

СЕМАНТИЧЕСКИЙ ПОИСК И ПОДБОР КНИГ С ПОМОЩЬЮ ИНТЕГРИРОВАННОГО ИИ-АССИСТЕНТА

Коришов М.И.

*Белорусский государственный университет информатики и радиоэлектроники,
г. Минск, Республика Беларусь*

Научный руководитель: Верхов К.А. – ассистент кафедры ПИКС

Аннотация. В статье рассматривается проблема семантического разрыва в электронных книжных каталогах, из-за которой традиционные поисковые алгоритмы не справляются с ассоциативными запросами пользователей. В качестве решения была предложена модель интеллектуального поиска, объединяющая векторное представление данных и диалогового ИИ-ассистента. Было выявлено, что такой подход позволяет системе перейти от лексического сопоставления слов к глубокому пониманию контекста, что, в свою очередь, избавляет пользователя от необходимости знать точные метаданные книги и существенно повышает релевантность рекомендаций.

Ключевые слова: семантический поиск, ИИ-ассистент, семантический разрыв, векторный поиск, электронный каталог, искусственный интеллект

Введение. При использовании электронных библиотек и книжных онлайн-каталогов пользователи часто сталкиваются с проблемой поиска литературы по неточному или ассоциативному описанию. На практике читатель далеко не всегда знает точное название книги или фамилию автора, формулируя свой запрос через детали сюжета, жанровые особенности или общую атмосферу произведения. Традиционные поисковые алгоритмы, работающие на основе прямого лексического совпадения введенных ключевых слов с атрибутами базы данных, в таких ситуациях показывают крайне низкую эффективность. Возникает семантический разрыв – ситуация, при которой информационная система не способна сопоставить смысловое намерение пользователя с жестко структурированными метаданными каталога.

В данной статье рассматривается подход к преодолению семантического разрыва путем внедрения интегрированного ИИ-ассистента в процесс подбора книг. Применение технологий искусственного интеллекта и алгоритмов семантического поиска позволяет сместить фокус с поиска по символам на поиск по смыслу. ИИ-ассистент способен анализировать неструктурированные запросы на естественном языке, извлекать из них ключевые смысловые признаки и формировать релевантную выдачу даже при полном отсутствии прямых лексических совпадений, что существенно повышает качество работы рекомендательных систем.

Основная часть. Современные информационно-поисковые системы в электронных книжных каталогах преимущественно опираются на лексический анализ (алгоритмы семейства TF-IDF, BM25). В таких системах релевантность документа оценивается по частоте совпадения ключевых слов из запроса с текстом описания или метаданными книги. Данный подход демонстрирует высокую эффективность при поиске по точно известным атрибутам (ФИО автора, ISBN, точное название). Однако при поиске по абстрактному смыслу, сюжетным тропам или эмоциональной окраске произведения классический лексический поиск оказывается полностью неработоспособным. Ввиду нулевого или нерелевантного результата пользователь вынужден покидать пределы электронного каталога и обращаться к сторонним источникам (внешним поисковым системам, тематическим форумам) для выяснения точных атрибутов искомой книги. В результате возникает проблема семантического разрыва – фундаментального несоответствия между ассоциативным мышлением пользователя и жесткой структурой индексов реляционных баз данных [1].

Для преодоления данного ограничения в современных электронных системах применяется парадигма семантического поиска. На технологическом уровне ее реализация базируется на алгоритмах векторного поиска и использовании плотных векторных представлений текста. Текстовые описания книг и пользовательские запросы трансформируются нейросетевыми моделями в многомерные векторы, а процесс поиска сводится к вычислению метрик пространственной близости (например, косинусного расстояния) между вектором запроса и векторами документов. Это позволяет находить литературу по смысловой релевантности, даже если в текстах нет ни одного общего слова [2].

Однако внедрение исключительно векторного поиска не решает проблему комплексной обработки сложных, неструктурированных запросов. Эффективность системы многократно возрастает при интеграции в архитектуру сервиса диалогового ИИ-ассистента на базе больших языковых моделей. В рамках процесса подбора литературы ИИ-ассистент выполняет три ключевые функции [3]:

1 Интерпретация пользовательского намерения. Интерпретация включает в себя глубокий анализ запроса на естественном языке, выделение неявных сущностей (жанр, сеттинг, архетипы персонажей) и очистка ввода от информационного шума.

2 Оптимизация поискового вектора. Что означает преобразование интерпретированного запроса в структурированный формат (запрос), оптимальный для обращения к векторному индексу каталога.

3 Генерация аргументированной выдачи. В первую очередь – анализ полученной выборки релевантных книг и формирование интерактивного ответа, в котором система логически обосновывает пользователю, почему конкретное произведение соответствует его изначальным ожиданиям.

Сравнительная характеристика подходов к организации информационного поиска в книжных сервисах представлена в таблице 1.

Таблица 1 – Сравнительный анализ парадигм поиска литературы

Критерий оценки	Лексический (традиционный) поиск	Семантический поиск с ИИ-ассистентом
Метод сопоставления данных	Поиск подстрок и частотный анализ совпадений терминов.	Вычисление векторной близости и контекстное понимание текста.
Обработка абстракций и синонимов	Требует постоянной поддержки и обновления словарей синонимов вручную.	Встроена в архитектуру нейросетевой модели по умолчанию.
Пользовательское взаимодействие	Однонаправленное (ввод запроса – получение статического списка).	Интерактивное (ведение диалога, уточнение контекста, аргументация выбора).

Таким образом, переход от прямого поиска по атрибутам к семантическому поиску с интегрированным ИИ-ассистентом меняет саму логику взаимодействия пользователя с электронным каталогом. Интеллектуальный агент берет на себя вычислительно сложную задачу перевода человеческих абстракций в машинные алгоритмы, выступая интеллектуальным шлюзом между неструктурированной потребностью клиента и базой данных.

Заключение. Проведенный анализ методов информационного поиска показал, что традиционные лексические алгоритмы обладают критической уязвимостью при обработке ассоциативных запросов в книжных онлайн-каталогах. Установлено, что применение технологий семантического векторного поиска в комплексе с диалоговыми ИИ-ассистентами позволяет эффективно решить проблему семантического разрыва. Данный подход обеспечивает выделение скрытого смысла из неструктурированных пользовательских формулировок и автоматизирует процесс подбора релевантной литературы.

Также определено, что внедрение ИИ-ассистента не только повышает точность поисковой выдачи, но и качественно трансформирует пользовательский опыт за счет

интерактивного взаимодействия и аргументации предлагаемых результатов. В практическом плане использование рассмотренных интеллектуальных технологий позволяет удерживать пользователя внутри экосистемы каталога, исключая необходимость обращения к сторонним поисковым ресурсам, что напрямую способствует повышению эффективности электронных систем продаж.

Список литературы

1. Маннинг, К. Введение в информационный поиск / К. Маннинг, П. Рагхаван, Х. Шютце. – Москва: Вильямс, 2011. – 528 с.
2. Шалагин, Н. Д. Обзор алгоритмов семантического поиска по текстовым документам / Н. Д. Шалагин // International Journal of Open Information Technologies. – 2024. – Т. 12, № 8. – С. 98–105.
3. Малышев, А. А. Использование метода RAG и больших языковых моделей в интеллектуальных образовательных экосистемах / А. А. Малышев, А. В. Смирнов // Информатика и образование. – 2024. – Т. 39, № 3. – С. 45–56.

UDC 004.891.2

SEMANTIC SEARCH AND BOOK SELECTION USING AN INTEGRATED AI ASSISTANT

Karshou M.I.

Belarusian State University of Informatics and Radioelectronics, Minsk, Republic of Belarus

Verkhov K.A. – assistant of the department of ICSD

Annotation. The article addresses the problem of the semantic gap in electronic book catalogs, which prevents traditional search algorithms from effectively processing associative user queries. As a solution, an intelligent search model combining vector data representation and a conversational AI assistant is proposed. It was found that this approach enables the system to transition from lexical word matching to deep context understanding, which, in turn, relieves the user from needing exact book metadata and significantly improves search relevance.

Keywords: semantic search, AI assistant, semantic gap, vector search, electronic catalog, artificial intelligence