

A. Dwornik, D. Fabisiak

**ZASTOSOWANIE SZTUCZNEJ INTELIGENCJI JAKO NARZĘDZIA
WSPOMAGAJĄCEGO WYSZUKIWANIE W KATALOGU BIBLIOTECZNYM
(APPLICATION OF ARTIFICIAL INTELLIGENCE AS A SUPPORTING TOOL
FOR SEARCH IN THE LIBRARY CATALOGUE)**

А. Дворнік, Д. Фабізіак

**ЗАСТОСУВАННЯ ШТУЧНОГО ІНТЕЛЕКТУ
ЯК ІНСТРУМЕНТУ ПІДТРИМКИ ПОШУКУ В БІБЛІОТЕЧНОМУ КАТАЛОЗІ**

Narzędzia oparte na tzw. sztucznej inteligencji zaczęły powstawać od lat 50-tych ubiegłego wieku. Powszechna świadomość ich działania, zastosowania i wpływu na pozyskiwanie informacji rozpoczęła się na początku XXI w., a najbardziej przyczyniła się do niego prezentacja przez firmę OpenAI 30 listopada 2022 r. zaawansowanego modelu językowego sztucznej inteligencji — ChatGPT.

Środowisko bibliotekarskie z uwagą obserwowało rozwój tej technologii. Po wdrożeniu systemu nowej generacji Alma z wyszukiwarką Primo w Bibliotece Uniwersyteckiej Uniwersytetu Mikołaja Kopernika w Toruniu (BU UMK), nowym doświadczeniem w zarządzaniu informacją bibliograficzną stało się zintegrowanie niejednorodnych metadanych o wszystkich źródłach wiedzy prezentowanych w jednym środowisku wyszukiwawczym (publikacje drukowane, elektroniczne, zdigitalizowane). Poza niekwestionowanymi korzyściami — jeden punkt wyszukiwawczy, zamiast korzystania z wielu dostępnych baz, skutkowało to niejednorodnością metadanych, a co z tym związane ograniczeniem pełnego wykorzystania możliwości wyszukiwarki — fasetowania. Z nadzieją zatem oczekiwano wspomaganie sztucznej inteligencji w korzystaniu z katalogów bibliotecznych.

We wrześniu 2024 r. firma Exlibris, producent oprogramowania Alma i Primo, zaprezentowała nowe narzędzie wspomagające wyszukiwanie zasobów bibliotecznych w języku naturalnym: Asystenta Wyszukiwania (Research Assistant), w którym zastosowano model językowy oparty na architekturze Retrieval Augmented Generation (RAG). Łączy on możliwości dużego modelu językowego (LLM — Large Language Model) z wiedzą zindeksowaną w Central Discovery Index (CDI) — źródle zawierającym ponad 5 miliardów rekordów dotyczących materiałów naukowych i akademickich w formie elektronicznej pochodzących od wydawców oraz z agregatorów i repozytoriów.

Chociaż wspomaganie wyszukiwania dotyczy publikacji elektronicznych (marzec 2025 r.), to w planach dostawcy oprogramowania jest dalszy rozwój tego narzędzia w kierunku przeszukiwania wszystkich metadanych zawartych w katalogu bibliotecznym. Dla bibliotekarzy tworzących metadane rozumienie jak działa Asystent Wyszukiwania jest kluczowe — metadane i narzędzia, które je interpretują, muszą rozwijać się symbiotycznie. W tym celu BU UMK od 2018 r. współpracuje z Biblioteką Narodową w Warszawie (od czerwca 2021 r. posiada wspólny katalog), m.in. w zakresie rozwoju metadanych. Ze względu na zmianę narzędzi wyszukiwawczych w kontekście sieci semantycznej w BU UMK zrezygnowano z charakterystyki rzeczowej w prekoordynowanym języku haseł przedmiotowych KABA, zastępując ją zasadami katalogowania Biblioteki Narodowej opartymi na języku postkoordynowanym — Deskryptorach Biblioteki Narodowej (DBN). Testy użyteczności prowadzone są w wersji testowej wyszukiwarki Primo BU UMK z aktywną funkcją Asystenta Wyszukiwania dla zalogowanych użytkowników, do czasu przeprowadzenia niezbędnych rekomendacji użytkownikom.

Opis działania

Po wpisaniu zapytania wyszukiwawczego Asystent Wyszukiwania przeszukuje metadane bibliograficzne oraz streszczenia publikacji. Komponuje odpowiedź w języku naturalnym na podstawie 5 źródeł oraz przekierowuje do katalogu biblioteki, gdzie zapytanie jest przekształcane na zapytanie z użyciem operatora logicznego (Boolean query) OR. Nie oferuje wsparcia dla bardziej zaawansowanych operatorów czy niestandardowych masek znakowych, które są szczególnie przydatne w językach fleksyjnych, jakim jest język polski. Umożliwia jednak edycję zapytania po przejściu do katalogu, co pozwala na jego uszczegółowienie. Daje to co prawda dużą precyzję wyszukiwania, ale może też znacząco je zawęzić i pominąć wartościowe dla użytkownika publikacje dostępne w katalogu. Z drugiej strony dzięki mapowaniu operatorów i tworzeniu alternatywnych zapytań w języku polskim i angielskim zmniejsza się ryzyko błędów tłumaczeniowych.

Przykład działania:

Zapytanie: „Jakie skutki dla społeczeństwa polskiego przyniósł Potop Szwedzki?”

Funkcja przekształciła w zapytanie w wyszukiwarce:

(Swedish Deluge impact on Polish society) OR (effects of the Swedish invasion on Poland) OR (Poland social consequences of the Great Northern War) OR (historical analysis of the Swedish Deluge in Poland) OR (cultural changes in Poland after the Swedish Deluge) OR (skutki Potopu Szwedzkiego dla społeczeństwa) OR (Potop Szwedzki a zmiany społeczne w Polsce) OR (wpływ Potopu Szwedzkiego na życie codzienne Polaków) OR (Potop Szwedzki i jego konsekwencje dla kultury polskiej) OR (historie społeczne z czasów Potopu Szwedzkiego) OR (Jakie skutki dla społeczeństwa polskiego przyniósł Potop Szwedzki?)

Otrzymanie trafnych odpowiedzi od Asystenta Wyszukiwania wymaga utworzenia właściwego tzw. prompta (zapytania). Dużym plusem jest możliwość zadawania pytań w dowolnym języku, który później zostanie przetłumaczony na język angielski w celu wygenerowania pełnego zapytania dwujęzycznego. Dzięki temu narzędzie jest intuicyjne i łatwe w użyciu dla początkujących. Jednakże zarówno zbyt ogólne, jak i zbyt szczegółowe zapytania mogą powodować problemy — zbyt wiele wyników lub brak wskazań, co utrudnia pracę typowo naukową. W tym przypadku wśród wskazanych 5 podstawowych publikacji trafne było tylko jedno źródło. Pozostałe dotyczyły m.in. relacji polsko-czeskich w XX wieku, czy współczesnego handlu zagranicznego w Szwecji.

Asystent, dzięki oparciu na architekturze RAG, świetnie radzi sobie z zapytaniami bazującymi na powszechnie przyjętej terminologii. Dlatego jest bardzo pomocny przy pytaniach dotyczących nauk ścisłych. Z kolei w pytaniach z zakresu humanistyki częściej zdarzają się problemy terminologiczne lub błędne, nieprecyzyjne skojarzenia. W podanym przykładzie Potop Szwedzki (1655–1660) został skojarzony z Great Northern War (III Wojna Północna 1700–1721), a wyrażenie «Swedish invasion on Poland» może prowadzić do nieporozumień, gdyż istnieje ryzyko odniesienia do innych konfliktów polsko-szwedzkich. Ponadto Asystentowi zdarzają się również nietrafne przekłady poszczególnych słów rzutujące na nacechowanie i interpretację wypowiedzi, np. użycie „effects” w kontekście skutków inwazji, zamiast „consequences” lub „repercussions”.

Następnym zidentyfikowanym problemem jest jakość i kompletność metadanych bibliograficznych w CDI. W badanym zakresie występują dane przekazane od dostawców publikacji elektronicznych. Obok bardzo szczegółowych opisów bibliograficznych stosunkowo często pojawiają się rekordy zawierające jedynie tytuł i ISBN, co może wpłynąć na umniejszenie wagi danej publikacji dla Asystenta lub nawet całkowite jej pominięcie.

Przeprowadzono również próbę porównania wyników dla tego samego zapytania, o wpływ japońskiego przywódcy, Ody Nobunagi, na historię Japonii, w trzech językach — polskim, angielskim i japońskim. Wykazała ona, że sztuczna inteligencja w każdej wersji językowej wybrała i uszeregowała 5 różnych źródeł, których podsumowania kładły nacisk na inne aspekty historyczne, co doprowadziło do przedstawienia faktów z odmiennych perspektyw. Choć z jednej strony daje to możliwość poszerzenia wiedzy, dzięki poznaniu różnych kontekstów kulturowych czy priorytetów badawczych w danych krajach, może również wpływać na percepcję prezentowanych faktów i brak obiektywizmu. Warto zwrócić też uwagę, że jedynie zapytanie w języku japońskim opierało się na pojedynczych pojęciach, podczas gdy zarówno wersje angielska jak i polska wyszukiwały głównie frazy zdaniowe (小田信長 日本歴史 影響 *Oda Nobunaga historia Japonii wpływ*) vs (*Oda Nobunaga influence on Japan history*) i (*Oda Nobunaga wpływ na historię*).

Oczywiście problemy z odpowiedziami na pytania związane z obszarem Polski lub historią Japonii mogą wynikać ze stosunkowo niewielkiej ilości dostępnych źródeł na te tematy w publikacjach zawartych w CDI. Jednak wyszukiwanie bezpośrednio w wyszukiwarce za pomocą słów kluczowych może zwracać bardziej adekwatne wskazania. Szczególnie, że obejmą wtedy wszystkie metadane w katalogu — również te dla publikacji drukowanych.

Podsumowując, Asystent Wyszukiwania zapewnia wsparcie dla wstępnej eksploracji tematów, dzięki przejrzystości cytowań i bogatej bazie CDI. Krótkie streszczenie otrzymywane po każdym zapytaniu jest przydatnym punktem wyjścia dla dalszych badań. Nadanie priorytetu źródłom, zamiast wygenerowanej parafrazie abstraktów ułatwia użytkownikom dotarcie do pełnych tekstów interesujących ich publikacji. Z drugiej strony Asystent wciąż boryka się z ograniczeniami językowymi i technicznymi, utrudniającymi wykorzystanie go do przeprowadzenia głębszej analizy. Narzędzie jednak bardzo szybko się rozwija, a firma stawia na komunikację z bibliotekami użytkującymi system. W opublikowanej “mapie drogowej” Primo na 2025 r. zasygnalizowano wiele poprawek w działaniu Asystenta Wyszukiwania. Narzędzie oparte jest na produkcie OpenAI — wdrażane są jego wyższe modele, ale dla właściwego efektu model powinien być trenowany na metadanych bibliograficznych z odpowiednią ich hierarchizacją — priorytet powinny mieć hasła przedmiotowe lub słowa kluczowe opisów bibliograficznych. Mamy nadzieję, że w przyszłości Asystent Wyszukiwania stanie się ogromną pomocą dla użytkowników bibliotek.

A. Nowak-Drzymała

**STOWARZYSZENIE BIBLIOTEKARZY POLSKICH: STRUKTURA I DZIAŁANIA
LOKALNE. NA PRZYKŁADZIE DZIAŁALNOŚCI W KOLE PRZY BIBLIOTECIE
UNIWERSYTECKIEJ ORAZ ODDZIALE W TORUNIU. (POLISH LIBRARIANS
ASSOCIATION: STRUCTURE AND LOCAL ACTIVITIES. BASED ON THE ACTIVITIES
OF THE UNIVERSITY LIBRARY CLUSTER AND THE BRANCH IN TORUŃ)**

A. Nowak-Drzymała

**СПІЛКА ПОЛЬСЬКИХ БІБЛІОТЕКАРІВ: СТРУКТУРА ТА МІСЦЕВА ДІЯЛЬНІСТЬ.
НА ПРИКЛАДІ ДІЯЛЬНОСТІ УНІВЕРСИТЕТСЬКОГО БІБЛІОТЕЧНОГО КЛАСТЕРУ
ТА ФІЛІЇ У ТОРУНІ**

Stowarzyszenie Bibliotekarzy Polskich (SBP) jest organizacją pożytku publicznego, zrzeszającą ponad 4 tysiące członków związanych zawodowo lub naukowo z bibliotekarstwem i informacją naukową. Od początku istnienia (1917) działania SBP przyczyniają