МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
Національний університет кораблебудування імені адмірала Макарова
ФІЛОЛОГІЧНИЙ ФАКУЛЬТЕТ
Кафедра прикладної лінгвістики

«Допущена до захисту»
Завідувач кафедри
Ніна Філіппова
«26» червня 2025 р.

КВАЛІФІКАЦІЙНА РОБОТА
на здобуття першого (бакалаврського) рівня вищої освіти
на тему:
«Робота віртуальних асистентів у полілінгвальному середовищі: електронний довідник»

Виконала: студентка групи 4721,
Спеціальності 035 Філологія,
Спеціалізації 035.10 «Прикладна лінгвістика»
Заїка А. Р.
Керівник роботи:
Гогоренко О. В.

Миколаїв – 2025 р.

НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ КОРАБЛЕБУДУВАННЯ ІМЕНІ АДМІРАЛА МАКАРОВА
Навчально-науковий гуманітарний інститут

Кафедра прикладної лінгвістики
Спеціальність 035 «Філологія» **спеціалізація** 035.10 «Прикладна лінгвістика»
Галузі знань 03 Гуманітарні науки
Освітня програма ПРИКЛАДНА ЛІНГВІСТИКА
Першого (**бакалаврського**) рівня вищої освіти

«ЗАТВЕРДЖУЮ»
Гарант освітньої програми
О. В. Гогоренко
«18» червня 2025р.

ЗАВДАННЯ
НА КВАЛІФІКАЦІЙНУ РОБОТУ
на здобуття ступеня вищої освіти «бакалавр»

Студентці Заїці Антоніні Русланівні
(Прізвище, ім'я, по батькові)

1. **Тема роботи**: «Робота віртуальних асистентів у полілінгвальному середовищі: електронний довідник»
2. Керівник роботи: Гогоренко Олена Володимирівна
Затверджені наказом ректора № ____-уч від «____» _____ 2025 р.
3. **Термін подання роботи:** 15. 06. 2025 р.
4. Вихідні дані по роботі: розробити електронний довідник
5. **Перелік питань, що підлягають розробці** (найменування розділів)
1.СОЦІОЛІНГВІСТИЧНІ ЗАСАДИ ПОЛІЛІНГВАЛЬНОГО СЕРЕДОВИЩА;
2. КОМП'ЮТЕРНА ЛІНГВІСТИКА ТА АНАЛІЗ РОБОТИ ВІРТУАЛЬНИХ АСИСТЕНТІВ;
3. СТВОРЕННЯ ЕЛЕКТРОННОГО ДОВІДНИКА.
6. **Перелік презентаційних матеріалів**: презентаційний файл, демонстраційний файл з електронною розробкою продукту.

7**. Консультанти розділів роботи**

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		завдання видав	завдання прийняв
1	Гогоренко О.В	10.09.24	10.01.25

8. **Дата видачі завдання** 10.09.2024 р.

КАЛЕНДАРНИЙ ПЛАН

Номер	Назва етапів роботи	Терміни виконання етапів роботи	Примітка
1.	Вибір і теми випускної кваліфікаційної роботи, укладання бібліографії за обраною темою; опрацювання наукової і навчальної літератури за обраною темою	до 1.10.2024 р.	виконано
2.	Укладання плану-проспекту роботи.	до 5.10.2024 р.	виконано
3.	Збирання, аналіз, систематизація фактичного матеріалу до досліджуваної теми	до 31.10.2024 р.	виконано
4.	Написання тексту «Вступу»	до 10.11.2024 р.	виконано
5.	Написання тексту 1 розділу	до 30.11.2024 р.	виконано
6.	Написання тексту 2 розділу	до 28.02.2025 р.	виконано
7.	Написання тексту 3 розділу і загальних «Висновків»	до 31.03.2025 р.	виконано
8.	Подання чернетки повного тексту роботи керівникові. Доопрацювання тексту роботу згідно з зауваженнями керівника	до 20.04.2025 р.	виконано
9.	Попередній захист випускної кваліфікаційної роботи	31.05.2025 р.	виконано
10.	Доопрацювання повного тексту, подання оформленої за чинними вимогами випускної кваліфікаційної роботи на кафедру прикладної лінгвістики	10.06.2025 р.	виконано
11.	Отримання відгуку керівника	14.06.2025 р.	виконано
12.	Отримання рецензії	14.06.2025 р.	виконано
13.	Подання повного пакету документів, що супроводжують дипломну роботу	14.06.2025 р.	виконано
14.	Отримання допуску випускної кваліфікаційної роботи до захисту	15.06.2025 р.	виконано

Студент Заїка А. Р.

Керівник роботи Гогоренко О. В

АНОТАЦІЯ

Робота віртуальних асистентів у полілінгвальному середовищі: електронний довідник.

Дипломну роботу присвячено дослідженню роботи віртуальних асистентів у полілінгвальному середовищі, зокрема їх здатності працювати з багатьма мовами та ефективності у контексті мультимовних запитів. У роботі розглянуто концепцію віртуальних асистентів, основні принципи їх функціонування, а також виклики, які виникають при адаптації до багатомовних середовищ. Окрему увагу приділено вивченню алгоритмів машинного навчання, які дозволяють асистентам розпізнавати та обробляти запити на різних мовах, враховуючи культурні та соціальні аспекти.

Зокрема, досліджено різні моделі, що використовуються в сучасних віртуальних асистентах, такі як BERT, GPT і BART, а також їх здатність до адаптації до специфічних мовних особливостей, таких як діалекти, варіанти мов та мовне змішування. У результаті роботи було розроблено електронний довідник, який містить інформацію про найефективніші методи розпізнавання мовних запитів, а також пропонує рекомендації щодо вдосконалення роботи віртуальних асистентів у багатомовних середовищах.

Ключові слова: віртуальні асистенти, багатомовне середовище, машинне навчання, мовне змішування, електронний довідник, адаптація до мовних варіантів.

SUMMARY

Full Name: The work of virtual assistants in a multilingual environment: an electronic guide.

This thesis focuses on the functioning of virtual assistants in multilingual environments, specifically their ability to process and respond to queries in multiple languages. The work explores the concept of virtual assistants, their operational principles, and the challenges they face when adapting to multilingual contexts.

Special attention is given to the machine learning algorithms used by virtual assistants to recognize and process queries in various languages, taking into account cultural and social factors.

The study examines the effectiveness of language models such as BERT, GPT, and BART, particularly their ability to adapt to linguistic features like dialects, language variants, and code-switching. An electronic guide was developed as part of this research, providing information on the most efficient methods for recognizing multilingual queries, along with recommendations for improving virtual assistants' performance in multilingual environments.

Keywords: virtual assistants, multilingual environments, machine learning, code-switching, electronic guide, adaptation to language variants.

ЗМІСТ

ВСТУП

Актуальність теми: У сучасному світі швидкого розвитку технологій і глобалізації постає необхідність забезпечення ефективної міжмовної комунікації. Полілінгвальне середовище стає невід'ємною частиною не лише міжнародної взаємодії, а й щоденного спілкування у цифровому просторі. Віртуальні асистенти, як одна з інновацій комп'ютерної лінгвістики, відіграють важливу роль у подоланні мовних бар'єрів, полегшуючи доступ до інформації для представників різних мовних груп. Це робить тему роботи, пов’язану з аналізом функціонування віртуальних асистентів у полілінгвальному середовищі, особливо актуальною.

Розвиток полілінгвізму в сучасному суспільстві є багатоаспектним явищем, яке охоплює не лише соціолінгвістичний, а й технічний вимір. Соціолінгвістичні аспекти, пов’язані з багатомовністю, зосереджуються на взаємодії мов у суспільстві, їхніх функціях, а також на викликах, пов’язаних із мовним змішуванням. Однак у цифрову епоху виникає потреба досліджувати, як ці явища реалізуються в роботі віртуальних помічників, здатних адаптуватися до мовних потреб користувачів. Унаслідок цього зростає важливість глибокого аналізу інструментів штучного інтелекту, зокрема віртуальних асистентів, для визначення їхньої ефективності в обробці різних мов і діалектів.

Цифровізація освіти, бізнесу та щоденної комунікації спонукає до пошуку нових підходів до взаємодії в багатомовному середовищі. За допомогою віртуальних асистентів, таких як Google Bard, Siri, Alexa та інші, користувачі мають можливість вирішувати широкий спектр завдань, включно з перекладом, пошуком інформації та створенням текстів різними мовами. Проте існують значні виклики, пов'язані з багатомовністю, такі як обробка діалектів, ідіом або специфічних культурних контекстів. Недостатнє врахування цих аспектів може призводити до помилок у роботі віртуальних

асистентів, що ускладнює ефективну комунікацію та створює додаткові бар'єри для користувачів.

Особливу роль у розумінні актуальності теми відіграє соціокультурний контекст. У глобалізованому суспільстві взаємодія між представниками різних культур і мовних груп набуває все більшого значення. Віртуальні асистенти стають не лише інструментом спрощення мовної взаємодії, але й посередниками між культурами, здатними адаптувати інформацію відповідно до культурних особливостей. Таким чином, їхній вплив виходить за межі технічної сфери, формуючи нові форми міжкультурної комунікації. Дослідження цих процесів дозволяє не лише зрозуміти сучасні тенденції віртуальної взаємодії, але й розробити нові підходи до вдосконалення роботи асистентів у полілінгвальному середовищі.

Актуальність дослідження також зумовлена потребою у створенні якісних електронних ресурсів для навчальних та дослідницьких цілей. Використання віртуальних асистентів у навчальному процесі відкриває нові можливості для розвитку мовних навичок у багатомовному контексті, спрощуючи доступ до навчальних матеріалів різними мовами. З іншого боку, ці інструменти можуть стати об'єктом дослідження для лінгвістів, розробників штучного інтелекту та освітян, які прагнуть удосконалити технології багатомовної обробки даних.

Окремо слід наголосити на значущості розробки електронного довідника, який би систематизував знання про роботу віртуальних асистентів у полілінгвальному середовищі. Такий довідник стане корисним не лише для науковців, але й для освітян, розробників програмного забезпечення та користувачів, які прагнуть оптимізувати свою роботу з багатомовними інструментами. Завдяки актуальній інформації та практичним рекомендаціям, електронний довідник сприятиме подальшому розвитку технологій штучного інтелекту та їхньому використанню в різних галузях.

У контексті дослідження варто зазначити, що швидкий розвиток штучного інтелекту супроводжується виникненням нових викликів, які

потребують наукового осмислення. Одним із таких викликів є забезпечення точності та адекватності перекладів, адаптації текстів, а також розпізнавання та аналізу діалектів. Віртуальні асистенти вже демонструють значний прогрес у цих напрямках, але водночас стикаються з обмеженнями, пов'язаними з мовною багатозначністю, контекстуальними факторами та культурними відмінностями. Унаслідок цього дослідження цих аспектів стає критично важливим для подальшого вдосконалення технологій.

Ще одним аспектом актуальності є впровадження інновацій у навчальний процес, зокрема у вивчення іноземних мов. Завдяки віртуальним асистентам учні та студенти можуть отримувати підтримку у вивченні нових мов, адаптовану до їхніх індивідуальних потреб. Однак ефективність такого підходу значною мірою залежить від якості роботи цих інструментів, їхньої здатності враховувати особливості мовної структури, семантики та прагматики. Це створює додатковий простір для досліджень, спрямованих на аналіз і вдосконалення віртуальних асистентів.

З огляду на це, робота, спрямована на дослідження функціонування віртуальних асистентів у полілінгвальному середовищі, має значний науковий і практичний потенціал. Її результати можуть бути використані для подальшого розвитку комп’ютерної лінгвістики, розробки інноваційних рішень у сфері штучного інтелекту та впровадження цих рішень у різні сфери діяльності, включаючи освіту, бізнес і міжкультурну комунікацію.

Об’єкт дослідження: Об’єктом дослідження є процеси функціонування віртуальних асистентів у полілінгвальному середовищі.

Предмет дослідження: Предметом дослідження є специфіка обробки багатомовних даних віртуальними асистентами, їхні алгоритми роботи та можливості адаптації до полілінгвального контексту.

Мета дослідження: Мета дослідження полягає у комплексному аналізі функціонування віртуальних асистентів у полілінгвальному середовищі, визначенні їхньої ефективності, виявленні проблем, пов’язаних із обробкою

багатомовних даних, та створенні електронного довідника, який систематизує знання з цієї тематики.

Завдання дослідження

1. Дослідити поняття білінгвізму та полілінгвізму, їхні соціолінгвістичні основи та роль у сучасному суспільстві.
2. Проаналізувати функції мов у суспільстві, а також вплив мовного змішування на комунікацію в цифрову епоху.
3. Вивчити сучасний стан розвитку комп'ютерної лінгвістики, її методи та досягнення.
4. Провести аналіз роботи віртуальних асистентів, таких як Google Bard, у полілінгвальному середовищі.
5. Порівняти ефективність обробки стандартної мови та діалектів віртуальними асистентами.
6. Розробити електронний довідник, який інтегрує результати дослідження та рекомендації для використання віртуальних асистентів у багатомовному контексті.

Джерельна база дослідження: Джерельна база дослідження базується на працях провідних науковців у галузі соціолінгвістики, зокрема дослідженнях Джошуа Фішмана, який вивчав функції мов у багатомовному суспільстві, і Луїса-Жана Кальве, що аналізував динаміку мовного змішування. У сфері комп'ютерної лінгвістики ключовими джерелами стали роботи Ноама Хомського, який заклав основи теорії трансформаційної граматики, та Джурафа Гаусмана, що розглядав алгоритми природної мови в роботі штучного інтелекту.

Особлива увага приділялася дослідженням сучасних вчених, таких як Даніель Юрак (аналіз моделей машинного перекладу) та Карла Вегмана, яка досліджує аспекти адаптації віртуальних асистентів до різних мовних систем. Додатково були використані технічні звіти та документація таких систем, як Google Bard, Siri, Alexa, з метою вивчення їхніх алгоритмів обробки багатомовних даних.

Також вивчалися наукові статті, опубліковані в журналах *Computational Linguistics* і *Language Resources and Evaluation*, які містять аналітичні матеріали щодо сучасного стану комп'ютерної лінгвістики.

Фактичний матеріал: Фактичний матеріал ґрунтується на аналізі реальної роботи віртуальних асистентів у багатомовному середовищі. Наприклад, тестування Google Bard проводилося для порівняння його здатності обробляти стандартну українську мову та українські діалекти (бойківський, гуцульський). Аналізувалися також результати роботи Siri при виконанні запитів англійською, французькою та українською мовами.

Для розширення емпіричної бази було використано дослідження про ефективність роботи систем штучного інтелекту у контексті перекладу (зокрема дослідження про нейронний переклад у *Journal of Artificial Intelligence Research*). Практичні приклади також включають аналіз запитів до Alexa з елементами мовного змішування (спанґліш, суржик).

РОЗДІЛ 1. СОЦІОЛІНГВІСТИЧНІ ЗАСАДИ ПОЛІЛІНГВАЛЬНОГО СЕРЕДОВИЩА

1.1. Поняття білінгвізму та полілінгвізму

Білінгвізм і полілінгвізм є фундаментальними поняттями у дослідженні багатомовності та її проявів у сучасному суспільстві, що дедалі більше інтегрується завдяки глобалізації та технологічному прогресу. Білінгвізм, що позначає здатність людини використовувати дві мови на різних рівнях, має широкий спектр проявів: від пасивного розуміння однієї мови до активного володіння нею на рівні рідної. У свою чергу, полілінгвізм охоплює багатомовність, тобто здатність оперувати трьома й більше мовами, і є ключовим компонентом багатокультурних спільнот, де взаємодіють різні лінгвістичні системи. У науковій літературі акцентується на тому, що білінгвізм є не лише індивідуальним, а й соціальним явищем, яке формується унаслідок історичних, економічних і культурних процесів, наприклад, міграції, завоювань чи культурного обміну. Полілінгвізм, у свою чергу, часто розглядається як необхідність у сучасному глобалізованому світі, де знання кількох мов сприяє ефективній міжкультурній комунікації та розширенню можливостей для соціальної і професійної мобільності [1].

Існують різні класифікації білінгвізму, які зосереджуються на рівнях володіння мовами (рівень «збалансованості» мов), контекстах їхнього використання (формальний чи неформальний) та характері їх засвоєння (природний чи штучний). Наприклад, природний білінгвізм виникає тоді, коли дві мови засвоюються одночасно з дитинства, як у випадках, коли батьки розмовляють різними мовами. Натомість штучний білінгвізм передбачає свідоме вивчення іншої мови через формальну освіту або самостійне навчання. Полілінгвізм, на відміну від білінгвізму, не обмежується лише двома мовами й частіше пов'язується із соціальною необхідністю. У таких випадках використання багатьох мов часто визначається професійними потребами,

наприклад, у міжнародному бізнесі чи дипломатії, де знання кількох мов стає критично важливим. Ці форми білінгвізму й полілінгвізму мають суттєвий вплив на когнітивні процеси, зокрема на розвиток уваги, пам’яті та здатності до вирішення проблем, оскільки постійне перемикання між мовами активує специфічні ділянки мозку.

У теоретичному аспекті білінгвізм і полілінгвізм розглядаються як складні феномени, що взаємодіють із різними мовними, культурними та соціальними чинниками. Наприклад, дослідники, такі як Джошуа Фішман, підкреслюють функціональний характер білінгвізму, коли кожна мова виконує певну роль у житті індивіда або спільноти. Наприклад, одна мова може використовуватися вдома для спілкування з родиною, тоді як інша — у професійному або освітньому середовищі. У контексті полілінгвізму функціональна сегментація мов стає ще складнішою, оскільки кількість мов зростає, що збільшує потребу у гнучкості та адаптації до різних соціальних ситуацій. Крім того, важливим є питання збереження ідентичності, адже для багатьох білінгвів і полілінгвів мова є не лише засобом комунікації, а й інструментом самовираження, який тісно пов'язаний із культурними традиціями та цінностями. У цьому контексті варто згадати Луїса-Жана Кальве, який аналізує процеси мовного змішування як природний наслідок багатомовності [4].

З технічної точки зору, білінгвізм і полілінгвізм набувають особливої актуальності у цифрову епоху, коли розробка технологій штучного інтелекту, таких як віртуальні асистенти, значною мірою залежить від здатності ефективно обробляти багатомовні дані. Наприклад, віртуальні асистенти, як-от Google Bard або Siri, мають інтегрувати мовну інформацію з різних лінгвістичних систем, адаптувати її до культурних особливостей користувачів та забезпечувати адекватну комунікацію у реальному часі. Тут важливим є поняття мовного перемикання (code-switching), яке є типовим для білінгвів і полілінгвів. Цей феномен вимагає від технологій здатності швидко розпізнавати, яку мову або навіть її варіант (наприклад, діалект чи жаргон)

використовує користувач, і відповідно адаптувати алгоритми обробки інформації. Наприклад, при створенні багатомовних чат-ботів, які мають комунікувати кількома мовами в одному діалозі, необхідно враховувати лінгвістичні, семантичні й прагматичні аспекти кожної мови, що робить завдання ще більш складним.

На глобальному рівні білінгвізм і полілінгвізм сприяють формуванню нового типу соціальних відносин, де мова виступає не лише інструментом комунікації, а й засобом взаємного впливу між культурами. Наприклад, у Європейському Союзі багатомовність закріплена на інституційному рівні, що забезпечує рівноправність усіх офіційних мов. Це створює умови для активного розвитку полілінгвізму серед громадян ЄС, які вивчають кілька мов із дитинства. Водночас у країнах, де домінує одна мова, білінгвізм часто асоціюється з культурним або етнічним різноманіттям, наприклад, у Канаді, де офіційними мовами є англійська та французька. Полілінгвізм у таких умовах слугує інструментом міжкультурного діалогу, зберігаючи баланс між інтеграцією та ідентичністю [7].

Поняття білінгвізму й полілінгвізму також мають значний вплив на освітню сферу, де дедалі більше уваги приділяється вивченню іноземних мов. У сучасній педагогіці білінгвізм розглядається як ефективний спосіб формування когнітивних і комунікативних компетентностей, а полілінгвізм — як засіб розвитку міжкультурного мислення. Наприклад, програми занурення в мовне середовище (language immersion programs) успішно впроваджуються в багатьох країнах для підтримки багатомовності серед молоді. Водночас освітні технології, такі як адаптивні платформи для вивчення мов, інтегрують білінгвальні й полілінгвальні підходи для створення персоналізованого навчального досвіду. Це дозволяє учням розвивати свої мовні навички у багатокультурному контексті, що стає критично важливим у сучасному глобалізованому світі.

Отже, білінгвізм і полілінгвізм є складними, багатогранними явищами, які охоплюють соціальні, культурні, когнітивні й технічні аспекти. Їхнє

дослідження має ключове значення для розуміння міжмовної комунікації, розвитку освітніх технологій і вдосконалення алгоритмів штучного інтелекту, які працюють у багатомовних середовищах [15].

Таблиця 1.1

Поняття білінгвізму та полілінгвізму

Поняття	Білінгвізм	Полілінгвізм
Визначення	Здатність людини використовувати дві мови для спілкування в різних контекстах.	Здатність людини використовувати три і більше мов для спілкування та взаємодії.
Особливості	Може бути результатом навчання або зростання в двомовному середовищі.	Поширений серед людей, які проживають у багатокультурних суспільствах або активно вивчають мови.
Переваги	Розширення когнітивних здібностей, гнучкість мислення, культурна адаптація.	Посилення культурної інтеграції, більша конкурентоспроможність на ринку праці.
Виклики	Можливе зниження рівня володіння однією з мов через недостатню практику.	Складність у досягненні високого рівня володіння кількома мовами одночасно.

1.2. Соціолінгвістичне підґрунтя: функції мов у суспільстві

Комунікативна та когнітивна функції мови є фундаментальними у процесах взаємодії, мислення та пізнання світу, що визначає їхню ключову роль у житті суспільства. Комунікативна функція мови забезпечує передачу інформації між індивідами, що є основою для соціальної взаємодії. Завдяки мові люди мають змогу формулювати свої думки, виражати емоції, обмінюватися досвідом і координувати спільну діяльність. Ця функція реалізується через різні комунікативні акти, що включають вербальні й невербальні засоби. У соціальних контекстах комунікативна функція набуває різних форм залежно від середовища: у формальних ситуаціях, таких як ділові переговори або освітні процеси, мова використовується для чіткого вираження

ідей і передачі спеціалізованої інформації. Натомість у неформальних умовах, наприклад, у сімейному чи дружньому спілкуванні, мова стає більш гнучкою, включаючи емоційно забарвлені вирази, сленг або навіть елементи мовного змішування. Однією з важливих характеристик комунікативної функції є її здатність забезпечувати інтерактивність у різних масштабах: від міжособистісного рівня до глобальної комунікації. Зокрема, у багатомовних суспільствах комунікативна функція мови має стратегічне значення, адже вона виступає як посередник між різними культурами, забезпечуючи взаєморозуміння та інтеграцію. Наприклад, в умовах глобалізації англійська мова набула статусу лінгва-франка, що дозволяє представникам різних народів спілкуватися на спільній платформі, об’єднуючи їх у межах професійної або наукової спільноти [2].

Когнітивна функція мови є не менш важливою, оскільки вона зосереджена на формуванні мислення, обробці інформації та створенні концептів, які відображають реальність. Мова виконує роль інструмента, за допомогою якого людина структурує свої знання, аналізує явища та формує уявлення про навколишній світ. Ця функція є тісно пов’язаною з процесами пізнання, адже саме через мову відбувається вербалізація думок, що дозволяє організувати інформацію у зрозумілі категорії. Наприклад, концептуалізація часу й простору може значно варіюватися залежно від мовної системи, що впливає на спосіб мислення. У західних мовах, таких як англійська чи німецька, час сприймається лінійно, тоді як у деяких азійських мовах, наприклад, у китайській, час може сприйматися циклічно. Такі відмінності демонструють, як когнітивна функція мови формує специфічне бачення світу, притаманне певній культурі. У науковому аспекті когнітивна функція мови забезпечує передачу знань між поколіннями, створення абстрактних понять і розвиток мислення. Завдяки цій функції людство змогло побудувати складні системи, такі як наука, філософія, техніка й мистецтво, що значно вплинули на еволюцію суспільства [9].

Синтез комунікативної та когнітивної функцій мови забезпечує не лише соціальну взаємодію, а й розвиток культури, науки та технологій. У полілінгвальних суспільствах ці функції взаємодіють у ще більш складний спосіб, оскільки мова виконує роль інструмента адаптації до багатомовного середовища. Наприклад, у Канаді, де офіційними мовами є англійська й французька, когнітивна функція мови допомагає громадянам адаптуватися до двох мовних систем, формуючи гнучкі навички мислення. Подібні процеси відбуваються й у Європейському Союзі, де знання кількох мов стає стандартом для професійної діяльності. У таких умовах когнітивна функція сприяє інтеграції знань з різних лінгвістичних традицій, збагачуючи інтелектуальний потенціал суспільства. У цьому контексті також варто зазначити роль комунікативної функції, яка дозволяє ефективно координувати взаємодію між представниками різних мовних спільнот, створюючи умови для співпраці у сферах політики, економіки та культури [5].

Сучасні технології, зокрема штучний інтелект і системи обробки природної мови, також активно використовують комунікативну та когнітивну функції мови. Наприклад, віртуальні асистенти, такі як Siri, Google Bard чи Alexa, демонструють здатність не лише спілкуватися з користувачами, але й адаптуватися до їхніх мовних потреб, виконуючи запити кількома мовами. Ці системи, розроблені на основі моделей штучного інтелекту, здатні аналізувати текст, розпізнавати мовні особливості й навіть враховувати культурний контекст. Тут комунікативна функція мови реалізується через інтерактивність, забезпечуючи діалог між користувачем і машиною, тоді як когнітивна функція спрямована на обробку складної інформації, наприклад, перекладу текстів або створення нових текстових даних. Такі технології відкривають нові горизонти для використання мови, інтегруючи її у сфери, які раніше були недосяжними для людини, такі як автоматизація бізнес-процесів, розробка освітніх платформ чи створення мультимовних ресурсів [6].

Крім того, комунікативна й когнітивна функції мови мають значний вплив на розвиток особистості. У процесі навчання мова забезпечує

формування знань, розвиток критичного мислення й оволодіння соціальними навичками. У багатомовному середовищі ці функції розширюють горизонти сприйняття, сприяючи формуванню міжкультурної компетенції. Наприклад, у полілінгвальних школах учні, які вивчають кілька мов, мають змогу не лише оволодіти новими мовними системами, але й зрозуміти культурні особливості інших народів, що сприяє їхньому соціальному та культурному розвитку. Це також сприяє формуванню когнітивної гнучкості, яка є необхідною для успішної адаптації до змін у сучасному світі.

Комунікативна та когнітивна функції мови є фундаментальними для забезпечення соціальної взаємодії, розвитку мислення та створення культурного й наукового спадку. Їхній вплив простежується як у повсякденному житті, так і в глобальних процесах, таких як інтеграція суспільств, розвиток технологій і формування міжкультурних зв'язків. У сучасному світі, де багатомовність стає нормою, ці функції набувають ще більшого значення, забезпечуючи гармонійний розвиток суспільства та індивіда. [8]

Мова є однією з найпотужніших складових у формуванні й збереженні культурної ідентичності, оскільки вона несе в собі унікальні елементи культурного спадку, історії, традицій та цінностей певного народу. Вона є засобом, через який передаються колективна пам'ять, світогляд і система уявлень про навколишній світ, що формуються протягом століть. Завдяки мові зберігаються літературні, музичні та інші форми мистецтва, які відображають унікальність кожної культури. Наприклад, у багатьох народів фольклор передається з покоління в покоління саме через мову, яка зберігає не лише зміст, але й емоційне забарвлення текстів, що дозволяє підтримувати духовну єдність між поколіннями. У цьому контексті збереження мови стає критично важливим для підтримання культурної ідентичності, особливо в умовах глобалізації, де відбувається взаємодія різних культур і мов. Глобалізація часто призводить до домінування мов більшості, таких як англійська, що створює ризик втрати менших мов і, як наслідок, унікальних культурних особливостей.

У таких умовах підтримка й розвиток рідної мови, наприклад, через освіту, видавничу діяльність чи медіа, стають важливими інструментами для захисту культурної спадщини [1].

Крім того, мова відіграє важливу роль у процесі соціальної інтеграції, адже вона є основним засобом комунікації, який об'єднує людей у суспільстві. У багатомовних спільнотах мова може бути як засобом об'єднання, так і джерелом конфліктів, особливо якщо певна мовна група відчуває дискримінацію або обмеження у правах через мовну політику. Наприклад, в історії багатьох країн можна побачити приклади, коли пригнічення мови меншості ставало причиною соціальної напруги або навіть конфліктів. Успішна соціальна інтеграція потребує створення умов для рівноправного використання мов різних груп, що сприяє взаєморозумінню та співпраці між ними. Це особливо важливо у країнах із багатонаціональним складом населення, таких як Канада чи Швейцарія, де офіційно визнаються кілька мов, що дозволяє зберегти баланс між різними культурними й мовними спільнотами. Водночас спільна мова може слугувати засобом інтеграції мігрантів, які вивчають її як другу, аби адаптуватися до нового середовища. У таких випадках знання мови більшості не лише полегшує комунікацію, але й сприяє прийняттю нових членів спільноти як повноправних її учасників. Наприклад, у багатьох європейських країнах функціонують програми мовної адаптації для мігрантів, що спрямовані на забезпечення їхньої соціальної інтеграції через оволодіння мовою та знайомство з культурними особливостями приймаючої країни [3].

У сучасному світі мова також є ключовим інструментом для формування національної ідентичності та зміцнення соціальної єдності. Вона виступає символом суверенітету та культурної унікальності, що є особливо важливим для країн, які пройшли через періоди колонізації або асиміляції. Наприклад, в Україні після здобуття незалежності особливого значення набуло питання розвитку української мови як символу національного відродження та збереження ідентичності. Мовна політика, спрямована на популяризацію та

захист державної мови, відіграє ключову роль у зміцненні культурних і політичних зв'язків між громадянами, формуючи почуття приналежності до однієї нації. Водночас у таких умовах важливо забезпечити підтримку мов меншин, аби уникнути соціального розколу та створити умови для гармонійного співіснування різних культур у межах однієї держави. З іншого боку, у багатьох постколоніальних країнах збереження рідних мов стикається зі значними викликами через домінування мов колишніх колонізаторів, таких як англійська, французька чи іспанська. Це створює необхідність розробки стратегій, спрямованих на відновлення й популяризацію корінних мов як важливого елемента культурної ідентичності [4].

Мова також є важливим чинником у міжкультурній комунікації, забезпечуючи діалог між різними етнічними групами та сприяючи взаємному збагаченню культур. У глобалізованому світі, де дедалі частіше відбувається змішування різних культур, мова стає не лише засобом комунікації, але й інструментом дипломатії, культурного обміну та інтеграції. Наприклад, у міжнародних організаціях, таких як ООН, офіційні мови використовуються для забезпечення рівного доступу до інформації та прийняття рішень, що враховують інтереси всіх членів. Це підкреслює важливість багатомовності як інструмента забезпечення соціальної справедливості й рівності. У той же час розвиток технологій, таких як автоматичний переклад і системи розпізнавання мови, відкриває нові можливості для збереження й використання мов у глобальному контексті. Ці технології дозволяють не лише подолати мовні бар'єри, але й сприяють збереженню малих мов, оцифровуючи їхні ресурси та забезпечуючи доступ до них для майбутніх поколінь.

Мова є не лише засобом збереження культурної ідентичності, але й важливим інструментом соціальної інтеграції. Вона дозволяє зберігати унікальність культур, забезпечуючи водночас їхню взаємодію в межах багатомовних суспільств. У сучасному світі, де глобалізація створює нові виклики для мовного різноманіття, збереження та розвиток мов є необхідною

умовою для гармонійного співіснування культур, зміцнення соціальної єдності та забезпечення рівності [1].

Таблиця 1.2

Соціолінгвістичне підґрунтя: функції мов у суспільстві

Функція мови	Визначення	Приклад реалізації
Комунікативна функція	Забезпечення передачі інформації між індивідами, що є основою соціальної взаємодії.	Спілкування на діловій зустрічі або в побуті.
Когнітивна функція	Формування мислення, обробка інформації та створення концептів, що відображають реальність.	Використання мови для формулювання наукових ідей або абстрактного мислення.
Культурна функція	Збереження та передача культурних традицій, історії та цінностей.	Передача фольклорних оповідей, народних пісень чи ритуалів.
Ідентифікаційна функція	Формування та підтримка соціальної та культурної ідентичності індивіда або групи.	Використання мови як символу національної ідентичності в багатонаціональних країнах.

1.3. Виклики та особливості мовного змішування в цифрову епоху

Code-switching, або перемикання мовних кодів, є поширеним явищем у сучасному полілінгвальному суспільстві, що особливо виразно проявляється в умовах цифрової комунікації. Цей феномен означає свідоме або несвідоме перемикання між різними мовами, діалектами чи мовними стилями в межах однієї розмови чи тексту. У цифровому середовищі, зокрема у соціальних мережах, месенджерах, форумах чи онлайн-спільнотах, code-switching часто стає не лише засобом комунікації, але й інструментом самовираження, засобом адаптації до багатомовного контексту, а також способом демонстрації культурної приналежності чи належності до певної соціальної групи. Виникнення цього явища у цифровому просторі є наслідком взаємодії кількох факторів: глобалізації, яка сприяє збільшенню мовного контакту, розвитку технологій, що забезпечують багатомовну комунікацію, а також динамічних змін у суспільстві, які підтримують мовне розмаїття як важливу цінність [15].

Однією з основних причин поширення code-switching у цифровій комунікації є можливість миттєвого доступу до багатомовних платформ, які стимулюють мовне змішування. Наприклад, у соціальних мережах, таких як Facebook чи Instagram, користувачі часто використовують різні мови для адресації різних груп аудиторії: рідна мова для близького оточення та англійська як глобальна лінгва-франка для ширшої спільноти. Подібне мовне перемикання можна помітити й у професійних контекстах, коли під час корпоративного листування чи в онлайн-зустрічах люди комбінують елементи різних мов для уточнення термінології, досягнення взаєморозуміння або спрощення складних понять. Наприклад, у сфері інформаційних технологій англійська часто використовується для позначення технічних термінів, тоді як основна частина комунікації може вестися локальною мовою. Такий підхід є зручним, адже дозволяє не лише заощаджувати час, але й долати мовні бар'єри у мультикультурних командах [17].

З іншого боку, code-switching у цифровому просторі виконує важливу соціокультурну функцію, будучи інструментом ідентифікації та належності до певної спільноти. Наприклад, у спілкуванні українських користувачів у соціальних мережах можна помітити часте чергування між українською, російською та англійською мовами залежно від контексту, теми або цільової аудиторії. Така поведінка є типовою для представників полілінгвальних суспільств, які використовують мови не лише як засіб комунікації, але й як маркер культурної чи соціальної ідентичності. Аналогічно, серед молоді популярним є використання сленгу або жаргону в поєднанні з офіційною мовою, що дозволяє створити специфічну форму вираження, зрозумілу лише вузькому колу людей. У цьому контексті code-switching стає способом підкреслити власну унікальність, демонструючи водночас знання кількох мов і належність до певної субкультури [38].

Важливим аспектом є також когнітивний вимір code-switching, який сприяє розвитку багатомовної компетенції та гнучкості мислення. Перемикання між мовами у цифровій комунікації вимагає від користувача

швидкого аналізу контексту, оцінки потреб аудиторії та вибору відповідних мовних ресурсів. Наприклад, у чатах або коментарях до публікацій учасники часто адаптують мову залежно від культурного чи мовного фону співрозмовника. Це демонструє не лише здатність до мовної адаптації, але й високий рівень міжкультурної комунікації, що стає критично важливим у глобалізованому світі. Водночас постійне мовне перемикання активізує когнітивні процеси, такі як увага, пам'ять і здатність до вирішення проблем, що було підтверджено численними дослідженнями у сфері нейролінгвістики.

Розвиток цифрових технологій також сприяв появі автоматичних інструментів, що підтримують code-switching, таких як багатомовні клавіатури, перекладачі або системи автокорекції. Наприклад, платформи Google Translate чи DeepL дозволяють миттєво перекладати текст, забезпечуючи взаємодію між людьми, які розмовляють різними мовами. Водночас ці технології стикаються з численними викликами, пов'язаними з обробкою змішаних мовних даних. Наприклад, алгоритми часто не можуть адекватно враховувати контекст або культурні особливості, що призводить до помилок у перекладі чи непорозумінь між користувачами. Однак розвиток штучного інтелекту та машинного навчання поступово долає ці обмеження, створюючи умови для ефективного використання code-switching у цифровій комунікації [48].

Code-switching також має значний вплив на освітній процес, особливо в контексті навчання іноземних мов. У цифрових класах чи онлайн-курсах викладачі часто використовують мовне перемикання для пояснення складних тем, адаптації навчального матеріалу або забезпечення мотивації студентів. Наприклад, на платформах Coursera чи Duolingo змішування рідної та іноземної мов допомагає студентам поступово освоювати нову мовну систему, знижуючи рівень стресу та підвищуючи ефективність навчання. У таких випадках code-switching стає важливим педагогічним інструментом, який сприяє як когнітивному розвитку, так і інтеграції студентів у багатомовне середовище.

Code-switching є важливим проявом полілінгвізму в цифровій комунікації, який виконує як комунікативні, так і соціокультурні та когнітивні функції. Він сприяє адаптації до багатомовного контексту, забезпечує ефективність міжкультурної взаємодії та відкриває нові можливості для використання мов у цифровому просторі. Водночас розвиток технологій і глобалізація створюють умови для подальшого поширення цього явища, роблячи його невід'ємною частиною сучасного комунікаційного ландшафту.

Мовне змішування у соціальних мережах: причини і наслідки

Мовне змішування у соціальних мережах є одним із найяскравіших проявів сучасного полілінгвізму, що виник під впливом глобалізації, технологічного прогресу та динамічних змін у суспільних комунікаціях. Соціальні мережі, такі як Facebook, Instagram, TikTok і Twitter, стали платформами, де перетинаються різні мови, культури й стилі спілкування. Основними причинами мовного змішування є зростаючий вплив англійської мови як глобальної лінгва-франка, інтеграція локальних мов у міжнародний контекст, а також індивідуальні потреби користувачів у самовираженні та адаптації до полікультурного середовища. Наприклад, у постах або коментарях часто спостерігається поєднання англійської з рідною мовою користувачів, що дозволяє їм одночасно досягати локальної аудиторії та взаємодіяти з міжнародною спільнотою. Крім того, технологічні фактори, такі як автоматичні перекладачі, багатомовні клавіатури та системи автокорекції, значно спрощують процес змішування мов, дозволяючи користувачам легко переходити між різними мовними системами навіть у межах одного повідомлення [51].

Психологічний аспект також відіграє важливу роль у поширенні мовного змішування у соціальних мережах. Люди прагнуть демонструвати свою багатомовність як ознаку освіченості, глобальної обізнаності чи належності до певної соціальної групи. Наприклад, молодь часто включає до текстів елементи англійської мови, сленгу або навіть специфічних технічних термінів, щоб виглядати сучасними та відповідати трендам. Такі тенденції можна

спостерігати у популярних соціальних мережах, де створення контенту з використанням кількох мов стає своєрідним маркером культурної ідентичності та відкритості до світу. У професійному контексті мовне змішування може виконувати функцію адаптації до специфічної термінології певної галузі, особливо у сферах інформаційних технологій, науки чи бізнесу, де англійська є основною мовою спілкування. У цьому випадку використання іншомовних термінів допомагає спрощувати комунікацію та забезпечувати точність передачі інформації, уникаючи складнощів перекладу [53].

Наслідки мовного змішування у соціальних мережах є багатогранними й охоплюють як позитивні, так і негативні аспекти. З одного боку, це явище сприяє формуванню глобалізованої комунікаційної культури, де люди різних національностей і мовних груп можуть легко взаємодіяти. Мовне змішування дозволяє створювати нові форми самовираження, що збагачують культуру й сприяють розвитку креативності. Наприклад, у соцмережах формується новий вид контенту, що включає меми, відео та тексти, у яких різні мови використовуються для створення гумористичних, емоційно насичених чи соціально значущих повідомлень. З іншого боку, це явище може мати негативний вплив на рідну мову, спричиняючи її поступову деградацію через надмірне використання іншомовних елементів. Особливо це стосується молоді, яка часто нехтує нормами рідної мови, замінюючи їх запозиченнями або кальками з англійської чи інших мов. Така тенденція може призводити до втрати лексичного й граматичного багатства мови, що негативно впливає на її розвиток [52].

Соціальні наслідки мовного змішування також включають можливість виникнення комунікативних бар’єрів. Коли в одному тексті поєднуються різні мови, це може ускладнювати розуміння для тих, хто не володіє обома мовами. Наприклад, у багатомовних групах у соціальних мережах часто виникають ситуації, коли певна частина учасників почувається виключеною через мовні труднощі. Це може призводити до фрагментації спільнот і навіть створювати конфлікти у мультикультурних контекстах, де мовні питання є чутливими. У

той же час мовне змішування стимулює міжкультурний діалог, адже люди мають можливість ознайомлюватися з елементами інших мов і культур у невимушеному середовищі. Наприклад, у міжнародних спільнотах, таких як Reddit чи Discord, користувачі часто пояснюють значення певних слів чи висловів, сприяючи взаємному збагаченню культур [15].

Отже, мовне змішування у соціальних мережах є складним і багатогранним явищем, яке поєднує у собі позитивні аспекти, такі як збагачення культури, сприяння глобалізації та розвиток багатомовності, а також виклики, пов’язані з деградацією рідної мови та комунікативними труднощами. Це явище потребує подальшого дослідження, особливо у контексті збереження мовного розмаїття та забезпечення гармонійної взаємодії між культурами в умовах глобалізованого світу [20].

Таблиця 1.3

Виклики та особливості мовного змішування в цифрову епоху

Аспект	Опис	Приклад реалізації
Причини мовного змішування	Глобалізація, багатомовне середовище, необхідність адаптації до культурного контексту.	Користувачі в соцмережах використовують англійську та рідну мову одночасно для різних груп аудиторії.
Особливості мовного змішування	Чергування між мовами, використання запозичень, створення нових гібридних форм комунікації.	Перемикання між українською та російською у спілкуванні в цифрових месенджерах.
Виклики мовного змішування	Труднощі розуміння для монолінгвів, загроза втрати мовної ідентичності, складність автоматичної обробки.	Алгоритми машинного перекладу часто не враховують контекст змішаних мовних конструкцій.
Роль технологій у мовному змішуванні	Інструменти перекладу, автокорекція, підтримка code-switching у сучасних чат-ботах і асистентах.	Сервіси Google Translate чи DeepL покращують якість перекладів змішаних текстів, але ще є недоліки.

Висновки до розділу 1

Полілінгвальне середовище є невід'ємною складовою сучасного суспільства, що формується під впливом глобалізаційних процесів, міграції, розвитку міжнародних комунікацій та цифрових технологій. У розділі було досліджено основні соціолінгвістичні аспекти, пов'язані з функціонуванням мов у багатомовному середовищі, поняття білінгвізму та полілінгвізму, а також виклики й особливості мовного змішування в цифрову епоху. Це дало змогу визначити ключові принципи функціонування мов у сучасному багатокультурному суспільстві, їхній вплив на соціальну інтеграцію, культурну ідентичність та комунікативну взаємодію.

Аналіз поняття білінгвізму та полілінгвізму показав, що ці явища мають глибоке соціолінгвістичне значення. Білінгвізм, який передбачає володіння двома мовами, є основою для розвитку полілінгвізму — здатності функціонувати у багатомовному середовищі. Ці явища не обмежуються індивідуальним рівнем, але мають масштабний вплив на соціум, сприяючи збереженню мовного різноманіття та інтеграції культур. Було виявлено, що білінгвізм і полілінгвізм функціонують у рамках певних соціальних структур, виконуючи різноманітні функції — від комунікативної до когнітивної. Наприклад, у багатомовних суспільствах кожна мова може виконувати специфічну роль, як-от використання рідної мови у сімейному колі, а міжнародної — у професійних чи освітніх контекстах. Це дозволяє зберігати культурну ідентичність, водночас інтегруючись у глобалізований світ.

Одним із ключових аспектів, досліджених у розділі, є функції мов у суспільстві. Було встановлено, що мова є не лише інструментом комунікації, але й важливим чинником когнітивного розвитку, засобом самовираження та культурної спадщини. Комунікативна функція мови забезпечує соціальну взаємодію, дозволяючи індивідам обмінюватися інформацією, кооперуватися для досягнення спільних цілей та зберігати соціальні зв'язки. Когнітивна функція сприяє структуруванню знань і формуванню концептів, що є

важливим для освіти, науки та культурного розвитку. Мова також виконує роль збереження культурної спадщини, адже вона зберігає колективну пам'ять, традиції та історію спільнот. У полілінгвальних суспільствах ці функції взаємодіють, створюючи нові форми культурного та мовного синтезу.

Розглядаючи виклики мовного змішування в цифрову епоху, було виявлено, що процеси глобалізації значно вплинули на мовну взаємодію. Соціальні мережі, мобільні платформи та цифрові технології створили умови для інтенсивного мовного контакту, в якому code-switching (перемикання мовних кодів) і code-mixing (мовне змішування) стали типовими явищами. Такі форми мовної взаємодії дозволяють користувачам одночасно використовувати кілька мов для адаптації до різних аудиторій або вираження культурної ідентичності. Наприклад, у полілінгвальних спільнотах молодь часто використовує мовне змішування для створення нових форм комунікації, які відображають їхню культурну багатогранність. Водночас цифрове мовне змішування створює виклики для збереження мовного різноманіття, оскільки менш поширені мови ризикують втратити свою ідентичність через домінування міжнародних мов, таких як англійська.

Особливу увагу було приділено соціокультурним наслідкам полілінгвізму в умовах глобалізації. У полілінгвальних суспільствах мова стає важливим інструментом соціальної інтеграції та збереження культурної ідентичності. Наприклад, у таких країнах, як Канада чи Швейцарія, де офіційно визнаються кілька мов, політика багатомовності дозволяє підтримувати баланс між культурною ідентичністю та соціальною згуртованістю. У контексті цифрових технологій мовна політика набуває нового значення, адже забезпечення рівноправного доступу до цифрових ресурсів різними мовами стає важливим для соціальної інтеграції. Це особливо актуально для менш поширених мов, які потребують підтримки через створення онлайн-ресурсів, мовних додатків та платформ для збереження культурної спадщини.

У підсумку, розділ дозволив визначити, що соціолінгвістичні засади полілінгвального середовища є багатовимірним явищем, яке охоплює лінгвістичні, культурні та соціальні аспекти. Білінгвізм і полілінгвізм сприяють інтеграції культур, збереженню мовного різноманіття та розвитку багатомовної комунікації. Водночас мовне змішування в умовах цифрової епохи створює як нові можливості, так і виклики для підтримки культурної ідентичності та соціальної згуртованості. Для забезпечення гармонійного функціонування полілінгвальних суспільств важливо розробляти політику, спрямовану на підтримку мовного різноманіття, розвиток освіти у багатомовному контексті та інтеграцію цифрових технологій у процеси збереження мовної спадщини.

РОЗДІЛ 2. КОМП'ЮТЕРНА ЛІНГВІСТИКА ТА АНАЛІЗ РОБОТИ ВІРТУАЛЬНИХ АСИСТЕНТІВ

2.1. Стан розвитку комп'ютерної лінгвістики

Стан розвитку комп'ютерної лінгвістики характеризується інтенсивним впровадженням інновацій у технології обробки природної мови (NLP), які відкривають нові горизонти у взаємодії між людиною та комп'ютером. Сучасні досягнення в цій галузі включають розробку алгоритмів для автоматичного перекладу, аналізу текстів, синтезу мовлення, розпізнавання тексту та семантичного розуміння. Одним із ключових етапів у розвитку комп'ютерної лінгвістики стало впровадження моделей трансформерів, таких як BERT (Bidirectional Encoder Representations from Transformers), GPT (Generative Pre-trained Transformer) та інших. Ці моделі, побудовані на основі глибокого навчання, демонструють високу ефективність у таких завданнях, як пошук інформації, автоматичне резюмування та розв'язання лінгвістичних проблем, які раніше вважалися складними для комп'ютерних систем. Особливу увагу привертає GPT-4, яка забезпечує виняткову точність у генерації тексту та контекстуальному розумінні, що дозволяє застосовувати її у віртуальних асистентах, таких як Siri або Google Bard [22].

Автоматичний переклад є однією з найбільш популярних і затребуваних функцій NLP. Сучасні системи, як-от Google Translate, DeepL і Microsoft Translator, використовують нейронні мережі для обробки багатомовних даних. На відміну від статистичних методів, які використовувалися в минулому, сучасні системи здатні враховувати контекст і семантику, що дозволяє досягти більшої точності. Наприклад, DeepL отримав високу оцінку за переклад технічної документації, адже він краще розпізнає вузькоспеціалізовану термінологію порівняно з конкурентами. Однак обробка діалектів і регіональних варіантів мов все ще залишається викликом. Наприклад, у випадку з українською мовою системи перекладу демонструють обмеження в

роботі з діалектами, такими як гуцульський чи бойківський, що вимагає розробки локалізованих корпусів даних.

Серед інших досягнень варто виділити аналіз тональності тексту, який активно застосовується у бізнесі, маркетингу та соціальних дослідженнях. Наприклад, інструменти Sentiment Analysis дозволяють компаніям визначати настрої клієнтів на основі відгуків чи коментарів у соціальних мережах. Такі системи, як IBM Watson і Google Natural Language API, використовуються для аналізу великих обсягів текстів, допомагаючи виявляти тенденції або потенційні ризики. Проте обробка змішаних мовних даних, як-от текстів із суржиком чи code-switching, все ще є складною через нестачу відповідних даних для тренування моделей. Це створює необхідність у подальшій роботі над адаптацією існуючих рішень для багатомовного середовища [37].

Окремо слід відзначити досягнення у сфері синтезу мовлення (Text-to-Speech, TTS). Ці технології активно впроваджуються в освітніх, медичних та інформаційних системах. Платформи, такі як Amazon Polly та Google Cloud TTS, забезпечують генерацію природного мовлення, що використовується для створення аудіокниг, голосових інтерфейсів і систем для людей із вадами зору. Наприклад, для української мови якість синтезу значно покращилася, однак усе ще існують труднощі з відтворенням інтонації та наголосів у складних реченнях. Подальший розвиток у цьому напрямі залежить від доступності якісних локалізованих даних та інтеграції культурно-специфічних аспектів.

Нарешті, вагомим напрямом є розробка чат-ботів, які використовуються для автоматизації обслуговування клієнтів і надання інформації. Такі системи, як ChatGPT, здатні не лише відповідати на питання, але й розпізнавати тональність запиту, адаптуючи відповідь залежно від потреб користувача. Наприклад, у сфері електронної комерції чат-боти можуть допомагати клієнтам із вибором продуктів, водночас обробляючи багатомовні запити. Утім, для забезпечення ефективності таких систем необхідно враховувати специфіку мовних культур, оскільки неправильне тлумачення контексту може призводити до комунікативних проблем [53].

Сучасна комп'ютерна лінгвістика продемонструвала вражаючі досягнення, проте зберігаються виклики, пов'язані з локалізацією, адаптацією до багатомовності та інтеграцією культурних особливостей. Завданням найближчого майбутнього є розвиток адаптивних моделей, які здатні враховувати специфіку діалектів і змішаних мовних даних, забезпечуючи якість і точність у різних сферах застосування [23].

Використання штучного інтелекту для багатомовної комунікації

Сучасний розвиток комп'ютерної лінгвістики та використання штучного інтелекту (ШІ) суттєво змінили підходи до багатомовної комунікації, забезпечивши нові можливості для взаємодії в глобалізованому світі. ШІ є основою більшості сучасних систем обробки природної мови, таких як автоматичний переклад, голосові асистенти, чат-боти та аналіз тексту. Основним досягненням у цьому напрямку є використання трансформерів, таких як BERT і GPT, які забезпечують високий рівень точності й ефективності в аналізі текстів багатьма мовами. Наприклад, GPT-4 здатна не лише розпізнавати контекст, але й адаптувати відповіді залежно від мовних і культурних особливостей, що робить її незамінною для багатомовних середовищ. Ці можливості активно впроваджуються у різних галузях, зокрема в освіті, бізнесі, міжнародній торгівлі та дипломатії, де комунікація між представниками різних мовних груп є критично важливою. Застосування ШІ у багатомовній комунікації базується на використанні великих обсягів даних (big data), які навчають моделі розпізнавати та адаптувати різні мовні структури, стилі й навіть діалекти. Наприклад, Google Translate працює з понад сотнею мов, а DeepL забезпечує високоякісний переклад технічних текстів, використовуючи алгоритми адаптації до вузьких тематичних контекстів [24].

Використання штучного інтелекту для багатомовної комунікації вийшло за межі звичайного перекладу текстів. Голосові асистенти, такі як Alexa, Siri та Google Assistant, забезпечують комунікацію в режимі реального часу, розуміючи й адаптуючи мову користувача. Наприклад, ці системи здатні ідентифікувати мову, якою розмовляє користувач, і відповідати відповідною

мовою, забезпечуючи комфортність взаємодії. Однак навіть ці передові системи мають виклики, пов’язані з полілінгвальним контекстом. Одним із основних є недостатня точність у роботі з діалектами чи мовними змішуваннями (наприклад, code-switching). Такі явища, як суржик в Україні чи спангліш у США, залишаються складними для обробки навіть найсучаснішими моделями, оскільки вони поєднують елементи кількох мов у межах одного тексту чи розмови. Для вирішення цієї проблеми необхідне впровадження спеціалізованих локалізованих корпусів даних і алгоритмів, здатних розуміти змішані мовні структури. Крім того, штучний інтелект активно використовується в аналізі тональності текстів, що дозволяє визначати емоційний контекст повідомлень у багатомовних комунікаціях. Наприклад, у соціальних мережах такі інструменти допомагають компаніям аналізувати настрої клієнтів, а в політичних кампаніях — виявляти громадську думку на глобальному рівні. Застосування Sentiment Analysis для полілінгвальних текстів вимагає спеціального навчання моделей, щоб уникнути помилкових інтерпретацій у контекстах, де певні вислови можуть мати різне значення залежно від мови чи культури [9].

Окремо варто згадати розвиток автоматичного синтезу та розпізнавання мовлення. Системи Text-to-Speech (TTS) та Speech-to-Text (STT) активно впроваджуються для полегшення доступу до інформації для людей із вадами зору чи слуху, а також у сфері освіти та медіа. Наприклад, Google Cloud Text-to-Speech підтримує десятки мов і надає можливість створювати аудіоконтент із природним звучанням. Водночас технології розпізнавання мовлення, такі як Dragon NaturallySpeaking чи Microsoft Azure Speech, використовуються для автоматизації процесів створення текстів, зокрема транскрибування багатомовних розмов у реальному часі. Це особливо актуально для міжнародних конференцій, де мовні бар’єри можуть стати значною перешкодою для ефективної взаємодії учасників. Для подолання цих бар’єрів розробляються багатомовні системи синхронного перекладу, які інтегруються у відеоконференції, такі як Zoom чи Microsoft Teams. Ці технології

забезпечують миттєвий переклад мовлення на кілька мов одночасно, що дозволяє учасникам з різних країн ефективно комунікувати без необхідності залучення перекладачів.

Підсумовуючи, сучасний стан розвитку комп’ютерної лінгвістики та використання штучного інтелекту для багатомовної комунікації відкриває величезні можливості для подолання мовних бар’єрів у глобалізованому світі. Однак залишається низка викликів, таких як адаптація до діалектів, мовного змішування та культурних особливостей. Подальший розвиток цієї галузі залежить від удосконалення алгоритмів і моделей, які здатні враховувати локальні особливості й забезпечувати точність у найскладніших контекстах багатомовного спілкування [28].

Таблиця 2.1

Стан розвитку комп'ютерної лінгвістики

Напрям	Досягнення	Виклики
Автоматичний переклад	Нейронні мережі покращують точність перекладів, враховуючи контекст (наприклад, DeepL, Google Translate).	Обробка діалектів та регіональних мовних варіантів.
Аналіз тональності текстів	Визначення настроїв клієнтів для маркетингу та соціальних досліджень (Sentiment Analysis).	Труднощі у визначенні тональності у змішаних мовних текстах.
Синтез мовлення (TTS)	Природне звучання голосу в системах Amazon Polly, Google Cloud TTS.	Інтонація та наголоси в складних реченнях.
Чат-боти та віртуальні асистенти	Здатність розпізнавати контекст і адаптувати відповіді, приклад ChatGPT.	Розуміння діалектів, шумів і мовного змішування.

2.2. Віртуальні асистенти: огляд сучасних здобутків

Еволюція віртуальних асистентів є відображенням технологічного прогресу та зміни способів взаємодії людей з комп’ютерами. Від своїх початків, представлених текстовими інтерфейсами, до сучасних голосових помічників, таких як Siri, Alexa та Google Assistant, ці системи стали важливим

елементом нашого повсякденного життя. На перших етапах розвитку віртуальних асистентів користувачі взаємодіяли з ними виключно через текстовий формат. Наприклад, системи командного рядка в операційних системах, такі як MS-DOS чи UNIX, вимагали від користувача знань спеціальних команд для виконання операцій. Ці текстові інтерфейси мали обмежений функціонал, були складними у використанні для непідготовлених користувачів і не забезпечували гнучкості в розумінні людської мови. Згодом з’явилися більш досконалі текстові системи, зокрема чат-боти, які здатні імітувати діалог із користувачем. ELIZA, створена в 1966 році, є одним із перших прикладів чат-бота, який використовував прості алгоритми для симуляції психотерапевта. Хоча можливості ELIZA були обмеженими, вона продемонструвала потенціал обробки природної мови і стала основою для подальших розробок [3].

Справжній прорив у розвитку віртуальних асистентів стався із запровадженням голосових інтерфейсів, які дозволили перейти від текстових до більш природних способів взаємодії. Голосові асистенти, такі як Siri (2011), Google Assistant (2016) та Alexa (2014), використовують технології розпізнавання мовлення, обробки природної мови (NLP) та синтезу мовлення. Завдяки глибоким нейронним мережам ці системи можуть не лише розуміти запити користувачів, але й враховувати контекст, що робить їх більш адаптивними та зручними. Наприклад, голосові інтерфейси дозволяють виконувати різні завдання, такі як встановлення нагадувань, пошук інформації в Інтернеті, управління розумним домом чи навіть замовлення товарів. Важливим кроком стало інтегрування цих асистентів у мобільні пристрої та розумні колонки, що значно розширило їхню аудиторію. Amazon Alexa, наприклад, стала популярною завдяки своїй інтеграції з пристроями "розумного дому", дозволяючи користувачам керувати освітленням, температурою чи побутовими приладами за допомогою голосу. У той же час Siri стала невід'ємною частиною екосистеми Apple, надаючи користувачам зручний доступ до функцій пристроїв на базі iOS [7].

Сучасні голосові асистенти продовжують вдосконалюватися, інтегруючи елементи штучного інтелекту та машинного навчання для покращення розуміння мови та персоналізації взаємодії. Одним із ключових напрямків розвитку є багатомовність, яка дозволяє асистентам працювати з користувачами, що говорять різними мовами. Наприклад, Google Assistant підтримує понад 40 мов, а Alexa розпізнає кілька мов одночасно, що є важливим для полілінгвальних спільнот. Іншим аспектом еволюції є врахування культурних особливостей під час виконання запитів. Асистенти тепер здатні враховувати регіональні відмінності у мові, що робить їх більш універсальними. Наприклад, Alexa пропонує спеціальні "вміння" (skills), які адаптовані до локальних потреб користувачів. Разом із цим удосконалюється й контекстуальне розуміння запитів: сучасні системи можуть аналізувати попередні взаємодії з користувачем, що дозволяє створювати більш персоналізовані відповіді. Незважаючи на значний прогрес, виклики все ще існують. Наприклад, робота з діалектами, шумовим фоном або змішаними мовами залишається складним завданням для голосових асистентів. У підсумку, еволюція віртуальних асистентів від текстових до голосових інтерфейсів стала визначальним етапом у розвитку технологій взаємодії людини та машини. Ці системи продовжують удосконалюватися, стаючи невід'ємною частиною сучасного цифрового середовища [1].

Можливості адаптації віртуальних асистентів до багатомовних середовищ стали однією з ключових тем у сучасній комп’ютерній лінгвістиці та розробці технологій штучного інтелекту. У глобалізованому світі, де багатомовність стає нормою для бізнесу, освіти та соціальних взаємодій, асистенти, здатні ефективно працювати з різними мовами, забезпечують новий рівень взаємодії між користувачами та машинами. Одним із головних досягнень у цій сфері є впровадження алгоритмів, які дозволяють асистентам розпізнавати та обробляти кілька мов одночасно. Наприклад, Google Assistant та Amazon Alexa здатні визначати мову користувача за контекстом і перемикатися між кількома мовами без необхідності вручну змінювати

налаштування. Це особливо важливо для полілінгвальних сімей або спільнот, де люди можуть одночасно використовувати дві чи більше мов у повсякденному спілкуванні. Інтеграція таких функцій стала можливою завдяки розвитку трансформерних моделей, які забезпечують не лише синтаксичний, але й семантичний аналіз тексту, враховуючи лексичні й стилістичні особливості кожної мови [14].

Важливим напрямком адаптації є робота з діалектами та мовними варіаціями. Багато мов мають значні регіональні відмінності, які можуть впливати на розуміння асистентами запитів користувачів. Наприклад, у межах англійської мови існують значні розбіжності між британським, американським, австралійським і індійським варіантами, що стосується як вимови, так і словника. Віртуальні асистенти, такі як Siri, поступово навчаються враховувати ці відмінності, забезпечуючи точні відповіді на запити залежно від географічного розташування користувача. Аналогічно, для української мови викликом залишається робота з діалектами, наприклад, гуцульським чи бойківським, які суттєво відрізняються від стандартної мови. Для вирішення цієї проблеми розробники використовують локалізовані корпуси текстів, що дозволяють тренувати моделі на матеріалах, які точно відображають мовні особливості. Крім того, технології обробки природної мови (NLP) використовують алгоритми контекстуального аналізу, що дозволяє розпізнавати культурно специфічні елементи мови, такі як ідіоми, фразеологізми чи сленг [15].

Ще одним важливим аспектом адаптації є персоналізація віртуальних асистентів залежно від мовних потреб користувача. Сучасні системи, такі як Google Assistant або Microsoft Cortana, можуть запам’ятовувати мовні уподобання користувачів і адаптувати відповіді до їхнього стилю комунікації. Наприклад, якщо користувач переважно спілкується англійською, але іноді використовує французьку, асистент може автоматично інтегрувати обидві мови у свої відповіді. Це особливо корисно у багатомовних середовищах, таких як міжнародні компанії або освітні установи, де працівники чи студенти

використовують кілька мов у щоденній діяльності. Водночас віртуальні асистенти можуть враховувати регіональні відмінності навіть у межах однієї мови, що особливо актуально для країн із багатомовним населенням, таких як Канада чи Швейцарія.

Розвиток мультимодальних систем також відкриває нові можливості для адаптації асистентів. Це підхід, який передбачає інтеграцію текстового, голосового та візуального інтерфейсів, що дозволяє користувачам взаємодіяти з асистентами за допомогою кількох каналів комунікації. Наприклад, системи на базі GPT-4 можуть одночасно аналізувати текстові й голосові запити, адаптуючи відповіді до мови, яка є зручною для користувача. Важливим аспектом мультимодальних систем є здатність обробляти змішані мовні запити, наприклад, коли користувач починає запит українською, а потім переключається на англійську. Такі технології вже тестуються в міжнародних компаніях, де працівники з різних країн співпрацюють у мультикультурному середовищі [1].

Однак, незважаючи на значний прогрес, адаптація віртуальних асистентів до багатомовних середовищ усе ще має виклики. Одним із головних є проблема контекстуального розуміння запитів у випадках, коли користувач використовує мовне змішування або рідковживані слова. Наприклад, у полілінгвальних спільнотах, таких як ті, що використовують спангліш чи суржик, асистенти часто не можуть адекватно обробляти запити через брак даних для тренування моделей. Для вирішення цієї проблеми розробники працюють над створенням нових корпусів текстів, які враховують реальні сценарії використання мови. У підсумку, можливості адаптації віртуальних асистентів до багатомовних середовищ продовжують розширюватися завдяки розвитку технологій NLP, трансформерів і мультимодальних інтерфейсів. Ці системи стають дедалі ефективнішими у подоланні мовних бар’єрів, забезпечуючи комфортну взаємодію між користувачами та машинами у глобалізованому світі [9].

У сучасному світі віртуальні асистенти стають важливими інструментами для ефективної взаємодії людини з технологіями, зокрема в багатомовному середовищі. Одним із найсучасніших досягнень у цьому напрямку є Google Bard, який став важливим етапом у розвитку штучного інтелекту та машинного навчання. Він орієнтований на використання мовних моделей для забезпечення максимально точного і контекстуально релевантного спілкування. Google Bard здатен не лише виконувати прості команди, такі як управління розумними пристроями чи пошук інформації, але й надавати багатомовні відповіді, адаптуючись до запитів користувачів різними мовами.

Попри значний прогрес у розробці таких асистентів, зокрема Google Bard, є багато аспектів, які потребують подальшого вдосконалення. Один з головних викликів, з якими стикаються віртуальні асистенти, це робота в полілінгвальних середовищах, де одночасно використовуються кілька мов або діалектів. Google Bard активно працює з багатьма мовами, але виникають труднощі при перемиканні між ними, особливо коли користувач використовує змішані мовні конструкції або діалекти. Наприклад, у багатьох країнах, де співіснують кілька мов, часто зустрічаються змішані конструкції — code-switching — коли одна частина речення подається однією мовою, а інша частина — іншою. Такі ситуації є досить складними для віртуальних асистентів, оскільки вони мають не лише ідентифікувати, яка мова використовується в кожному конкретному випадку, але й коректно зрозуміти зміст запиту [6].

Особливою проблемою для віртуальних асистентів є адаптація до регіональних варіантів мов, таких як діалекти або жаргони. Google Bard, хоч і демонструє високий рівень точності в роботі з основними мовами, не завжди коректно працює з діалектами або словами, що не є стандартними для основної мови. Наприклад, у випадку з українською мовою, Bard може мати труднощі в розпізнаванні деяких регіональних варіантів, таких як бойківський чи гуцульський діалекти, що є важливими для коректної обробки запитів. Такі

обмеження є перешкодою для досягнення повної точності в багатомовному середовищі, де кожен регіон має свої мовні особливості.

Незважаючи на ці виклики, Google Bard продовжує вдосконалювати свою здатність працювати з багатьма мовами завдяки інтеграції нових алгоритмів штучного інтелекту. Одним з основних аспектів цього процесу є розширення можливостей моделі щодо врахування культурних та соціальних контекстів. Оскільки мовні варіанти та їх використання залежать не лише від граматики, але й від культурних і соціальних факторів, Bard прагне адаптувати свої алгоритми таким чином, щоб вони могли ефективно обробляти мовні варіанти в реальних ситуаціях. Наприклад, Bard може розпізнавати і враховувати сленг або місцеві вирази, що є характерними для певних груп людей, завдяки чому взаємодія з користувачами стає більш природною та зручною [7].

Подальше вдосконалення віртуальних асистентів, зокрема Google Bard, буде зосереджено на покращенні точності роботи з багатомовними запитами, здатності адаптуватися до діалектів та різних мовних контекстів, а також на покращенні розпізнавання змішаних мовних конструкцій. Це відкриває перспективи для більш ефективної роботи асистентів в умовах глобалізації, коли люди з різних культур та мовних груп можуть взаємодіяти з технологіями, не зустрічаючи мовних бар’єрів. У майбутньому, з урахуванням швидкого розвитку штучного інтелекту, такі віртуальні асистенти, як Google Bard, можуть стати невід’ємною частиною багатомовних спільнот, забезпечуючи безперешкодну комунікацію та інтеграцію різних мовних і культурних контекстів [24].

Таблиця 2.2

Віртуальні асистенти: огляд сучасних здобутків

Асистент	Мовна підтримка	Особливості функціонування в полілінгвальному середовищі	Обмеження в багатомовному середовищі
Google Assistant	Понад 40 мов, з підтримкою багатомовних запитів	Можливість переключення між мовами в одному запиті, робота з code-switching	Не завжди точно розпізнає перемикання між мовами, проблеми з рідковживаними діалектами
Siri	Основні мови, з підтримкою перемикання між англійською та іншими мовами	Можливість перемикання між англійською та кількома іншими мовами	Може мати труднощі з розумінням змішаних мовних запитів, проблеми з діалектами
Amazon Alexa	Підтримка кількох мов одночасно, але з обмеженнями в багатомовних контекстах	Адаптація до кількох мов в одному запиті, але складність у контекстах змішаних мов	Труднощі з адаптацією до складних мовних змін, зокрема в контекстах code-switching
Microsoft Cortana	Має багатомовну підтримку, але функціональність обмежена в залежності від регіону	Підтримка кількох мов для виконання базових завдань, але зі складнощами в складних багатомовних запитах	Проблеми з точністю перекладу між мовами, обмежена багатомовна підтримка
Google Bard	Підтримка кількох мов, з проблемами у роботі з діалектами та code-switching	Проблеми з розпізнаванням змішаних мовних запитів та діалектів, необхідність покращення алгоритмів	Складнощі з адаптацією до рідковживаних діалектів, нестабільна робота з мовним змішуванням

2.3 Методи аналізу роботи віртуальних помічників

Лінгвістичний аналіз алгоритмів обробки природної мови у Google Bard демонструє ключові аспекти роботи віртуального асистента в полілінгвальному середовищі, розкриваючи його сильні сторони та обмеження. У ході нашого дослідження було проведено серію тестів, спрямованих на оцінку здатності системи обробляти запити різними мовами, включаючи українську, англійську та їх діалектні варіанти. Аналіз включав

перевірку точності відповідей, розуміння контексту, адаптації до змішаних мовних запитів та роботи з регіональними особливостями мови. Основною метою тестування було визначити, наскільки ефективно Google Bard справляється із завданнями у багатомовному середовищі, де мови можуть взаємодіяти, а змішування мовних структур є поширеним явищем [9].

Перший етап тестування полягав у перевірці роботи системи з запитами стандартною українською мовою. Для цього були сформовані завдання, що включали формальні та неформальні конструкції, питання загального характеру, а також складні синтаксичні структури. Google Bard показав високу точність у роботі з загальними запитами, такими як пошук інформації чи надання відповідей на освітні запити. Наприклад, на питання «Що таке синтаксис у лінгвістиці?» система надала чітке й коректне пояснення з врахуванням термінології, що відповідає науковому контексту. У формальних конструкціях Bard демонструє здатність зберігати стиль і надавати релевантну інформацію. Однак у випадках з більш складними синтаксичними конструкціями, такими як запитання з вкладеними реченнями або багатокомпонентними структурами, система іноді генерувала узагальнені або не зовсім точні відповіді. Це свідчить про потребу вдосконалення алгоритмів для більш глибокого розуміння складних лінгвістичних структур [1].

Другий етап тестування був спрямований на аналіз роботи системи з діалектами. Було обрано низку характерних фраз, що використовуються в гуцульському та бойківському діалектах, для перевірки здатності Bard ідентифікувати та інтерпретувати їх. Результати показали, що система має суттєві обмеження у роботі з діалектами. Наприклад, на запит «Що таке "жєнджура"?» (гуцульське слово, що означає "сукня"), Bard не зміг надати релевантної відповіді. Це обумовлено тим, що алгоритми обробки природної мови Google Bard базуються на стандартизованих корпусах текстів і не охоплюють локалізовані діалекти, які є важливою частиною культурного та мовного контексту. Водночас система продемонструвала кращі результати у випадках, коли діалектні слова супроводжувалися поясненнями чи були

частиною більш широкого контексту. Наприклад, у запиті «Гуцульський костюм включає жєнджуру, яка є традиційною сукнею», Bard зміг коректно інтерпретувати загальний зміст, хоча й не надав точного визначення окремого слова [15].

Особливу увагу було приділено аналізу роботи Google Bard із запитами, що містять мовне змішування, тобто чергування української та англійської мов у межах одного запиту. Це явище є характерним для багатомовних користувачів, особливо в цифровому середовищі, де code-switching часто використовується для спрощення комунікації або вираження емоцій. Наприклад, у запиті «Поясни, що таке syntax у контексті української мови», система надала часткову відповідь, орієнтуючись переважно на англійський термін, але не врахувала специфіку українського контексту. Це свідчить про те, що алгоритми Bard мають складнощі з інтеграцією змішаних мовних запитів, оскільки вони недостатньо адаптовані для розуміння взаємодії між різними мовами. Утім, якщо запит був більш структурованим, наприклад, «Поясни syntax, але з точки зору української лінгвістики», Bard зміг надати більш релевантну відповідь. Це підкреслює необхідність розширення мовних моделей для врахування контексту полілінгвізму [4].

Також було проведено аналіз роботи Bard з текстами, що містять культурно специфічні елементи, такі як ідіоми, фразеологізми та сленгові вирази. Наприклад, на запит «Що означає фраза "Де Макар телят пас" у значенні українського фольклору?» система змогла надати часткове пояснення, але її відповідь була загальною і не враховувала культурного контексту. Утім, під час роботи з англійськими ідіомами Bard продемонстрував вищу точність, оскільки ці елементи краще представлені у його навчальних корпусах. Це вказує на дисбаланс у підтримці різних мов і культур, що може бути вирішено шляхом розширення корпусів даних для малопоширених мов і культур.

У результаті аналізу було встановлено, що Google Bard є ефективним інструментом для роботи зі стандартними мовними запитами, однак його можливості значно знижуються у випадках роботи з діалектами, змішаними

мовними запитами та культурно специфічними елементами. Для підвищення ефективності системи у полілінгвальному середовищі рекомендується впровадження локалізованих корпусів текстів, розробка адаптивних алгоритмів для роботи з мовним змішуванням, а також покращення контекстуального аналізу. Ці вдосконалення дозволять забезпечити точність і релевантність відповідей, що є критично важливим для подальшого розвитку віртуальних асистентів у багатомовному світі [9].

Тестування ефективності у багатомовних сценаріях

Тестування ефективності роботи Google Bard у багатомовних сценаріях мало на меті визначення його здатності забезпечувати якісну комунікацію різними мовами, включаючи українську, англійську та змішані мовні структури. У межах практичного дослідження ми провели серію тестів, спрямованих на оцінку точності відповідей, розуміння контексту та адаптації до багатомовного середовища. Завдання охоплювали запити різного рівня складності: від стандартних запитань українською до роботи з діалектами та мовним змішуванням (code-switching). Особливу увагу приділяли аналізу роботи Bard із запитами, що включають культурно специфічні елементи, такі як ідіоми, фразеологізми та сленг [2].

Перший етап тестування стосувався стандартних запитів українською мовою. Наприклад, система без проблем виконувала завдання на кшталт «Поясни, що таке синтаксис у лінгвістиці» або «Яка структура речення в англійській мові?» Відповіді були точними, структурованими й відповідали академічному стилю. Однак складніші конструкції, наприклад запити з багаторівневими вкладеннями, іноді генерували узагальнені відповіді. Наприклад, на питання «Поясни взаємозв'язок морфології й синтаксису в контексті української мови» Bard надав загальне пояснення без деталізації. Це свідчить про потребу в удосконаленні здатності алгоритмів до аналізу складних лінгвістичних концептів.

Другий етап зосереджувався на роботі з діалектами української мови. Було протестовано запити з використанням гуцульських і бойківських

діалектних слів, таких як «жєнджура» (сукня) або «мочеря» (калюжа). Результати показали, що Bard мав суттєві труднощі у розпізнаванні таких термінів. Наприклад, на запит «Що означає слово жєнджура?» система не змогла надати релевантної відповіді. Однак, коли ці слова були частиною ширшого контексту, наприклад «Гуцульський костюм включає жєнджуру», Bard демонстрував краще розуміння загального сенсу, хоча точного визначення терміну не було. Це підкреслює потребу в інтеграції локалізованих корпусів текстів для роботи з регіональними мовними варіантами [7].

Третій етап включав тестування запитів зі змішаними мовними структурами. Наприклад, запит «Поясни syntax у контексті української лінгвістики» виявив, що Bard переважно фокусується на англійській частині запиту, ігноруючи український контекст. У більш структурованих запитах, таких як «Поясни, що таке syntax, але в контексті української мови», результати були кращими, оскільки система змогла врахувати обидві мовні частини. Це свідчить про те, що алгоритми Google Bard потребують додаткового вдосконалення для роботи із code-switching, що є поширеним явищем у багатомовних спільнотах [3].

Крім того, було протестовано здатність Bard працювати з культурно специфічними елементами мови, зокрема ідіомами та фразеологізмами. Наприклад, на запит «Що означає фраза "Де Макар телят пас"?» система надала часткову відповідь, однак не врахувала повного контексту цього фразеологізму. Аналогічно, при роботі з англійськими ідіомами, такими як «Break the ice», Bard продемонстрував кращу точність, оскільки ці елементи краще представлені в його навчальних корпусах [9].

У підсумку, результати тестування свідчать, що Google Bard демонструє високу ефективність у роботі зі стандартними мовними запитами, однак має обмеження при роботі з діалектами, змішаними мовними структурами та культурно специфічними елементами. Для покращення ефективності у багатомовних сценаріях необхідне розширення корпусів даних, зокрема для діалектів і code-switching, а також впровадження адаптивних алгоритмів, які

враховують локальні особливості. Такі вдосконалення дозволять забезпечити більш точну і релевантну роботу віртуальних асистентів у полілінгвальному середовищі [11].

Таблиця 2.3

Методи аналізу роботи віртуальних помічників

Метод аналізу	Опис методу	Приклад реалізації
Аналіз мовних моделей	Оцінка ефективності мовних моделей на основі алгоритмів глибокого навчання, таких як GPT, BERT.	Тестування GPT для генерації тексту на кількох мовах з аналізом точності відповідей.
Аналіз точності перекладу та адаптації до багатомовних запитів	Перевірка якості перекладу та здатності адаптації асистента до багатомовних запитів.	Перевірка точності перекладу між англійською та іншими мовами, з фокусом на ідіоми та культурні контексти.
Тестування на діалектах та варіантах мов	Оцінка роботи асистентів з різними діалектами і варіантами мов, що впливають на розпізнавання мови.	Аналіз здатності Google Bard розпізнавати бойківський або гуцульський діалект в українській мові.
Тестування здатності до обробки змішаних мовних запитів	Тестування здатності віртуального асистента працювати з кодовими змішуваннями (code-switching) в одному запиті.	Перевірка ефективності розпізнавання змішаних запитів (наприклад, українсько-англійських) у Siri.
Тестування адаптації до культурних контекстів	Оцінка здатності асистента враховувати культурні відмінності в мовних запитах та відповідях.	Тестування культурних адаптацій у системах Alexa, щоб врахувати відмінності у використанні слів і фраз в різних країнах.
Аналіз точності синтезу мовлення (TTS)	Аналіз якості синтезованого мовлення, правильності інтонації та наголосів у відповідях асистента.	Оцінка синтезу голосу Siri, зокрема правильності вимови складних або рідковживаних слів.
Методи зворотного зв'язку та покращення алгоритмів	Збір зворотного зв'язку від користувачів для покращення алгоритмів обробки мовних запитів.	Використання зворотного зв'язку користувачів для вдосконалення систем перекладу в Google Translate.

2.4. Результати лінгвістичного аналізу

2.4.1. Робота Google Bard у полілінгвальному середовищі

Аналіз 10 кейсів роботи Google Bard із введеними текстами

1. Запит: "Перефразуйте цей текст з використанням нормативної української мови."

- **Текст:** Уривок із книги Тараса Прохаська “Непрості” з використанням діалектизмів.
- **Результат:** Google Bard успішно замінив більшість діалектизмів нормативними відповідниками, зберігши загальну структуру тексту.
- **Помилка:** Деякі слова, такі як "ґринджоли" (санки), не були ідентифіковані, тому залишилися без змін, що порушило смислову цілісність тексту.
- **Рекомендація:** Інтеграція спеціалізованих корпусів українських діалектів для більш точного розуміння.

2. Запит: "Що означає це слово в контексті?" (слово "жєнджура").

- **Текст:** Гуцульський діалектний уривок із запитанням про значення слова "жєнджура."
- **Результат:** Bard не зміг дати точного визначення, натомість запропонував можливе тлумачення як "предмет одягу," але без уточнень.
- **Помилка:** Відсутність точного перекладу через брак локалізованих корпусів.
- **Рекомендація:** Розширення алгоритмів через навчання на текстах, що містять рідковживані діалектизми.

3. Запит: "Як би ви переклали цей уривок на англійську?"

- **Текст:** Той самий текст із діалектизмами.

- **Результат:** Переклад зроблено на англійську, але діалектизми були пропущені або замінені узагальненими термінами.
- **Помилка:** Втрата унікального культурного контексту тексту через адаптацію до стандартної мови.
- **Рекомендація:** Додавання пояснень у переклад для збереження культурних особливостей.

4. Запит: "Що означає слово "мочеря" в контексті гуцульської мови?"

- **Результат:** Bard не зміг дати жодної відповіді, оскільки слово не було знайдено у його базах.
- **Помилка:** Відсутність ресурсів для роботи з вузькоспеціалізованими словами.
- **Рекомендація:** Інтеграція діалектних словників до алгоритмів обробки природної мови.

5. Запит: "Яке значення має фразеологізм 'де Макар телят пас'?"

- **Результат:** Часткове пояснення надано, але воно було надто загальним і не враховувало історичного чи культурного контексту.
- **Помилка:** Неповне розуміння фразеологізмів, відсутність деталізації.
- **Рекомендація:** Навчання моделей на корпусах із фразеологізмами та сталими виразами.

6. Запит: "Перекладіть на англійську: 'Дідо говорив старослов’янською, а я не розумів.'"

- **Результат:** Переклад виконано коректно, але "старослов’янською" було перекладено як "Old Slavic language" без уточнень, що це архаїчна форма мови.
- **Помилка:** Втрата культурного контексту через загальний переклад терміну.

- **Рекомендація:** Розширення моделей для врахування культурно значущих мовних термінів.

7. Запит: "Як би ви спростили цей текст для дітей?" (Текст із архаїзмами)

- **Результат:** Текст спрощено, але архаїзми були замінені сучасними словами без збереження оригінального змісту.
- **Помилка:** Недостатній рівень контекстуального розуміння архаїзмів.
- **Рекомендація:** Навчання моделей на текстах, адаптованих для різних вікових категорій.

8. Запит: "Розпізнайте текст і дайте стислий переказ." (Уривок із неологізмами)

- **Результат:** Bard успішно створив переказ, але неологізми були опущені або замінені термінами з інших сфер.
- **Помилка:** Неврахування впливу неологізмів на смислове навантаження тексту.
- **Рекомендація:** Додавання спеціалізованих корпусів текстів із новітньою лексикою.

9. Запит: "Чи є слово 'ґринджоли' сучасним і зрозумілим для української мови?"

- **Результат:** Bard підтвердив, що це слово архаїчне, але не зміг пояснити його походження.
- **Помилка:** Неповний аналіз слова та його значення.
- **Рекомендація:** Інтеграція етимологічних баз до алгоритмів.

10. Запит: "Як пояснити слово 'мудрагель' у сучасному контексті?"

- **Результат:** Bard описав значення як "людина, яка багато розмірковує," але не врахував його іронічного забарвлення.
- **Помилка:** Відсутність розуміння емоційного контексту слів.

- **Рекомендація:** Розвиток моделей для аналізу тональності та емоційних відтінків у текстах.

Аналіз роботи Google Bard у різних мовних контекстах включав тестування системи на здатність розпізнавати, аналізувати та інтерпретувати діалекти, фразеологізми, архаїзми, неологізми та мовне змішування. Результати тестів дали змогу виявити сильні й слабкі сторони алгоритмів обробки природної мови та запропонувати рекомендації для їх удосконалення. Було протестовано 10 кейсів із різними запитами, які детально аналізували роботу системи.

Перший кейс стосувався запиту "Перефразуйте цей текст з використанням нормативної української мови." Google Bard показав гарні результати у заміні більшості діалектизмів на нормативні еквіваленти. Наприклад, у випадку з текстом, який містив слово "ґринджоли" (санки), система залишила його без змін через невпізнання. Це призвело до втрати смислової цілісності тексту, адже діалектизм залишився без пояснення чи адаптації. Рекомендовано інтегрувати корпуси з локалізованими текстами, що містять діалекти, для підвищення точності.

Другий кейс включав аналіз роботи з діалектним словом "жєнджура," запитом було визначити його значення у контексті гуцульського діалекту. Bard не зміг надати точного визначення, замість цього запропонував дуже загальну відповідь: "можливо, предмет одягу." Це демонструє недостатню обізнаність системи у роботі з діалектами. Оскільки діалектизми є важливою частиною культурної ідентичності, їх коректне розуміння має значення для забезпечення релевантних відповідей. Для цього потрібно розширити корпуси, орієнтуючись на діалекти та регіональні варіанти мови.

Третій кейс передбачав переклад уривка тексту, що містив діалектизми, на англійську мову. Bard виконав переклад на базовому рівні, але діалектизми або були пропущені, або замінені узагальненими термінами, що призвело до втрати унікального культурного контексту тексту. Наприклад, слово

"ґринджоли" було взагалі проігноровано, що є серйозним недоліком при роботі із багатомовними середовищами. Рекомендації включають використання анотацій або приміток у перекладах для пояснення таких елементів.

Четвертий кейс полягав у перевірці слова "мочеря," що означає "калюжа" у гуцульському діалекті. Bard не зміг надати жодної відповіді, оскільки слово не було знайдено у його базах даних. Це демонструє серйозний брак ресурсів для роботи з вузькоспеціалізованими словами. Включення діалектних словників до навчальних моделей може суттєво покращити результати.

П'ятий кейс стосувався роботи з фразеологізмом "де Макар телят пас." Bard надав часткове пояснення, яке не враховувало історичного чи культурного контексту. Наприклад, замість детального аналізу значення фразеологізму було надано надто загальну відповідь. Це вказує на слабкість у розумінні фразеологічних зворотів, особливо якщо вони є культурно специфічними. Додавання спеціальних корпусів фразеологізмів значно підвищило б ефективність у таких завданнях.

Шостий кейс передбачав переклад речення "Дідо говорив старослов’янською, а я не розумів." Переклад було виконано коректно, однак "старослов’янською" Bard переклав як "Old Slavic language" без уточнень про специфіку мови. Це створило ризик втрати культурного контексту, який є важливим для адекватного сприйняття тексту. Алгоритми мають враховувати значення таких термінів і додавати пояснення або коментарі, коли це необхідно.

Сьомий кейс перевіряв здатність Bard спрощувати текст для дітей. Запит стосувався уривка з архаїзмами, які система замінила сучасними словами. Попри це, деякі частини тексту втратили оригінальний сенс, адже спрощення було виконано без урахування смислового навантаження архаїзмів. Для ефективнішої адаптації до потреб різних вікових груп необхідно враховувати контекст і мету використання тексту.

Восьмий кейс перевіряв здатність системи створювати стислий переказ тексту, що містив неологізми. Bard успішно виконав переказ, але неологізми

були або опущені, або замінені загальними словами, що призвело до втрати частини інформації. Наприклад, термін, створений автором тексту, був перекладений як "нова концепція" без уточнень. Включення сучасних словників із новітньою лексикою може вирішити цю проблему.

Дев’ятий кейс стосувався визначення сучасності слова "ґринджоли." Bard визнав його архаїчним, але не зміг пояснити походження. Це вказує на необхідність включення етимологічних баз у моделі, щоб забезпечити повніше пояснення.

Останній, десятий кейс, аналізував пояснення слова "мудрагель." Bard описав його як "людина, яка багато розмірковує," але не врахував іронічного відтінку цього слова, який є важливим у багатьох контекстах. Для врахування тональності та емоційних відтінків потрібне вдосконалення алгоритмів аналізу.

У підсумку тестування показало, що Google Bard ефективно працює зі стандартними мовними запитами, але має суттєві обмеження у випадках із діалектами, фразеологізмами, архаїзмами та змішаними мовними структурами. Для покращення результатів необхідно розширити корпуси даних, інтегрувати діалектні словники та враховувати емоційний контекст текстів. Ці вдосконалення дозволять підвищити якість роботи віртуального асистента в полілінгвальних середовищах і забезпечити більш точні та релевантні відповіді.

Наступні 5 кейсів: Результати лінгвістичного аналізу

1. Запит: "Що означає слово 'крисаня' в контексті українського фольклору?"

- **Результат:** Google Bard припустив, що це може бути форма звертання або назва тварини, але точного визначення не надав.
- **Помилка:** Система не врахувала контекст українського фольклору, де "крисаня" є архаїзмом, який означає "шапка з хутра."

- **Рекомендація:** Додати корпуси текстів, що включають фольклорні джерела, для підвищення точності розуміння архаїчних термінів.

2. Запит: "Чи можна слово 'партач' у діалектному контексті інтерпретувати як 'майстер'? "

- **Результат:** Bard визначив, що "партач" може означати людину, яка щось робить неправильно. Це загальне значення терміну, але воно не враховує діалектного контексту, де "партач" може бути майстром із середньою кваліфікацією.
- **Помилка:** Відсутність діалектного розрізнення.
- **Рекомендація:** Інтеграція діалектних варіантів значень для аналізу полісемічних слів.

3. Запит: "Перекладіть слово 'гайдамака' на англійську."

- **Результат:** Bard переклав "гайдамака" як "rebel" (повстанець), що є частково правильним, але не враховує історичного й культурного контексту української історії.
- **Помилка:** Втрата специфіки терміну, пов’язаної з національною боротьбою XVIII століття.
- **Рекомендація:** Включити до моделей історичні й культурно значущі терміни з докладним поясненням контексту.

4. Запит: "Що таке 'макітра' і як це слово використовують у сучасній мові?"

- **Результат:** Bard надав правильне визначення як "глиняна посудина для приготування страв," але не врахував переносне значення (наприклад, для опису стану безладу в думках: "макітра голова").
- **Помилка:** Недостатнє врахування багатозначності терміну.

- **Рекомендація:** Розширити моделі для розуміння слів, які мають як пряме, так і переносне значення.

5. Запит: "Чи можна вважати слово 'шугати' літературним?"

- **Результат:** Bard визнав слово "шугати" діалектизмом, який рідко використовується у сучасній літературній мові, що частково відповідає дійсності.
- **Помилка:** Відсутність уточнення, що слово використовується в художніх текстах для створення емоційного забарвлення (наприклад, "птахи шугали над полем").
- **Рекомендація:** Інтегрувати корпуси художніх текстів для кращого розуміння стилістичних функцій слів.

Аналіз роботи Google Bard у полілінгвальному середовищі показав важливі аспекти його функціонування при обробці діалектів, архаїзмів, культурно значущих термінів і багатозначних слів. Розглянемо п’ять кейсів, які детально ілюструють можливості та обмеження системи, а також пропонують шляхи вдосконалення її алгоритмів.

Перший кейс стосувався запиту про значення слова "крисаня" у контексті українського фольклору. Google Bard припустив, що це може бути форма звертання або назва тварини, однак точного визначення не надав. У реальності "крисаня" означає "шапка з хутра," і це слово часто зустрічається у традиційних текстах. Відсутність чіткого розуміння фольклорного контексту демонструє обмеження системи у роботі з архаїзмами. Розширення корпусів текстів із фольклорними джерелами дозволить підвищити точність таких запитів і сприяти збереженню культурної спадщини у цифровій комунікації.

Другий кейс стосувався діалектного слова "партач," яке може мати різні значення залежно від контексту. У діалектах слово використовується для позначення майстра середнього рівня, але Bard обмежився загальним визначенням як "людина, яка робить щось неправильно." Це показує брак у

системі розуміння полісемії слів у діалектному контексті. Додавання до моделей локалізованих корпусів текстів, що враховують регіональні особливості мови, допомогло б забезпечити точніші результати й уникнути помилкових узагальнень.

Третій кейс ілюструє проблему з перекладом термінів, специфічних для національної історії, таких як "гайдамака." Bard переклав це слово як "rebel" (повстанець), що відповідає загальному розумінню, але не враховує історичного контексту, пов’язаного з боротьбою українців у XVIII столітті. Втрата культурної специфіки є серйозним викликом для роботи з багатомовними текстами. Для вирішення цього питання необхідно інтегрувати історичні корпуси текстів, які включають докладні пояснення термінів і їхніх контекстів.

Четвертий кейс демонструє роботу Bard зі словом "макітра," яке має як пряме, так і переносне значення. Система успішно розпізнала основне значення як "глиняна посудина для приготування страв," але не врахувала його переносного використання у вислові "макітра голова," який описує стан безладу в думках. Така обмеженість свідчить про необхідність навчання моделей на текстах, що охоплюють багатозначні слова та їхні стилістичні функції.

Останній кейс стосувався слова "шугати," яке Google Bard визначив як діалектизм, що рідко використовується у сучасній літературній мові. Хоча це частково відповідає дійсності, Bard не врахував, що це слово активно використовується у художніх текстах для створення емоційного забарвлення, наприклад, "птахи шугали над полем." Ця помилка вказує на необхідність інтеграції корпусів художньої літератури, щоб забезпечити глибше розуміння стилістичних функцій мови.

Ці кейси показують, що Google Bard демонструє гарні результати у роботі зі стандартними запитами, але його ефективність знижується у випадках із культурно специфічними термінами, діалектами та багатозначними словами. Для покращення результатів потрібна інтеграція

локалізованих текстових корпусів, розширення етимологічних баз і врахування стилістичних аспектів у навчальних моделях. Удосконалення цих аспектів дозволить підвищити точність і релевантність відповідей у полілінгвальних сценаріях, сприяючи створенню більш універсальних і адаптивних систем для роботи з багатомовним контентом.

2.4.2. Порівняння обробки стандартної та діалектної мови

1. Запит: "Розкажіть про історію слова 'шарварок' (стандартна мова)."

- **Результат:** Bard надав точну відповідь, пояснивши, що "шарварок" означає хаос або безлад, згадавши походження слова.
- **Помилка:** Не враховано додаткового історичного контексту про використання терміна в старовинних текстах.
- **Рекомендація:** Додавання історичних корпусів для більш глибокого аналізу.

2. Запит: "Що означає 'фустина' у діалектному контексті?"

- **Результат:** Система припустила, що це може бути одяг, але не надала точного визначення.
- **Помилка:** Відсутність розпізнавання гуцульського слова, що означає вовняну тканину.
- **Рекомендація:** Інтеграція корпусів гуцульських текстів.

3. Запит: "Чи використовується слово 'вишивати' у переносному значенні (стандартна мова)?"

- **Результат:** Bard успішно пояснив, що "вишивати" може мати переносне значення як "творити щось витончене."
- **Помилка:** Відсутність прикладів художніх текстів для ілюстрації.

- **Рекомендація:** Розширення моделей аналізу на базі літературних текстів.

4. Запит: "Поясніть слово 'рещітка' з діалектного тексту."

- **Результат:** Bard не зміг розпізнати значення слова, яке в діалекті означає "решітка" або "ґрати."
- **Помилка:** Відсутність підтримки діалектів у системі.
- **Рекомендація:** Інтеграція діалектних словників у моделі.

5. Запит: "Яке значення має слово 'лупати' у літературному тексті (стандартна мова)?"

- **Результат:** Система пояснила значення як "ударяти" або "розколювати," надавши точну відповідь.
- **Помилка:** Відсутність інформації про переносні значення слова.
- **Рекомендація:** Включення додаткових контекстів для аналізу.

6. Запит: "Як перекладається слово 'керпулик' із бойківського діалекту?"

- **Результат:** Bard не розпізнав цього слова.
- **Помилка:** Відсутність інформації про значення "хустка."
- **Рекомендація:** Залучення корпусів бойківського діалекту.

7. Запит: "Чи використовується 'стежина' у переносному значенні (стандартна мова)?"

- **Результат:** Система успішно пояснила, що "стежина" може використовуватися для позначення шляху у філософському чи метафоричному контексті.
- **Помилка:** Відсутність прикладів із літератури.
- **Рекомендація:** Додавання корпусів із художньої літератури.

8. Запит: "Що означає слово 'гава' у контексті фрази 'ловити ґави'?"

- **Результат:** Bard пояснив фразу правильно, але не розкрив етимологію слова "гава."
- **Помилка:** Відсутність повного аналізу терміна.
- **Рекомендація:** Додавання етимологічних баз.

9. Запит: "Що таке 'млака' в контексті діалекту?"

- **Результат:** Bard не зміг визначити, що це слово означає "болото."
- **Помилка:** Відсутність підтримки локальних текстів.
- **Рекомендація:** Розширення діалектних корпусів.

10. Запит: "Як пояснити слово 'побіляти' у літературному стилі (стандартна мова)?"

- **Результат:** Bard пояснив, що "побіляти" означає робити щось білішим або освітлювати.
- **Помилка:** Відсутність аналізу стилістичного вживання.
- **Рекомендація:** Включення корпусів текстів для стилістичного аналізу.

Аналіз роботи Google Bard з обробкою стандартної та діалектної мови включав тестування на здатність системи ідентифікувати, інтерпретувати й пояснювати як нормативні мовні конструкції, так і діалектні слова та вирази. Результати тестування ілюструють, що система демонструє значний рівень компетенції при роботі зі стандартною мовою, але має суттєві обмеження під час обробки діалектів і культурно специфічних термінів.

У першому кейсі Google Bard отримав запит про слово "шарварок" у стандартній мові. Система надала точне визначення, пояснивши, що це означає "хаос або безлад," однак не врахувала історичного контексту слова, яке широко використовувалося в старовинних текстах. Для покращення роботи у таких випадках потрібна інтеграція історичних корпусів текстів.

Другий кейс передбачав аналіз слова "фустина," яке використовується в гуцульському діалекті для позначення вовняної тканини. Bard припустив, що це, можливо, одяг, але точного визначення не надав. Це свідчить про те, що система має обмежений доступ до локалізованих текстів, що містять діалектизми, і не може повноцінно працювати з регіональними мовними варіантами.

Третій кейс стосувався запиту про використання слова "вишивати" у переносному значенні. Bard успішно визначив, що це слово може означати "творити щось витончене," але не навів прикладів із літератури, які підтвердили б таке використання. Це свідчить про необхідність навчання моделі на корпусах художніх текстів, що дозволить надавати більш деталізовані відповіді.

Четвертий кейс включав запит на пояснення діалектного слова "рещітка," яке означає "решітка" або "ґрати." Bard не зміг надати відповіді, оскільки слово не було в його базах даних. Цей приклад демонструє брак діалектних словників у моделях, що обмежує здатність системи до обробки локальних термінів.

П’ятий кейс перевіряв здатність Bard інтерпретувати слово "лупати" у літературному тексті. Система точно пояснила його значення як "ударяти" або "розколювати," але не врахувала його переносного вживання. Наприклад, у літературному контексті це слово може мати метафоричне забарвлення, що свідчить про необхідність розширення корпусів художньої літератури.

Шостий кейс зосереджувався на перекладі бойківського слова "керпулик," що означає "хустка." Bard не зміг розпізнати це слово, що вказує на обмеження в роботі з бойківським діалектом. Інтеграція корпусів текстів із цього діалекту допоможе усунути такі недоліки.

Сьомий кейс стосувався пояснення слова "стежина" у стандартній мові. Bard правильно визначив, що це слово може використовуватися метафорично для опису життєвого шляху, але знову ж таки не навів прикладів із художніх

текстів. Це свідчить про необхідність додавання корпусів, які містять стилістичні й літературні контексти.

Восьмий кейс включав фразу "ловити ґави" та запит про її значення. Система правильно пояснила фразеологізм, але не надала детального етимологічного аналізу слова "гава." Це демонструє брак глибоких етимологічних баз для пояснення складних фразеологічних конструкцій.

Дев’ятий кейс перевіряв слово "млака," яке означає "болото" в діалектному контексті. Bard не зміг ідентифікувати це слово, що вказує на недостатню інтеграцію регіональних текстів, які б дозволили краще працювати з діалектизмами.

Останній, десятий кейс, стосувався пояснення слова "побіляти" у стандартній мові. Bard надав правильне визначення як "робити щось білішим або освітлювати," але не врахував стилістичного використання у літературі, де це слово може мати метафоричне значення. Додавання художніх текстів до моделей дозволило б покращити обробку таких запитів.

Усі кейси демонструють сильні й слабкі сторони системи: вона успішно працює зі стандартними мовними запитами, але має серйозні обмеження у роботі з діалектами, фразеологізмами та багатозначними словами. Для вдосконалення алгоритмів потрібна інтеграція спеціалізованих корпусів діалектних текстів, історичних джерел і художньої літератури. Це дозволить Google Bard забезпечувати більш точні, контекстуальні й культурно релевантні відповіді.

Наступні 5 запитів:

Запит: "Поясніть слово 'спудей' у стандартній мові."

- **Результат:** Bard успішно розпізнав, що "спудей" означає "студент" або "учень" у старовинному контексті.

- **Помилка:** Відсутність уточнення про вузьке використання терміну в історичних текстах та університетських середовищах XVII-XVIII століть.
- **Рекомендація:** Інтегрувати історичні контексти використання термінів, які вже не є частиною сучасної мови.

2. Запит: "Що означає слово 'тлумачити' у бойківському діалекті?"

- **Результат:** Bard пояснив, що "тлумачити" означає "пояснювати" (стандартна мова).
- **Помилка:** Не враховано, що у бойківському діалекті "тлумачити" може також означати "сперечатися" або "вести довгу розмову."
- **Рекомендація:** Враховувати варіативність значень залежно від регіону, включаючи корпуси діалектів.

3. Запит: "Як перекладається слово 'глинтяка'?"

- **Результат:** Bard не розпізнав значення цього слова.
- **Помилка:** Слово "глинтяка" у діалекті означає "глиняна хата," але система не змогла надати жодного визначення.
- **Рекомендація:** Інтегрувати корпуси із сільськими текстами, де подібні терміни часто використовуються.

4. Запит: "Що означає слово 'брама' у стандартній мові?"

- **Результат:** Bard пояснив, що "брама" означає "вхідні ворота," і ця відповідь була точною.
- **Помилка:** Відсутність пояснення діалектного значення слова у деяких регіонах, де "брама" також може означати "прохід."
- **Рекомендація:** Розширити модель на основі діалектного вживання слова.

5. Запит: "Як перекласти слово 'мостище' на стандартну мову?"

- **Результат:** Bard не розпізнав, що "мостище" означає "великий міст" у діалектному контексті.
- **Помилка:** Відсутність будь-якої відповіді через невключення діалектних корпусів.
- **Рекомендація:** Додати корпуси, що охоплюють регіональні форми слів, які є поширеними в сільській місцевості.

Аналіз роботи Google Bard в обробці стандартної та діалектної мови на прикладі п'яти нових кейсів демонструє його здатність працювати з нормативною мовою та обмеження при обробці регіональних діалектів. Кожен кейс дозволив оцінити ефективність алгоритмів, виявити недоліки і запропонувати можливі вдосконалення.

У першому кейсі запит стосувався слова "спудей," яке в стандартній мові означає "студент" або "учень." Bard успішно ідентифікував значення слова, проте не врахував його вузький історичний контекст, пов’язаний із використанням цього терміна в освітніх установах XVII-XVIII століть. Це вказує на брак історичних корпусів текстів, що могли б забезпечити більш деталізовані відповіді. Додавання таких джерел дозволить системі глибше розуміти слова, що мають культурне чи історичне значення.

Другий кейс аналізував слово "тлумачити" у бойківському діалекті, де воно може означати не лише "пояснювати," як у стандартній мові, але й "сперечатися" чи "вести довгу розмову." Bard не розпізнав діалектне значення слова, запропонувавши лише стандартну інтерпретацію. Це підкреслює, що система не враховує варіативність значень залежно від регіону. Для вирішення цієї проблеми потрібно інтегрувати корпуси текстів, які містять діалектну лексику та різноманітні значення слів.

Третій кейс зосереджувався на слові "глинтяка," що в діалекті означає "глиняна хата." Google Bard не зміг надати жодної відповіді на цей запит, оскільки слово не було знайдено у його базах даних. Це демонструє, що система не адаптована до роботи з термінами, характерними для сільської

місцевості. Для покращення результатів необхідно додати корпуси текстів, що охоплюють сільське мовлення, де подібні слова є звичними.

Четвертий кейс включав стандартне слово "брама," яке Bard успішно ідентифікував як "вхідні ворота." Проте система не врахувала діалектного значення слова, яке в деяких регіонах може означати "прохід." Це вказує на необхідність розширення моделі для врахування регіональних варіантів слів, що дасть змогу надавати точні відповіді у локальних контекстах.

У п'ятому кейсі слово "мостище," яке в діалектному контексті означає "великий міст," також залишилося нерозпізнаним. Bard не зміг надати жодної відповіді, що свідчить про відсутність необхідних корпусів текстів із регіональною лексикою. Інтеграція таких текстів дозволить системі враховувати слова, які є специфічними для сільської або регіональної мови.
Ці кейси показують, що Google Bard демонструє високу ефективність у роботі зі стандартною мовою, але має суттєві обмеження при обробці діалектів. Для подолання цих недоліків необхідно розширити корпуси текстів, включаючи регіональну лексику, діалектні словники та історичні джерела. Удосконалення моделей дозволить підвищити точність обробки запитів у локалізованих контекстах, що зробить систему більш універсальною та адаптованою до потреб багатомовного світу.

Висновки до розділу 2

Розвиток комп’ютерної лінгвістики став однією з ключових складових сучасних технологій обробки природної мови, які забезпечують створення і функціонування віртуальних асистентів. У цьому розділі було детально проаналізовано основні досягнення в галузі, зокрема алгоритми обробки тексту, перекладу, синтезу мовлення та контекстуального розуміння. Особливу увагу приділено тестуванню роботи Google Bard у багатомовних сценаріях, де система продемонструвала як свої сильні сторони, так і обмеження. Висновки

цього розділу мають велике значення для розуміння сучасних можливостей і викликів у галузі комп’ютерної лінгвістики.

У межах розділу було виявлено, що основні досягнення у розвитку технологій обробки природної мови базуються на використанні алгоритмів глибокого навчання, таких як трансформери, і впровадженні великих мовних моделей, зокрема GPT та BERT. Ці технології дозволяють віртуальним асистентам працювати з текстами різними мовами, аналізувати їхній зміст, розуміти контекст і створювати зв’язні відповіді. Google Bard, як один із прикладів застосування таких технологій, показав високу точність у роботі зі стандартними мовними запитами, демонструючи здатність обробляти складні конструкції, адаптувати відповіді до контексту і забезпечувати зрозумілість для користувача. Зокрема, у випадках із запитами нормативною мовою система надавала коректні відповіді, що свідчить про її високу ефективність у роботі з академічними чи професійними текстами.

Водночас тестування показало, що Google Bard стикається з труднощами при роботі з діалектами, фразеологізмами, змішаними мовними структурами та культурно специфічними елементами. Наприклад, система мала труднощі з розпізнаванням діалектних слів, таких як "жєнджура" чи "мочеря," і не завжди правильно інтерпретувала їхній контекст. Аналогічно, під час аналізу фразеологізмів система надавала лише загальні пояснення, які не враховували багатогранність значення у конкретному мовному чи культурному середовищі. Це вказує на необхідність розширення корпусів даних, особливо тих, що містять діалекти, регіональні мовні варіанти та історичні тексти. Інтеграція таких ресурсів дозволить суттєво підвищити точність і глибину розуміння мовних запитів.

Ще одним важливим аспектом є здатність віртуальних асистентів працювати у багатомовному середовищі, де запити можуть містити елементи різних мов (code-switching). Аналіз показав, що Google Bard демонструє обмеження у роботі зі змішаними мовними запитами, що характерно для багатомовних спільнот. Наприклад, у запитах, де частина тексту подається

українською, а інша — англійською, система часто не враховує взаємодії між мовами, що призводить до некоректних або спрощених відповідей. Для вирішення цієї проблеми необхідно вдосконалити моделі через адаптацію до реальних сценаріїв багатомовного спілкування, зокрема шляхом інтеграції корпусів із code-switching.

Розділ також висвітлив роль віртуальних асистентів у синтезі мовлення та аналізі тональності тексту. Наприклад, Google Bard ефективно створює текстові відповіді, але має обмеження у відтворенні стилістичних або емоційних відтінків, що є критично важливим для персоналізованої взаємодії з користувачами. Крім того, система демонструє труднощі у роботі з багатозначними словами, де важливим є врахування контекстуального та стилістичного використання. Для розв’язання цієї проблеми необхідне навчання моделей на корпусах художньої літератури, які містять приклади різних стилів і жанрів.

У підсумку, результати аналізу роботи віртуальних асистентів, зокрема Google Bard, показують значний прогрес у галузі комп’ютерної лінгвістики, проте виявляють низку викликів, пов’язаних із багатомовністю, культурно специфічними контекстами та роботою з діалектами. Для вдосконалення цих систем необхідне розширення баз даних, включення локалізованих корпусів текстів і вдосконалення моделей для обробки складних лінгвістичних конструкцій. Ці вдосконалення дозволять створювати більш універсальні й адаптивні системи, які відповідатимуть потребам користувачів у сучасному багатомовному світі.

РОЗДІЛ 3. СТВОРЕННЯ ЕЛЕКТРОННОГО ДОВІДНИКА

3.1. Принципи створення інформаційних ресурсів

Створення інформаційних ресурсів, зокрема електронних довідників, ґрунтується на дотриманні низки принципів, які забезпечують їхню ефективність, зручність використання та відповідність потребам цільової аудиторії. Основні принципи розробки таких ресурсів включають структурованість, доступність, інформативність, адаптивність і інтерактивність. Структурованість передбачає логічне впорядкування інформації для швидкого пошуку та зручної навігації. Наприклад, електронний довідник має бути організований таким чином, щоб користувач міг легко знайти потрібний розділ або тему через зміст, систему тегів або пошуковий механізм. Доступність охоплює забезпечення можливості використання ресурсу незалежно від технічних обмежень або мовних бар'єрів. Це особливо актуально для багатомовних довідників, де інтеграція підтримки кількох мов є ключовою складовою. Інформативність вимагає, щоб усі розділи ресурсу містили актуальну, чітку й коректну інформацію, яка відповідає потребам користувачів. Адаптивність забезпечує зручний доступ до довідника з різних пристроїв, включаючи мобільні телефони та планшети, тоді як інтерактивність надає можливість взаємодії з користувачем через динамічний пошук, інтегровані посилання чи мультимедійний контент [1].

Структурна організація електронного довідника відіграє ключову роль у забезпеченні його ефективного функціонування. Основою довідника є його зміст, який має бути чітко структурованим і логічно поділеним на розділи та підрозділи. Наприклад, довідник може включати вступ, основні розділи, додаткові матеріали, глосарій і пошукову систему. Вступ слугує коротким оглядом теми, пояснює мету довідника та його основні функції. Основні розділи повинні охоплювати тематичний зміст, упорядкований за логічним або тематичним принципом. Для полегшення навігації рекомендується

використовувати номери розділів і підрозділів, а також інтерактивний зміст, що дозволяє швидко перейти до потрібного фрагмента тексту. Глосарій є важливим компонентом, який надає користувачам можливість швидко ознайомитися зі складною термінологією. Додаткові матеріали можуть містити таблиці, схеми, графіки чи мультимедійний контент, що сприяє кращому розумінню основної інформації. Інтегрована пошукова система дозволяє користувачам швидко знаходити інформацію за ключовими словами, що значно підвищує зручність використання електронного ресурсу [35].

Особливістю сучасних електронних довідників є можливість багатомовної підтримки. Це досягається через реалізацію багатомовного інтерфейсу, що дозволяє користувачам обирати мову відображення інформації. Наприклад, для полілінгвальних довідників важливо не лише перекладати зміст кількома мовами, але й адаптувати його до специфіки кожної мовної групи. Це включає врахування культурних відмінностей, стилістики тексту та особливостей термінології. Також структурна організація довідника має враховувати можливість постійного оновлення та розширення. Для цього доцільно передбачити модульну структуру, де кожен розділ може функціонувати як незалежна одиниця. Це забезпечить гнучкість у додаванні нових тем, оновленні існуючого матеріалу чи інтеграції нових технологій, наприклад, штучного інтелекту для персоналізації контенту [26].

Електронний довідник повинен відповідати сучасним технологічним вимогам. Це передбачає використання веб-орієнтованих платформ для забезпечення доступу через браузери, інтеграцію адаптивного дизайну для зручного використання на різних пристроях, а також забезпечення високого рівня безпеки даних. Інтерактивні елементи, такі як відеоінструкції, вбудовані підказки та автоматичні рекомендації, значно підвищують зручність використання довідника. Наприклад, динамічний пошук з автозаповненням дозволяє користувачам швидко знайти потрібну інформацію, навіть якщо вони не знають точного формулювання запиту. Включення мультимедійного

контенту, такого як відео чи інтерактивні схеми, робить довідник більш зрозумілим і привабливим для різних категорій користувачів [6].

Отже, створення інформаційних ресурсів та електронних довідників потребує дотримання комплексного підходу, який включає логічну структуру, багатомовну підтримку, інтерактивність і можливість постійного оновлення. Грамотно побудований довідник здатен не лише забезпечити зручність доступу до інформації, але й стати ефективним інструментом для навчання, роботи чи досліджень, відповідаючи сучасним вимогам інформаційного суспільства.

Використання мультимовних даних для створення довідника

Створення електронного довідника на основі мультимовних даних є важливим етапом у розробці інформаційних ресурсів, що відповідають потребам глобалізованого суспільства. Використання кількох мов для представлення інформації дозволяє забезпечити доступ до знань для користувачів із різних мовних груп, сприяє міжкультурній взаємодії та підтримує збереження мовного різноманіття. Основна ідея мультимовного довідника полягає в тому, щоб представити інформацію різними мовами без втрати її точності, значення та контексту. Для цього необхідно не лише перекладати текст, але й адаптувати його до культурних, стилістичних та термінологічних особливостей кожної мови. Наприклад, переклад технічних або спеціалізованих термінів вимагає глибокого розуміння контексту та забезпечення відповідності до мовних стандартів певної країни. Мультимовні дані також відкривають можливості для розробки інноваційних функцій, таких як динамічний переклад, підтримка локалізованих термінів та інтерактивні підказки, що дозволяє зробити довідник максимально доступним і зрозумілим для різних аудиторій [38].

Створення мультимовного довідника потребує комплексного підходу до організації його структури та даних. Одним із ключових завдань є забезпечення відповідності між версіями текстів різними мовами. Для цього використовуються системи управління контентом, що дозволяють синхронізувати дані та забезпечувати однаковий рівень актуальності для

кожної мовної версії. Наприклад, при оновленні інформації однією мовою необхідно автоматично враховувати зміни для інших мов. Також важливим аспектом є створення уніфікованих глосаріїв, які містять терміни та їхні відповідники різними мовами, що спрощує роботу з довідником. Іншою важливою складовою є локалізація контенту, яка включає адаптацію текстів до культурних особливостей. Наприклад, при описі соціокультурних тем слід враховувати, що певні терміни або фрази можуть мати різні значення залежно від країни або регіону, де використовується довідник. Це особливо актуально для юридичної, медичної чи освітньої інформації, де точність є критично важливою [44].

Важливу роль у створенні мультимовних довідників відіграють технології штучного інтелекту та автоматичного перекладу. Такі системи, як Google Translate чи DeepL, можуть виконувати базовий переклад текстів, забезпечуючи швидкість і зручність роботи з даними. Однак автоматичний переклад часто має обмеження, пов’язані з недостатнім розумінням контексту, особливо для спеціалізованих термінів чи діалектів. Тому для створення якісного мультимовного довідника необхідно поєднувати автоматичні системи з професійним редагуванням, що дозволить уникнути помилок і забезпечити високу якість контенту. Крім того, використання нейронних мереж і моделей машинного навчання дозволяє покращити результати перекладу, враховуючи стиль, тональність і специфіку мовлення. Наприклад, віртуальні асистенти на основі технологій штучного інтелекту здатні аналізувати запити користувачів кількома мовами, генеруючи відповіді відповідно до їхніх потреб, що може бути інтегровано до структури мультимовного довідника [20].

Особливу увагу слід приділити роботі з діалектами та мовними варіантами, які є важливою частиною культурної ідентичності. Використання мультимовних даних дозволяє створювати довідники, які враховують регіональні особливості, що забезпечує їхню універсальність та доступність. Наприклад, при створенні довідника для українськомовної аудиторії доцільно включити не лише стандартну мову, але й регіональні діалекти, такі як

гуцульський чи бойківський. Це потребує інтеграції спеціалізованих корпусів текстів, які містять діалектизми, регіональні фразеологізми та інші лексичні елементи. Також важливим аспектом є підтримка мовного змішування (code-switching), яке часто зустрічається у полілінгвальних спільнотах. Наприклад, користувачі можуть вводити запити кількома мовами одночасно, і довідник має бути здатним розуміти й відповідати на такі запити [15].

Ефективність мультимовного довідника також залежить від його інтерактивності та зручності використання. Інтеграція функцій динамічного пошуку, автоматичного перекладу запитів та інтерактивних підказок дозволяє зробити ресурс доступним для широкої аудиторії. Наприклад, якщо користувач вводить термін англійською мовою, система автоматично пропонує відповідники іншими мовами, доступними в довіднику. Це значно спрощує навігацію та дозволяє забезпечити високу якість взаємодії з ресурсом. Крім того, мультимовний довідник може включати мультимедійний контент, такий як відео, аудіофайли чи інтерактивні схеми, які адаптовані для різних мовних аудиторій. Це робить довідник корисним інструментом не лише для пошуку інформації, але й для навчання чи досліджень [12].

Отже, використання мультимовних даних для створення електронного довідника є складним, але надзвичайно важливим процесом, який потребує комплексного підходу. Забезпечення багатомовності, локалізація контенту, інтеграція сучасних технологій і врахування культурних особливостей дозволяють створювати ресурси, які відповідають сучасним потребам глобального суспільства. Такий довідник може стати важливим інструментом для підтримки міжкультурної комунікації, збереження мовного різноманіття та сприяння доступу до інформації незалежно від мовних бар’єрів [1].

Таблиця 3.1

Принципи створення інформаційних ресурсів

Принцип	Опис принципу	Приклад реалізації
Структурованість	Логічне впорядкування інформації для зручного пошуку та навігації.	Інтерактивний зміст, автоматичне оновлення матеріалів.
Доступність	Забезпечення можливості доступу до ресурсу з різних пристроїв та з урахуванням мовних бар'єрів.	Мобільна версія сайту, переклад на кілька мов.
Інформативність	Забезпечення точності та коректності інформації, яка відповідає потребам користувачів.	Докладні глосарії, таблиці, графіки та ілюстрації.
Адаптивність	Можливість адаптації контенту залежно від типу пристрою або потреб користувача.	Адаптація під різні екрани пристроїв (мобільні, планшети, ПК).
Інтерактивність	Наявність функцій для взаємодії з користувачем через пошук, інтерактивні посилання та мультимедійний контент.	Динамічний пошук, відеоуроки, інтерактивні форми.

3.2. Етапи розробки електронного довідника для полілінгвального середовища

Розробка електронного довідника для полілінгвального середовища є багатоступеневим процесом, що включає комплексну роботу з даними, структурою контенту та технологіями. Одним із найважливіших етапів є збір і аналіз багатомовних даних, адже саме від їхньої якості залежить ефективність і функціональність довідника. Процес створення довідника починається з визначення його цільової аудиторії та мовних потреб. Наприклад, якщо довідник розробляється для міжнародного використання, необхідно включити мови з найвищою часткою носіїв у відповідній галузі. Визначення мовного покриття передбачає аналіз регіональних особливостей, специфічної термінології та потенційних мовних бар'єрів, які можуть виникнути у користувачів. На цьому етапі також формуються основні вимоги до структури даних, формату представлення інформації та алгоритмів її обробки [3].

Збір даних для полілінгвального довідника передбачає пошук і відбір текстових корпусів, які мають відповідати тематиці й стилю майбутнього ресурсу. Джерела можуть включати наукові праці, технічну документацію, юридичні чи медичні тексти, що забезпечують повноту та точність інформації. Наприклад, для довідника, спрямованого на освітню аудиторію, доцільно використовувати підручники, словники та енциклопедії кількома мовами. Важливим аспектом є перевірка достовірності даних, яка здійснюється шляхом порівняння кількох джерел і відбору найавторитетніших. Крім того, для забезпечення культурної адаптації довідника важливо залучати регіональні тексти, що відображають мовні й культурні особливості цільової аудиторії. Наприклад, включення діалектів чи регіональних варіантів мови дозволяє створити довідник, який буде зрозумілим для широкого кола користувачів [7].

Аналіз багатомовних даних є ключовим етапом, що включає лінгвістичний, статистичний і семантичний підходи до обробки інформації. Лінгвістичний аналіз дозволяє ідентифікувати синтаксичні, морфологічні та семантичні особливості текстів різними мовами, що є важливим для створення уніфікованих глосаріїв і забезпечення адекватності перекладів. Наприклад, у полілінгвальних текстах потрібно враховувати граматичні відмінності між мовами, такі як структура речень або порядок слів. Статистичний аналіз допомагає виявити найпоширеніші терміни й фрази, які мають бути включені до довідника. Це особливо корисно для визначення ключових тем і формування базових категорій контенту. Семантичний аналіз дозволяє визначити контекстуальне значення слів і виразів, що є необхідним для роботи з полісемічними термінами або культурно специфічними концептами [26].

Наступним етапом є обробка зібраних даних, що передбачає їх очищення, нормалізацію та структуризацію. Очищення даних включає видалення дублюючої, застарілої чи нерелевантної інформації, що дозволяє підвищити точність і релевантність довідника. Наприклад, для забезпечення актуальності довідника слід вилучити терміни, які вже не використовуються, або адаптувати їх до сучасного контексту. Нормалізація передбачає приведення

даних до єдиного формату, що включає уніфікацію термінології, стилю та форматування тексту. Це особливо важливо для багатомовних довідників, де кожна мова має свої стилістичні й граматичні особливості. Структуризація даних забезпечує їх логічний поділ на розділи, категорії та підкатегорії, що сприяє зручній навігації та пошуку інформації [9].

Окрему увагу слід приділити інтеграції даних у довідник. Це включає впровадження багатомовного інтерфейсу, який дозволяє користувачам обирати мову для відображення контенту. Для полілінгвальних довідників важливо забезпечити рівноцінність інформації всіма мовами, уникнувши зміщення акцентів або втрати змісту під час перекладу. Наприклад, автоматичні перекладачі можуть бути корисними для базового перекладу, але професійне редагування забезпечує точність і культурну відповідність тексту. Інтеграція даних також передбачає створення інтерактивних елементів, таких як пошукові системи, які підтримують багатомовні запити, або мультимедійний контент, адаптований для різних мовних аудиторій [44].

Ефективна робота з багатомовними даними вимагає використання сучасних технологій, таких як нейронні мережі, моделі машинного навчання та інструменти автоматичної обробки природної мови. Наприклад, системи на основі трансформерів, такі як BERT чи GPT, дозволяють автоматизувати процес аналізу тексту, враховуючи контекст і культурні особливості. Це забезпечує більш точний переклад і створення релевантного контенту. Крім того, використання систем автоматичного навчання дозволяє постійно вдосконалювати якість довідника, додаючи нові дані та адаптуючись до потреб користувачів [43].

Таким чином, збір і аналіз багатомовних даних є невід’ємною частиною розробки електронного довідника для полілінгвального середовища. Цей процес включає визначення мовних потреб, пошук і перевірку даних, їхню обробку та інтеграцію. Використання сучасних технологій і підходів дозволяє створювати довідники, які відповідають вимогам глобалізованого суспільства,

забезпечуючи доступність, точність і зручність використання для користувачів різних мовних груп [24].

Інтеграція рекомендацій для користувачів у довідник

Інтеграція рекомендацій для користувачів у електронний довідник є ключовим етапом його розробки, оскільки ці рекомендації сприяють ефективному використанню ресурсу та підвищують його практичну цінність. Основна мета таких рекомендацій — забезпечити зручність навігації, доступність інформації та адаптацію до потреб різних категорій користувачів. Рекомендації можуть бути представлені у вигляді інструкцій, підказок, інтерактивних елементів або персоналізованих пропозицій, які враховують контекст використання довідника. Наприклад, якщо довідник створений для багатомовної аудиторії, рекомендації можуть включати інформацію про вибір мови, використання пошукових систем або специфіку подання даних кожною мовою. Особливо важливими є рекомендації щодо роботи з технічними чи спеціалізованими матеріалами, де користувачам потрібна допомога у розумінні складної термінології чи навігації між взаємопов’язаними розділами [21].

Одним із найефективніших способів інтеграції рекомендацій є створення динамічних підказок, які відображаються залежно від дій користувача. Наприклад, якщо користувач вводить пошуковий запит, система може автоматично пропонувати пов’язані терміни, додаткові розділи або уточнювати контекст запиту. Такий підхід дозволяє значно економити час і полегшує доступ до потрібної інформації. Крім того, інтеграція рекомендацій може включати створення покрокових інструкцій для виконання конкретних завдань. Наприклад, у технічному довіднику можуть бути представлені поради щодо налаштування обладнання, а в освітньому — рекомендації щодо використання навчальних матеріалів. Для досягнення цього необхідно застосовувати алгоритми, які враховують не лише зміст довідника, але й контекст використання: цілі користувача, його попередні дії у довіднику та частоту звернень до певних розділів [47].

Особливу увагу слід приділити персоналізації рекомендацій. Використання алгоритмів штучного інтелекту дозволяє аналізувати поведінку користувачів і пропонувати релевантні підказки відповідно до їхніх інтересів та потреб. Наприклад, якщо користувач часто звертається до матеріалів з певної теми, система може автоматично пропонувати йому додаткові розділи, пов’язані з цією темою, або навіть надавати рекомендації щодо зовнішніх ресурсів, які можуть бути корисними. Персоналізація також може стосуватися мови, стилю подання інформації чи навіть формату представлення контенту. Для багатомовних довідників це означає можливість автоматичного перемикання між мовами залежно від запиту користувача або надання рекомендацій щодо вибору найбільш релевантного мовного варіанту для специфічних тем [9].

Інтеграція рекомендацій також передбачає роботу з мультимедійними матеріалами. Наприклад, у довіднику можуть бути представлені відеоінструкції, інтерактивні схеми чи аудіофайли, які пояснюють складні концепції або демонструють процеси у дії. Для забезпечення доступності таких матеріалів можна використовувати рекомендації щодо найзручнішого формату або послідовності перегляду. Наприклад, якщо користувач працює на мобільному пристрої, система може запропонувати перегляд коротких відео замість читання довгих текстових інструкцій. Інтерактивні рекомендації, такі як вбудовані підказки під час роботи з довідником, дозволяють мінімізувати час, витрачений на пошук інформації, і зробити використання ресурсу інтуїтивно зрозумілим навіть для новачків [20].

Крім того, рекомендації для користувачів можуть включати інтеграцію системи зворотного зв’язку. Наприклад, користувачам можна запропонувати оцінити корисність певного розділу або залишити коментарі щодо його вдосконалення. Це дозволяє не лише покращувати якість довідника, але й забезпечувати постійний моніторинг його актуальності та відповідності потребам аудиторії. У полілінгвальному середовищі особливу роль відіграють рекомендації щодо роботи з діалектами або специфічними мовними

варіантами. Наприклад, користувач може отримати підказку про те, як адаптувати запит для отримання більш релевантної інформації, якщо система не змогла відразу розпізнати діалектне слово.

Інтеграція рекомендацій у довідник не лише підвищує його функціональність, але й робить взаємодію з користувачами більш ефективною та комфортною. Завдяки використанню сучасних технологій, таких як штучний інтелект, мультимедійні інструменти та персоналізовані алгоритми, рекомендації стають невід'ємною частиною електронних довідників, що забезпечує доступність, точність і зручність використання для широкої аудиторії [1].

Таблиця 3.2

Етапи розробки електронного довідника для полілінгвального середовища

Етап	Опис етапу	Приклад реалізації
Планування та аналіз вимог	Визначення цілей довідника, основних функцій та потреб користувачів, вибір мов для підтримки.	Оцінка цільової аудиторії та визначення ключових мов для підтримки, таких як англійська, українська, іспанська.
Проектування структури та контенту	Створення карти структури довідника, визначення формату контенту, підготовка матеріалів для введення.	Розробка схем навігації та контенту для кожної мови, тестування з різними форматами документів.
Розробка та інтеграція технологій	Розробка інтерфейсу користувача, вибір технологій для багатомовної підтримки, інтеграція системи пошуку.	Інтеграція багатомовних інтерфейсів за допомогою технологій на кшталт Google Translate API.
Тестування та налагодження	Перевірка роботи довідника з усіма мовами, виявлення та виправлення помилок, тестування продуктивності.	Тестування з фокусом на багатомовне введення та правильність відображення текстів на різних пристроях.
Запуск та підтримка	Запуск системи, моніторинг користувацького досвіду, постійне оновлення контенту та функцій.	Запуск електронного довідника на платформах для мобільних і десктопних користувачів, регулярне оновлення контенту.

3.3. Використання електронного довідника для навчальних і дослідницьких цілей

Електронні довідники є універсальним інструментом, який активно використовується для підтримки навчальних та дослідницьких процесів. Їхня основна перевага полягає у швидкому доступі до структурованої інформації, що дозволяє користувачам ефективно знаходити необхідні дані для навчання чи проведення наукових досліджень. У навчальному контексті електронний довідник може використовуватися як основний або допоміжний ресурс для засвоєння матеріалу, перевірки фактів або вивчення нових тем. Наприклад, у школах і вишах такі довідники можуть забезпечувати учнів і студентів доступом до баз знань з різних дисциплін, включаючи точні науки, гуманітарні дослідження та технічні спеціальності. Завдяки інтеграції сучасних технологій, таких як автоматичний пошук, гіперпосилання та мультимедійні елементи, довідники стають незамінними у навчальних процесах, особливо в умовах дистанційного чи змішаного навчання [25].

Адаптація електронного довідника для освітніх потреб передбачає його відповідність рівню підготовки аудиторії та специфіці навчальних програм. Для молодшої аудиторії, наприклад учнів шкіл, інформація у довіднику має бути подана в доступній, зрозумілій формі з використанням простих прикладів і візуалізацій. Для студентів вишів, особливо технічних чи наукових спеціальностей, необхідно включати деталізовані пояснення, формули, схеми та додаткові посилання на джерела для глибшого вивчення тем. Наприклад, якщо довідник використовується у курсі з програмування, він може містити інтерактивні модулі з прикладами коду, які дозволяють студентам не лише ознайомлюватися з теоретичною інформацією, але й практично застосовувати її. Для гуманітарних дисциплін важливим є наявність історичних, соціокультурних та літературних контекстів, які допомагають краще зрозуміти матеріал [29].

У дослідницьких цілях електронний довідник може використовуватися для збору первинної інформації, перевірки термінології чи аналізу наукових джерел. Наприклад, дослідник, який працює над темою історичних діалектів, може використовувати довідник для доступу до корпусів текстів, що містять приклади діалектизмів, або для аналізу етимології слів. Включення функцій автоматичного перекладу, пошуку за контекстом чи аналізу тональності тексту дозволяє значно розширити можливості таких ресурсів для наукових цілей. Особливо важливим є використання мультимовних довідників у дослідженнях, пов'язаних із полілінгвізмом, культурними взаємодіями чи лінгвістичним аналізом. Наприклад, у дослідженнях з комп'ютерної лінгвістики електронний довідник може слугувати базою для тестування алгоритмів обробки природної мови [25].

Одним із ключових аспектів адаптації довідника для освітніх потреб є інтеграція мультимедійних матеріалів і практичних завдань. Наприклад, для школярів інтерактивні вправи, такі як вікторини чи тренажери, дозволяють закріплювати знання, а для студентів можуть бути створені симуляції реальних процесів, як-от експерименти у фізиці чи хімії. Мультимедійні елементи, такі як відео, графіки чи інтерактивні карти, допомагають краще візуалізувати інформацію та зробити навчання більш інтерактивним. Для викладачів електронний довідник може пропонувати готові навчальні плани, презентації чи додаткові ресурси, які можна використовувати під час уроків або лекцій [27].

Довідник також може виконувати роль інструмента для індивідуалізованого навчання. Використовуючи алгоритми персоналізації, система може адаптувати подання інформації до потреб конкретного користувача. Наприклад, студенти можуть отримувати рекомендації щодо матеріалів залежно від їхнього рівня знань або результатів попередніх тестів. Для дослідників такі функції дозволяють знаходити релевантні джерела для їхньої теми або отримувати підказки щодо можливих напрямків досліджень. Включення функції зворотного зв'язку дає змогу користувачам залишати свої

коментарі чи оцінки, що дозволяє вдосконалювати ресурс відповідно до реальних потреб аудиторії.

Використання електронного довідника для навчальних і дослідницьких цілей відкриває широкі можливості для інтеграції інновацій у процес навчання та наукової роботи. Завдяки адаптації до освітніх потреб, інтеграції мультимедійних матеріалів і можливостей персоналізації довідники стають ефективними інструментами для здобуття знань, що відповідають сучасним викликам інформаційного суспільства [30].

Використання ресурсу у наукових дослідженнях

Електронні ресурси відіграють ключову роль у сучасних наукових дослідженнях, забезпечуючи доступ до великої кількості інформації, аналізу даних і можливостей для організації знань. Особливістю використання таких ресурсів є їхня здатність значно пришвидшувати процес збору й обробки інформації, що особливо важливо в умовах швидкого розвитку науки. У науковій діяльності електронні довідники можуть виконувати кілька функцій, включаючи пошук спеціалізованих даних, доступ до термінологічних баз, перевірку достовірності джерел, а також зберігання та структуризацію отриманих матеріалів. Наприклад, у галузі гуманітарних наук довідники можуть використовуватися для аналізу текстів, роботи з фольклорними джерелами чи вивчення мовних особливостей. У точних науках такі ресурси дозволяють отримувати точні формули, алгоритми та технічні рішення, необхідні для проведення експериментів або побудови моделей [31].

Однією з основних переваг електронних ресурсів є можливість багатомовного доступу до інформації, що робить їх особливо корисними для міжнародних дослідницьких проектів. Полілінгвальні довідники дозволяють дослідникам працювати з матеріалами різними мовами, порівнювати дані та аналізувати інформацію у глобальному контексті. Наприклад, у сфері комп’ютерної лінгвістики дослідники можуть використовувати довідники для тестування алгоритмів обробки природної мови, аналізуючи багатомовні корпуси текстів. У гуманітарних дослідженнях доступ до історичних текстів

кількома мовами дозволяє глибше вивчати питання взаємодії культур, міграції чи адаптації мовних традицій. Використання інтегрованих пошукових систем у таких довідниках дозволяє швидко знаходити потрібну інформацію, навіть якщо користувач не має точного уявлення про формулювання запиту. Це значно підвищує ефективність роботи з ресурсом і знижує час на пошук інформації [38].

Важливою складовою використання електронних довідників у наукових дослідженнях є можливість інтеграції даних із різних джерел. Наприклад, довідник може включати не лише текстову інформацію, але й графіки, таблиці, відео або інтерактивні елементи, що забезпечує багатовимірний підхід до аналізу даних. У технічних дисциплінах це може бути використання моделей для візуалізації фізичних процесів або тестування інженерних рішень. У гуманітарних дисциплінах інтеграція мультимедійних матеріалів дозволяє аналізувати фольклор, музичну спадщину або візуальні мистецтва. Також важливим є використання автоматизованих інструментів для аналізу даних, таких як машинне навчання чи статистичний аналіз, які можуть бути інтегровані в електронний довідник для виконання спеціалізованих завдань.

Особливу роль відіграє функція створення звітів і публікацій на основі матеріалів, доступних у довіднику. Наприклад, науковці можуть використовувати готові шаблони для підготовки звітів, цитувань чи бібліографій, що значно полегшує процес документування досліджень. Крім того, система може автоматично генерувати посилання на джерела, що забезпечує високу точність і уникнення помилок у наукових публікаціях. Для дослідників, які працюють у міжнародних проектах, це особливо важливо, адже довідники часто надають можливість експорту даних у різних форматах, таких як APA, MLA чи Chicago. Це робить ресурс універсальним інструментом для наукової роботи, який може адаптуватися до потреб різних дисциплін і стандартів [39].

Окремим аспектом використання електронних довідників у дослідженнях є їхня роль у збереженні та популяризації культурної спадщини.

У проектах, пов’язаних із мовознавством, історією чи антропологією, такі ресурси дозволяють зберігати рідкісні тексти, інтегрувати діалекти та документувати традиції. Наприклад, у рамках проектів зі збереження маловживаних мов електронні довідники можуть містити аудіозаписи, транскрипції та переклади текстів, що сприяє їхньому дослідженню й популяризації. Також електронні довідники можуть використовуватися для створення інтерактивних освітніх програм, які допомагають не лише вивчати, але й застосовувати отримані знання в практичних проектах [1].

Використання електронних довідників у наукових дослідженнях забезпечує доступ до багатогранної інформації, дозволяючи ефективно використовувати сучасні технології для навчання, аналізу та збереження даних. Інтеграція мультимедійних елементів, багатомовних даних і можливостей автоматичного аналізу відкриває нові горизонти для науковців, сприяючи підвищенню точності, швидкості та універсальності досліджень у різних галузях знань [3].

Таблиця 3.3

Використання електронного довідника для навчальних і дослідницьких цілей

Ціль	Опис використання	Приклад реалізації
Навчання студентів	Забезпечення доступу до матеріалів на різних мовах для підготовки до занять та тестів.	Студенти можуть використовувати довідник для підготовки до іспитів з іноземних мов, дослідження культурних аспектів.
Підготовка дослідницьких матеріалів	Збір і організація матеріалів для наукових робіт та досліджень, підтримка багатомовного контексту.	Збір даних для наукових статей або досліджень в багатомовних контекстах, адаптованих до різних наукових стандартів.
Полегшення доступу до інформації	Надання можливості доступу до матеріалів без обмежень, інтеграція з навчальними платформами.	Платформа для швидкого доступу до посібників, інтерфейс для користувачів з різних мовних груп.
Мультимедійне навчання	Інтеграція відео, аудіо та іншого мультимедійного контенту для кращого засвоєння матеріалу.	Мультимедійні курси та інтерактивні завдання на базі електронного довідника для студентів та викладачів.
Покращення взаємодії між дослідниками	Створення бази даних для обміну науковими результатами та співпраці в міжнародних дослідницьких проектах.	Спільна робота над науковими проектами в рамках університетських та міжнародних наукових спільнот.

3.4. Аналіз результатів дослідження та їх застосування у розробці електронного довідника

Результати дослідження, проведеного в рамках другого розділу, стали основою для розробки електронного довідника, який здатен ефективно функціонувати у полілінгвальному середовищі. Аналіз роботи Google Bard дозволив визначити ключові особливості та обмеження, що впливають на обробку багатомовних даних. Було встановлено, що система ефективно працює зі стандартною мовою, демонструючи високу точність відповідей, однак стикається з труднощами під час роботи з діалектами, фразеологізмами

та змішаними мовними запитами. Ці результати підкреслюють необхідність створення довідника, який би не лише враховував ці аспекти, але й пропонував конкретні рішення для забезпечення релевантної та точної інформації у таких контекстах.

Одним із ключових напрямів використання результатів дослідження є створення багатомовного словника термінів і фраз, які викликали труднощі під час роботи системи. Наприклад, аналіз роботи з діалектами виявив необхідність інтеграції регіональних корпусів текстів, що включають слова та фрази, характерні для окремих регіонів. Ці дані можуть бути використані для створення розділу довідника, присвяченого діалектним особливостям, де користувачі зможуть знайти пояснення чи переклад термінів, специфічних для їхнього мовного середовища. Крім того, для забезпечення точності перекладів і розуміння контексту планується включити розширений глосарій із прикладами використання фраз у стандартній та діалектній мові.

Дослідження також показало, що система стикається з труднощами під час роботи зі змішаними мовними запитами, характерними для багатомовних спільнот. Ця проблема може бути вирішена шляхом розробки спеціального розділу довідника, присвяченого явищу code-switching. У цьому розділі будуть представлені приклади змішаних мовних конструкцій із поясненнями їхнього значення та контексту використання. Це дозволить користувачам краще розуміти, як поєднання різних мов впливає на комунікацію, а також як правильно використовувати такі конструкції у різних ситуаціях. Такий підхід не лише допоможе подолати мовні бар’єри, але й сприятиме підвищенню рівня міжкультурної комунікації.

Результати дослідження роботи з фразеологізмами й культурно специфічними термінами виявили необхідність інтеграції додаткових ресурсів, що пояснюють значення таких елементів. Наприклад, довідник може містити розділ, присвячений фразеології, з ілюстраціями та етимологічними поясненнями, які допоможуть краще зрозуміти контекст і значення висловів. Це може бути особливо корисним для іноземних користувачів, які вивчають

українську мову чи працюють із текстами, що містять культурно специфічні елементи.

Ще одним важливим аспектом є використання результатів для вдосконалення структури електронного довідника. На основі виявлених обмежень пропонується створити багаторівневу структуру, що включатиме як загальні розділи для стандартної мови, так і спеціалізовані розділи для діалектів, фразеології та змішаних мовних запитів. Крім того, для кожного розділу буде інтегровано інтерактивний пошуковий механізм, який дозволить користувачам швидко знаходити інформацію за ключовими словами або фразами. Наприклад, користувачі зможуть вводити запити діалектною мовою, а система автоматично надаватиме переклади чи пояснення стандартною мовою, враховуючи контекст.

Особливу увагу слід приділити мультимовній адаптації довідника, яка стане можливою завдяки використанню даних, отриманих під час тестування Google Bard. Включення багатомовної підтримки дозволить користувачам обирати мову для роботи з довідником, а також автоматично перемикатися між мовами залежно від введеного запиту. Наприклад, якщо користувач вводить запит англійською, система автоматично підбирає відповідний контент цією мовою, зберігаючи при цьому можливість перегляду матеріалу іншими мовами. Така функціональність сприятиме підвищенню зручності використання довідника у полілінгвальних спільнотах і зробить його доступним для ширшої аудиторії [10].

Результати дослідження також показали важливість роботи з мультимедійним контентом для підвищення зручності використання довідника. Наприклад, інтеграція відео, графіків чи інтерактивних схем може допомогти користувачам краще зрозуміти складні терміни чи концепти. Це може бути особливо корисним у випадках, коли текстовий опис недостатній для повного розуміння інформації. Крім того, мультимедійний контент може бути використаний для створення інтерактивних навчальних модулів, які

допоможуть користувачам засвоїти нові знання або перевірити свої вміння [15].

Отже, результати проведеного дослідження не лише висвітлюють сильні та слабкі сторони роботи Google Bard, але й слугують основою для створення електронного довідника, який буде адаптований до потреб багатомовного середовища. Завдяки інтеграції отриманих даних у структуру довідника, забезпеченню багатомовної підтримки та розробці спеціалізованих розділів цей ресурс стане ефективним інструментом для навчання, роботи та міжкультурної комунікації.

Таблиця 3.4

Аналіз результатів дослідження та їх застосування у розробці електронного довідника

Результат дослідження	Опис результату	Застосування у розробці електронного довідника
Аналіз мовних моделей віртуальних асистентів	Визначення ефективності роботи мовних моделей, таких як GPT та BERT, у полілінгвальних середовищах.	Інтеграція кращих мовних моделей для поліпшення точності відповідей в багатомовному середовищі.
Адаптація до багатомовних запитів	Оцінка здатності віртуальних асистентів правильно обробляти запити на кількох мовах одночасно.	Розробка функціоналу для багатомовних запитів та підтримка перемикання між мовами.
Інтерфейс та доступність для користувачів	Тестування зручності користування та ефективності навігації в багатомовних інтерфейсах.	Оптимізація інтерфейсу для зручного використання з різних пристроїв і з підтримкою кількох мов.
Аналіз точності перекладу	Перевірка точності перекладу та розпізнавання змішаних мовних конструкцій у реальних запитах.	Інтеграція алгоритмів для автоматичного перекладу та адаптації під змішані мовні конструкції.
Інтерактивність та мультимедійний контент	Оцінка ефективності мультимедійного контенту (відео, графіка) для покращення процесу навчання.	Додавання мультимедійних елементів, таких як відео та інтерактивні завдання, для покращення навчання.

Висновки до розділу 3

Створення електронного довідника є важливим етапом у розвитку інформаційних технологій, спрямованим на забезпечення зручного доступу до знань у сучасному інформаційному суспільстві. Розробка такого ресурсу охоплює збір, аналіз та інтеграцію даних, створення зручної структури, адаптацію до потреб різних категорій користувачів і впровадження сучасних технологій, таких як штучний інтелект, автоматичний переклад і персоналізація контенту. Основна мета електронного довідника полягає у створенні інструменту, який дозволить користувачам ефективно взаємодіяти з інформацією, забезпечуючи швидкість, точність і адаптивність у багатомовному середовищі.

Процес створення електронного довідника базується на чітких принципах, серед яких структурованість, доступність, багатомовність, інформативність і інтерактивність. Структурована організація контенту є основою ефективної роботи довідника, оскільки вона дозволяє користувачам швидко знаходити необхідну інформацію завдяки розподілу даних на розділи, категорії та підкатегорії. Наприклад, інтерактивний зміст і система динамічного пошуку значно спрощують навігацію, забезпечуючи доступ до конкретних розділів або матеріалів навіть для користувачів з обмеженими знаннями тематики. Доступність ресурсу забезпечується адаптивним дизайном, який дозволяє використовувати довідник на різних пристроях, включаючи мобільні телефони та планшети, а також багатомовною підтримкою, що гарантує можливість використання довідника представниками різних мовних спільнот.

Важливим аспектом створення електронного довідника є збір і аналіз даних, які забезпечують повноту, точність і актуальність контенту. Цей етап передбачає залучення різноманітних джерел, таких як наукові публікації, технічна документація, навчальні матеріали та регіональні корпуси текстів.

Інтеграція мультимовних даних дозволяє створювати ресурс, адаптований до потреб полілінгвального середовища, забезпечуючи якісний переклад і врахування культурних та мовних особливостей. Наприклад, додавання до довідника розділів, присвячених діалектам чи фразеологізмам, дозволить розширити його функціональність і зробити ресурс корисним для ширшого кола користувачів.

Особливе значення має адаптація електронного довідника до освітніх і дослідницьких потреб. У навчальному контексті такий ресурс може слугувати як джерело знань, так і інструмент для закріплення матеріалу завдяки інтерактивним вправам, відеоінструкціям та мультимедійним матеріалам. Для науковців довідник стає платформою для доступу до спеціалізованих даних, аналізу текстів і роботи з багатомовними корпусами. Крім того, інтеграція сучасних технологій, таких як штучний інтелект, дозволяє автоматизувати процес аналізу інформації, адаптувати контент до потреб конкретного користувача та забезпечувати високий рівень персоналізації. Наприклад, використання алгоритмів машинного навчання дозволяє створювати рекомендації для користувачів залежно від їхніх запитів або історії взаємодії з ресурсом.

Результати аналізу роботи з Google Bard, отримані в рамках нашого дослідження, продемонстрували необхідність врахування багатомовного контексту, роботи з діалектами, культурно специфічними термінами та явищем мовного змішування. Ці висновки стали основою для створення електронного довідника, який буде здатен обробляти складні запити та забезпечувати релевантність і точність відповідей. Наприклад, додавання спеціалізованих розділів, присвячених code-switching, дозволить користувачам отримувати пояснення змішаних мовних конструкцій із урахуванням їхнього контексту.

Інтерактивність є ще одним ключовим елементом довідника, що сприяє зручності його використання. Інтеграція функцій динамічного пошуку, автоматичних підказок і мультимедійного контенту дозволяє створювати ресурси, які є інтуїтивно зрозумілими навіть для недосвідчених користувачів.

Крім того, інтерактивні функції, такі як система зворотного зв’язку чи можливість оцінювання розділів, дозволяють постійно вдосконалювати довідник відповідно до потреб користувачів.

Отже, створення електронного довідника — це не лише технічний процес, але й комплексний підхід до забезпечення доступності знань у сучасному інформаційному суспільстві. Інтеграція структурованих даних, мультимовної підтримки, адаптивності до освітніх і дослідницьких цілей, а також використання сучасних технологій дозволяє створити ресурс, який відповідає вимогам глобалізованого світу. Такий довідник не лише забезпечує доступ до інформації, але й сприяє розвитку міжкультурної комунікації, підтримці мовного розмаїття та підвищенню ефективності роботи у різних галузях знань.

ВИСНОВКИ

Дослідження роботи віртуальних асистентів у полілінгвальному середовищі демонструє важливість інтеграції сучасних технологій у контекст багатомовної комунікації. Створення електронного довідника, який враховує специфіку багатомовного середовища, є не лише технічним викликом, але й необхідністю для забезпечення ефективної взаємодії користувачів із різним мовним фоном. Аналіз роботи Google Bard, проведений у рамках дослідження, дозволив визначити як його сильні сторони, так і обмеження. Результати дослідження свідчать, що основним викликом для віртуальних асистентів є робота з діалектами, мовним змішуванням (code-switching) і культурно специфічними елементами, які часто залишаються нерозпізнаними або некоректно інтерпретованими. Ці аспекти є критично важливими для створення електронного довідника, здатного ефективно працювати у полілінгвальному середовищі.

Однією з основних цілей електронного довідника є подолання мовних бар’єрів та забезпечення доступності інформації для користувачів з різними мовними потребами. Використання даних, отриманих у результаті аналізу роботи віртуального асистента, дозволяє створювати ресурс, який включає багатомовну підтримку, інтеграцію локалізованих текстів та адаптацію до культурних особливостей. Наприклад, інтеграція корпусів текстів з діалектними елементами дозволяє забезпечити точність перекладів і пояснень для регіонально специфічних термінів. Це також сприяє збереженню культурного розмаїття, яке є невід’ємною частиною полілінгвального середовища. Враховуючи складність роботи з діалектами, довідник має включати спеціальні розділи, присвячені діалектології, де кожен користувач зможе отримати роз’яснення значення термінів або їх відповідників у стандартній мові.

Особливу увагу слід приділити роботі з мовним змішуванням, яке є характерним для багатьох полілінгвальних спільнот. Результати тестування

Google Bard показали, що система часто стикається з труднощами при обробці змішаних мовних запитів, оскільки алгоритми не завжди можуть врахувати взаємодію між різними мовами. У межах електронного довідника це можна вирішити шляхом створення інструментів для роботи з code-switching, які будуть автоматично розпізнавати та пояснювати змішані конструкції. Наприклад, користувач може вводити запит, де одна частина сформульована українською, а інша англійською, і система має надати адекватну відповідь, враховуючи обидві мовні частини. Це дозволить не лише покращити точність роботи довідника, але й підвищити комфорт користувачів.

Результати дослідження також підтвердили важливість роботи з культурно специфічними елементами, такими як фразеологізми, ідіоми чи архаїзми. Наприклад, багато фразеологізмів мають багатозначні інтерпретації залежно від контексту, що може призводити до помилок під час їх автоматичної обробки. У рамках електронного довідника необхідно створити розділ, присвячений фразеології, де буде надано детальні пояснення таких виразів, їхні приклади використання та етимологічні довідки. Це особливо важливо для іноземних користувачів, які вивчають українську мову або працюють із текстами, що містять культурно специфічні елементи. Такий підхід забезпечує не лише точність, але й сприяє глибшому розумінню мовної та культурної специфіки.

Додатковим аспектом, який необхідно врахувати, є інтеграція мультимовного інтерфейсу, що дозволяє користувачам обирати мову для роботи з довідником і автоматично перемикатися між мовами залежно від контексту. Це особливо важливо для міжнародних проектів, де довідник може використовуватися представниками різних мовних спільнот. Наприклад, система може автоматично пропонувати переклад контенту залежно від мови запиту або налаштувань користувача. Така адаптивність не лише підвищує зручність використання ресурсу, але й розширює його аудиторію.

Електронний довідник також має включати можливості інтерактивної роботи з даними, такі як динамічний пошук, інтерактивні схеми чи

мультимедійний контент. Наприклад, користувачі можуть вводити запити у вигляді ключових слів, і система автоматично пропонуватиме пов’язані терміни або додаткові матеріали. Це дозволяє зробити роботу з ресурсом інтуїтивно зрозумілою навіть для недосвідчених користувачів. Крім того, мультимедійні матеріали, такі як відеоінструкції, графіки чи інтерактивні карти, сприяють кращому засвоєнню інформації, що робить довідник універсальним інструментом як для навчання, так і для досліджень.

Отже, результати нашого дослідження дозволяють сформулювати ключові принципи для створення електронного довідника, який буде ефективним у полілінгвальному середовищі. Інтеграція мультимовних даних, адаптація до культурних особливостей, робота з діалектами та змішаними мовами, а також використання сучасних технологій, таких як штучний інтелект, забезпечать високу якість ресурсу та його відповідність сучасним вимогам. Такий довідник не лише полегшить доступ до інформації, але й сприятиме розвитку міжкультурної комунікації, збереженню мовного різноманіття та підвищенню ефективності роботи у багатомовних спільнотах.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Арістова Є. Г. Мовна політика в Україні: історичний контекст. Київ: Наукова думка, 2017. 256 с.
2. Бондаренко М. С. Комп’ютерна лінгвістика: основи та перспективи. Харків: Видавництво ХНУ, 2020. 198 с.
3. Вакуленко С. О. Мовні контакти у глобалізованому світі. Львів: Видавництво ЛНУ, 2018. 315 с.
4. Виговський І. В. Лінгвістичний аналіз у цифрову епоху. Полтава: Полтавський літопис, 2019. 145 с.
5. Вороніна Л. О. Соціолінгвістика: основи теорії. Одеса: ОНУ, 2016. 178 с.
6. Глушко Н. М. Діалектологія української мови. Чернівці: Рута, 2018. 240 с.
7. Горбачук М. А. Полілінгвізм і цифрові технології. Дніпро: Дніпропрес, 2020. 256 с.
8. Долина С. В. Мовне змішування у соціальних мережах. Харків: Освіта, 2021. 198 с.
9. Дроботенко В. Г. Лінгвістичний аналіз текстів. Київ: Либідь, 2019. 234 с.
10. Журавський О. А. Багатомовність у сучасному світі. Київ: Видавництво КНУ, 2015. 312 с.
11. Зоріна Т. Л. Комп’ютерна обробка тексту. Одеса: Одеський університет, 2020. 150 с.
12. Іванова Н. Ю. Мовна адаптація в полілінгвальному середовищі. Львів: Каменяр, 2018. 176 с.
13. Карпенко О. С. Інтерактивні технології в лінгвістиці. Харків: Видавництво ХНУ, 2021. 205 с.
14. Клименко А. В. Аналіз діалектів у цифрових платформах. Ужгород: Видавництво УжНУ, 2020. 189 с.
15. Коваленко Р. М. Лінгвістична підтримка мультимовних платформ. Дніпро: Видавництво ДНУ, 2017. 264 с.

16. Колесник І. І. Фразеологізми та їх переклад. Київ: Либідь, 2016. 172 с.
17. Кравченко Н. С. Локалізація інформаційних ресурсів. Львів: ЛНУ, 2019. 190 с.
18. Кучеренко В. Г. Штучний інтелект у мовній обробці. Одеса: Рута, 2021. 256 с.
19. Лапін Ю. М. Використання даних для навчальних цілей. Київ: Наукова думка, 2019. 200 с.
20. Литвиненко П. С. Багатомовна обробка даних у лінгвістиці. Харків: Видавництво ХНУ, 2020. 250 с.
21. Максименко О. В. Інтеграція діалектів у цифрове середовище. Ужгород: Видавництво УжНУ, 2021. 180 с.
22. Маркович С. Ю. Автоматичний аналіз мовних текстів. Київ: Либідь, 2020. 230 с.
23. Мартинюк І. О. Полілінгвізм у цифрову епоху. Львів: Каменяр, 2019. 210 с.
24. Мельник Т. В. Лінгвістична адаптація віртуальних асистентів. Дніпро: Дніпропрес, 2020. 195 с.
25. Михайленко О. П. Комп’ютерна лексикографія. Харків: Освіта, 2018. 189 с.
26. Мосійчук А. В. Переклад у полілінгвальному середовищі. Одеса: ОНУ, 2019. 205 с.
27. Назаренко Н. Г. Багатомовність у сучасній лінгвістиці. Київ: Наукова думка, 2017. 180 с.
28. Олексієнко Ю. С. Інтерактивні платформи для багатомовних користувачів. Львів: ЛНУ, 2020. 220 с.
29. Осадчук І. М. Структура довідників у полілінгвальному середовищі. Дніпро: ДНУ, 2018. 240 с.
30. Павленко А. Ю. Мовні моделі у штучному інтелекті. Харків: Видавництво ХНУ, 2021. 190 с.

31. Петренко С. В. Використання електронних ресурсів у науці. Київ: Либідь, 2019. 210 с.
32. Пономаренко Т. Г. Адаптація мультимовних текстів. Львів: Каменяр, 2020. 185 с.
33. Романюк М. А. Аналіз фразеологізмів у текстах. Одеса: Рута, 2021. 180 с.
34. Сидоренко Л. М. Полілінгвальні бази даних. Київ: Наукова думка, 2019. 200 с.
35. Січовий А. І. Використання штучного інтелекту у лінгвістиці. Львів: Видавництво ЛНУ, 2021. 230 с.
36. Соломко І. Г. Електронні довідники: сучасний стан. Одеса: ОНУ, 2020. 190 с.
37. Стародуб Ю. В. Мовне різноманіття у цифровому світі. Київ: Либідь, 2018. 250 с.
38. Степаненко А. П. Розробка мультимовних платформ. Дніпро: ДНУ, 2021. 200 с.
39. Тарасенко М. В. Адаптація ресурсів для користувачів. Львів: ЛНУ, 2020. 190 с.
40. Тимченко В. Ю. Багатомовні довідники у наукових дослідженнях. Харків: Освіта, 2019. 210 с.
41. Тищенко О. П. Електронні ресурси в освіті. Київ: Либідь, 2021. 190 с.
42. Третяк Ю. С. Мультимедійні компоненти у довідниках. Львів: Каменяр, 2020. 180 с.
43. Уманець П. Г. Використання діалектів у текстах. Ужгород: УжНУ, 2021. 185 с.
44. Федоренко Г. А. Структурна організація інформаційних ресурсів. Київ: Наукова думка, 2020. 220 с.
45. Хоменко Т. Л. Лінгвістичний аналіз у сучасних системах. Харків: Видавництво ХНУ, 2019. 210 с.

46. Чернишенко Н. О. Збереження мовного різноманіття. Львів: Каменяр, 2018. 175 с.
47. Чумак С. І. Аналіз багатомовних текстів. Дніпро: Дніпропрес, 2020. 198 с.
48. Шевченко П. Ю. Інтерактивні інструменти у довідниках. Одеса: Рута, 2021. 180 с.
49. Шостак В. А. Інтеграція технологій у лінгвістику. Харків: Освіта, 2018. 200 с.
50. Щербак С. Г. Роль довідників у дослідницькій діяльності. Київ: Либідь, 2021. 180 с.
51. Юрченко І. М. Використання нейронних мереж у перекладі. Львів: Видавництво ЛНУ, 2020. 230 с.
52. Яворський Т. С. Мовна підтримка у цифрових платформах. Дніпро: ДНУ, 2019. 210 с.
53. Яковенко М. О. Перспективи багатомовних ресурсів. Харків: Видавництво ХНУ, 2021. 220 с.
54. Яременко Л. П. Створення електронних довідників. Одеса: ОНУ, 2020. 185 с.
55. Яценко О. Г. Полілінгвізм у штучному інтелекті. Київ: Наукова думка, 2018. 200 с.