

ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ

УДК 004.9

DOI <https://doi.org/10.32782/EIS/2024-106-14>

АНАЛІЗ НЕСТРУКТУРОВАНИХ ДАНИХ КОНТАКТ-ЦЕНТРУ ДЛЯ ПІДТРИМКИ ПРИЙНЯТТЯ РІШЕНЬ

Гнатушенко Володимир Володимирович,

доктор технічних наук, професор,
завідувач кафедри інформаційних технологій та комп'ютерної інженерії
Національного технічного університету «Дніпровська політехніка»
ORCID: 0000-0003-3140-3788
SCOPUS-AUTHOR ID: 6505609275

Каштан Віта Юрїївна,

кандидат технічних наук, доцент,
доцент кафедри інформаційних технологій та комп'ютерної інженерії
Національного технічного університету «Дніпровська політехніка»
ORCID: 0000-0002-0395-5895
SCOPUS-AUTHOR ID: 57201902879

Іванько Артем,

аспірант кафедри інформаційних технологій та комп'ютерної інженерії
Національного технічного університету «Дніпровська політехніка»
ORCID: 0009-0002-0491-5374

Овчаренко Максим,

аспірант кафедри інформаційних технологій та комп'ютерної інженерії
Національного технічного університету «Дніпровська політехніка»
ORCID: 0009-0006-0730-0913

Мета – розроблення методу аналізу неструктурованих даних контакт-центру, який забезпечує інтеграцію й ефективно оброблення запитів клієнтів із різних каналів комунікації, зокрема аудіо- і текстових даних.

Методи. У дослідженні запропоновано схему оброблення неструктурованих даних в інформаційних системах контакт-центру, яка включає можливості аналізу інформації в аудіо- (Voice Data) та текстовому (Text Data) форматах. Система забезпечує обслуговування запитів, отриманих через різні канали, як-от телефонний зв'язок, системи VoIP, аудіо-конференц-зв'язок, а також текстові повідомлення з онлайн-чатів, месенджерів та вбудованих чатів. Для подальшого оброблення голосові дані конвертуються в текст за допомогою модуля Speech To Text. Інформацію про емоційний фон розмови виділяють та зберігають за допомогою модулів Speech Emotion Recognition та Sentiment Analysis. Для реалізації повнотекстового пошуку, а також пошуку даних за емоціями і тональністю використовуються модулі Full Text Search та Emotion Sentiment Search. Модуль Text Classification дозволяє створювати класифікатори та організовувати пошук відповідно до отриманих моделей.

Результати. Дослідження показало, що середній час розмови становив 15 хвилин, з яких 65% часу було активним мовленням, а 35% – часом тиші. Мовлення оператора становило приблизно 60% від загального часу активного мовлення, тоді як мовлення клієнта – 40%. Середня швидкість мовлення оператора становила 150 слів на хвилину, водночас швидкість мовлення клієнта була 130 слів на хвилину. Також визначено кількість перебувань: у середньому фіксувалося 8 перебувань за розмову, з яких 5 належало оператору, а 3 – клієнту. Ці результати дозволяють зробити висновки щодо якості взаємодії під час дзвінків і дотримання стандартів комунікації. Аналіз тональності та емоцій в аудіо- і текстових даних виявив, що 70% розмов мали нейтральний тон, 20% – позитивний, а 10% – негативний. У 30% випадків у мовленні клієнтів спостерігалися ознаки роздратування, тоді як емоційна нейтральність у мовленні операторів відзначалася у 80% випадків.

Новизна. У роботі запропоновано новий метод оброблення неструктурованих даних контакт-центру, що дозволяє аналізувати інформацію як в аудіо, так і в текстовому форматах та забезпечує інтеграцію запитів із різних каналів зв'язку. Використання модуля перетворення голосових даних на текст, а також модулів для емоційного аналізу сприяє отриманню більш детальної інформації про психоемоційний

стан учасників розмови. Аугментція аудіоданих розширює параметри аналізу взаємодії між операторами та клієнтами, що відкриває нові можливості для підвищення ефективності бізнес-процесів контакт-центрів та вдосконалення клієнтського обслуговування.

Цінність. Результати дослідження мають практичну цінність для компаній, що працюють у сфері обслуговування клієнтів, оскільки дозволяють покращити якість обслуговування, швидкість реагування на запити, а також підвищити рівень задоволеності клієнтів. Запропоновані рішення можуть бути адаптовані до різних інформаційних систем контакт-центру, що робить їх універсальними для впровадження в різні організації.

Ключові слова: контакт-центр, автоматичне розпізнавання мовлення, розпізнавання мовних емоцій, аналіз настроїв, семантична подібність, мовна модель.

Hnatushenko Volodymyr, Kashtan Vita, Ivanko Artem, Ovcharenko Maksym. Analysis of unstructured contact center data to support decision-making

Purpose. To develop a method for analyzing unstructured contact center data that enables the integration and efficient processing of customer inquiries from various communication channels, including both audio and text data.

Methods. The study proposes a data processing framework for contact center information systems that can analyze audio (Voice Data) and text (Text Data) formats. The system processes inquiries from multiple channels, such as phone calls, VoIP systems, audio conferencing, and text messages from online chats, messaging apps, and embedded chats. Voice data is converted into text using a Speech-to-Text module for further processing. Emotional tone in conversations is identified and stored using Speech Emotion Recognition and Sentiment Analysis modules. Full-text and emotion sentiment searches are implemented to retrieve data based on keywords or emotions. The Text Classification module allows the creation of classifiers to organize search results according to the generated models.

Results. The study found that the average conversation lasted 15 minutes, with 65% of the time involving active speech and 35% being periods of silence. The operator accounted for approximately 60% of the active speaking time, while the customer accounted for 40%. The operator's average speech rate was 150 words per minute, whereas the customer's was 130 words per minute. Additionally, there was an average of 8 interruptions per conversation, 5 by the operator and 3 by the customer. These results provide insights into the quality of interactions during calls and adherence to communication standards. Sentiment and emotion analysis of audio and text data revealed that 70% of conversations had a neutral tone, 20% were positive, and 10% were negative. Signs of frustration were observed in 30% of customer speech, while operator speech was emotionally neutral in 80% of cases.

Originality. The study introduces a new method for processing unstructured contact center data, enabling the analysis of audio and text formats and integrating inquiries from different communication channels. Speech-to-text conversion and emotion analysis modules provide deeper insights into conversation participants' psychological and emotional states. The augmentation of audio data expands the parameters for analyzing interactions between operators and customers, opening up new opportunities for enhancing the efficiency of contact center business processes and improving customer service.

Practicality. The research findings have practical value for companies in the customer service industry, as they offer ways to improve service quality, response time, and overall customer satisfaction. The proposed solutions can be adapted to various contact center information systems, making them versatile for implementation across different organizations.

Key words: contact center, automatic speech recognition, speech emotion recognition, sentiment analysis, semantic similarity, language model.

Актуальність. Реалізація ефективних комунікацій є необхідною складовою частиною бізнес-процесів суб'єктів будь-якої форми власності. Особливої актуальності це набуває у сферах клієнтського обслуговування, маркетингу та управління персоналом, де важливо отримати інсайди з великого обсягу голосових і текстових даних. Сучасний бізнес висуває високі вимоги до якості обслуговування клієнтів і необхідності сталого покращення сервісу та культури обслуговування за допомогою персональної взаємодії з клієнтами. Одним з інструментів організації персонального контакту з клієнтом є контакт-центр.

Для покращення якості обслуговування клієнтів важливими питаннями є зростання показників результативності та скорочення вартості оброблення звернень до контакт-центру, але

відповідна автоматизація реалізована лише для дуже обмеженого набору бізнес-процесів. Це пояснюється великою кількістю різноманітних факторів, що впливають на процеси оброблення різнорідних даних контакт-центру.

Серед головних складнощів щодо оброблення та аналізу даних сучасного контакт-центру виділяють: велику кількість неструктурованих текстових та голосових даних; відсутність об'єктивної оцінки розмови або значні похибки, можливість аналізувати випадкові розмови; оцінювання KPI операторів здійснюється за непрямими параметрами, а не за їх безпосередньою активністю під час розмов; неавтоматизований аналіз розмов може зайняти багато часу та бути дорогим через людський фактор, залучений до такого оцінювання.

Подальший розвиток інформаційних систем підтримки прийняття рішень під час оброблення неструктурованих даних контакт-центрів неможливий без використання технологій перетворення мовлення в текст (Speech-to-Text, далі – STT) для проведення транскрибування всього об'єму мовних даних.

Аналіз літературних джерел. Тенденція зростання як кількості учасників, що потребують поглибленого аналізу даних під час комунікації, так і зростання об'єму даних (голосових, відео, текстових) визначає актуальність саме автоматизованого оброблення даних. Подолання цих комплексних проблем є складною окремою задачею, яка потребує технічного і методологічного розв'язання [1]. Використання технологій штучного інтелекту дає можливість підвищення ефективності всіх бізнес-процесів, у тому числі прискорюючи прийняття рішень та підвищуючи продуктивність підприємств [2, с. 824] завдяки розв'язанню проблеми трудомісткого процесу ручного оброблення великого обсягу неструктурованих даних, що потребує багато часу та фінансових ресурсів [3].

Головним параметром продуктивності моделі перетворення мовлення в текст STT, що застосовується в системі автоматичного розпізнавання мовлення (Automatic Speech Recognition, далі – ASR), є показник частоти помилок слів word error rate (далі – WER) [4, с. 1]. Стрімкий розвиток моделей ASR, як комерційних, серед яких Amazon Transcribe, Azure Speech-to-Text,

Google Speech-to-Text, IBM Watson Speech-to-Text, так і моделей з відкритим кодом, як-от Deep Speech, Conformer, HuBERT, SpeechBrain, WhisperX та SpeechStew, дозволяють досягати значень показника WER до 5% включно [4, с. 6; 5, с. 7]. Це у свою чергу створює передумови широкого застосування моделей ASR для оброблення неструктурованих мовних даних контакт-центру в складі інформаційних систем.

Мета роботи – розроблення методу аналізу неструктурованих даних контакт-центру, який забезпечує інтеграцію й ефективне оброблення запитів клієнтів із різних каналів комунікації, зокрема аудіо– та текстових даних. Досягнення мети пов'язано з розв'язанням таких задач, як:

- розроблення способу аугментації аудіо-даних контакт-центру, що дозволить отримати додаткові параметри для аналізу;

- аналіз тональності та емоцій у розмовах для оцінювання якості взаємодії між операторами та клієнтами.

Матеріали та методи. Контакт-центри відіграють вирішальну роль в обслуговуванні та підтримці клієнтів різних підприємств. Вони обробляють широкий спектр запитів клієнтів, починаючи від технічної підтримки на запити щодо виставлення рахунків. Ефективність цих взаємодій безпосередньо впливає на задоволеність клієнтів і, зрештою, на успіх у бізнесі.

У роботі запропоновано схему аналізу неструктурованих даних контакт-центру для підтримки прийняття рішень (рис. 1).

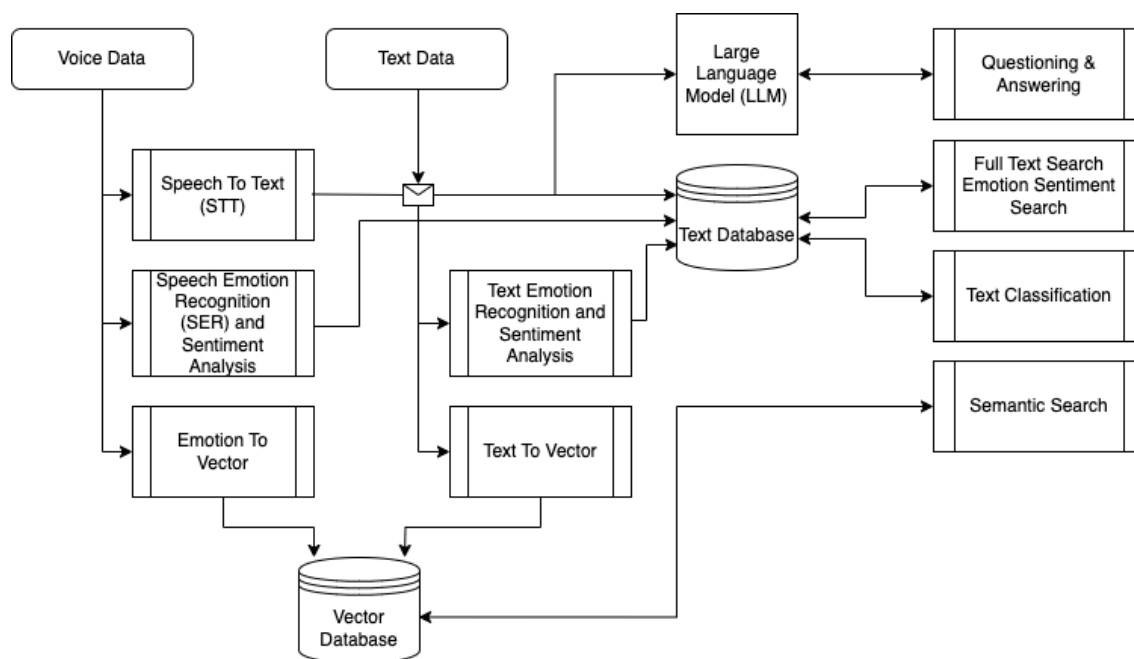


Рис. 1. Схема оброблення даних контакт-центру

Запропонована схема аналізу неструктурованих даних в інформаційних системах контакт-центру включає можливості оброблення інформації як у форматі аудіо (Voice Data), так і в текстовому (Text Data). Це дозволяє обслуговувати запити, отримані різними каналами: голосом (телефонний зв'язок, системи VoIP, та аудіо-конференц-зв'язку), текстовими повідомленнями (онлайн-чати, месенджери, вбудовані чати тощо). Для подальшої роботи голосові дані перетворюються на текст за допомогою модуля Speech-to-Text. За допомогою модулів Speech Emotion Recognition та Sentiment Analysis виділяється та зберігається інформація, щодо емоційного фону розмови. Для реалізації повнотекстового пошуку, пошуку даних за емоціями та тональністю, використовується відповідні модулі Full Text Search Emotion Sentiment Search. Модуль Text Classification дозволяє створювати моделі класифікаторів та організувати пошук згідно з отриманими моделями. Класифікація тексту є важливою проблемою оброблення природної мови. Деякі проблеми розпізнавання тексту показали, що використання глибоко навчених нейронних мереж не завжди є переконливим сучасним методом. Тому для використання пропонується використання методів, що зменшують на кілька порядків величини складності пам'яті певних текстових класифікаторів без шкоди для точності та швидкості [6, с. 8].

Текстові дані після перетворення з аудіоданих або дані текстових повідомлень додатково обробляються модулями Text Emotion Recognition та Sentiment Analysis, що дозволяє отримувати більш достовірні дані щодо психоемоційного стану співрозмовників завдяки об'єднанню результатів декількох рівнів – голосового та текстового, а надалі ще й аналіз відео, що може збагатити дані щодо емоцій за допомогою аналізу емоцій обличчя, жестів та пози тіла. Текстові дані направляються до сервісу Text Database, що забезпечує повнотекстове індексування та можливість пошуку. Для розв'язання задач порівняння семантичної подібності, кластеризації та семантичного пошуку необхідно перетворення текстових даних у векторний формат за допомогою модуля Text to Vector з використанням, наприклад, технології глибокого навчання нейронних мереж типу трансформер – Sentence Bidirectional Encoder Representations from Transformers (SBERT), яка демонструє обчислювальну ефективність у порівнянні з іншими перетвореннями [7, с. 8]. Для зберігання даних після перетворення до векторного вигляду використовується векторна база даних, позначена модулем Vector Database.

Інтеграція мовних моделей (Languages Models) в процеси контакт-центру можуть революціонізувати галузь [8, с. 1]. Використання модуля великої мовної моделі LLM, що навчені на величезній кількості даних, продемонстрували ефективність під час розв'язання низки завдань природного оброблення мови, включаючи машинний переклад, сентимент-аналіз і підсумовування тексту, а можливість взаємодії з LLM у форматі запитання – відповідь за допомогою модуля Questioning Answering дозволяє отримувати додаткову інформацію з використанням природної мови.

Розроблення, реалізація та використання інформаційних систем оброблення неструктурованих даних контакт-центру розширеного рівня складності потребує створення нових комплексних показників ефективності бізнес-процесів контакт-центрів, з урахуванням різноманітності параметрів діалогу, в тому числі оцінок тональності та емоційного забарвлення протягом розмови, характеристик мовлення клієнта й оператора, відповідності вимогам сценаріїв та релевантних відповідей з наявними базами знань контакт-центра.

Результати дослідження. У ході дослідження було протестовано запропонований метод оброблення неструктурованих даних контакт-центру. Набір додаткових параметрів, отриманих за результатами аугментації, наведено на рисунку 2. Аугментація аудіоданих значно підвищила кількість метрик і параметрів, які можуть бути використані для детального аналізу взаємодії між операторами та клієнтами. Аудіодані проходять етапи аугментації, після чого вони розбиваються на аудіо– та текстові сегменти (chunks). Це дозволяє виділити як технічні характеристики аудіо (наприклад, тривалість, бітрейт тощо), так і змістові параметри розмов, що є важливими для подальшого аналізу.

Аналіз аудіофайлів включав визначення основних технічних характеристик, як-от формат файлу, кодек, бітрейт, частота вибірки, глибина бітів, кількість каналів, тривалість файлу та його розмір. У процесі аугментації додатково було отримано такі кількісні показники: загальна тривалість розмови, час активного мовлення (з поділом на оператора та клієнта), відсоткове співвідношення мовлення оператора і клієнта, час тиші та кількість перебивань.

Відсоток активного мовлення можна розрахувати за формулою (1):

$$P_{\text{мовлення}} = \frac{T_{\text{мовлення}}}{T_{\text{розмови}}} \times 100\% , \quad (1)$$

де $P_{\text{мовлення}}$ – відсоток активного мовлення, $T_{\text{розмови}}$ – загальна тривалість розмови, а $T_{\text{мовлення}}$ – час активного мовлення.

Співвідношення часу мовлення оператора до клієнта розраховується за формулою (2):

$$R_{\text{оператор / клієнт}} = \frac{T_{\text{оператор}}}{T_{\text{клієнт}}}, \quad (2)$$

де $R_{\text{оператор/клієнт}}$ – співвідношення часу мовлення оператора до клієнта, $T_{\text{оператор}}$ – час мовлення оператора, а $T_{\text{клієнт}}$ – час мовлення клієнта.

Середню швидкість мовлення (кількість слів на хвилину) можна обчислити за формулою (3):

$$S_{\text{слів/хв}} = \frac{W}{T_{\text{хвилини}}}, \quad (3)$$

де W – кількість слів, сказаних за час, $T_{\text{хвилини}}$ – середня швидкість мовлення (кількість слів на хвилину).

Якщо за час розмови $T_{\text{розмови}}$ було зафіксовано $I_{\text{оператор}}$ перебивань оператором та $I_{\text{клієнт}}$ перебивань клієнтом, то загальну кількість перебивань можна розрахувати за формулою (4):

$$I_{\text{загальне}} = I_{\text{оператор}} + I_{\text{клієнт}}. \quad (4)$$

У табл. 1 наведено метрики, отримані в процесі аналізу аудіо- та текстових даних.

Таким чином, середній час розмови становив 15 хвилин, з яких 65% часу було активним мовленням, а 35% – часом тиші. Мовлення оператора становило приблизно 60% від загального часу активного мовлення, тоді як мовлення клієнта – 40%. Середня швидкість мовлення оператора становила 150 слів на хвилину, а клієнта – 130 слів на хвилину.

Крім того, було визначено кількість перебивань, що дозволяє оцінити динаміку розмови. У середньому під час розмови фіксувалося 8 перебивань, з яких 5 належало оператору, а 3 – клієнту. Це дозволяє зробити висновки

Таблиця 1

Кількісні результати аналізу розмов між операторами та клієнтами контакт-центру

Метрика	Значення
Середній час розмови	15 хвилин
Час активного мовлення	65% від загального часу розмови
Час тиші	35% від загального часу розмови
Мовлення оператора	60% від загального часу активного мовлення
Мовлення клієнта	40% від загального часу активного мовлення
Швидкість мовлення оператора	150 слів на хвилину
Швидкість мовлення клієнта	130 слів на хвилину
Кількість перебивань (загальна)	8 перебивань
Кількість перебивань оператора	5 перебивань
Кількість перебивань клієнта	3 перебивання
Тональність розмов (сентимент-аналіз)	Нейтральна – 70%, Позитивна – 20%, Негативна – 10%
Емоційний стан клієнтів (роздратування)	30% випадків
Емоційна нейтральність операторів	80% випадків

щодо якості взаємодії під час дзвінків та дотримання стандартів комунікації.

Дискусія. Результати представленого дослідження підтверджують ефективність запропонованого методу аналізу неструктурованих даних контакт-центру. Під час експерименту було отримано параметри (тривалість розмови, час активного мовлення, співвідношення мовлення між оператором і клієнтом, швидкість мовлення, кількість перебивань, а також тональність і емоційний стан учасників), що дозволили значно поглибити аналіз розмов

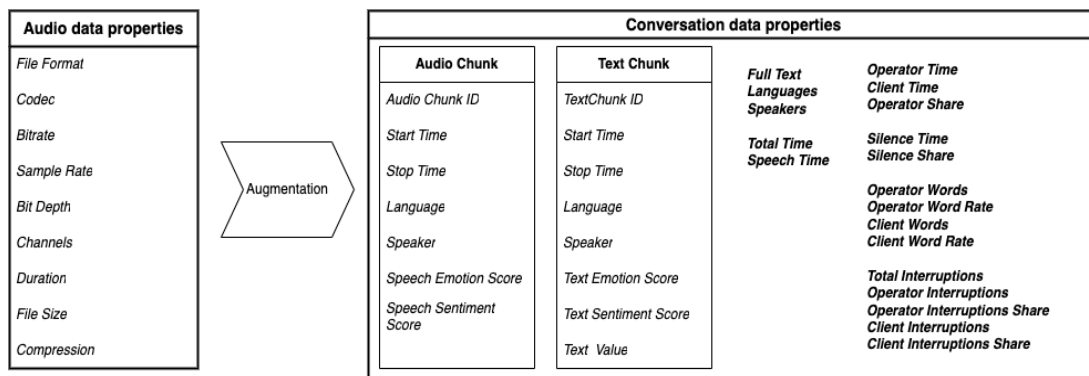


Рис. 2. Розширення параметрів розмови за результатами аугментації аудіоданих контакт-центру

між операторами та клієнтами. Аугментація аудіоданих не тільки розширила обсяг отриманої інформації, але й забезпечила можливість детальнішого аналізу комунікаційних процесів.

Отримані результати демонструють, що 65% часу розмови було витрачено на активне мовлення, що вказує на ефективність взаємодії між сторонами. Цей показник, разом із відсотковим співвідношенням мовлення оператора та клієнта, дає можливість оцінити не тільки динаміку розмови, але й можливі проблеми в комунікації. Виявлені середні швидкості мовлення (150 слів на хвилину для операторів та 130 слів на хвилину для клієнтів) свідчать про те, що оператори зберігають контроль над розмовою, однак деякі клієнти можуть відчувати труднощі у висловленні своїх думок. Аналіз тональності вказує на домінування нейтрального тону розмов, проте наявність 20% позитивних та 10% негативних інтервенцій потребує подальшого дослідження. Важливо розуміти, що емоційний стан клієнтів, виражений через ознаки роздратування у 30% випадків, може негативно вплинути на загальне сприйняття сервісу. З огляду на це доцільно ввести додаткові тренінги для операторів з метою покращення емоційного сприйняття та реакції на клієнтські запити. Серед виявлених перебивань, більшість з яких належить операторам, можна зробити висновок, що це може свідчити про прагнення операторів контролювати хід розмови. Однак надмірна кількість перебивань може бути сприйнята клієнтами як неввічливість

і зменшити їх задоволеність від обслуговування. Тому важливо знайти баланс між активним веденням розмови та можливістю клієнта вільно висловлювати свої думки.

Загалом, результати цього дослідження вказують на перспективи подальшого розвитку технологій оброблення неструктурованих даних у сфері контакт-центрів. Використання аугментації даних може стати основою для розроблення нових інструментів аналітики, що дозволять більш точно оцінювати якість обслуговування та покращувати взаємодію між клієнтами та компаніями.

Висновки. Дослідження, проведене в межах цієї роботи, підтвердило ефективність запропонованої методики оброблення неструктурованих даних контакт-центру. Проведені експериментальні дослідження свідчать про можливість детального вивчення якості взаємодії між учасниками розмови, а також про підвищення рівня обслуговування клієнтів. Високий відсоток активного мовлення (65%) та зменшена кількість перебивань (у середньому 8, з яких 5 належали оператору і 3 – клієнту) вказують на наявність ефективної комунікації, що є критично важливим для досягнення бізнес-цілей у сфері обслуговування.

Подальші дослідження будуть спрямовані на інтеграцію даних, отриманих за допомогою цієї методики, з іншими системами управління бізнесом для забезпечення комплексного підходу до аналізу і покращення обслуговування.

ЛІТЕРАТУРА:

1. Dogan O. A process-centric performance management in a call center. *Appl. Intell.* 53, 3304–3317 2023, URL: <https://doi.org/10.1007/s10489-022-03740-9>
2. Shah S., Ghomeshi H., Vakaj E. et al. A review of natural language processing in contact centre automation. *Pattern Anal Applic*, 26, 823–846, 2023, <https://doi.org/10.1007/s10044-023-01182-8>
3. Chicu D., del Mar P'amies M., Ryan G., Cross C. Exploring the influence of the human factor on customer satisfaction in call centres. *BRQ Business Res Quart*, 2019, 22(2): 83–95. URL: <https://doi.org/10.1016/j.brq.2018.08.004>
4. Automatic Speech Recognition System-Independent Word Error Rate Estimation Chanhon Park, Mingjie Chen, Thomas Hain. URL: <https://doi.org/10.48550/arXiv.2404.16743>
5. Benchmarking open source and paid services for speech to text: an analysis of quality and input variety Antonino Ferraro, Antonio Galli, Valerio La Gatta, Marco Postiglione Volume 6 – 2023. <https://doi.org/10.3389/fdata.2023.1210559>
6. FastText.zip: Compressing text classification models Armand Joulin, Edouard Grave, Piotr Bojanowski, Matthijs Douze, Herve Jégou, Tomas Mikolov. URL: <https://doi.org/10.48550/arXiv.1612.03651>
7. N. Reimers, I. Gurevych, Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks (2019), Proceedings of the 2019 Conference on Empirical Methods in 2019. URL: <https://doi.org/10.48550/arXiv.1908.10084>
8. Towards Probing Contact Center Large Language Models Varun Nathan, Ayush Kumar, Digvijay Ingle, Jithendra Vepa. URL: <https://doi.org/10.48550/arXiv.2312.15922>

REFERENCES:

1. Dogan, O. (2023). A process-centric performance management in a call center. *Appl. Intell.* 53, 3304–3317 Retrieved from <https://doi.org/10.1007/s10489-022-03740-9>.
2. Shah, S., Ghomeshi, H., Vakaj, E., et al. (2023). A review of natural language processing in contact centre automation. *Pattern Anal Applic* 26, 823–846 Retrieved from <https://doi.org/10.1007/s10044-023-01182-8>
3. Chicu, D., del Mar P`amies, M., Ryan, G., Cross, C. (2019). Exploring the influence of the human factor on customer satisfaction in call centres. *BRQ Business Res Quart* 22(2): 83–95. Retrieved from <https://doi.org/10.1016/j.brq.2018.08.004>
4. Automatic Speech Recognition System-Independent Word Error Rate Estimation Chanh Park, Mingjie Chen, Thomas Hain Retrieved from <https://doi.org/10.48550/arXiv.2404.16743>
5. Benchmarking open source and paid services for speech to text: an analysis of quality and input variety Antonino Ferraro, Antonio Galli, Valerio La Gatta, Marco Postiglione Volume 6 – 2023 Retrieved from <https://doi.org/10.3389/fdata.2023.1210559>
6. FastText.zip: Compressing text classification models Armand Joulin, Edouard Grave, Piotr Bojanowski, Matthijs Douze, Herve Jégou, Tomas Mikolov. Retrieved from <https://doi.org/10.48550/arXiv.1612.03651>
7. N. Reimers, I. Gurevych, Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks (2019), Proceedings of the 2019 Conference on Empirical Methods in 2019. Retrieved from <https://doi.org/10.48550/arXiv.1908.10084>
8. Towards Probing Contact Center Large Language Models Varun Nathan, Ayush Kumar, Digvijay Ingle, Jithendra Vepa. Retrieved from <https://doi.org/10.48550/arXiv.2312.15922>