

БІБЛІОМЕТРИЧНИЙ АНАЛІЗ НАВЧАННЯ ЧАТ-БОТІВ

Ляшенко Р. О., Семеріков С. О.

Криворізький державний педагогічний університет, Кривий Ріг, Україна

Анотація. У роботі проведено бібліометричний аналіз джерел з проблеми навчання чат-ботів з використанням бази даних Scopus. За результатами аналізу виявлено період стійкого зростання публікацій з даної тематики починаючи з 2018 року, зокрема після появи ChatGPT у відкритому доступі. Крім того, було згруповано ключові слова у кластери та визначено найбільш значущі поняття для кожного з них. Аналіз показав, що основні напрями досліджень включають опрацювання природної мови, застосування технологій штучного інтелекту, машинне навчання та використання чат-ботів у навчанні та сфері послуг.

Ключові слова: чат-боти, навчання, бібліометричний аналіз, природна мова, штучний інтелект.

BIBLIOMETRIC ANALYSIS OF CHATBOT TRAINING

Liashenko R., Semerikov S.

Kryvyi Rih State Pedagogical University, Kryvyi Rih, Ukraine

Abstract. The paper presents a bibliometric analysis of sources on the training of chatbots using the Scopus database. The analysis revealed a period of sustained growth in publications on this topic starting from 2018, particularly after the emergence of ChatGPT in open access. Furthermore, key words were grouped into clusters and the most significant concepts for each of them were identified. The analysis showed that the main research directions include natural language processing, the application of artificial intelligence technologies, machine learning, and the use of chatbots in education and services.

Keywords: chatbots, training, bibliometric analysis, natural language, artificial intelligence.

Вступ. У цій роботі представлено бібліометричний аналіз 549 джерел з бази даних Scopus з метою визначення ключових понять, пов'язаних з навчанням чат-ботів. Дослідження мотивоване стрімким зростанням інтересу та публікацій у галузі створення та навчання чат-ботів, що спостерігається з 2018 року. Провідні ІТ-компанії активно працюють над розробкою чат-ботів для різних застосувань, включаючи навчання студентів та побудову систем типу “питання-відповідь”.

Мета роботи. Метою дослідження є визначення та візуалізація ключових понять, що відносяться до навчання чат-ботів.

Постановка проблеми. Основними завданнями дослідження є: 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.

Результати. Дослідження включало кілька етапів. Спочатку був виконаний пошук у базі Scopus за ключовими словами “chatbot” та “training”. Результати пошуку були експортовані у форматі CSV та BibTeX для подальшої обробки. Зокрема, у файлі CSV було виконано нормалізацію ключових слів до однини та заміну аббревіатур повними назвами термінів.

Далі результати були проаналізовані з використанням інструменту VOSviewer для побудови карти зв'язків ключових слів та їх групування у кластери. Карта створювалася з параметрами: тип аналізу за одночасною появою терміну, одиницею аналізу були всі ключові слова, а метод підрахунку – повний підрахунок. Загальна кількість ключових слів (3705) була зменшена до 58 відкиданням тих, що зустрічалися менше ніж 12 разів.

Для кожного кластера були визначені найбільш значущі ключові слова та їх кількісні характеристики. Зокрема, розглядалися такі показники: кількість зв'язків слова з іншими словами, загальна сила цих зв'язків (кількість публікацій, де два терміни зустрічаються разом), кількість документів із заданим ключовим словом, середній рік публікації цих документів, середня кількість цитувань та нормалізована кількість цитувань. Для відібраних ключових слів була обчислена їх значущість від 1 до 6 на основі їх рангу за зазначеними показниками. На рис. 1 подано карту зв'язків ключових слів із значущістю не менше 1.

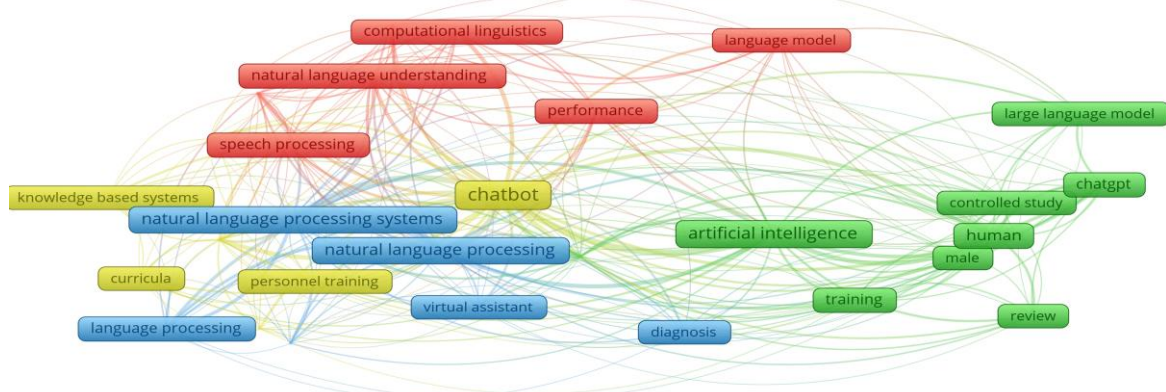


Рис. 1. Карта зв'язків найбільш значущих ключових слів

Аналіз виявив чотири кластери напрямів досліджень: 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. У кожному кластері були визначені найбільш значущі ключові поняття, що дозволило окреслити сучасні пріоритети досліджень у відповідних галузях.

Зокрема, у першому кластері (опрацювання природної мови) найбільш значущими виявилися 7 ключових слів (41%): computational linguistics, natural language understanding, performance, language model, speech processing. Це вказує на фокус досліджень на обчислювальній лінгвістиці, моделюванні та розумінні природної мови. У другому кластері (застосування технологій опрацювання природної мови у суспільстві) виділилися 10 ключових слів (63%): artificial intelligence, large language model, chatgpt, human, training, review, conversational agent, controlled study, education, male. Отже, ключовими є відповідальне використання великих мовних моделей та діалогових агентів, їх вплив на людей та суспільство. Третій кластер (застосування машинного навчання для опрацювання природної мови) охопив 6 найбільш значущих слів (50%): natural language processing, natural language processing systems, virtual assistant, language processing, user interfaces, diagnosis. Фокус тут робиться на розробці віртуальних помічників, природномовних інтерфейсів та систем діагностики. У четвертому кластері (чат-боти у навчанні та сфері послуг) ключовими виявилися 7 слів (58%): chatbot, learning systems, curricula, learn+, knowledge based systems,

personnel training, customer service. Отже, ключові застосування чат-ботів визначено в адаптивних навчальних системах, системах управління знаннями та обслуговуванні клієнтів.

Висновки. Результати дослідження надали можливість визначити провідні напрями досліджень у галузі навчання чат-ботів та пов'язаних технологій. У сфері опрацювання природної мови фокус робиться на обчислювальній лінгвістиці, розумінні та моделюванні мови. У суспільних застосуваннях ключовими є відповідальне використання великих мовних моделей та діалогових агентів. Основні техніки машинного навчання охоплюють віртуальних помічників, інтерфейси та діагностику. Ключові застосування чат-ботів визначено в адаптивних навчальних системах, управлінні знаннями та обслуговуванні клієнтів.

Список використаних джерел

1. Centre for Science and Technology Studies, Leiden University. (2024) VOSviewer - Visualizing scientific landscapes. Retrieved. URL: <https://www.vosviewer.com/>