджерів та країн за рівнем їх використання
(кросс-матриця) [32]**

<i>Країна</i>	<i>1-ше місце</i>	<i>2-ге місце</i>	<i>3-тє місце</i>
Росія	WhatsApp	Viber	Telegram
Великобританія	WhatsApp	Facebook Messenger	Skype
США	Facebook Messenger	Snapchat	WhatsApp
Китай	WeChat	QQ	Momo

*Джерело: складено на основі [32]

Як правило, рейтинги складаються без урахування особливостей додатків (протоколи, функції), навіть без сортування на захищені та класичні. Тому, для повноти картини та відповідності темі, варто навести вивірені списки захищених месенджерів, які використовують шифрування за замовчуванням.

Найкращі месенджери із шифруванням за версією сайту AVG (компанії Avast) від березня 2020 року: 1. Signal, 2. Wickr Me, 3. Dust, 4. WhatsApp, 5. Telegram. Найкращі месенджери із шифруванням за версією онлайн-видання TechRadar від квітня 2020 року: 1. Signal, 2. WhatsApp, 3. Telegram, 4. Threema, 5. Silence.

Навіть захищені месенджери стикаються із проблемами безпеки. Багатьом на ринку відома технологія блокчейн і вона вважається безпечним рішенням. Звичайно, компанії намагаються впроваджувати блокчейн всюди. Тільки потім виявляється, що диво-технологія не здатна забезпечити абсолютно все, знизити витрати та одночасно підвищити дохід.

Теоретично у блокчейна багато сфер застосування, але на практиці він поки що приносить користь лише у сфері фінансових послуг, торгівлі та виробництві. Якщо звузити коло інтересів до месенджерів, то тут виникають дві перешкоди, які призводять до несумісності трендової технології з програмами обміну повідомленнями.

Першою перешкодою при спробі використання блокчейна в месенджерах стає сховище для всього, що супроводжує повідомлення, які відправляються і одержуються. Давно минули ті часи, коли повідомлення було чистим текстом. Зараз це більше фотографій, голосових повідомлень, відеороликів, документів. Усьому цьому потрібно місце для зберігання, про що не замислюються користувачі неблокчейн-додатків [29].

Автор блокчейн-месенджера повинен вирішити, де розміщувати всі дані з повідомлень. Адже блокчейну потрібні фулноди - повні вузли з синхронізованою копією реєстру. Ними можуть бути сервери власника месенджера, але тоді не буде децентралізації. А чи захочуть люди зберігати у себе гігабайти сторонньої інформації, тим більше смартфонів? У них вбудована пам'ять лише 32–128 ГБ; картка microSDXC на 256 ГБ коштує близько 1000 грн. Для порівняння: у травні 2020 року розмір блокчейну Біткоїна досяг 280 ГБ, а Ефіріуму - 140 ГБ. Це лише інформація про транзакції, без фото, аудіо-, відеофайлів. Крім того, у захищених месенджерів виникне конфлікт інтересів: блокчейн зберігатиме всі повідомлення (включаючи вкладення), або метадані, і видалити їх не є можливим. А це порушення двох принципів із чотирьох, на яких будується захист [29].

Другою перешкодою стає алгоритм консенсусу, який необхідний будь-якому блокчейну для створення довіри між учасниками мережі. Недостатньо просто розподілити реєстр на тисячі пристроїв для зберігання інформації. Частина повинна постійно підтверджувати дії користувачів перед тим, як додати записи в блок. Наприклад, у пірингових платіжних системах такими діями є криптовалюти транзакції. За валідацію транзакцій майнери, стейкери чи делегати беруть комісію, за додавання блоку до ланцюга отримують від системи монети.

Валідаторам потрібен стимул у вигляді винагороди, інакше їм немає сенсу підтримувати мережу.

Якщо комусь здасться цікавою ідея створити захищений месенджер на блокчейні, то новатор має бути готовим ділитися чимось із головними учасниками мережі. Тоді вони теж ділитимуться своїми ресурсами і часом на валідацію повідомлень, що проходять через мережу. Водночас доведеться створювати токеноміку там, де спочатку вона не була потрібна. Все взаємопов'язано: без оплати за дії не з'являться валідатори, без валідаторів не виникне довіри, без довіри не знайдеться стільки користувачів, які готові тримати фулноди активними.

Крім того, для захищених месенджерів нікуди не подінеться порушення принципів захисту конфіденційності: сотні валідаторів плюс тисячі інших учасників мережі будуть бачити метадані повідомлень. Поціновувачів приватності турбує, що одна компанія (наприклад, Facebook) володіє їх даними, а тут вся мережа знатиме, коли і кому вони надсилали повідомлення.

Отже, на ринку мобільного зв'язку у 2020 році також спостерігалось зниження витрат багато в чому через уповільнене продовження контрактів, скорочення витрат поза пакетними тарифами та різке падіння доходів операторів від роумінгу. У 2020 році глобальний телекомунікаційний ринок пройшов серйозне випробування, і воно успішно впоралося з ним. IDC вважає, що уроки, що були зроблені в 2020 році, допоможуть галузі забезпечити стабільне зростання в майбутньому. Найкращі месенджери із шифруванням за версією сайту AVG (компанії Avast) від березня 2020 року: 1. Signal, 2. Wickr Me, 3. Dust, 4. WhatsApp, 5. Telegram. Найкращі месенджери із шифруванням за версією онлайн-видання TechRadar від квітня 2020 року: 1. Signal, 2. WhatsApp, 3. Telegram, 4. Threema, 5. Silence. У різних країнах сформувалися свої рейтинги месенджерів на основі статистики активних користувачів протягом місяця. У 169 країнах перше місце посідає WhatsApp, але ще в 25 він поступився першістю. З ряді країн у лідирує Facebook Messenger, і лише у 10 країнах переважають месенджери, які не належать корпорації Facebook.

2.3 Основні елементи екосистеми мобільного меседжингу

16 червня 1993 року вважається датою, коли в Україні було запроваджено мобільний зв'язок. Першою компанією на ринку мобільного зв'язку стала компанія UMC (Ukrainian Mobile Communications — Український мобільний зв'язок, потім МТС, тепер — NEQSOL Holding під брендом Vodafone).

Напередодні лібералізації ринку мобільного зв'язку створювалися юридичні особи, які бажають конкурувати з "УМЗ" (зараз - "МТС-Україна") - єдиним у країні оператором мобільного зв'язку, який надає послуги у застарілому стандарті NMT.

Першим GSM-оператором в Україні стала компанія "Банкомзв'язок" (ТМ Golden Telecom), яка відкрила мережу в Києві наприкінці 1996 року у форматі GSM-1800.

У ході тендеру на частоти GSM-900 переможцями були оголошені три компанії, у тому числі вже працюючий "Український мобільний зв'язок" та два новачки: "Київстар" та "Українські радіосистеми" (УРС (пізніше Beeline), яка пізніше влилася в "Київстар"). .

МТС/Водафон. 2003 року МТС купує UMC. У цей час оператор запускає новий стандарт GSM-1800 (з GPRS) та віртуального оператора JEANS, який орієнтований на молодь – він швидко стає популярним. У лютому 2015 року оператор купує 3G ліцензію. 16 жовтня 2015 року МТС стає Vodafone і до кінця року запускає 3G майже у всіх обласних центрах.

Київстар. Нинішній лідер ринку мобільного зв'язку формально запустився 1994 року, коли було зареєстровано компанію «Bridge», але її перейменували на «Київстар». Послуги мобільного зв'язку вона почала представляти лише з 9 грудня 1997 року у стандарті GSM. Оператор уже 1998-го першим в Україні запустив послугу SMS. На даний момент Київстар має найбільший спектр Частот 3 і 4 покоління. Лідерство в частотах компанія здобула в результаті покупки в 2009 році оператора «Білайн».

life:)/lifecell — бренд білоруського оператора мобільного зв'язку ЗАТ «БеСТ», третій за чисельністю абонентів оператор мобільного зв'язку. ЗАТ «Білоруська мережа телекомунікацій» одразу ж запустилася із мережею EDGE. Після цього почалися спроби відвоювати у двох гігантів частину ринку. У грудні 2008 року компанія зробила черговий ребрендинг і почала надавати послуги під торговою маркою life:), яку використовує дочірня компанія Turkcell в Україні. Після ребрендингу – з 15 січня 2016 року lifecell – торгова марка, під якою український оператор мобільного зв'язку ТОВ «Лайфсел» надає свої послуги.

Нині в Україні можливості та потенціал технології 4G, яка була запущена операторами у 2018 році, ще не до кінця освоєно. 4G має дуже широкий потенціал розвитку з погляду розширення покриття та впровадження додаткових сервісів для приватних абонентів, бізнесу та держави. Мільйони українських абонентів досі не почали користуватися смартфонами та швидкісним мобільним інтернетом.

На ринку стільникового зв'язку та дистрибуції у сегменті реалізації мобільних телефонів в Україні присутні великі гравці, які присутні у всіх регіонах країни (рис. 2.4).

Ринкові частки на ринку мобільного зв'язку за абонентами

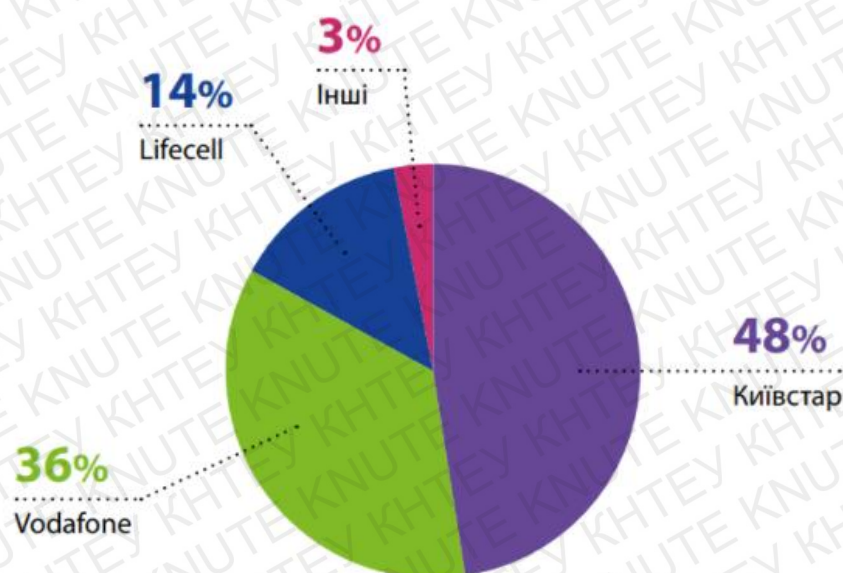


Рис. 2.4. Структура ринку мобільного зв'язку за основними операторами, %
(станом на 2020 рік)

*Джерело: складено на основі [2]

Найбільші оператори охоплюють практично всі області України, і конкурувати з ними в масштабах усієї України недоцільно, оскільки оператори мають велику фінансову владу (у разі змови можуть дозволити собі демпінгувати для витіснення конкурента). Варто виходити на вибіркові ринки – міста з інтенсивним зростанням населення, а також з розвинуеною інфраструктурою та промисловістю, де рівень заробітної плати вищий за середньоукраїнський.

За підсумками 2020 року українські мобільні оператори заробили 46,9 млрд грн - це 63,7% від загального обсягу ринку телекомунікаційних послуг (рис. 2.5).



Рис. 2.5. Доходи операторів мобільного зв'язку, (станом на 2020 рік)

*Джерело: складено на основі [12]

Більшу частину доходів (59,9%) оператори заробили від надання послуг передачі даних - 28,1 млрд грн. На другому місці опинилися доходи від послуг голосової телефонії - 18,1% або 8,5 млрд грн. Від послуг пропуску трафіку оператори отримали 5,4 млрд грн - 11,6%. На роумінгу в пандемічний 2020 рік оператори заробили 725 млн - це 1,5% від загальної суми. За оцінкою телеком-регулятора, доходи операторів від надання послуг роумінгу зменшилися приблизно вдвічі. Така ситуація, в першу чергу, пов'язана з обмежувальними

заходами на пересування громадян, які вводили країни в зв'язку з поширенням пандемії.

На ринку стільникового зв'язку та дистрибуції у сегменті реалізації мобільних телефонів в Україні існують певні ризики:

1. Ризики зниження платоспроможності: пандемія, та заходи, вжиті у зв'язку з нею, змусили багатьох підприємців відправити робітників у відпустку за свій рахунок, або скоротити виплати у зв'язку зі зниженням активності ринку. У разі другого стрибка заражень, карантинні заходи можуть посилитися, що призведе до зниження купівельної спроможності багатьох потенційних споживачів.

2. Ризики монополізації ринку: висока влада головних гравців ринку створює ризик демпінгування ціни послуги для посилення свого монопольного становища.

3. Валютні ризики: ринок стільникового зв'язку на 100% залежить від імпорту. Будь-які різкі валютні стрибки відображаються на вартості товару.

4. Ризики фінансової нестабільності: збільшення процентних ставок на кредити та інфляційні процеси. Розвиток торгових мереж (салонів зв'язку) вимагає значних інвестицій, яких у дрібних і середніх компаній немає – такі компанії користуються кредитами, і будь-які коливання країни підвищують ризикованість ринку.

5. Ризики розвитку ринку «сірого» імпорту: на даний момент через поштові відправлення, недобросовісне декларування товарів та ін. ринок сірого імпорту мобільних телефонів складає близько 29%

Якщо розглядати мобільні оператори: Київстар, Водафон і Лайфселл, то в сегменті зв'язку 3-го покоління оператори розвиваються порівняно рівномірно, проте світова тенденція свідчить про відмову споживачів від цього покоління зв'язку. Більшість споживачів користується зв'язком 4-го покоління – близько 50% світового ринку бездротового зв'язку. А також помітно бурхливий розвиток зв'язку 5-го покоління, який призначений насамперед для потреб бізнесу (B2B).

У сегменті зв'язку 4-го покоління на даний момент лідирує Київстар (у 2018 році набув найвигідніших частот, сплативши найвищу ціну) – у оператора

сконцентровано у користуванні найбільший діапазон частот, які навіть у разі перерозподілу переважно належать Київстару. За попередніми прогнозами, у цьому сегменті найактивніше розвиватиметься Київстар.

Також гравцям на ринку стільникового зв'язку слід стежити за розвитком ринку зв'язку 4-го покоління в діапазоні частот 800 МГц, оскільки даний діапазон дозволить покрити автомобільні та залізничні колії, а також надасть зв'язок для віддалених малозаселених населених пунктів. Але головне – цей діапазон зв'язку дозволить активніше розвиватися аграрному бізнесу (системи точного землеробства).

Усі оператори з випередженням графіка забезпечують розгортання покриття LTE в Україні, цей процес розвитку ще триватиме кілька років. А існування якісного 4G покриття також є однією з необхідних умов і для запуску 5G.

Зростання бізнесу - виступає головним замовником та користувачем 5G. Запуск технології 5G в Україні може бути успішним лише після того, як буде:

- повністю використаний потенціал LTE-мереж;
- сформовано чітку державну стратегію розвитку наступного покоління зв'язку, який би враховувала позиції операторів, бізнесу та ринкових потреб;
- проведено заходи щодо виділення та ліцензування частот під запуск та розвиток 5G;
- з'явиться реальний попит бізнесу на послуги та рішення, для яких необхідний 5G, адже цей стандарт орієнтований в основному саме на промислові та інфраструктурні потреби b2b-ринку, а не абонентів;
- з'являться доступні за ціною термінали та обладнання, які підтримують нову технологію;
- буде створено екосистему (бізнес – оператори – провайдери сервісів) для успішного застосування нової технології, яка зробить її економічно обґрунтованою.

Використання різних діапазонів частот забезпечує максимально повне та стабільне покриття.

Отже, серед тенденцій ринку мобільного зв'язку України можна виділити виділив:

- отримання операторами ліцензії на впровадження нової радіотехнології у смузі 900 МГц;
- значне розширення 4G-покриття на території країни і зростання кількості користувачів мобільного інтернету;
- зростання кількість операторів, які пропонують користувачам eSIM;
- збільшення обсягів межоператорського трафіку на тлі зниження розрахункових такс;
- скорочення телекомунікаційних мереж операторів ТриМоб і Інтертелеком;
- поступове зростання кількості споживачів послуг межмашинного взаємодії (Machine-to-machine, M2M) і служб речей (Internet of Things, IoT).

Висновки до розділу 2

Розглянуто особливості мобільного меседжингу як бізнес-екосистеми. У бізнес-операціях спілкування стало важливим елементом підтримки ефективного та динамічного робочого процесу. Оскільки бізнесмени та їхні працівники завантажені безліччю завдань, месенджери перетворилися на їхніх головних помічників. Однак так само, як допомогти, вони можуть нашкодити. Основним принципом безпечного обміну повідомленнями є наскрізне шифрування.

Проаналізовано тенденції ринку мобільного зв'язку у світі. Визначено, що у 2020 році також спостерігалось зниження витрат багато в чому через уповільнене продовження контрактів, скорочення витрат поза пакетними тарифами та різке падіння доходів операторів від роумінгу. У 2020 році глобальний телекомунікаційний ринок пройшов серйозне випробування, і воно успішно впоралося з ним.

Проаналізовано ринок стільникового зв'язку України. Найбільші оператори охоплюють практично всі області України, і конкурувати з ними в масштабах

усієї України недоцільно, оскільки оператори мають велику фінансову владу (у разі змови можуть дозволити собі демпінгувати для витіснення конкурента). Варто виходити на вибіркові ринки – міста з інтенсивним зростанням населення, а також з розвинутою інфраструктурою та промисловістю, де рівень заробітної плати вищий за середньоукраїнський.

РОЗДІЛ 3. BIG DATA РІШЕННЯ ОПЕРАТОРА МОБІЛЬНОГО МЕСЕДЖИНГУ НА ПРИКЛАДІ БІЗНЕС-КЕЙСУ ДЛЯ АВТОМОБІЛЬНОГО ЦЕНТРУ ТОВ «ГРАНД АУТОС»

3.1 Формування основних функціональних вимог та моделювання бізнес-процесів задачі просування продукції ТОВ «Гранд Аутос».

У даному питанні варто розглянути особливості застосування Big Data як інструменту вирішення бізнес-завдань підприємства. З цією метою необхідно сформувати бізнес-кейс, який потребуватиме обґрунтування рішення щодо впровадження відповідної інформаційної системи або доопрацювання вже існуючої.

Об'єктом бізнес-кейсу є бізнес-замовник – компанія автомобільний центр ТОВ «Гранд Аутос», що потребує оптимізації процесу просування своєї продукції на ринку.

Предметом бізнес-кейсу є інформаційне рішення з використанням Big Data, що дозволить вирішити бізнес-задачі замовника, а саме дослідження цільової аудиторії для проведення рекламних кампаній та таргетованих розсилок.

Для більш детально формулювання основних параметрів бізнес-кейсу, пропонується розглянути стратегічний контекст функціонування бізнес-замовника.

Офіційна назва досліджуваного підприємства – Товариство з обмеженою відповідальністю «Гранд Аутос». Предметом діяльності ТОВ «Гранд Аутос» є імпорт автомобілів та їх реалізації на території України, надання агентських послуг з пошуку автомобілів для замовника, оформлення договорів оренди на авто, купівлі-продажу авто. Підприємство здійснює пошук авто на замовлення клієнтам у таких країнах: США, Німеччина, Латвія, Литва, Чехія, Угорщина, Польща, Італія.

Місією підприємства є задоволення споживачів у якісних послугах з пошуку та доставки в Україну автомобілів, які були у використанні.

Опис структури бізнесу підприємства можна навести з використанням бізнес-моделі, яка виконана у формі стратегічної діаграми Business Model Canvas (рис. 3.1). Для побудови стратегічної канви було використано підхід до формування бізнес-моделі за О. Остервальдером та І. Пінье, який дозволяє формувати різноманітні артефакти для оцінки систем менеджменту компаній.

Ключові види діяльності	Ключові цінності	Взаємовідносини з клієнтом	Сегменти споживачів
Пошук авто для клієнта; Імпорт авто в Україну; Допомога у оформленні авто	Оперативність та швидкість реакції на запити клієнтів; Глибоке опрацювання замовлень клієнтів; Професійний підхід до виконання замовлень клієнтів	Клуб постійного клієнта (програма лояльності); Організація івентів для корпоративних клієнтів	1) Фізичні особи / юридичні особи (бізнес) 2) Постійні клієнти / клієнти, що здійснюють разові замовлення;
Ключові ресурси		Канали комунікацій	
Персонал – менеджери по роботі з клієнтами, фахівці ринку авто; Орендований офіс; Засоби зв'язку та ІТ-інфраструктури		Сайт компанії; Соціальні мережі Facebook та Twitter; Вивіска компанії	
Ключові партнери			
Посередники у зарубіжних країнах Агенти з просування послуг компанії			
Структура ключових витрат (збутова частина)		Потоки доходів	
Витрати на з/п, оренду приміщення, оренду авто, послуги сторонніх організацій, на збут, утримання офісу		Від продажу авто клієнтам; Від надання послуг з доставки авто, його митного оформлення	

Рис. 3.1. Діаграма Business Model Canvas досліджуваного підприємства ТОВ «Гранд Ауто»

*складено автором за даними підприємства

Підприємство працює на таких основних сегментах ринку торгівлі автомобілями, що були у використанні:

1) ринок індивідуального споживача (клієнти придбавають авто для власних

потреб);

2) ринок корпоративного клієнта (автомобілі замовляються компаніями/ організаціями для бізнес-потреб).

Водночас до переліку клієнтів підприємства відносяться як фізичні особи, так і корпоративні клієнти (юридичні особи). Ключовими цінностями компанії є наступні: оперативність та швидкість реакції на запити клієнтів; якісне опрацювання замовлень клієнтів; професійний підхід до пошуку авто.

Наступним кроком обґрунтування бізнес-вимог було визначення основних проблем розвитку бізнесу, які можуть бути вирішені шляхом автоматизації та використанням Big Data.

PROBLEMS/ OPPORTUNITIES	ROOT CAUSES	GOALS	SOLUTIONS	OBJECTIVES
-Досить низька конверсія сайту підприємства. -Відсутність бази даних клієнтів для потенційного здійснення рекламних розсилок	-Відсутність таргетингу при розсилання рекламних повідомлень -Неможливість коректно розмістити контекстну та банерну реклами через незнання веб-ресурсів, які використовуються цільовою аудиторією	-Збільшити ефективність процесу просування бренду компанії	збільшити рівень впізнаваності, залучити нових клієнтів, отримати перед замовлення. FEATURES • База даних клієнтів • Статистика наявності трагетингу цільової аудиторії	-Збільшити рівень конверсії сайту на 50% -Розширити базу даних клієнтів на 200%
BUSINESS REQUIREMENTS <ul style="list-style-type: none"> • можливість введення інформації про поточних клієнтів компанії; • можливість відправлення пошукового запиту на компанію оператора мобільного зв'язку; • можливість отримання даних про цільову аудиторію та запис їх у базу даних клієнтів; • виведення статистики про обсяги постачання та продажу автозапчастин • оцінка KPI сайту 				STAKEHOLDERS - власники підприємства; - директор підприємства; - комерційний директор; старший менеджер.

Рис. 3.2. Діаграма виявлення бізнес-вимог основних стейкхолдерів проектованої системи

*складено автором за даними підприємства

Основними стейкхолдерами, з якими було проведено комунікації для збору вимог, були:

- власники підприємства;
- директор підприємства;
- комерційний директор;
- старший менеджер;
- аналітик.

За результатами комунікацій зі стейкхолдерами та збору відповідних вимог було розроблено моделлю Solution Vision Canvas, яка формалізує основні групи бізнес-вимог для запиту.

Наступним кроком було визначення системних вимог – функціональних та нефункціональних вимог. З метою визначення функціональних вимог було використано методику створення Use-case діаграми.

Приклад Use-case діаграми, що описує основні функціональні вимоги до проєктованої системи з точки зору користувачів (акторів), наведено на рис. 3.3.

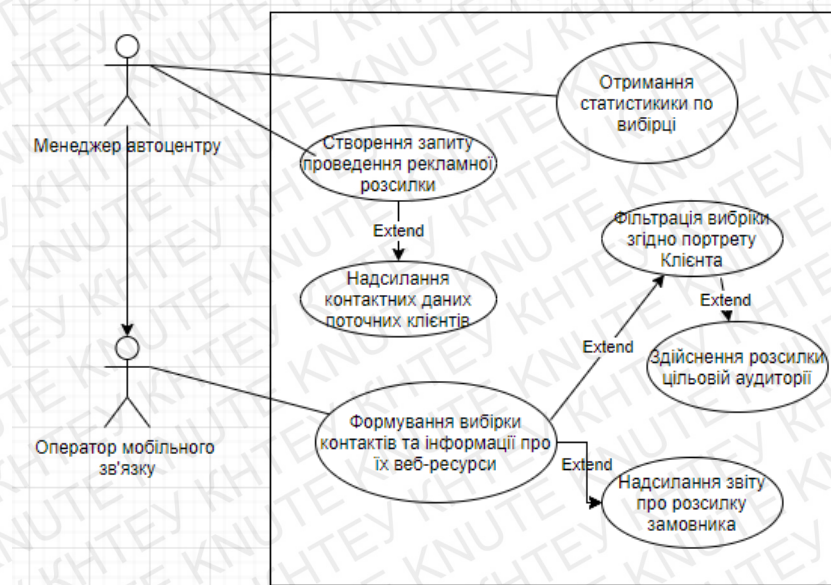


Рис. 3.3. Use-case діаграма основних функціональних вимог до проєктованої системи

*складено автором за даними підприємства

Коротко опишемо пропонований бізнес-кейс на діаграмі юз-кейсів.

Програмне рішення має виступати своєрідним веб-орієнтованим додатком.

Клієнт може авторизуватися в цьому додатку. Послідовність дій є наступною:

1. Авторизація користувача.
2. Створення запиту та визначення інтересів, критеріїв відбору.
3. Внесення даних у формі створення запиту:
 - додавання переліку контактних номерів вже існуючих клієнтів автоцентру;
 - додавання параметрів таргетингу цільової аудиторії.
4. Програмне забезпечення на стороні операторів, провайдеру здійснює пошук даних обраних клієнтів, аналізує поведінку зазначених абонентів та знаходить закономірності (зона Big Data).
5. Відбувається побудова загального портрету абонентів у розрізі значимих показників (на прикладі автоцентру додатково визначаються популярні веб-ресурси для подальшого розміщення реклами).
6. Валідація отриманих контактних даних на їх актуальність.
7. Здійснюється розсилка рекламних повідомлень від Альфа-імені Автоцентру на контактні номери, що вже провалідовані.
8. В особистому кабінеті клієнт знаходить звіт про виконання розсилки та загальну статистику.

Таким чином, виходячи з наведеного сценарію, Автоцентр формує зразок рекламного повідомлення, в якому він розміщує інформацію про власний сайт (потенційний клієнт, отримавши повідомлення, може перейти на сайт і одразу зареєструватися у два кліки). Також додатково у звіті від мобільного оператора бізнес-замовник отримує перелік веб-ресурсів, які найчастіше відвідуються його клієнтами. За цим переліком замовник зможе здійснювати розміщення рекламної інформації (банерна реклама, розміщення посилань на сайт), чим також збільшить рівень конверсії сайту.

Відповідно до профілю клієнта переважну більшість відповідно статі складала чоловіки вони становили 77%, жінки 23%. Важливим був критерій наявності автомобіля та становив 96,95%. Кількість володарів смартфонів, що використовують мобільний інтернет становив 86%. Далі важливим критерієм був рівень достатку клієнта за цим критерієм маємо наступні показники: високий

32%, вище середнього 48%, середній 8%, нижче середнього 9% та низький 3%. Цінова категорія мобільного телефону: найдорожчі 16%, дорогі 20%, середньої вартості 56%, найдешевші 8%. Бренд-виробник мобільного телефону Apple 41%, Samsung 25%, Nokia 5%, Xiaomi 7%, Lenovo 4% , інші 18%. Користування інтернетом в день: Facebook 11.03 МБ, Google 6.59 МБ, Instagram 32.70 МБ, Messenger 2.37 МБ, YouTube 29.85 МБ. Частка абонентів, що подорожували за кордон за останні півроку 31,45%. Середня кількість грошей, що абонент витрачає на зв'язок в місяк 350 грн. Схильність до e-commerce 18%. Додаткові інтереси: автосервіс 68%, автозапчастини 43%, автосалони 27%. Відвідують АЗС WOG 24, ОККО 28%, БРСМ-Нафта 17%, інші 31%.

Відповідно до наведених даних було побудовано наступну діаграму BPMN для верхньорівневого опису бізнес-процесу взаємодії між основними акторами бізнес-кейсу (рис. 3.4):

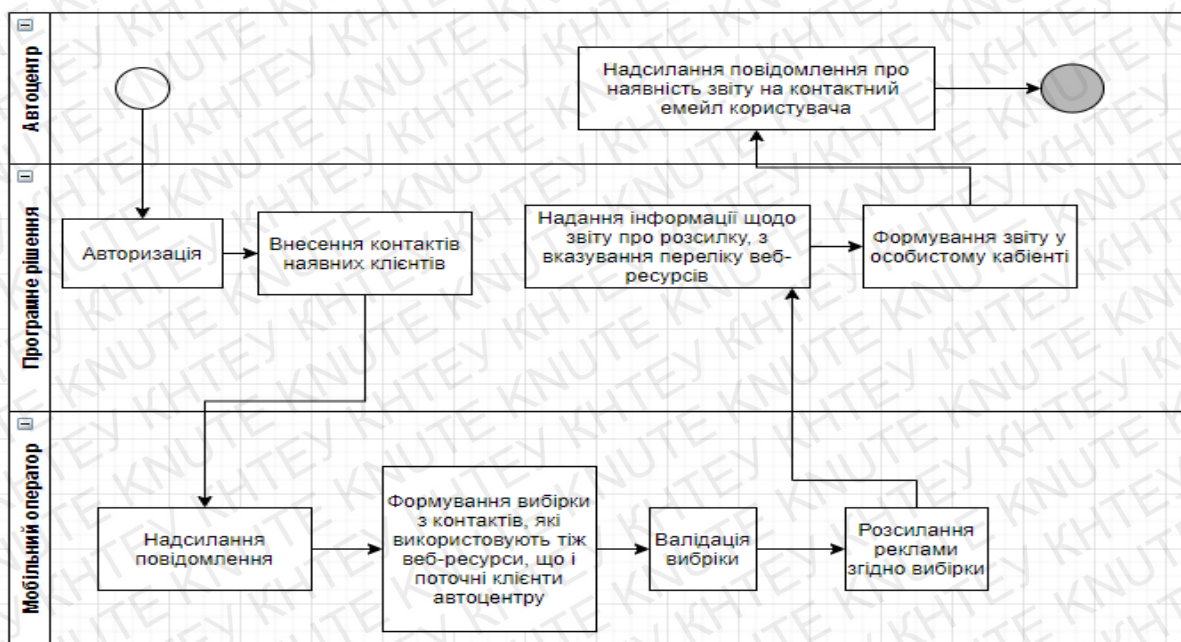


Рис. 3.4 Діаграма BPMN для верхньорівневого опису бізнес-процесу взаємодії між основними акторами бізнес-кейсу

*складено автором

При цьому відмітимо, що для опису бізнес-процесу формування розсилок на основі запропонованого алгоритму «as-is» було використано нотацію моделювання бізнес процесі BPMN 2.0 (business process modelling notation 2.0).

Для характеристики окремих процесів було надано їх детальний опис у табл. 3.1.

Таблиця 3.1

Операції процесу створення рекламних розсилок на основі пропонуваного рішення

<i>N</i>	<i>Елементи програмного рішення, що беруть участь в процесі</i>	<i>Операції</i>	<i>Виконавець</i>
1.	Етап авторизації	Користувач автоцентру авторизується в додатку	Користувач ПЗ
2.	Інтерфейс особистого кабінету користувача	Введення фактичної інформації про контактні дані клієнтів автоцентру	Користувач ПЗ
3.	Інтерфейс особистого кабінету користувача	Ініціація запиту на обробку до мобільного оператора, провайдера	Користувач ПЗ
4.	Бекенд проєктованого програмного рішення	Надсилання запиту на обробку до мобільного оператора, провайдера	Проєктована система
5.	База даних мобільного оператора та скрипти на обробку запитів	Обробка запиту, валідація контактів, розсилка рекламних повідомлень згідно переліку	Інформаційна система мобільного оператора
6	Бекенд проєктованого програмного рішення	Отримання даних про виконану розсилку	Проєктована система
7	Бекенд проєктованого програмного рішення	Формування звіту про розсилку	Проєктована система

*складено автором

Згідно проведеного аналізу основних бізнес-процесів було встановлено основні проблеми та завдання інформатизації, які дозволять ці проблеми вирішити (табл. 3.2).

Таблиця 3.2

Основні проблеми та завдання інформатизації

<i>№</i>	<i>Виявлені проблеми</i>	<i>Причини проблем</i>	<i>Можливі рішення впровадження інформаційних систем, які дозволяють вирішити проблеми</i>
1	Неефективна реклама, неможливість здійснювати ефективне просування	Відсутність можливості таргетованої реклами у бізнес-замовника	Програмне рішення дозволить здійснювати розсилання рекламних повідомлень трагетованій цільовій аудиторії
2	Низька конверсія сайту	Відсутність інформації про веб-ресурси, які використовують потенційні клієнти	Рішення дозволить отримати перелік веб-ресурсів, на яких можна розміщувати банерну та контекстну рекламу, що підвищить рівень конверсії сайту

*складено автором

Отже, за результатами комунікацій зі стейкхолдерами та збору відповідних вимог було виявлено основні проблеми бізнес-замовника та розроблено моделлю Solution Vision Canvas, яка формалізує основні групи бізнес-вимог до створюваного програмного рішення. До ключових проблем, які необхідно вирішити за допомогою запиту на операторів, провайдерів, відносяться такі: відсутність можливості таргетованої реклами у бізнес-замовника; відсутність інформації про веб-ресурси, які використовують потенційні клієнти. Програмне рішення дозволить здійснювати розсилання рекламних повідомлень таргетованій цільовій аудиторії. Рішення також дозволить отримати перелік веб-ресурсів, на яких можна розмішувати банерну та контекстну рекламу, що підвищить рівень конверсії сайту. Виходячи з логіки запропонованого програмного рішення, бізнес-замовник (автоцентр) формує зразок рекламного повідомлення, в якому він розміщує інформацію про власний сайт (потенційний клієнт, отримавши повідомлення, може перейти на сайт і одразу зареєструватися у два кліки). Також додатково у звіті від мобільного оператора бізнес-замовник отримує перелік веб-ресурсів, які найчастіше відвідуються його клієнтами. За цим переліком замовник зможе здійснювати розміщення рекламної інформації (банерна реклама, розміщення посилань на сайт), чим також збільшить рівень конверсії сайту. Відповідно, будуть досягнені цілі по зростанні конверсії сайту та збільшенні кількості замовлень.

3.2. Веб-орієнтований додаток особистого кабінету для оптимізації процесу проведення таргетованих розсилок.

У даному питанні варто спочатку навести основні характеристики, за якими має створюватися запит на обробку даних на стороні мобільного оператора, провайдера. Після підписання договору з провайдером, замовнику буде сформований та наданий персональний доступ до особистого кабінету, що передбачає можливість внесення даних у поля відповідної форми, призначений для зручного та ефективного використання послуг.

У роботі було розроблено веб-орієнтований додаток у вигляді особистого кабінету.

Спочатку при переході за посиланням в особистий кабінет користувачу необхідно авторизуватись.

The screenshot shows a login interface for 'Big Data Messages'. At the top, there is a dark blue header with the text 'Big Data Messages'. Below the header, the main content area has a light blue background. It starts with a greeting 'Доброго Дня!' (Good Day!) in bold. Underneath, there is a prompt: 'Увійдіть, щоб продовжити роботу з сервісом' (Log in to continue working with the service). The login form includes a label 'Ваша електронна пошта' (Your email) above an input field, a label 'Пароль' (Password) above another input field, and a blue button labeled 'Підтвердити' (Confirm).

Рис. 3.5. Вікно авторизації для введення облікових даних користувача

*складено автором

На рис. 3.6 наведено головна сторінка особистого кабінету, до якого користувач потрапляє одразу після авторизації:

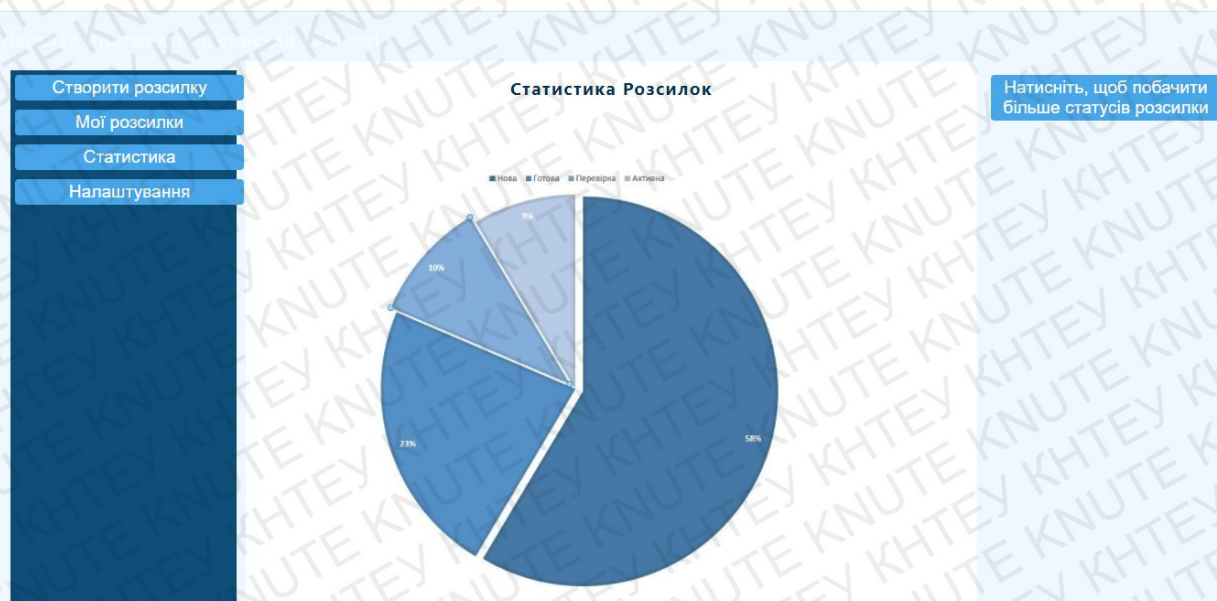


Рис. 3.6. Головна сторінка особистого кабінету

*складено автором

Далі клієнт при натиску на кнопку «Створити розсилку» потрапляє на сторінку «Критерії Вибірки» рис.3.7., саме тут клієнт має можливість самостійно обрати ключові критерії відповідно до свого портрета клієнта.

Рис. 3.7. Сторінка «Критерії вибірки»

*складено автором

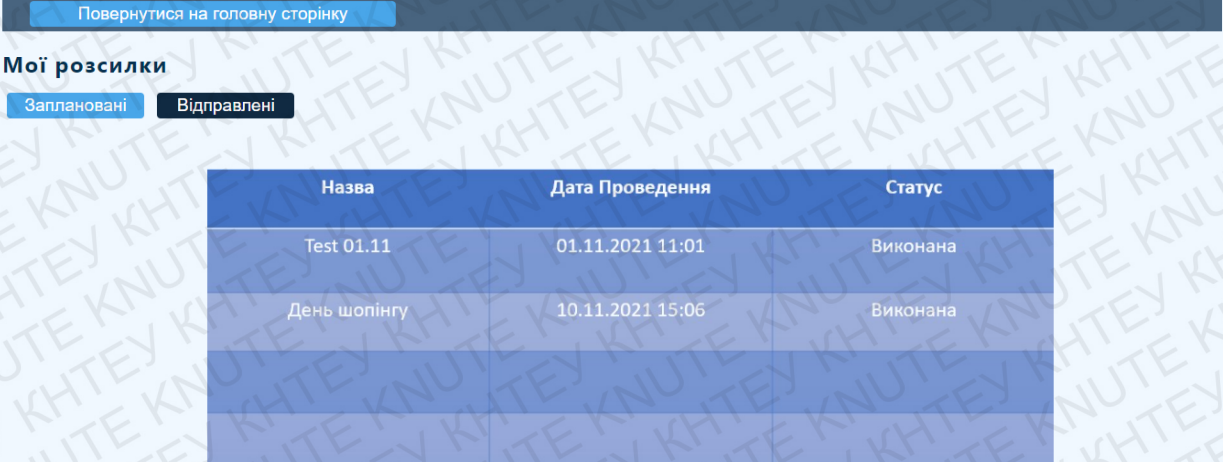
Наступним кроком є створення самої розсилки на сторінці «Запланувати розсилку» рис.3.8.

Рис. 3.8. Сторінка «Запланувати розсилку»

*складено автором

Необхідно зазначити назву розсилки. Обрати канал відправки, залежно від аудиторії, клієнт має можливість проводити розсилки в декілька каналів доставки, в такому випадку повідомлення буде по черзі будуть відправлятися. В залежності від того в який канал доставки та коли повідомлення було доставлено будуть продовжуватися відправки. Наприклад, клієнт може обрати 2 канали доставки Viber та SMS повідомлення може бути одразу успішно доставленим у Viber, проте якщо у абонента вимкнений Інтернет повідомлення автоматично буде доставлено в SMS-канал. Далі треба вказати «термін життя повідомлення» - це час під час якого повідомлення буде намагатися бути доставленим до абонента, оскільки не завжди абонент може отримати повідомлення одразу в момент відправки, тому що він може бути поза зоною дії мережі, з виключеним Інтернетом або в певних соц. мережах вказаний сервіс може бути заблокований, тому для того, щоб повідомлення було актуальним необхідно зазначити часові обмеження. Далі клієнт вводить безпосередньо текст самого повідомлення та обрати час відправки: в момент створення або ж запланувати розсилку на певний час. Справа можна побачити як повідомлення буде відображатись.

Також клієнт має можливість продивитись на сторінці «Мої розсилки» рис.3.9. заплановані розсилки у випадку, якщо попередньо вже створив розсилку на певний час та проведені розсилки відповідно до назви, дати проведення та статусу.



The screenshot shows a web interface for managing newsletters. At the top, there is a navigation bar with a link to 'Return to the main page'. Below it, the title 'My newsletters' is displayed. There are two tabs: 'Scheduled' (active) and 'Sent'. The main content is a table with three columns: 'Name', 'Date of execution', and 'Status'.

Назва	Дата Проведення	Статус
Test 01.11	01.11.2021 11:01	Виконана
День шопінгу	10.11.2021 15:06	Виконана

Рис. 3.9. Сторінка «Мої розсилки»

*складено автором

На сторінці «Статистика» рис. 3.10 клієнт має можливість подивитись та проаналізувати більш детальніше інформацію на основі даних коли проводилися розсилки, які канали доставки використовувалися під час певної розсилки, кількість доставлених та недоставлених повідомлень залежно від каналів доставки та загальний обсяг розсилки. Це важлива інформація, яку надалі клієнт може використовувати при формуванні розсилок в майбутньому, залежно від ефективності певних розсилок можна вносити певні корективи в канали доставки чи загальні обсяги розсилки.

[Повернутися на головну сторінку](#)

Статистика

Дата розсилки	Канал доставки	Кількість повідомлень		Всього
		доставлені	не доставлені	
01.11.2021	Telegram	1001	100	1101
	Viber	589	20	609
	SMS	2050	30	2080
10.11.2021	Telegram	699	107	806
	Viber	789	19	808

Рис. 3.10. Сторінка «Статистика»

*складено автором

Також в Особистому кабінеті є сторінка «Налаштування» рис. 3.11 на вказаній сторінці можна перевіряти кому належить доступ, електронну адресу, пароль, можна обрати мову особистого кабінету та часовий пояс. Останнє є важливим, тому що у випадку запланованих розсилок при некоректному виборі часового поясу фактичний час відправки розсилки може відрізнятись від очікуваного та запланованого.

Повернутися на головну сторінку

Налаштування

Ім'я
Анастасія

Ваша електронна пошта
Anastasiia21@gmail.com

Мова Українська Російська Англійська

Часовий пояс
UTC+2

Пароль

Зберегти

Рис. 3.11. Сторінка «Налаштування»

*складено автором

Логіка створення розсилок буде виглядати наступним чином:

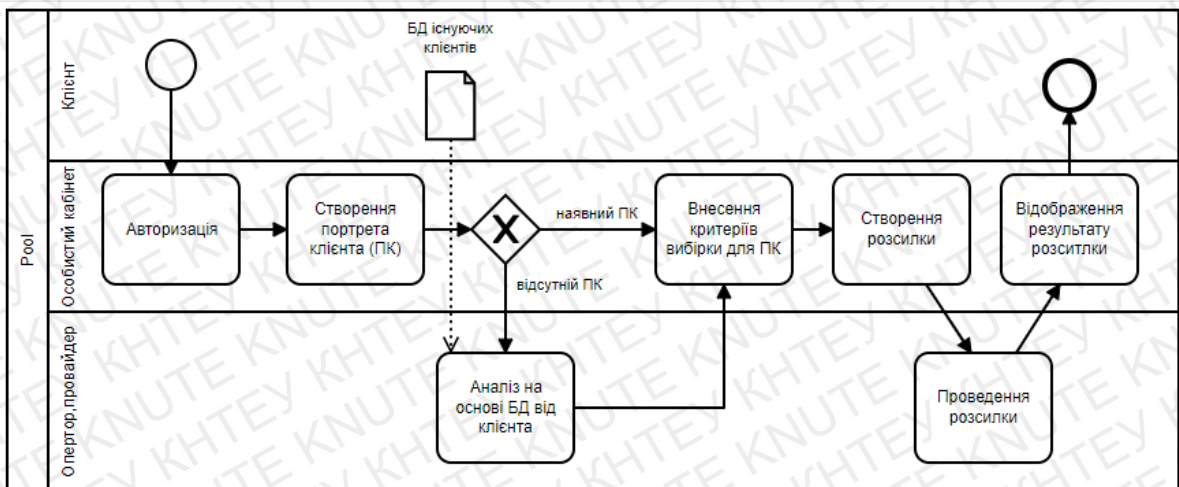


Рис. 3.12. BPMN діаграма бізнес-процесу створення рекламної розсилки через особистий кабінет

*складено автором

Також надалі варто обґрунтувати деякі технічні параметри створення особистого кабінету.

Доцільно виділити такі етапи розробки та впровадження проектного програмного забезпечення (рис. 3.13) [4]:

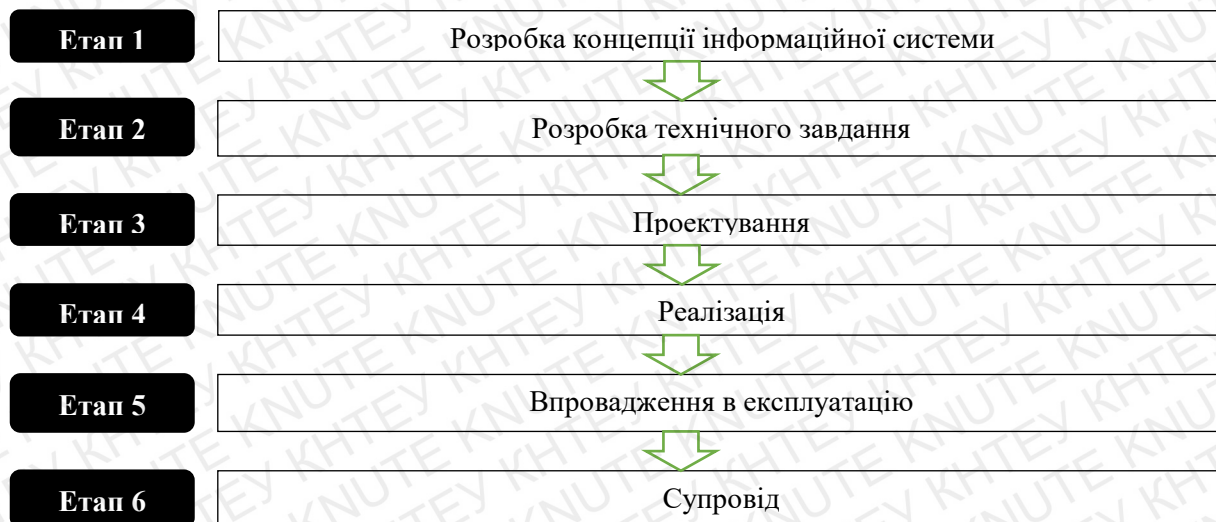


Рис. 3.13. Схема етапів впровадження проектного програмного забезпечення

*складено автором [27]

Дана розробка та впровадження особистого кабінету зробить досконаліший та простіший процес надсилання та обробку запиту. Якщо раніше для проведення розсилки, аналізу бази клієнту необхідно було заповнювати файл Excel (рис. 3.14), надсилати на обробку на певну ел. адресу або відповідального менеджера, то завдяки створенню особистого кабінету процес надсилання, обробки запиту стає швидшим, прозорішим та ефективнішим.

The screenshot shows an Excel spreadsheet with a form titled "Запит на послугу" (Request for service). The form is organized into several sections, each with a list of criteria and checkboxes:

- Критерії виборки** (Selection criteria): Includes "Сторона" (Side) with checkboxes for "Позитивна" (Positive) and "Негативна" (Negative).
- Профіль** (Profile): Includes "Рівень доступу" (Access level) with checkboxes for "Високий" (High), "Середній" (Medium), and "Низький" (Low); "Тип інформації" (Information type) with checkboxes for "Публічна" (Public) and "Приватна" (Private); "Використання веб-адресами" (Use of web addresses) with checkboxes for "Так" (Yes) and "Ні" (No).
- Локалізація** (Localization): Includes "Район навколо певної адреси" (Area around a specific address) with a dropdown menu.
- Цілісні групи** (Target groups): Includes "Перелік сайтів (включити)" (List of sites (include)) and "Перелік номерів (включити)" (List of numbers (include)).
- Індивідуальні та аудиторські додаткові критерії** (Individual and audit additional criteria): Includes "Перелік адрес (включити)" (List of addresses (include)).

The spreadsheet also has a header row with columns A through O and a footer row with the text "Заповнену форму надсилайте на пошту request@bigdata.com".

Рис. 3.14. Форма запиту

*складено автором

До переваг створення та впровадження особистого кабінету можна віднести:

- автоматизація процесу направлення запиту;
- проста та швидка обробка запитів;
- можливість у реальному часі відслідковувати статус запиту розсилки;
- доступ до особистого кабінету 24/7 у будь-який зручний час;
- збереження та доступ до перегляду статистики;
- аналіз статистики у розрізі каналів доставки;
- можливість за необхідністю вносити зміни;
- доступний попередній перегляд;
- планування розсилок;
- зручний інтерфейс.

До недоліків можна віднести:

- дороге обслуговування платформи;
- придбання хмарного сховища для зберігання даних;
- проведення інструктажу для клієнтів;
- необхідність проведення доробок та впровадження оновлень.

Отже, було описано основні принципи обробки даних мобільними операторами, провайдерами в рамках запропонованого бізнес-кейсу з розсилання рекламних повідомлень через проектоване програмне рішення – веб-орієнтований додаток. Наведені принципи можуть бути покладені в основу функціональних та системних вимог при впровадженні особистого кабінету, що має значну кількість переваг.

Техніко-економічне обґрунтування передбачає визначення можливого економічного ефекту для бізнесу, який можна отримати при розробці та впровадженні наведеного програмного рішення. При цьому для обґрунтування вигідності розробленого рішення пропонуємо використати такий алгоритм:

- 1) Визначення витрат часу на створення рішення та його впровадження, враховуючи навчання персоналу;
- 2) Визначення вартості витрат на створення програмного продукту;
- 3) Визначення можливого приросту доходів шляхом визначення

потенційних доходів та можливих витрат за умов врахування запропонованих напрямків;

4) Визначення економічного ефекту як різниці між приростом доходів та вартістю розробки програмного забезпечення. Даний розрахунок можна зробити порівнявши різницю між перспективним та ретроспективним прогнозами операційного прибутку, а отримані значення продисконтувати.

Важливим етапом також є обґрунтування необхідного методичного забезпечення впровадження проекту. Перелік методів та підходів для виконання кожної фази проекту викладено у Додатку А.

Економічна оцінка заходів з реалізації програми завдань проекту може бути проведена виходячи з графіку на реалізацію заходів (Додаток Б). Варто визначити можливі витрати робочого часу на реалізацію кожного завдання. Результати оцінювання фонду робочого часу на впровадження заходів з реалізації проекту наведено у Додатку Б. Проаналізувавши можна зробити висновок, що найоптимальнішим варіантом є використання Big Data на стороні оператора мобільного меседжингу за допомогою веб-орієнтованого додатку особистого кабінету.

Висновки до розділу 3

У даному розділі було розглянуто стратегічний контекст функціонування бізнес-замовника ТОВ «Гранд Ауто». Описано структуру бізнесу підприємства ТОВ «Гранд Ауто» за допомогою побудови стратегічної діаграми Business Model Canvas. На прикладі ТОВ «Гранд Ауто» було визначено основні проблеми розвитку бізнесу, які були зображені у вигляді моделі Solution Vision Canvas, що формалізує основні групи бізнес-вимог для запиту. Рішенням зазначених проблем підприємства є використання Big Data для дослідження цільової аудиторії, проведення таргетованих розсилок та оптимізації процесу просування продукції ТОВ «Гранд Ауто».

Було розроблено та запропоновано рішення у формі веб-орієнтованого додатку особистого кабінету, що дозволить вирішити бізнес-задачі замовника - дослідження цільової аудиторії для проведення рекламних кампаній та таргетованих розсилок. На побудованій діаграмі BPMN зображено верхньорівневий опис бізнес-процесу взаємодії оператора мобільного меседжингу або провайдера, програмного рішення, а саме веб-орієнтованого додатку особистого кабінету, та підприємства. Розроблено, описані та зображені основні сторінки особистого кабінету. Проаналізовані переваги та недоліки впровадження даного рішення, до основних переваг можна віднести автоматизацію процесу направлення запиту на обробку великих даних оператору мобільного меседжингу або провайдеру, просту та швидку обробку запиту та зручне проведення таргетованих розсилок.

ВИСНОВКИ ТА ПРОПОЗИЦІЇ:

В ході виконання випускної кваліфікаційної роботи було проведено аналіз теоретичних положень поняття великих даних, ключових характеристик, роль та функції великих даних в інформаційних технологіях. Наведено характеристику методів, техніки аналізу та визначенні основні принципи роботи з великими даними. На сьогоднішній день великі дані набули свого поширення практично у всіх сферах людського життя, які мають відношення до автоматизації та наявності певних систем, які здатні фіксувати інформацію. Кожні 5 років обсяг великих даних мав тенденцію до значного розширення, починаючи з 2000-х років. Використання Big Data дозволяє побачити певні та непомітні закономірності, що в свою чергу дозволяє оптимізувати всі сфери нашого життя – від державного управління до виробництва та телекомунікацій.

Було проаналізовано та охарактеризовано мобільний меседжинг як окрему бізнес-екосистему. Оскільки в бізнес-операціях спілкування за допомогою використання мобільного меседженгу стало важливим елементом підтримки ефективного та динамічного робочого процесу. Досліджено ринок телекомунікаційних послуг та закономірності його функціонування. Даний ринок характеризується великою ємністю, оскільки тією чи іншою мірою охоплює майже все населення нашої держави та світу, також для нього характерним є процеси глобалізації та інтеграції. Окремо було досліджено ринок мобільного зв'язку України, розглянуто його значення для національної економіки. Проаналізовано структуру ринку, стан конкурентного середовища та виявлено частку основних гравців на ньому.

Розглянуто особливості застосування Big Data, як інструменту вирішення бізнес-задач підприємств. На прикладі автомобільного центру ТОВ «Гранд Аутос» розглято процес формування запиту на стороні клієнта, шляхи його реалізації. Наведено способи виявлення та зазначені основні проблеми бізнес-замовника, розроблено модель Solution Vision Canvas, яка формалізує основні групи бізнес-вимог клієнта. Наведені принципи, що можуть бути покладені в

основу функціональних та системних вимог при проектуванні веб-орієнтованого додатку. Важливими питаннями, які додатково мають бути визначені при проектуванні веб-орієнтованого додатку, є вимоги до архітектури, вибір середовища розробки, вибір середовища для хостингу. Були зазначені особливості принципів обробки даних мобільними операторами в рамках запропонованого бізнес-кейсу.

У випускній кваліфікаційній роботі було розроблено веб-орієнтований додаток особистого кабінету, що забезпечить оптимізацію процесу надсилання запиту на обробку великих даних та проведення таргетованих розсилок. Дане рішення забезпечить збільшення ефективності, керованості процесом комунікації та буде доречним для впровадження для будь-якої сфери бізнесу.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ:

1. Барсегян А.А. Технології аналізу даних: Data Mining, Visual Mining, Text Mining, OLAP, 2007. 384 с.
2. Болгов В. Є. Ринок мобільного зв'язку України: тенденції, проблеми та перспективи / В. Є. Болгов, М. Р. Рязанов. // Економіка і організація управління. – 2017. – №2. – С. 51–55.
3. Від icq до whatsapp: еволюція мобільних месенджерів. URL: <https://uip.me/2015/01/ot-aski-do-whatsapp-evolyuciya-mobilnyx-messendzherov/>
4. Гнітецький Є. В. Big Data в маркетингу: орієнтація на споживача.//Економічний вісник НТУУ «КПІ».- 2017.- № 14. - С. 281–285.
5. Думікян М.М Геопросторова та секторальна структуризація ринку телекомунікаційних послуг // Причорноморські економічні студії. – 2016. – №10. – С. 12–17.
6. Історія розвитку технологій стільникового зв'язку. URL: <https://inventure.com.ua/analytics/investments/rynok-sotovoj-svyazi-ukrainy>
7. Ковальчук О. Я., Гайда Т. Ю., Жонца С. Я. Технології Big Data в інноваційному маркетингу.//Український журнал прикладної економіки. - 2018.- Том 3.- № 1 - С. 36-52.
8. Майєр-Шенбергер, Великі дані. Революція, яка змінила те, як ми живемо, працюємо та мислимо. М.:Манн, Іванов и Фербер, 2014. 240с.
9. Полякова А., Telegram, Viber, WhatsApp, Signal — яким месенджером можна довіряти Економічна правда 2017. URL: <https://www.epravda.com.ua/publications/2017/12/15/632183/>
10. Резниченко В.А. Що таке BigData. Проблеми програмування. 2019. № 3. С. 86-98
11. Силен Д, Мейсман А., Али М. Основи Data Science и BigData. Python і наука про дані, 2017. 336с.
12. Стільниковий зв'язок у розрізі. На чому заробляють мобільні оператори. Мобільний зв'язок України. URL: <https://mzu.in.ua/category/news/>

13. Big Data от А до Я. Часть 1: Принципы работы с великими данными, парадигма MapReduce. URL: <https://habr.com/company/dca/blog/267361/>
14. BigData and the Public Sector. A Survey of IT Decision Makers in Federal and State Public Sector organizations. TechAmerica Foundation, 2012. URL: <https://www.splunk.com/content/dam/splunk2/pdfs/fact-sheets/sap-publicsector-big-data-report-fiscal-2012.pdf>
15. BigData. 2021. URL: <https://www.it.ua/knowledge-base/technology-innovation/big-data-bolshie-dannye>
16. Dion Hinchcliffe. Big Data, The Moving Parts: Fast Data, Big Analytics, and Deep Insight. URL: <https://www.flickr.com/photos/dionh/7550578346/in/photostream/>
17. E. Begoli and J. Horey, "Design principles for effective knowledge discovery from big data," in Proceedings of the 10th Working IEEE/IFIP Conference on Software Architecture (ECSA '12). P. 215-218, August 2012.
18. Eileen McNulty. Understanding Big Data: The Seven V's (англ.). Dataconomy URL: <http://dataconomy.com/2014/05/seven-vs-big-data/>
19. George Firican. The 10 Vs of Big Data. URL: <https://tdwi.org/articles/2017/02/08/10-vs-of-big-data.aspx>
20. Gil Press. Cleaning Big Data: Most Time Consuming, Least Enjoyable Data Science Task, Survey Says. URL: <http://www.forbes.com/sites/gilpress/2016/03/23/data-preparation-most-time-consuming-least-enjoyable-data-science-task-survey-says/>
21. Growth in the Far East – United States. February, 2013. URL: <https://www.emc.com/collateral/analyst-reports/idcdigital-universe-united-states.pdf>
22. IDC Digital Universe Study. The digital universe in 2020: BigData, Bidder Digital Shadows, and Biggest
23. James Manyika, Michael Chui, Brad Brown, Jacques Bughin, Richard Dobbs, Charles Roxburgh, Angela Hung Byers. Big data: The next frontier for innovation, competition, and productivity. McKinsey Global Institute, 2011. URL: https://bigdatawg.nist.gov/pdf/MGI_big_data_full_report.pdf

24. Jennifer Dutcher. What Is Big Data? URL: <https://datascience.berkeley.edu/what-is-big-data/>
25. Laney Doug (2001) 3D Data Management: Controlling Data Volume, Velocity, and Variety. Technical Report 949, METAGroup (now Gartner). URL: <https://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf>
26. Nawsher Khan, Ibrar Yaqoob, Ibrahim Abaker Targio Hashem, Zakira Inayat, Waleed KamaleldinMahmoud Ali, Muhammad Alam, Muhammad Shiraz, and Abdullah Gani1. Big Data: Survey, Technologies, Opportunities, and Challenges // Hindawi Publishing Corporation The Scientific World Journal Volume 2014, Article ID 712826, 18 pages, URL: <http://dx.doi.org/10.1155/2014/712826>
27. Nawsher Khan, Ibrar Yaqoob, Ibrahim Abaker Targio Hashem, Zakira Inayat, Waleed KamaleldinMahmoud Ali, Muhammad Alam, Muhammad Shiraz, and Abdullah Gani. Big Data: Survey, Technologies, Opportunities, and Challenges. Hindawi Publishing Corporation The Scientific World Journal, Volume 2014, Article ID 712826, 18 pages. URL: <http://dx.doi.org/10.1155/2014/712826>
28. Neil Biehn. The Missing V's in Big Data: Viability and Value (англ.). URL: <https://www.wired.com/insights/2013/05/the-missing-vs-in-big-data-viability-and-value/>
29. The Four V's of Big Data (англ.). IBM (2011). URL: http://www.ibmbigdatahub.com/sites/default/files/infographic_file/4-Vs-of-big-data.jpg
30. Tom McNeill. The Eight V's of Supercomputing and Big Data. URL: <https://www.nimbix.net/eight-vs-supercomputing-big-data/>
31. Zikopoulos P., Parasuraman K., Deutsch T., Giles J., Corrigan D. (2013) Harness the power of big data The IBM big data platform. McGraw Hill Professional, New York, NY. URL: ftp://public.dhe.ibm.com/software/pdf/at/SWP10/Harness_the_Power_of_Big_Data.pdf
32. GSMA Intelligence Shares 'Global Mobile Trends 2021' URL: <https://www.gsma.com/newsroom/press-release/gsma-intelligence-shares-global-mobile-trends-2021/>

Перелік методів та підходів для виконання кожної фази проекту для
впровадження розсилок, через особистий кабінет

№	Фази проекту	Назва методу / групи методів
1	2	3
1	Ініціація проекту (організаційна частина)	- проведення переговорів;
2	Бізнес-аналіз	- методи стратегічного аналізу; - методи фінансового аналізу; - метод анкетного опитування
3	Розробка програмного продукту – системи електронного документообороту (технічна частина)	- використання технічних приймів та підходів до створення програмного забезпечення
4	Навчання персоналу та запуск програмного забезпечення в роботу	- метод опитування; - проведення тренінгів
5	Завершення проекту	- підготовка звітів, узагальнення

*Джерело: складено автором

Оцінювання фонду робочого часу на реалізацію заходів з реалізації проекту на підприємстві ТОВ «Гранд Аутос» у 2022 р.

Завдання (етапи) впровадження веб-орієнтованого додатку	Календарні строки реалізації завдань	Загальна тривалість часу на виконання завдань, людино-годин	Розподіл часу за ресурсами, людино-годин		
			персонал підприємства	Менеджер проекту, бізнес-аналітик	Технічний ІТ-персонал
1	2	3	4	5	6
Визначення основних проблем та загроз роботі підприємства	Лютий 2022	60	60	-	-
Обґрунтування можливих шляхів вирішення проблем	Лютий 2022	60	20	40	-
Проведення стартових комунікацій зі стейкхолдерами (переважно це власники та топ-менеджмент підприємства)	Лютий 2022	40	-	40	-
Визначення відповідальних за реалізацію проекту	Лютий 2022	10	5	5	-
Проведення стратегічного аналізу роботи підприємства	Березень 2022	80	-	80	-
Збір та узгодження бізнес-вимог від основних стейкхолдерів проекту	Березень 2022	60	20	40	-
Формування та узгодження функціональних та системних вимог до веб-орієнтованого додатку	Березень - квітень 2022	80	-	40	40
Складання плану імплементації проекту, формування його бюджету	Квітень 2022	20	-	20	-
Розробка необхідного програмного забезпечення	Травень – червень 2022	280	-	30	250
Проведення тестування програмного забезпечення		60	-	10	50
Знаходження помилок, їх виправлення, внесення необхідних змін для доведення програмного забезпечення до функціональних та системних вимог	Червень 2022	60	-	10	50
Підготовка документації для користувачів веб-орієнтованого додатку	Червень 2022	20	-	10	10
Проведення тренінгів для персоналу	Липень 2022	30	-	30	-

* Джерело: складено автором