

УДК 004.89:004.912(043.3)

Ляшенко Роман Олегович

студент

Криворізький державний педагогічний університет, м. Кривий Ріг, Україна

ORCID ID 0009-0000-2614-6997

*romanslss777@gmail.com***Семеріков Сергій Олексійович**

доктор педагогічних наук, професор, старший дослідник

Криворізький державний педагогічний університет, м. Кривий Ріг, Україна

Інститут цифровізації освіти НАПН України, м. Київ, Україна

Державний університет «Житомирська політехніка», м. Житомир, Україна

Криворізький національний університет, м. Кривий Ріг, Україна

Академія когнітивних та природничих наук, м. Кривий Ріг, Україна

ORCID ID 0000-0003-0789-0272

semerikov@gmail.com

БІБЛІОМЕТРИЧНИЙ АНАЛІЗ ДОСЛІДЖЕНЬ З НАВЧАННЯ ЧАТ-БОТІВ: КЛЮЧОВІ ПОНЯТТЯ ТА ТЕНДЕНЦІЇ

Анотація. Бібліометричний аналіз проведено з метою визначення актуальних напрямків і пріоритетів досліджень у галузі навчання чат-ботів – програмних агентів, здатних вести природний діалог. Дослідження виконане на основі опрацювання 549 наукових джерел з бази даних Scopus, присвячених зазначеній проблематиці. Аналіз виявив стійке зростання кількості відповідних публікацій, починаючи з 2018 року, що свідчить про підвищення актуальності цієї тематики в останні роки. За результатами кластерного аналізу ключових слів виокремлено чотири основні напрямки досліджень: опрацювання природної мови, застосування відповідних технологій у різних сферах життя суспільства, використання методів машинного навчання для опрацювання природної мови, а також застосування чат-ботів у навчанні та сфері послуг. У сфері опрацювання природної мови фокус сучасних досліджень робиться на проблемах обчислювальної лінгвістики, моделюванні та машинному розумінні природної мови, зокрема на задачах розпізнавання мовлення. Провідні дослідження застосувань штучного інтелекту в цій галузі стосуються відповідального та етичного використання сучасних великих мовних моделей та розмовних агентів, як-от ChatGPT, Gemini, Claude, у галузях освіти та охорони здоров'я. Активно розробляються методи машинного навчання для створення віртуальних інтелектуальних помічників, природномовних інтерфейсів користувача та інших систем опрацювання природної мови, зокрема для задач діагностики в медицині. Ключові застосування чат-ботів виявлено в системах адаптивного навчання, управління знаннями та обслуговування клієнтів. За результатами аналізу визначено найбільш значущі поняття в кожній з досліджуваних сфер для окреслення пріоритетів подальших досліджень у галузі навчання чат-ботів. У перспективі планується провести систематичний літературний огляд з автоматизацією окремих етапів за допомогою великих мовних моделей. Зокрема, передбачається використати такі моделі для автоматичної класифікації анотацій досліджень за критеріями включення/виключення на етапі скринінгу. Автоматизація етапів систематичного огляду за допомогою штучного інтелекту відкриває значні перспективи для прискорення наукових досліджень, зокрема в освітній галузі на засадах доказового підходу.

Ключові слова: навчання чат-ботів; опрацювання природної мови; машинне навчання; бібліометричний аналіз; систематичний огляд літератури; великі мовні моделі.

1. ВСТУП

Постановка проблеми. Розвиток штучного інтелекту та технологій опрацювання природної мови зумовлює стрімке зростання інтересу до створення програмних агентів, здатних вести природний діалог – так званих чат-ботів. Провідні ІТ-компанії, такі як

Google [1] – [6], Microsoft [7], [8], Meta (Facebook) [9] – [11], OpenAI [12] – [16] активно працюють над розробкою чат-ботів та великих мовних моделей, що їх підтримують. Успішні проекти на кшталт ChatGPT від OpenAI [17] демонструють значний потенціал застосування таких систем у різноманітних сферах людської діяльності. У вказаних дослідження можна виділити дві основні групи задач: 1) навчання (викладання та учіння) студентів; 2) побудова чат-ботів – систем типу «питання-відповідь», які навчають власні моделі за даними, отриманими від користувача. Водночас постає низка важливих наукових та практичних завдань щодо ефективного навчання таких систем, забезпечення їх надійності, безпеки та етичності.

Чат-боти як природномовні системи штучного інтелекту мають величезний потенціал для підвищення ефективності та якості різноманітних видів людської діяльності шляхом автоматизації рутинних процесів, надання інтелектуальної підтримки користувачам, забезпечення персоналізованого навчання тощо. Успішні розробки у цій галузі здатні у перспективі докорінно змінити характер взаємодії людини з комп'ютерними системами, підвищити продуктивність праці та навчання. Разом з тим, вирішення етичних та безпекових проблем є критично важливим для впровадження цих технологій в інтересах суспільства.

Аналіз останніх досліджень і публікацій. У галузі створення та навчання чат-ботів активно працюють як науковці, так і розробники ІТ-компаній. Окремі аспекти цієї проблематики висвітлюються у низці публікацій. Зокрема результати пошуку у блозі DeepLearning.AI [18] за 2019-2023 рр. надають можливість виокремити наступні практично розв'язані задачі:

- з 2020 року чат-боти (Generative BST від Facebook та Google Meena) можна використовувати для *коротких діалогів на загальні теми* [19], [20];
- з 2023 року Microsoft надає можливість використання чат-ботів у Office 365 та Windows для *покращення продуктивності* [21];
- з 2020 року чат-боти активно *застосовуються в бізнесі* для обслуговування клієнтів, продажу тощо [22];
- з 2023 року Google та Microsoft вводять чат-боти у *пошукові системи* [23].

Статті у блозі DeepLearning.AI вказують також на прогалини знань та відкриті проблеми, що існують у даній галузі:

- медичні чат-боти можуть ставити помилкові діагнози [24];
- такі чат-боти, як-то BlenderBot 3 та Galactica, можуть виявляти упередженість, токсичність та перекручувати факти [25];
- зависока вартість розгортання великих мовних моделей [26];
- недостатня прозорість моделей чат-ботів може викликати етичні питання [27].

Аналіз останніх досліджень і публікацій виявляє необхідність консолідації зусиль для визначення актуальних напрямів і пріоритетів досліджень з метою ефективного навчання чат-ботів.

Мета дослідження: на основі бібліометричного аналізу визначити ключові поняття, актуальні напрямки та тенденції досліджень у галузі навчання чат-ботів – програмних агентів, здатних вести природний діалог.

2. МЕТОДИКА ДОСЛІДЖЕННЯ

Відповідно до мети визначено такі **завдання дослідження:**

- 1) провести бібліометричний аналіз джерел з проблеми навчання чат-ботів для визначення ключових понять дослідження;
- 2) визначити найбільш значущі ключові поняття дослідження та побудувати карту

їх зв'язків.

Методи дослідження:

- бібліометричний аналіз результатів пошуку для а) виявлення хронологічної межі, починаючи з якої спостерігається стійке зростання кількості публікацій за темою дослідження; б) для створення карти зв'язків ключових слів; в) для групування ключових слів у кластери; г) для визначення ключових понять дослідження;
- використання великих мовних моделей для генерації припущень щодо назв кластерів ключевих слів.

3. РЕЗУЛЬТАТИ

Для бібліометричного аналізу було обрано наукометричну базу даних Scopus видавництва Elsevier як одну з найбільших реферативних баз даних, що індексує більше 42 тис. періодичних видань та містить переважно якісні наукові джерела з технічних, природничих, медичних та суспільних наук, що робить її репрезентативною для аналізу публікацій з проблематики навчання чат-ботів на стику кількох наукових галузей. Збалансоване покриття різних наукових дисциплін у Scopus забезпечує необхідну повноту для виявлення ключових тенденцій та напрямків досліджень в обраній галузі.

Для визначення, в яких роботах зустрічаються ключові поняття дослідження «навчання» та «чат-боти», 18 листопада 2023 року було виконано пошук у базі Scopus [28]. Відбирались ті документи, у назвах, анотаціях та ключових словах яких одночасно зустрічались слова «chatbot» і «training» (рис. 1).

The screenshot shows the Scopus search results page. At the top, there is a search bar with the query 'chatbot AND training'. Below the search bar, it indicates '549 documents found'. The results are displayed in a table with columns for Document title, Authors, Source, Year, and Citations. Two results are visible:

Document title	Authors	Source	Year	Citations
1 Exploring the design and utility of an integrated web-based chatbot for young adults to support healthy eating: a qualitative study	Ashton, L.M., Adam, M.T., Whatnall, M., ...Hansen, V., Collins, C.E.	International Journal of Behavioral Nutrition and Physical Activity, 20(1), 119	2023	0
2 Investigating ChatGPT and cybersecurity: A perspective on topic modeling and sentiment analysis	Okey, O.D., Udo, E.U., Rosa, R.L., Rodriguez, D.Z., Kleinschmidt, J.H.	Computers and Security, 135, 103476	2023	0

Рис. 1. Результати пошуку в базі Scopus за словами «chatbot» і «training»

Загальна кількість документів – 549; розподіл документів за роками подано на рис. 2.

Перегляд ключових слів показує, що найбільш часто серед них зустрічаються такі: Chatbot (244 рази), Chatbots (221), Artificial Intelligence (126), Natural Language Processing Systems (90), Natural Language Processing (90). Впорядкування ключових слів за алфавітом надає можливість визначити, які з них відрізняються лише за числом (в однині та у множині), та призвести всі до однини:

- Chatbots є множиною від Chatbot;

- Conversational Agents є множиною від Conversational Agent;
- Convolutional Neural Networks є множиною від Convolutional Neural Network;
- Customer Services є множиною від Customer Service (Customer-service);
- Dialogue Generations є множиною від Dialogue Generation;
- Humans є множиною від Human;
- Large Language Models є множиною від Large Language Model;
- LSTM є абрєвіатурою Long Short-term Memory;
- Machine-learning є синонімом Machine Learning;
- Mobile Applications є множиною від Mobile Application;
- NLP є абрєвіатурою від Natural Language Processing (NAtural Language Processing);
- NLU є абрєвіатурою від Natural Language Understanding;
- Reinforcement Learnings є множиною від Reinforcement Learning;
- Virtual Assistants є множиною від Virtual Assistant.

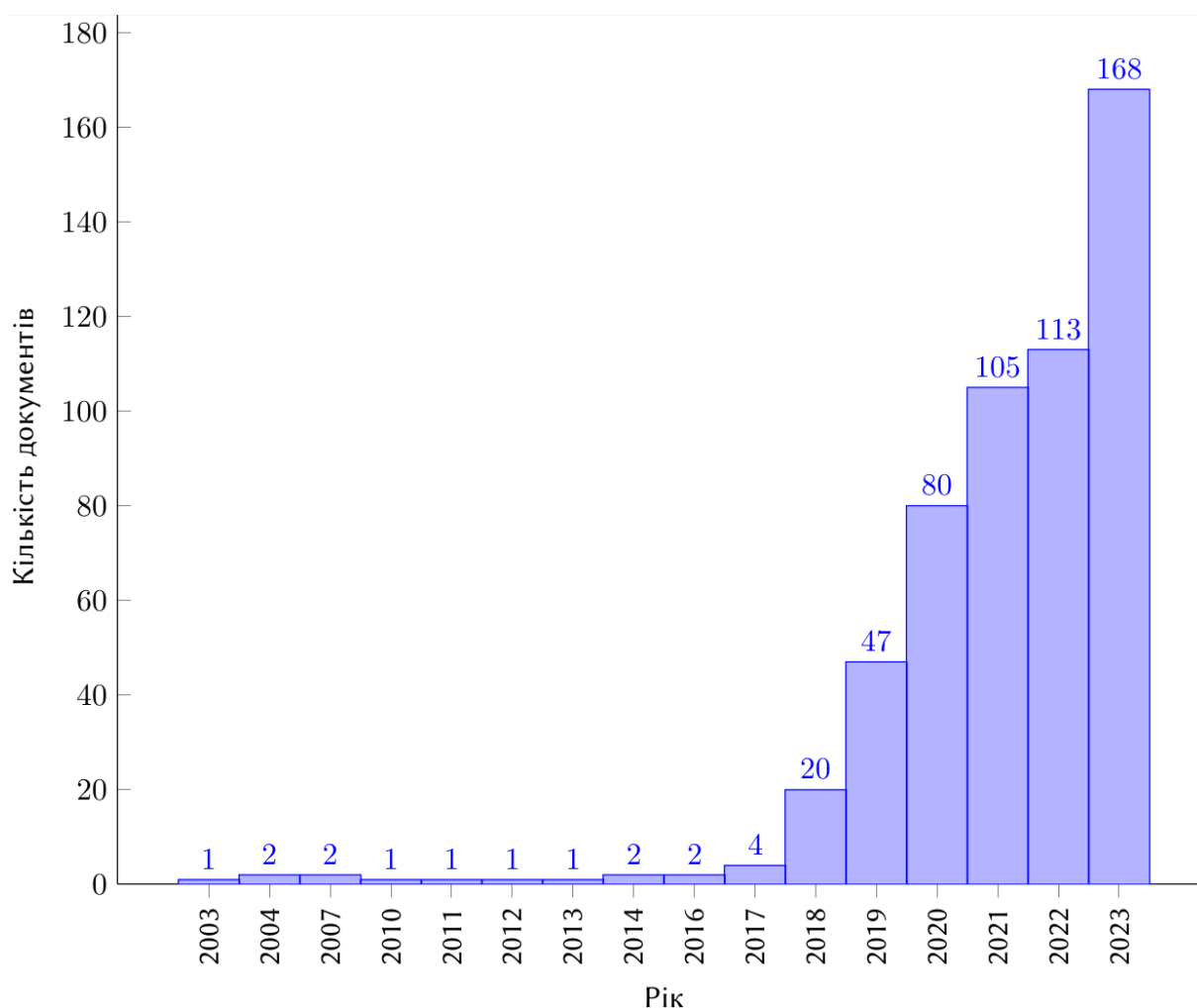


Рис. 2. Розподіл результатів пошуку за роками

Знайдені документи були експортовані у форматах CSV [29, с. 30] та BibTeX. У файлі CSV були виконані заміни ключових слів у множині та абрєвіатур на ключові слова в однині.

Для виконання бібліометричного аналізу був використаний VOSviewer 1.6.20 [30]: Create... → Create a map based on bibliographic data → Read data from bibliographic database file

→ *Scopus*.

Карта створюється з такими параметрами: тип аналізу – за одночасною появою терміну (co-occurrence); одиниця аналізу – усі ключові слова; метод підрахунку – повний підрахунок (full counting). VOSviewer автоматично групує ключові слова у кластери на основі їх семантичної близькості та візуалізує результати на карті, використовуючи різні кольори для позначення окремих кластерів.

Загальна кількість ключових слів – 3705 – може бути зменшена до 58 відкиданням тих з них, що зустрічаються менше ніж 12 разів. Із відібраних слів додатково вилучено слово «article» як таке, що не є специфічним для пошукового запиту.

Побудовану карту (табл. 1) подано на рис. 3.

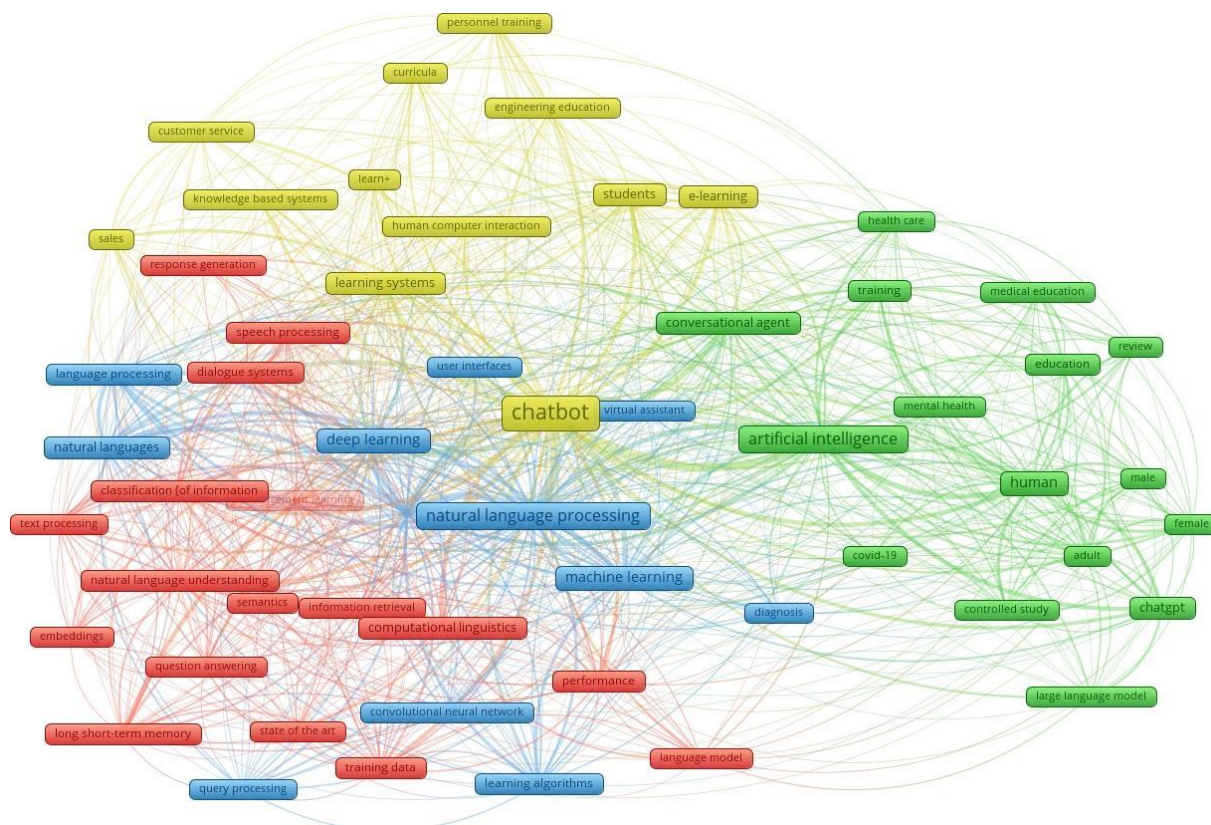


Рис. 3. Карта зв'язків ключових слів

Таблиця 1

Розподіл ключових слів за кластерами

Ключове слово	Кластер	W_{Links}	$W_{Total\ link\ strength}$	$W_{Occurrences}$	$S_{Avg.\ pub.\ year}$	$S_{Avg.\ citations}$	$S_{Avg.\ norm.\ citations}$	Значущість
classification (of information)	1	37	122	24	2021.2083	3.7917	0.5908	3
training data	1	36	120	27	2021.1111	12.8519	1.8358	3
speech processing	1	36	102	26	2020.4615	9.4231	1.1619	1
performance	1	36	88	24	2021.5833	6.2917	1.02	2
semantics	1	34	90	18	2020.7222	8.8333	0.9379	
dialogue systems	1	33	96	23	2020.4348	8.4348	1.0096	
natural language understanding	1	32	123	32	2020.9375	6.0938	0.8055	3
computational linguistics	1	32	114	33	2020.1515	16.1515	2.0909	6

Ключове слово	Кластер	W_{Links}	$W_{Total\ link\ strength}$	$W_{Occurrences}$	$S_{Avg.\ pub.\ year}$	$S_{Avg.\ citations}$	$S_{Avg.\ norm.\ citations}$	Значущість
text processing	1	31	97	16	2021.25	4.4375	0.6438	
long short-term memory	1	30	113	24	2020.7917	4.8333	0.6383	
state of the art	1	29	54	14	2021.0714	4.0714	0.5852	
language model	1	28	61	14	2021.9286	3.5	0.6501	2
question answering	1	28	60	12	2021.1667	1.4167	0.2476	
information retrieval	1	27	51	12	2020.4167	9.3333	0.9633	
embeddings	1	24	53	13	2020.8462	7.5385	0.6933	
reinforcement learning	1	23	52	15	2020.0667	9.0667	1.1396	
response generation	1	17	37	12	2020.75	11.5833	1.343	
artificial intelligence	2	52	452	126	2021.3333	9.8571	1.8428	6
conversational agent	2	47	203	50	2020.18	18.68	1.9112	1
human	2	37	266	60	2022.1167	15.7667	2.7326	2
adult	2	29	117	17	2021.9412	17	2.3017	
training	2	29	84	22	2021.7727	16.7273	3.5069	2
male	2	27	94	13	2021.4615	23.3846	3.0896	1
medical education	2	27	66	12	2022	3.5833	0.557	
female	2	26	101	16	2021.375	22.6875	2.6942	
controlled study	2	26	76	15	2022.2	9.2	1.6989	1
health care	2	26	67	15	2021.4	3.6	0.51	
mental health	2	25	56	13	2021.3846	22.6154	2.2811	
chatgpt	2	24	124	35	2023	4.8	1.4989	2
education	2	24	77	19	2021.1053	12.4211	3.3792	1
covid-19	2	23	47	13	2021.9231	3.5385	0.884	
review	2	19	70	12	2021.9167	34.25	3.3067	2
large language model	2	19	50	12	2023	2.9167	0.9108	2
natural language processing	3	54	505	100	2021.28	6.29	1.1837	6
natural language processing systems	3	52	484	90	2020.7667	5.5667	0.6711	3
machine learning	3	50	276	51	2021.4314	8.9412	1.0081	
deep learning	3	46	265	56	2021.1429	5.6607	0.6799	
natural languages	3	42	235	40	2021.5	4.1	0.6042	
language processing	3	37	165	25	2022.56	1.6	0.5094	2
learning algorithms	3	37	153	26	2021.0769	3.4231	0.4283	
virtual assistant	3	35	67	15	2021.0667	13.3333	1.5093	3
user interfaces	3	29	61	14	2020.8571	8.9286	1.7491	2
convolutional neural network	3	26	53	12	2020.9167	6.4167	0.7242	
query processing	3	24	82	14	2021.1429	1.4286	0.3373	
diagnosis	3	24	53	13	2021.5385	9.8462	1.3824	2
chatbot	4	56	1065	365	2021.137	6.3014	0.901	6
learning systems	4	46	223	43	2020.2326	7.186	0.4899	3
students	4	42	176	38	2020.9737	8.2105	0.9839	
e-learning	4	35	128	37	2020.8378	5.0541	0.6268	
learn+	4	27	53	13	2021.7692	3.9231	0.6622	2
sales	4	26	64	17	2021	5.5882	0.6926	
human computer interaction	4	26	53	14	2019.8571	6.0714	0.5181	
knowledge based systems	4	26	51	13	2019.5385	15.5385	0.7796	2
personnel training	4	25	62	18	2020.8889	6.5556	1.4229	1
curricula	4	24	58	12	2021	9.5833	2.41	3
customer service	4	24	49	15	2021.2	4.2667	0.4819	1
engineering education	4	24	46	12	2020.75	4.5833	0.4115	

У табл. 1 використовуються такі позначення [29, с. 6, 38]:

- ключове слово – визначений автором або системою індексування термін, характерний для публікації;
- кластер – номер групи ключових слів;
- W_{Links} – кількість зв'язків слова з іншими словами;
- $W_{Total\ link\ strength}$ – загальна сила зв'язків слова з іншими словами (тут кількість публікацій, у яких два терміни зустрічаються разом);
- $W_{Occurrences}$ – кількість документів, у яких зустрічається ключове слово;
- $S_{Avg.\ pub.\ year}$ – середній рік публікації документів, у яких зустрічається ключове слово;
- $S_{Avg.\ citations}$ – середня кількість цитувань документів, у яких зустрічається ключове слово;
- $S_{Avg.\ norm.\ citations}$ – середня нормалізована кількість цитувань документів, у яких зустрічається ключове слово;
- значущість – обчислюється як сума, де для кожного ключового слова обирається 2, якщо воно має найбільше значення за колонкою (позначені золотим кольором), та 1, якщо воно має значення, що передує найбільшому (позначені срібним кольором).

У додатку подано результати застосування генеративних чат-ботів ChatGPT (GPT 3.5), Bing (GPT 4.0, 3 режими), Claude 2 та Google Bard для визначення імен кластерів (табл. 2).

Таблиця 2

Імена кластерів ключових слів

№	Пропозиції від чат-ботів	Обране ім'я
1	Advanced Language Processing and AI Applications (A.1); Language Processing and Information Retrieval (A.2); Natural Language Processing (NLP) (A.3); Natural Language Understanding and Generation (A.4); Natural Language Processing Applications (A.5); Technical NLP (A.6)	Опрацювання природної мови / Natural Language Processing (A.3)
2	AI in Healthcare and Social Context (A.1); AI in Healthcare and Education (A.2); Artificial Intelligence (A.3); Artificial Intelligence and Health Care (A.4); Conversational AI and Impacts (A.5); Applications of NLP (A.6)	Застосування технологій опрацювання природної мови в суспільстві / Applications of Natural Language Processing in Society (A.6, A.5, A.1)
3	Core Technologies in Natural Language Processing and Machine Learning (A.1); Machine Learning and User Interaction (A.2); Machine Learning (A.3); Natural Language Processing and Machine Learning: (A.4); Core AI and ML Techniques (A.5); Machine Learning and Deep Learning for NLP (A.6)	Застосування машинного навчання для опрацювання природної мови / Application of Machine Learning for Natural Language Processing (A.6, A.4)
4	Educational and Service-oriented AI Applications (A.1); AI in Education and Customer Service (A.2); Learning Systems (LS) (A.3); Chatbot and Education (A.4); AI for Training and Customer Service (A.5); Chatbots and Learning Systems (A.6)	Чат-боти в навчанні та сфері послуг / Chatbots in Education and Services (A.1, A.2, A.4, A.5)

Залучені великі мовні моделі генерували декілька варіантів назв та описів кластерів на основі аналізу ключових слів, які входять до їх складу. Для кожного кластеру наведені всі запропоновані моделями варіанти з позначеннями A.1, A.2 і т.ін.

Вибір остаточних назв кластерів, представлених у стовпці «Обране ім'я» табл. 2 як українською, так і англійською мовами, здійснювався шляхом ретельного аналізу та критичної оцінки авторами всіх згенерованих різними моделями пропозицій. Ключовими критеріями відбору стали змістовність, чіткість, лаконічність назв з одночасним максимально повним охопленням ключових термінів у складі кластерів.

Обрані назви кластерів не лише узагальнюють їх зміст, а й логічно відображають виявлені за результатами бібліометричного аналізу основні напрями досліджень у галузі

навчання чат-ботів: 1) базові методи опрацювання природної мови; 2) застосування відповідних технологій у різних сферах суспільної діяльності; 3) використання машинного навчання як провідного інструментарію розробки систем опрацювання природномовних даних; 4) спектр ключових застосувань чат-ботів, зокрема в освітній галузі та сфері послуг.

Отже, остаточний вибір назв кластерів став результатом поєднання можливостей великих мовних моделей з експертною думкою авторів і ґрунтується на чітких обґрунтованих позиціях, спрямованих на уникнення невизначеностей у трактуванні отриманих результатів.

На рис. 4 представлено карту зв'язків ключових слів зі значущістю не менше 1, тобто найбільш значущих ключових слів, що належать до четвірки виділених кластерів. Дана карта дозволяє проаналізувати взаємозв'язки між центральними поняттями дослідження та візуалізувати провідні напрями в межах кожного кластера.

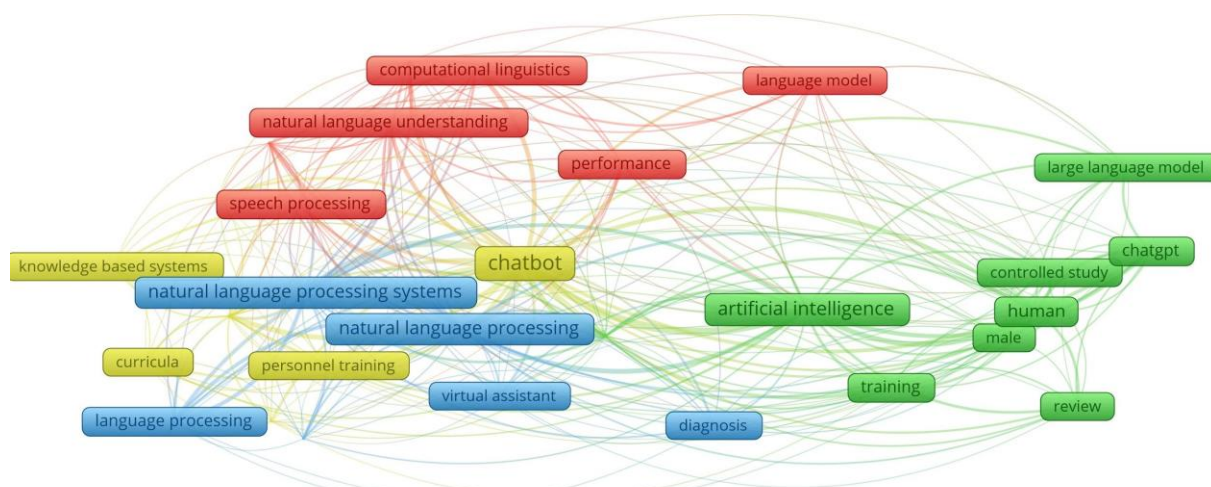


Рис. 4. Карта зв'язків найбільш значущих ключових слів

Побудова цієї фокусної карти здійснювалась з метою концентрації уваги на найбільш важливих ключових словах, відібраних за принципом максимальних та передмаксимальних значень низки показників (кількість зв'язків, загальна вага зв'язків, кількість згадувань у публікаціях, середні значення року публікації, цитованості тощо). Візуалізація найбільш значущих термінів та їх взаємозв'язків дає змогу ідентифікувати пріоритетні напрями у кожній з досліджуваних сфер для окреслення перспективних напрямків майбутніх наукових пошуків.

Такий підхід дозволяє деталізувати і поглибити висновки, зроблені на основі загальної карти ключових слів, концентруючись безпосередньо на центральних поняттях навчання чат-ботів як провідних агентів штучного інтелекту.

4. ВИСНОВКИ

1. Бібліометричний аналіз 549 джерел із бази Scopus з проблеми навчання чат-ботів надав можливість:

а) виявити нижню хронологічну межу (2018 рік), починаючи з якої спостерігається стійке зростання кількості публікацій з навчання чат-ботів;

б) висловити припущення про те, що причиною суттєвого зростання кількості робіт у 2023 році (на 55 порівняно із 2022 роком) є поява у відкритому доступі ChatGPT та пов'язаний з цим сплеск інтересу наукової спільноти до тематики великих мовних

моделей та чат-ботів: для підтвердження чи спростування цього припущення можна проаналізувати контекст ключових слів у публікаціях 2023 року на предмет безпосередніх згадувань ChatGPT або близьких до нього систем, а також розглянути динаміку цитувань праць, присвячених ChatGPT, в інших роботах за цей період;

в) 4 кластери авторські та індексні ключові слова джерел з навчання чат-ботів: 1) опрацювання природної мови; 2) застосування технологій опрацювання природної мови у суспільстві; 3) застосування машинного навчання для опрацювання природної мови; 4) чат-боти у навчанні та сфері послуг.

2. Для визначення найбільш значущих ключових понять дослідження у кожному кластері були визначені ключові слова, що мають найбільше та наступне за ним значення за такими показниками: кількість зв'язків, загальна сила зв'язків, кількість документів із заданими ключовим словом та середні: рік публікації, кількість цитувань, нормалізована кількість цитувань. Для відібраних ключових слів була обчислена їх значущість від 1 до 6:

а) у першому кластері (опрацювання природної мови) найбільш значущими є 7 ключових слів (41\%): computational linguistics (6), natural language understanding (3), performance (2), language model (2), speech processing (1);

б) у другому кластері (застосування технологій опрацювання природної мови у суспільстві) найбільш значущими є 10 ключових слів (63\%): artificial intelligence (6), large language model (2), chatgpt (2), human (2), training (2), review (2), conversational agent (1), controlled study (1), education (1), male (1);

в) у третьому кластері (застосування машинного навчання для опрацювання природної мови) найбільш значущими є 6 ключових слів (50\%): natural language processing (6), natural language processing systems (3), virtual assistant (3), language processing (2), user interfaces (2), diagnosis (2);

г) у четвертому кластері (чат-боти у навчанні та сфері послуг) найбільш значущими є 7 ключових слів (58\%): chatbot (6), learning systems (3), curricula (3), learn+ (2), knowledge based systems (2), personnel training (1), customer service (1).

3. Під час аналізу карти ключових слів було виявлено, що третій кластер (застосування машинного навчання для опрацювання природної мови) у певний спосіб перетинається зі змістом першого (опрацювання природної мови) та другого (застосування технологій опрацювання природної мови у суспільстві) кластерів. Це пояснюється тим, що методи машинного навчання, зокрема глибинного навчання, є фундаментальними інструментами для розробки систем опрацювання природної мови та їх застосувань. Тому деякі ключові терміни третього кластера, як-от опрацювання природної мови, віртуальні асистенти, нейронні мережі тощо, тісно пов'язані з концепціями інших кластерів. Це свідчить про тісну інтеграцію різних напрямків досліджень у єдиному науковому полі створення інтелектуальних розмовних агентів. Подальший аналіз перетинів між кластерами може стати перспективним напрямком майбутніх наукових пошуків для більш чіткого розмежування предметних областей.

4. Аналіз карти зв'язків найбільш значущих ключових слів надав можливість визначити провідні напрями досліджень:

а) в опрацюванні природної мови засобами обчислювальної лінгвістики – розуміння природної мови, побудова мовних моделей та розпізнавання мовлення;

б) у застосуванні технологій штучного інтелекту для опрацювання природної мови – контрольоване використання великих мовних моделей та чат-ботів (зокрема ChatGPT) в освіті;

в) у застосуванні машинного навчання для опрацювання природної мови – використання віртуальних асистентів, природномовних інтерфейсів користувача та інших систем опрацювання природної мови, зокрема для діагностики;

г) у застосуванні чат-ботів у навчанні та сфері послуг – використання чат-ботів, систем навчання та систем управління знаннями для доповненого та адаптивного навчання.

5. Результати проведеного бібліометричного аналізу можуть знайти застосування: – в освіті:

а) визначені пріоритетні напрямки досліджень у галузі навчання чат-ботів можуть бути використані для формування тематики курсів, програм підготовки фахівців з опрацювання природної мови та розробки систем штучного інтелекту;

б) виявлені ключові поняття можуть бути взятими за основу для розробки навчальних матеріалів, покликаних висвітлити найважливіші концепції та технології створення розмовних агентів;

в) огляд застосувань чат-ботів в освітньому процесі окреслює перспективні шляхи їх інтеграції в системи адаптивного і персоналізованого навчання;

– у наукових дослідженнях:

а) карта зв'язків ключових понять може бути базою для формування наукових гіпотез та побудови концептуальних моделей під час проектування досліджень;

б) результати аналізу відкривають перспективи для подальших бібліометричних студій динаміки та взаємозв'язків у предметному полі;

в) виявлені пріоритетні напрямки визначають актуальні вектори майбутніх наукових розвідок у галузі штучного інтелекту та його використання в наукових дослідженнях у галузі соціальних наук.

Визначені напрямки можуть бути корисними для організацій при розробці стратегій використання штучного інтелекту та інтеграції розмовних агентів у виробничу діяльність.

5. ПЕРСПЕКТИВИ ПОДАЛЬШИХ ДОСЛІДЖЕНЬ

Продовженням бібліометричного аналізу може бути систематичний літературний огляд, що містить багато трудомістких і тривалих етапів, серед яких:

1. Визначення обсягу огляду та критеріїв включення/виключення.

2. Пошук у наукометричних базах даних для виявлення потенційно релевантних досліджень.

3. Перевірка назв, анотацій та повних текстів для визначення, які дослідження відповідають критеріям включення.

4. Видобування даних із відібраних досліджень.

5. Синтез та аналіз отриманої інформації.

Наразі кожен із цих кроків виконується переважно вручну. У нашій роботі ми використовуємо класичні та сучасні методи опрацювання природної мови для напівавтоматичного виконання якомога більшої кількості кроків. Це може значно скоротити час і зусилля дослідника при проведенні систематичних оглядів.

Конкретна проблема, яку ми плануємо дослідити, полягає в автоматизації етапу скринінгу (крок 3 вище) шляхом використання великих мовних моделей (large language model – LLM) для оцінки анотацій досліджень за попередньо визначеними критеріями включення/виключення. Цей етап скринінгу є критичним вузьким місцем, оскільки вимагає індивідуального перегляду тисяч анотацій.

Формально описати цю проблему можна у такий спосіб.

Дано:

- база даних з N наукових робіт (статей, дисертацій, книжок та їх частин, звітів), представлених їхніми анотаціями $\{A_1, A_2, \dots, A_N\}$;
- критерії включення C_1, C_2, \dots, C_m ;

- критерії виключення E_1, E_2, \dots, E_n .

Нам потрібна модель M , яка може приймати на вхід кожен анотацію A_i та критерії $\{C_1, C_2, \dots, C_m, E_1, E_2, \dots, E_n\}$ і виводити мітку L_i , яка вказує, чи відповідає дослідження критеріям, чи ні:

$$M(A_i, \{C_1, C_2, \dots, C_m, E_1, E_2, \dots, E_n\}) = L_i,$$

де L_i може набувати одного з k значень:

- відповідає всім критеріям включення та не відповідає критеріям виключення;
- порушує критерії включення C_j ;
- порушує критерії виключення E_k ;
- порушує деякі критерії включення та виключення тощо.

Ми припускаємо, що великі мовні моделі, попередньо навчені на великих текстових корпусах, можуть ефективно виконувати цю класифікацію, урахувавши семантичне значення анотацій і критеріїв. Для перевірки цього припущення ми створили програму з використанням OpenAI API для доступу до LLM GPT-3.5 Turbo, яка зчитує набір із 347 статей з їх анотаціями, ідентифікаторами статей та двома критеріями включення:

C_1 : фокус – стаття присвячена підготовці вчителів або їх професійному розвитку в галузі персоналізованого навчання на основі штучного інтелекту;

C_2 : популяція – у статті розглядаються вчителі або майбутні вчителі – студенти, що навчаються за освітніми програмами підготовки вчителів.

До кожної статті було надано текст анотації як підказку для LLM разом з інструкціями аналізу критеріїв: чи були вони дотримані, чи були порушені або чи було недостатньо інформації. У відповіді LLM мала класифікувати кожен статтю за однією з чотирьох категорій: відповідає обом критеріям (1 1), порушує критерії C_1 (0 1), порушує критерії C_2 (1 0) або порушує обидва критерії (0 0).

LLM класифікувала 173 з 347 статей як такі, що відповідають обом критеріям включення. Крім того, 3 статті були помилково визначені як виключені через відсутність анотацій у джерелі даних, загалом 176 статей було відібрано для повнотекстового пошуку після врахування цієї помилки:

- виключено через порушення критерію C_1 – 44;
- виключено через порушення критерію C_2 – 109;
- виключено через порушення обох критеріїв – 21.

Для оцінки точності ми вручну переглянули випадкову вибірку з 50 статей у всіх чотирьох категоріях. Для статей, які відповідали критеріям, точність становила 67%. Було виявлено більше хибнопозитивних результатів (44%), ніж хибнонегативних (16%). Через велику частку хибнопозитивних результатів у вибірці було виконано повний ручний перегляд відібраних 176 статей, у результаті якого було виявлено, що 97 статей (55%), включених LLM, насправді порушують один із критеріїв включення.

Слід зазначити кілька обмежень цього дослідження. По-перше, формулювання критеріїв включення були виконано у відносно простих семантичних одиницях щодо фокусу дослідження та популяції. Більш складні критерії, що включають дизайн дослідження, деталі втручання (впливу на популяцію), статистичний аналіз тощо, можуть стати викликом для поточних можливостей LLM, однак їх ефективність можна покращити за допомогою інженерії запитів або доопрацювання моделі на основі даних систематичних оглядів, виконаних без застосування LLM. Крім того, ми не проводили оцінки ефективності LLM у різних галузях дослідження: критерії включення/виключення були надані моделі як підказки без прив'язки до конкретної галузі. Налаштування або доопрацювання LLM для кожної дослідницької галузі може підвищити ефективність. Незважаючи на вказані обмеження, дане дослідження підкреслює величезний потенціал використання передових моделей штучного інтелекту

для прискорення роботи із синтезу доказового дослідження. Оскільки LLM швидко розвиваються, вони, ймовірно, відіграватимуть дедалі помітнішу роль в оптимізації проведення не лише систематичних оглядів, а й інших важливих видів дослідницької діяльності.

Перспективні напрямки передбачають використання LLM для інших етапів систематичного огляду, таких як повнотекстовий скринінг, видобування даних із досліджень, аналіз ризику упередженості та навіть інтелектуальний синтез результатів досліджень. Вивчення того, як LLM можуть напівавтоматизувати процес систематичного огляду в надійному, валідованому конвеєрі машинного навчання (LLMOps), є захоплюючою сферою для майбутніх досліджень із глибокими наслідками для прискорення наукових відкриттів на основі синтезу та прийняття рішень на основі доказової бази. Для вітчизняних освітніх досліджень це – ще одна можливість переходу до *доказової освіти*, яка є подібною до доказової медицини тим, що надає можливість прийняття обґрунтованих рішень про вибір конкретних дій (впливів, втручань тощо) щодо освітнього процесу, ефективність яких у заданому контексті є науково доведеною.

ПОДЯКА

Автори щиро вдячні Ірині Сергіївні Мінтій за проведення експериментальної роботи, покладеної в основу розділу «Перспективи подальших досліджень», а також Ользі Павловні Пінчук та Олександрі Миколаївні Соколюк за чудове спрямування у найкращому поданні результатів цього дослідження. Користуючись випадком, поздоровляємо видавця, редакторів, рецензентів, авторів та читачів одного з кращих світових наукових видань – «Інформаційні технології і засоби навчання» – з ювілейним сотим томом!

ДОДАТОК: РЕЗУЛЬТАТИ ЗАСТОСУВАННЯ ГЕНЕРАТИВНИХ ЧАТ-БОТІВ ДЛЯ ВИЗНАЧЕННЯ ІМЕН КЛАСТЕРІВ

Запит був створений англійською мовою 30.11.2023. У запиті містились відомості про кількість кластерів та ключові слова, що стосувались кожного кластеру. Від чат-боту очікувались пропозиції щодо імен та стислого опису кластерів. Запит мав наступний вигляд:

“There are 4 clusters of keywords related to chatbot training research:

Cluster 1: classification (of information), training data, speech processing, performance, semantics, dialogue systems, natural language understanding, computational linguistics, text processing, long short-term memory, state of the art, language model, question answering, information retrieval, embeddings, reinforcement learning, response generation

Cluster 2: artificial intelligence, conversational agent, human, adult, training, male, medical education, female, controlled study, health care, mental health, chatgpt, education, covid-19, review, large language model

Cluster 3: natural language processing, natural language processing systems, machine learning, deep learning, natural languages, language processing, learning algorithms, virtual assistant, user interfaces, convolutional neural network, query processing, diagnosis

Cluster 4: chatbot, learning systems, students, e-learning, learn+, sales, human computer interaction, knowledge based systems, personnel training, curricula, customer service, engineering education

Please suggest cluster names and brief descriptions for each of the 4 clusters based on the keywords.”

ChatGPT (GPT 3.5)

1. *Advanced Language Processing and AI Applications*: This cluster focuses on advanced language processing techniques and various AI applications related to text and speech. It encompasses topics like natural language understanding, computational linguistics, information retrieval, dialogue systems, and the utilization of cutting-edge technologies such as long short-term memory (LSTM), embeddings, and reinforcement learning in language models and question answering systems.

2. *AI in Healthcare and Social Context*: This cluster centers around the applications of artificial intelligence in healthcare, mental health, and education with a focus on human interaction. It explores the use of conversational agents, particularly in medical education, mental health care, and addressing issues like COVID-19. Additionally, it touches on the gender aspect in AI studies by mentioning controlled studies with adults and gender variations.

3. *Core Technologies in Natural Language Processing and Machine Learning*: This cluster revolves around foundational technologies crucial in natural language processing and machine learning. It includes topics such as deep learning, machine learning algorithms, language processing, neural networks (like convolutional neural networks), query processing, and the development of virtual assistants and user interfaces.

4. *Educational and Service-oriented AI Applications*: This cluster focuses on AI applications in education, learning systems, and service-oriented domains. It highlights topics like e-learning, knowledge-based systems, human-computer interaction, customer service, and sales. It also touches upon the use of AI in student learning, engineering education, and personnel training, emphasizing knowledge acquisition and application in various educational and service-based contexts.

Bing, precise mode (GPT 4.0)

1. *Language Processing and Information Retrieval*: This cluster focuses on the technical aspects of natural language understanding and computational linguistics. It includes topics such as speech processing, semantics, dialogue systems, text processing, language modeling, question answering, and reinforcement learning for response generation.

2. *AI in Healthcare and Education*: This cluster revolves around the application of artificial intelligence, particularly conversational agents like ChatGPT, in fields such as medical education, healthcare, and mental health. It also touches on demographic factors (male, female, adult) and current issues like COVID-19.

3. *Machine Learning and User Interaction*: This cluster is centered on machine learning, deep learning, and natural language processing systems. It also includes user interfaces and virtual assistants, indicating a focus on how users interact with these AI systems.

4. *AI in Education and Customer Service*: This cluster pertains to the use of AI systems like chatbots in educational settings (e-learning, engineering education, curricula) and customer service. It also includes topics related to human-computer interaction and knowledge-based systems.

Bing, balanced mode (GPT 4.0)

1. *Natural Language Processing (NLP)*: This cluster includes topics related to the processing of human language by computers. It covers a wide range of subfields such as computational linguistics, text processing, information retrieval, and question answering. The cluster also includes techniques such as embeddings, long short-term memory, and reinforcement learning.

2. *Artificial Intelligence (AI)*: This cluster includes topics related to the development of

intelligent agents that can perform tasks that typically require human intelligence. It covers a wide range of subfields such as conversational agents, large language models, and chatbots. The cluster also includes topics related to health care, education, and COVID-19.

3. *Machine Learning (ML)*: This cluster includes topics related to the development of algorithms that can learn from data and make predictions or decisions. It covers a wide range of subfields such as deep learning, learning algorithms, and convolutional neural networks. The cluster also includes topics related to virtual assistants, user interfaces, and diagnosis.

4. *Learning Systems (LS)*: This cluster includes topics related to the development of systems that can learn from data and improve over time. It covers a wide range of subfields such as e-learning, personnel training, and customer service. The cluster also includes topics related to sales, human-computer interaction, and engineering education.

Bing, creative mode (GPT 4.0)

1. *Natural Language Understanding and Generation*: This cluster contains keywords related to the tasks and methods of natural language understanding and generation, such as classification, dialogue systems, question answering, and response generation. It also includes keywords that indicate the performance and state of the art of these tasks, such as semantics, long short-term memory, embeddings, and reinforcement learning.

2. *Artificial Intelligence and Health Care*: This cluster contains keywords related to the applications and implications of artificial intelligence in health care, such as conversational agent, medical education, mental health, and covid-19. It also includes keywords that indicate the characteristics and challenges of these applications, such as human, adult, male, female, controlled study, review, and large language model.

3. *Natural Language Processing and Machine Learning*: This cluster contains keywords related to the general concepts and techniques of natural language processing and machine learning, such as natural language processing systems, machine learning, deep learning, learning algorithms, and convolutional neural network. It also includes keywords that indicate the interfaces and domains of these techniques, such as virtual assistant, user interfaces, query processing, and diagnosis.

4. *Chatbot and Education*: This cluster contains keywords related to the design and evaluation of chatbot systems for education, such as chatbot, learning systems, students, e-learning, learn+, and curricula. It also includes keywords that indicate the contexts and outcomes of these systems, such as sales, human computer interaction, knowledge based systems, personnel training, customer service, and engineering education.

Claude 2

1. *Natural Language Processing Applications*: This cluster focuses on various NLP applications like classification, speech processing, dialogue systems, question answering, etc. Central themes include semantics, computational linguistics, LSTMs, performance, and building practical natural language systems.

2. *Conversational AI and Impacts*: This cluster deals with conversational AI, its use in healthcare/education, and analyzing its impacts. Key themes look at chatbots/assistants, how they interact with and impact humans, with analysis around ethics, proper training, and responsible AI.

3. *Core AI and ML Techniques*: This cluster covers fundamental AI and ML techniques like natural language processing, machine learning, deep learning, algorithms, etc. It focuses on the key technical methods and capabilities that power many AI systems.

4. *AI for Training and Customer Service*: The fourth cluster deals with leveraging AI for use cases like education/training and customer service. Central ideas look at e-learning,

knowledge systems, training curriculums, human-computer interaction, and providing automation for repetitive customer service queries.

Google Bard

1. *Technical NLP*: This cluster focuses on the technical aspects of natural language processing (NLP), including classification, training data, speech processing, semantics, dialogue systems, and language models.

2. *Applications of NLP*: This cluster focuses on the applications of NLP in various domains, including conversational agents, healthcare, education, and customer service.

3. *Machine Learning and Deep Learning for NLP*: This cluster focuses on the machine learning and deep learning techniques used in NLP.

4. *Chatbots and Learning Systems*: This cluster focuses on the development and evaluation of chatbots and other learning systems.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

- [1] K. Peyton and S. Unnikrishnan, “A comparison of chatbot platforms with the state-of-the-art sentence BERT for answering online student FAQs,” *Results in Engineering*, vol. 17, 2023. doi: <https://doi.org/10.1016/j.rineng.2022.100856>.
- [2] V. Gadiraju, S. Kane, S. Dev, A. Taylor, D. Wang, E. Denton, and R. Brewer, ““I wouldn’t say offensive but...”: Disability-Centered Perspectives on Large Language Models,” in *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency*, ser. FAccT ‘23. New York, NY, USA: Association for Computing Machinery, 2023, p. 205–216. doi: <https://doi.org/10.1145/3593013.3593989>.
- [3] N. S. Patil, R. S. Huang, C. B. van der Pol, and N. Larocque, “Comparative Performance of ChatGPT and Bard in a Text-Based Radiology Knowledge Assessment,” *Canadian Association of Radiologists Journal*, 2023. doi: <https://doi.org/10.1177/08465371231193716>.
- [4] B. Kim, J. Seo, and M.-W. Koo, “Randomly Wired Network Based on RoBERTa and Dialog History Attention for Response Selection,” *IEEE/ACM Transactions on Audio Speech and Language Processing*, vol. 29, pp. 2437–2442, 2021. doi: <https://doi.org/10.1109/TASLP.2021.3077119>.
- [5] J. Zhang, J. Zhang, S. Ma, J. Yang, and G. Gui, “Chatbot design method using hybrid word vector expression model based on real telemarketing data,” *KSII Transactions on Internet and Information Systems*, vol. 14, no. 4, pp. 1400–1418, 2020. doi: <https://doi.org/10.3837/TIIS.2020.04.001>.
- [6] S. S. Abdullahi, S. Yiming, A. Abdullahi, and U. Aliyu, “Open Domain Chatbot Based on Attentive End-to-End Seq2Seq Mechanism,” in *Proceedings of the 2019 2nd International Conference on Algorithms, Computing and Artificial Intelligence*, ser. ACAI ‘19. New York, NY, USA: Association for Computing Machinery, 2020, p. 339–344. doi: <https://doi.org/10.1145/3377713.3377773>.
- [7] A. G. Usigan, M. I. Salomeo, G. J. L. J. Zafe, C. J. Centeno, A. A. R. C. Sison, and A. G. Bitancor, “Implementation of an Undergraduate Admission Chatbot Using Microsoft Azure’s Question Answering and Bot Framework,” in *Proceedings of the 2022 5th Artificial Intelligence and Cloud Computing Conference*, ser. AICCC ‘22. New York, NY, USA: Association for Computing Machinery, 2023, p. 240–245. doi: <https://doi.org/10.1145/3582099.3582135>.
- [8] E. Ruane, R. Young, and A. Ventresque, “Training a Chatbot with Microsoft LUIS: Effect of Intent Imbalance on Prediction Accuracy,” in *Companion Proceedings of the 25th International Conference on Intelligent User Interfaces*, ser. IUI ‘20 Companion. New York, NY, USA: Association for Computing Machinery, 2020, p. 63–64. doi: <https://doi.org/10.1145/3379336.3381494>.
- [9] N. Bhartiya, N. Jangid, S. Jannu, P. Shukla, and R. Chapaneri, “Artificial Neural Network Based University Chatbot System,” in *2019 IEEE Bombay Section Signature Conference*, IBSSC 2019, vol. 2019January. Institute of Electrical and Electronics Engineers Inc., 2019, Conference paper. doi: <https://doi.org/10.1109/IBSSC47189.2019.8973095>.
- [10] B. Hancock, A. Bordes, P.-E. Mazaré, and J. Weston, “Learning from dialogue after deployment: Feed yourself, Chatbot!” in *ACL 2019 - 57th Annual Meeting of the Association for Computational Linguistics, Proceedings of the Conference*. Association for Computational Linguistics (ACL), 2020, Conference paper, pp. 3667–3684.
- [11] M. N. Sreedhar, K. Ni, and S. Reddy, “Learning improvised chatbots from adversarial modifications of natural language feedback,” in *Findings of the Association for Computational Linguistics Findings of ACL: EMNLP 2020*, 2020, pp. 2445–2453.

- [12] S.-W. Lee and W.-J. Choi, "Utilizing ChatGPT in clinical research related to anesthesiology: a comprehensive review of opportunities and limitations," *Anesthesia and Pain Medicine*, vol. 18, no. 3, pp. 244–251, 2023. doi: <https://doi.org/10.17085/apm.23056>.
- [13] C. K. Lo, "What Is the Impact of ChatGPT on Education? A Rapid Review of the Literature," *Education Sciences*, vol. 13, no. 4, p. 410, 2023. doi: <https://doi.org/10.3390/educsci13040410>.
- [14] A. Pack and J. Maloney, "Using Generative Artificial Intelligence for Language Education Research: Insights from Using OpenAI's ChatGPT," *TESOL Quarterly*, vol. 57, no. 4, pp. 1571–1582, 2023. doi: <https://doi.org/10.1002/tesq.3253>.
- [15] J. Chervenak, H. Lieman, M. Blanco-Breindel, and S. Jindal, "The promise and peril of using a large language model to obtain clinical information: ChatGPT performs strongly as a fertility counseling tool with limitations," *Fertility and Sterility*, vol. 120, no. 3, pp. 575–583, 2023. doi: <https://doi.org/10.1016/j.fertnstert.2023.05.151>.
- [16] A. V. Riabko and T. A. Vakaliuk, "Physics on autopilot: exploring the use of an AI assistant for independent problem-solving practice," *Educational Technology Quarterly*, vol. 2024, no. 1, p. 56–75, Mar. 2024. doi: <https://doi.org/10.55056/etq.671>.
- [17] OpenAI, "Introducing ChatGPT," Nov. 2022. [Електронний ресурс]. Доступно: <https://openai.com/blog/chatgpt>
- [18] DeepLearning.AI, "Search | The Batch | AI News & Insights," Dec. 2023. [Електронний ресурс]. Доступно: <https://www.deeplearning.ai/search/>
- [19] "Big Bot Makes Small Talk: A research summary of Facebook's Generative BST chatbot," May 2020. [Електронний ресурс]. Доступно: <https://www.deeplearning.ai/the-batch/big-bot-makes-small-talk/>
- [20] "Bot Comic: How Google's Meena chatbot developed a sense of humor," Feb. 2020. [Електронний ресурс]. Доступно: <https://www.deeplearning.ai/the-batch/bot-comic/>
- [21] "Chatbots for Productivity," Sep. 2023. [Електронний ресурс]. Доступно: <https://www.deeplearning.ai/the-batch/microsoft-extends-copilot-365-windows/>
- [22] "China Chases Chatbots," Mar. 2023. [Електронний ресурс]. Доступно: <https://www.deeplearning.ai/the-batch/chinese-tech-companies-race-to-cash-in-on-chatgpt-fever/>
- [23] "Search War!" Feb. 2023. [Електронний ресурс]. Доступно: <https://www.deeplearning.ai/the-batch/google-and-microsoft-both-announce-ai-powered-search/>
- [24] "Chatbots Disagree on Covid-19: Medical chatbots offered conflicting Covid advice," Apr. 2020. [Електронний ресурс]. Доступно: <https://www.deeplearning.ai/the-batch/chatbots-disagree-on-covid-19/>
- [25] "Language Models, Extended: Large language models grew more reliable and less biased in 2022," Dec. 2022. [Електронний ресурс]. Доступно: <https://www.deeplearning.ai/the-batch/language-models-grew-more-reliable-and-less-biased-in-2022/>
- [26] "Cost Containment for Generative AI: Microsoft's quest to reduce the size and cost of language models," Oct. 2023. [Електронний ресурс]. Доступно: <https://www.deeplearning.ai/the-batch/microsofts-quest-to-reduce-the-size-and-cost-of-language-models/>
- [27] "What We Know — and Don't Know — About Foundation Models: A new Stanford index to assess the transparency of leading AI models," Nov. 2023. [Електронний ресурс]. Доступно: <https://www.deeplearning.ai/the-batch/a-new-stanford-index-to-assess-the-transparency-of-leading-ai-models/>
- [28] Elsevier B.V., "Scopus - Document search | Signed in," 2023. [Електронний ресурс]. Доступно: <https://www.scopus.com/search/form.uri?display=basic#basic>
- [29] N. J. Van Eck and L. Waltman, VOSviewer Manual. Universiteit Leiden, 2023. [Електронний ресурс]. Доступно: https://www.vosviewer.com/documentation/Manual_VOSviewer_1.6.20.pdf
- [30] Centre for Science and Technology Studies, Leiden University, The Netherlands, "VOSviewer - Visualizing scientific landscapes," 2023. [Електронний ресурс]. Доступно: <https://www.vosviewer.com/>

Матеріал надійшов до редакції 20.02.2024 р.

BIBLIOMETRIC ANALYSIS OF CHATBOT TRAINING RESEARCH: KEY CONCEPTS AND TRENDS

Roman Liashenko

student

Kryvyi Rih State Pedagogical University, Kryvyi Rih, Ukraine

ORCID ID 0009-0000-2614-6997

romanslss777@gmail.com

Serhiy Semerikov

D.Sc. (Computer Science Education), professor, senior researcher
Kryvyi Rih State Pedagogical University, Kryvyi Rih, Ukraine
Institute for Digitalisation of Education of the NAES of Ukraine, Kyiv, Ukraine
Zhytomyr Polytechnic State University, Zhytomyr, Ukraine
Kryvyi Rih National University, Kryvyi Rih, Ukraine
Academy of Cognitive and Natural Sciences, Kryvyi Rih, Ukraine
ORCID ID 0000-0003-0789-0272
semerikov@gmail.com

Abstract. This bibliometric analysis aims to identify current research directions and priorities in the field of chatbot training – software agents capable of natural language dialogue. The study is based on the analysis of 549 scientific sources from the Scopus database on this topic. The analysis revealed a steady increase in relevant publications starting from 2018, indicating a growing relevance of this subject area in recent years. Based on a cluster analysis of keywords, four main research areas were identified: natural language processing, application of relevant technologies in various spheres of society, use of machine learning methods for natural language processing, and application of chatbots in education and services. In the field of natural language processing, the focus of current research is on computational linguistics, language modeling and machine comprehension, particularly speech recognition tasks. Leading research on artificial intelligence applications in this area concerns the responsible and ethical use of modern large language models and conversational agents, such as ChatGPT, in education and healthcare. Machine learning methods are actively being developed for creating virtual intelligent assistants, natural language user interfaces, and other natural language processing systems, including for diagnostic tasks in medicine. Key applications of chatbots are identified in adaptive learning systems, knowledge management, and customer service. Based on the analysis, the most significant concepts in each of the studied areas are defined to outline priorities for further research in the field of chatbot training. Future work involves conducting a systematic literature review with the automation of certain stages using large language models. In particular, these models will be employed to automatically classify study abstracts according to inclusion/exclusion criteria during the screening phase. Automating systematic review stages using artificial intelligence opens up significant prospects for accelerating scientific research, particularly in the education field based on an evidence-based approach.

Keywords: chatbot training; natural language processing; machine learning; bibliometric analysis; systematic literature review; large language models.

REFERENCES (TRANSLATED AND TRANSLITERATED)

- [1] K. Peyton and S. Unnikrishnan, “A comparison of chatbot platforms with the state-of-the-art sentence BERT for answering online student FAQs,” *Results in Engineering*, vol. 17, 2023. doi: <https://doi.org/10.1016/j.rineng.2022.100856>.
- [2] V. Gadiraju, S. Kane, S. Dev, A. Taylor, D. Wang, E. Denton, and R. Brewer, ““I wouldn’t say offensive but...”: Disability-Centered Perspectives on Large Language Models,” in *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency*, ser. FAccT ‘23. New York, NY, USA: Association for Computing Machinery, 2023, p. 205–216. doi: <https://doi.org/10.1145/3593013.3593989>.
- [3] N. S. Patil, R. S. Huang, C. B. van der Pol, and N. Larocque, “Comparative Performance of ChatGPT and Bard in a Text-Based Radiology Knowledge Assessment,” *Canadian Association of Radiologists Journal*, 2023. doi: <https://doi.org/10.1177/08465371231193716>.
- [4] B. Kim, J. Seo, and M.-W. Koo, “Randomly Wired Network Based on RoBERTa and Dialog History Attention for Response Selection,” *IEEE/ACM Transactions on Audio Speech and Language Processing*, vol. 29, pp. 2437–2442, 2021. doi: <https://doi.org/10.1109/TASLP.2021.3077119>.
- [5] J. Zhang, J. Zhang, S. Ma, J. Yang, and G. Gui, “Chatbot design method using hybrid word vector expression model based on real telemarketing data,” *KSII Transactions on Internet and Information Systems*, vol. 14, no. 4, pp. 1400–1418, 2020. doi: <https://doi.org/10.3837/TIIS.2020.04.001>.
- [6] S. S. Abdullahi, S. Yiming, A. Abdullahi, and U. Aliyu, “Open Domain Chatbot Based on Attentive End-to-End Seq2Seq Mechanism,” in *Proceedings of the 2019 2nd International Conference on Algorithms, Computing and Artificial Intelligence*, ser. ACAI ‘19. New York, NY, USA: Association for Computing

- Machinery, 2020, p. 339–344. doi: <https://doi.org/10.1145/3377713.3377773>.
- [7] A. G. Usigan, M. I. Salomeo, G. J. L. J. Zafe, C. J. Centeno, A. A. R. C. Sison, and A. G. Bitancor, “Implementation of an Undergraduate Admission Chatbot Using Microsoft Azure’s Question Answering and Bot Framework,” in *Proceedings of the 2022 5th Artificial Intelligence and Cloud Computing Conference*, ser. AICCC ‘22. New York, NY, USA: Association for Computing Machinery, 2023, p. 240–245. doi: <https://doi.org/10.1145/3582099.3582135>.
- [8] E. Ruane, R. Young, and A. Ventresque, “Training a Chatbot with Microsoft LUIS: Effect of Intent Imbalance on Prediction Accuracy,” in *Companion Proceedings of the 25th International Conference on Intelligent User Interfaces*, ser. IUI ‘20 Companion. New York, NY, USA: Association for Computing Machinery, 2020, p. 63–64. doi: <https://doi.org/10.1145/3379336.3381494>.
- [9] N. Bhartiya, N. Jangid, S. Jannu, P. Shukla, and R. Chapaneri, “Artificial Neural Network Based University Chatbot System,” in *2019 IEEE Bombay Section Signature Conference, IBSSC 2019*, vol. 2019January. Institute of Electrical and Electronics Engineers Inc., 2019, Conference paper. doi: <https://doi.org/10.1109/IBSSC47189.2019.8973095>.
- [10] B. Hancock, A. Bordes, P.-E. Mazaré, and J. Weston, “Learning from dialogue after deployment: Feed yourself, Chatbot!” in *ACL 2019 - 57th Annual Meeting of the Association for Computational Linguistics, Proceedings of the Conference*. Association for Computational Linguistics (ACL), 2020, Conference paper, pp. 3667–3684.
- [11] M. N. Sreedhar, K. Ni, and S. Reddy, “Learning improvised chatbots from adversarial modifications of natural language feedback,” in *Findings of the Association for Computational Linguistics Findings of ACL: EMNLP 2020*, 2020, pp. 2445–2453.
- [12] S.-W. Lee and W.-J. Choi, “Utilizing ChatGPT in clinical research related to anesthesiology: a comprehensive review of opportunities and limitations,” *Anesthesia and Pain Medicine*, vol. 18, no. 3, pp. 244–251, 2023. doi: <https://doi.org/10.17085/apm.23056>.
- [13] C. K. Lo, “What Is the Impact of ChatGPT on Education? A Rapid Review of the Literature,” *Education Sciences*, vol. 13, no. 4, p. 410, 2023. [Online]. Available: <https://doi.org/10.3390/educsci13040410>
- [14] A. Pack and J. Maloney, “Using Generative Artificial Intelligence for Language Education Research: Insights from Using OpenAI’s ChatGPT,” *TESOL Quarterly*, vol. 57, no. 4, pp. 1571–1582, 2023. doi: <https://doi.org/10.1002/tesq.3253>.
- [15] J. Chervenak, H. Lieman, M. Blanco-Breindel, and S. Jindal, “The promise and peril of using a large language model to obtain clinical information: ChatGPT performs strongly as a fertility counseling tool with limitations,” *Fertility and Sterility*, vol. 120, no. 3, pp. 575–583, 2023. doi: <https://doi.org/10.1016/j.fertnstert.2023.05.151>.
- [16] A. V. Riabko and T. A. Vakaliuk, “Physics on autopilot: exploring the use of an AI assistant for independent problem-solving practice,” *Educational Technology Quarterly*, vol. 2024, no. 1, p. 56–75, Mar. 2024. doi: <https://doi.org/10.55056/etq.671>.
- [17] OpenAI, “Introducing ChatGPT,” Nov. 2022. [Online]. Available: <https://openai.com/blog/chatgpt>
- [18] DeepLearning.AI, “Search | The Batch | AI News & Insights,” Dec. 2023. [Online]. Available: <https://www.deeplearning.ai/search/>
- [19] “Big Bot Makes Small Talk: A research summary of Facebook’s Generative BST chatbot,” May 2020. [Online]. Available: <https://www.deeplearning.ai/the-batch/big-bot-makes-small-talk/>
- [20] “Bot Comic: How Google’s Meena chatbot developed a sense of humor,” Feb. 2020. [Online]. Available: <https://www.deeplearning.ai/the-batch/bot-comic/>
- [21] “Chatbots for Productivity,” Sep. 2023. [Online]. Available: <https://www.deeplearning.ai/the-batch/microsoft-extends-copilot-365-windows/>
- [22] “China Chases Chatbots,” Mar. 2023. [Online]. Available: <https://www.deeplearning.ai/the-batch/chinese-tech-companies-race-to-cash-in-on-chatgpt-fever/>
- [23] “Search War!” Feb. 2023. [Online]. Available: <https://www.deeplearning.ai/the-batch/google-and-microsoft-both-announce-ai-powered-search/>
- [24] “Chatbots Disagree on Covid-19: Medical chatbots offered conflicting Covid advice,” Apr. 2020. [Online]. Available: <https://www.deeplearning.ai/the-batch/chatbots-disagree-on-covid-19/>
- [25] “Language Models, Extended: Large language models grew more reliable and less biased in 2022,” Dec. 2022. [Online]. Available: <https://www.deeplearning.ai/the-batch/language-models-grew-more-reliable-and-less-biased-in-2022/>
- [26] “Cost Containment for Generative AI: Microsoft’s quest to reduce the size and cost of language models,” Oct. 2023. [Online]. Available: <https://www.deeplearning.ai/the-batch/microsofts-quest-to-reduce-the-size-and-cost-of-language-models/>
- [27] “What We Know — and Don’t Know — About Foundation Models: A new Stanford index to assess the transparency of leading AI models,” Nov. 2023. [Online]. Available: <https://www.deeplearning.ai/the-batch/a-new-stanford-index-to-assess-the-transparency-of-leading-ai-models/>

- [28] Elsevier B.V., “Scopus - Document search | Signed in,” 2023. [Online]. Available: <https://www.scopus.com/search/form.uri?display=basic#basic>
- [29] N. J. Van Eck and L. Waltman, VOSviewer Manual. Universiteit Leiden, 2023. [Online]. Available: https://www.vosviewer.com/documentation/Manual_VOSviewer_1.6.20.pdf
- [30] Centre for Science and Technology Studies, Leiden University, The Netherlands, “VOSviewer - Visualizing scientific landscapes,” 2023. [Online]. Available: <https://www.vosviewer.com/>

