International Journal of Scientific Research and Engineering Development-– Volume 7 Issue 6, Nov-Dec 2024
Available at www.ijsred.com

RESEARCH ARTICLE OPEN ACCESS

Evolution of Search From Lexical Methods to Semantic Technologies With AI-Driven Vector Models

Andrei Ivanouski
Lead Software Engineer, Contexxt Ltd.
Brest, Belarus

Abstract:

The article is devoted to the evolution of search technologies, which is caused by the development of technological processes. Because of this, a transition was made from simple lexical methods to modern semantic technologies based on vector models controlled by artificial intelligence. Initially, search engines relied on keyword selection, which, despite its simplicity, often did not allow them to assess the contextual significance of queries. The main purpose of the work is to consider the historical process of development of search technologies, as well as to study the role of artificial intelligence, which is used in vector models.

The article provides an overview of various search models, starting with logical, vector spatial, and ending with modern semantic search technologies based on artificial intelligence. It also describes how these achievements have improved the accuracy and relevance of search results.

The information presented in the framework of the work will be useful for consideration by other scientific authors, as well as programmers and employees working in the field of information retrieval, machine learning and artificial intelligence. Due to the fact that it gives an idea of the evolution of search engines, thereby provides a basis for understanding future trends in the development of search technologies.

Keywords: Lexical Search, Semantic Technologies, AI-Driven Vector Models, Information Retrieval, Deep Learning, Natural Language Processing (NLP), Machine Learning.

INTRODUCTION

Technological advancements have driven changes in the mechanisms of search systems. Initially, search systems relied on approaches based on recognizing specific words. However, with the improvement of technology, the demands of both users and organizations increased, necessitating the modernization of search technologies.

Currently, search is based on the use of vector models, deep learning, artificial intelligence algorithms, and natural language processing, which have significantly improved accuracy. The lexical approach to information retrieval, which relied on direct word matching, was limited by its inability to interpret synonyms and nuances in query meanings. Semantic search, by contrast, focuses on understanding user intent by analyzing relationships between words. This shift has accelerated technological progress.

Transformer models, such as BERT, operating on artificial intelligence algorithms, have enabled search systems to comprehend and process language. This progress underscores the role of AI in solving complex information retrieval tasks. The integration of deep learning methods with traditional search algorithms has enhanced accuracy and expanded applications in fields such as personalized recommendations. As a result, modern search systems have become tools that transform the way users interact with information, paving the way for future innovations in knowledge discovery.

The objective of this work is to examine the historical evolution of search technologies, with a focus on the transformational role of AI-driven vector models.

MATERIALS AND METHODS

This study employs interdisciplinary approaches, including a literature review, comparative analysis of algorithms and their applications, and modeling based on recent advancements in machine learning and natural language processing.

The work by Muniyappa C. and Kim E. [1] describes an approach based on a universal sentence encoder combined with algorithms to enhance search accuracy, particularly relevant for complex textual data.

For multitask optimization, Liu S. et al. [2] propose an accelerated evolutionary search strategy that leverages neural networks to improve efficiency. This approach reduces data processing time and achieves high-quality solutions in fewer iterations.

Regarding neural network search, noteworthy solutions are presented in the works of Kobayashi M. and Nagao T. [3], as well as Awad N. and Mallik N. [7]. The former describes a method combining modern algorithms with techniques for effectively searching differentiable architectures, resulting in enhanced performance. Awad N. and Mallik N. [7] emphasize the use of evolutionary strategies, highlighting their ability to deliver consistent results.

The framework presented in the study by Xu J., Jin Y., and Du W. [4] enables data optimization in distributed systems, facilitating work with large datasets without requiring centralized data collection.

To improve the convergence and optimization characteristics of differential evolution, Xu J. et al. [5] introduce an algorithm based on chaotic search strategies. This approach enhances the effectiveness of evolution in complex optimization tasks by reducing the likelihood of being trapped in local minima.

In the context of generative models and unsupervised learning, Chen S. et al. [6] propose using networks to search for architectures suitable for zero-shot learning tasks. According to the authors, this opens new prospects for algorithms, enhancing their ability to generate new data under limited training conditions.

Finally, approaches based on continuous evolution, as proposed by Yang Z. et al. [8], offer innovative solutions for neural architecture search. These methods ensure efficient parameter utilization and improve results over successive generations.

An interview published on the website seb.e.lanbook.com [9] was also analyzed, providing insights into future trends in search technology development.

The studies presented demonstrate a wide range of applications for evolutionary algorithms in optimization and search tasks. Their combination can improve solution quality, ultimately paving the way for promising directions in both scientific research and practical applications.

RESULTS AND DISCUSSION

Information retrieval has undergone significant transformations over the past few decades, evolving from primitive keyword matching algorithms to sophisticated AI-based systems capable of understanding context and user intent. In the early stages of search engine development, lexical methods dominated, focusing on exact term matching across vast document collections. As data complexity increased, search technologies advanced, incorporating models and semantic structures that transformed search from a mechanical term-matching task into a complex process where algorithms predict user intent. Figure 1 illustrates the evolution of search, highlighting the shift from lexical methods to semantic technologies with AI-driven vector models.

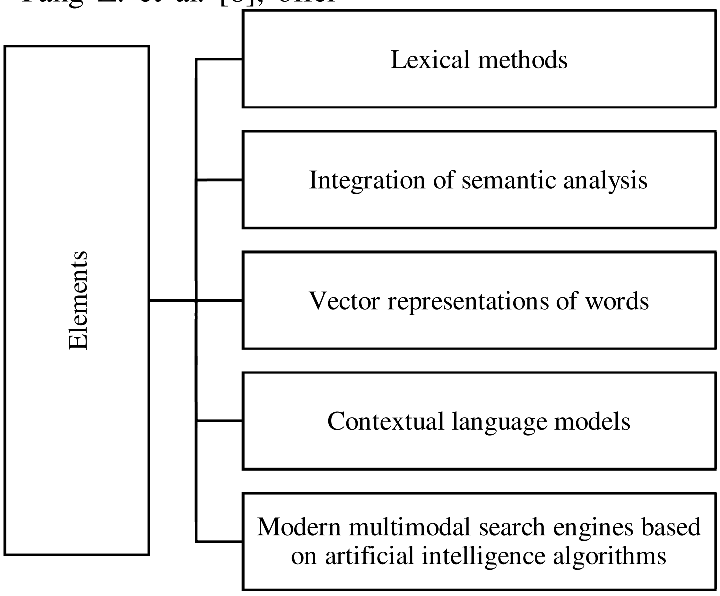

Fig. 1. Elements reflecting the evolution of search from lexical methods to semantic technologies with AI-driven vector models (compiled independently).

Early information retrieval methods were primarily lexical, relying on simple logical operators. These systems were based on the concept of an inverted index, which linked each term to a list of documents. While effective for basic data extraction, this approach did not account for language nuances. Its strict dependence on term matching led to a superficial understanding of content and limited search performance in complex scenarios.

With the emergence of large digital repositories and growing demand for advanced search capabilities, lexical models became increasingly inadequate. This gap drove the development of statistical and probabilistic search models, which aimed to measure the likelihood of a document's relevance to a given query based on statistical patterns rather than direct term matching. A breakthrough in this field was the introduction of the vector space model, representing both queries and documents as vectors in a multidimensional space. In this model, terms were treated as dimensions, and the relevance of a document was determined by the cosine similarity between the query vector and the document vector.

The emergence of semantic search marked a significant departure from traditional lexical and statistical models, aiming to improve the quality of search results by focusing not only on the query terms themselves but also on their meanings and relationships. This shift in focus was driven by advancements in computational linguistics and formal knowledge representation methods, including ontologies and knowledge graphs.

Ontologies provided a formalized structure for representing relationships between concepts. A well-structured ontology enabled search systems to understand that terms such as "car" and "automobile" are synonyms and should be treated interchangeably. Knowledge graphs, which model information as networks of interconnected entities and their relationships, allowed search systems to effectively capture and represent complex real-world knowledge [6].

A breakthrough in search technology occurred with the advent of deep learning and neural networks. Artificial intelligence technologies enabled a new approach to information classification. Previously, search and result generation were based on semantics, focusing on keywords. Next-generation search mechanisms aim not merely to locate sources containing the desired keywords but to comprehend the query's meaning and identify its context. This involves finding text within a database and determining whether it aligns semantically with the user's query [2].

This approach eliminates the need for extensive lists of links to sources related to the query, even with advanced relevance sorting, filtering, and similar options. Instead, the search result is a single or a few generated responses directly addressing the query.

This new method of information retrieval transforms the process into a fundamentally different experience, resembling a conversation with an intelligent entity. Such a system can recognize a broader range of meanings and user needs embedded in the query. In other words, interaction with a next-generation search engine occurs in natural human language.

Vector representations of words, such as Word2Vec, GloVe, and FastText, represent words as dense elements in continuous spaces where similar terms are located closer to each other. These representations capture semantic relationships between terms by modeling their usage across large datasets. Figure 2 illustrates the components of vector representations of words in search queries.

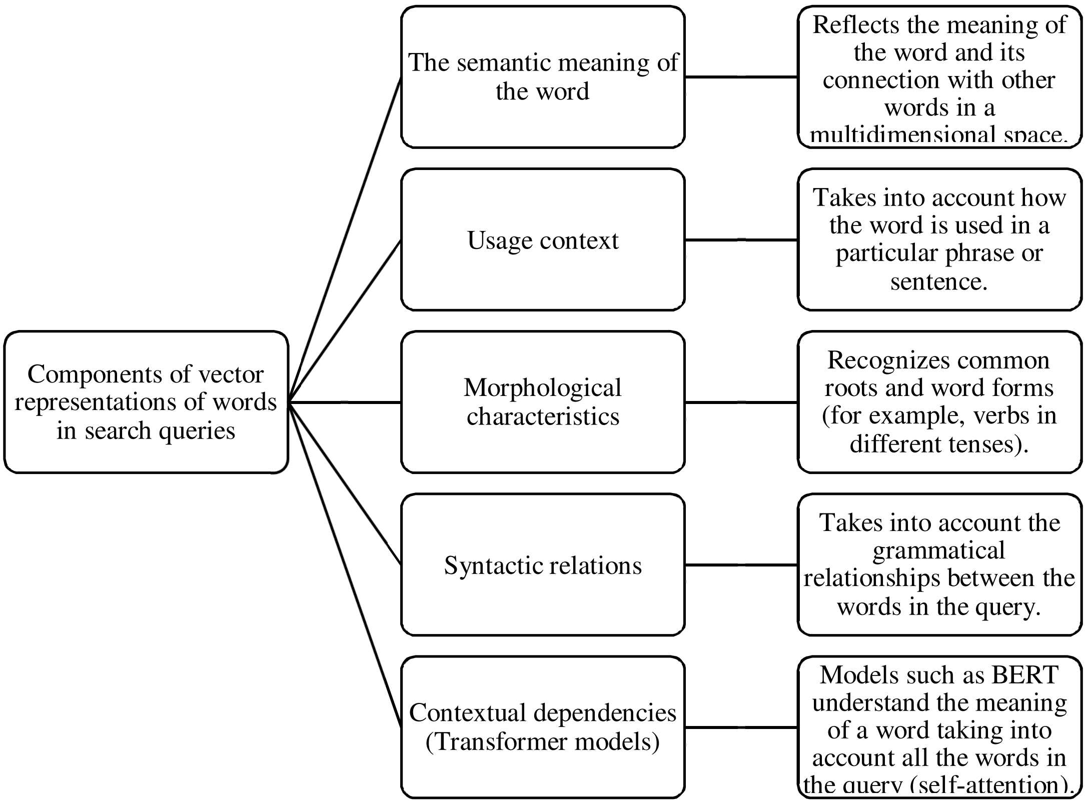

Fig. 2. Components of vector representations of words in search queries (compiled independently).

Moreover, transformer-based models, such as Generative Pretrained Transformers (hereafter GPT), play a critical role in enabling search systems to generate human-like text, facilitating advanced forms of interactive search. GPT technologies serve as a tool to enhance search capabilities rather than replacing traditional search mechanisms entirely. This is comparable to the evolution of communication methods: while video calls became possible with smartphones, traditional mobile calls and text messaging did not disappear. Instead, a new form of communication was added, complementing the existing ones. The use of large pre-trained models in natural language processing makes semantic search more adaptive to language nuances, improving query interpretation and real-time response generation [8].

It is likely that under the influence of these technologies, the user experience will evolve over time, with the balance between traditional and new search methods shifting toward increased interaction with artificial intelligence. However, this does not preclude the preservation of the traditional format for presenting search results, alongside the introduction of additional options for obtaining answers.

Even global technology giants like Google are expected to retain the classic version of their search engines. An expert quoted by Vedomosti noted, "The company [Google] plans not only to update its search system but also to create an entirely new one. The main search engine is likely to remain as it is now, while the new one will be in a chat format to retain users who are accustomed to the current version and not ready to immediately switch" [9].

These models do not simply focus on retrieving documents based on keyword matching; they represent both queries and documents as vectors in multidimensional spaces, allowing the system to measure similarity. By encoding representations of both query terms and documents, these models enable systems to find results that match user intent, even when the exact query terms are absent from the document.

The transition to vector-based search has introduced the use of vector spaces, where entire documents are represented as points in continuous vector spaces. Similarity between documents or between a query and a document is determined by calculating the distance between them, typically using cosine similarity or other distance metrics.

Semantic search operates directly on text. If a query or its synonymous terms are not present in the text, the source of that text will not appear at the top of the search results. The structure of semantic search is based on metadata, including source titles, publication years, authors, and education levels for which the materials are intended. This metadata is manually entered into the database, introducing human factors such as slower processing speed and the risk of errors. A well-trained vector search algorithm performs the same tasks faster, on a larger scale, and without time constraints. It "sees" the entire content of a book at once and finds the answer to a user's query within the content itself, rather than simply providing a source based on its title, publication year, author, or catalog category.

Text vectorization fully transitions search to the next generation. In vectorized text, input data is converted into vectors that machine learning models can process, allowing artificial intelligence to work seamlessly with the text [7].

Below, Table 1 describes the advantages and disadvantages of semantic technologies with vector models driven by artificial intelligence algorithms.

TABLE I . Advantages and Disadvantages of Semantic Technologies with AI-Driven Vector Models (compiled independently).

Category	Advantages	Disadvantages
Search Quality	- Understanding query meaning and context, not just keywords.	- Potential errors in interpreting complex or ambiguous queries.
Usability	- User-friendly, no need for precise query formulation.	- Possibility of oversimplifying queries, leading the system to miss important details.
Performance	- Fast query processing through optimized vector operations on modern platforms.	- High computational resource requirements for building and running models.
Learning and Adaptation	- Models can learn and adapt to new data, improving search quality.	- Dependence on large volumes of high-quality training data.
Scalability	- Capability to handle large datasets, including unstructured sources.	- Integration challenges with existing infrastructure, especially in distributed systems.
Multimodality Support	- Ability to integrate textual, audio, and visual data into a unified search model.	- Complexity in processing multimodal data and the need for specialized hardware.
Data Quality	- Can process "dirty" or incomplete data through training on large datasets.	- Risk of low accuracy when training data contains incorrect or conflicting information.
Economic Aspects	- Cost savings on developing traditional search algorithms due to the versatility of vector models.	- High initial costs for model training and periodic retraining.
Ethics and Control	- Capability for automatic filtering of undesirable content (e.g., spam).	- Potential issues with transparency and explainability of AI decisions.

Search technologies are expected to continue evolving, integrating into all levels of personalization based on user behavior and preferences.

Beyond processing textual queries, modern systems are now capable of handling multimodal data, including images, videos, and voice. Technologies such as Contrastive Language-Image Pretraining combine these representations to facilitate search, allowing users to input queries not only through text but also via other formats.

Another promising area is the growth of conversational artificial intelligence. Leveraging large language models, search systems now engage in natural, dynamic, context-aware conversations with users. Instead of relying on single, isolated queries, users can conduct ongoing dialogues with search systems, where the system anticipates follow-up questions and refines results based on the conversation [3].

Thus, the evolution of search technologies, from simple lexical models to sophisticated AI-based systems, represents a monumental shift in information retrieval. As these technologies continue to advance, the integration of personalization, multimodal data processing, and conversational AI will transform interactions with information, creating a future where search becomes not merely a tool for data retrieval but a dynamic, intelligent assistant capable of anticipating and responding to complex user needs.

CONCLUSION

This study explored the development of search technologies, starting from the use of basic lexical methods and culminating in modern semantic technologies based on AI-driven vector models.

The implementation of machine learning methods, including vector spaces and transformer models such as BERT, has significantly improved search accuracy and relevance. This progress was enabled by the ability of new models to account for context, polysemy, and relationships between concepts. Additionally, the integration of deep learning and natural language processing has allowed search systems to adapt to user query structures, expanding their functionality in areas such as personalization, multimodal search, and conversational interfaces.

The evolution of search has not only demonstrated technological progress but also established a new paradigm for processing and accessing essential information.

V. REFERENCES

[1] Muniyappa C., Kim E. Evolutionary Algorithms Approach For Search Based On Semantic Document Similarity //Proceedings of the 2023 11th International Conference on Computer and Communications Management, 2023, pp. 123-128.

[2] Liu S. et al. Learning to accelerate evolutionary search for large-scale multiobjective optimization //IEEE Transactions on Evolutionary Computing, 2022, vol. 27, No. 1, pp. 67-81.

[3] Kobayashi M., Nagao T. An evolution-based approach for efficient differentiated architecture search //Proceedings of the 2020 Genetic and Evolutionary Computing Conference Companion, 2020, pp. 131-132.

[4] Xu J., Jin Y., Du W. A federated data-driven evolutionary algorithm for expensive multi-/many-objective optimization //Complex & Intelligent Systems, 2021, vol. 7. No. 6, pp. 3093-3109.

[5] Xu J. et al. A federated data-driven evolutionary algorithm //Knowledge-based systems, 2021, vol. 233, p. 107532.

[6] Chen S. et al. Egans: Evolutionary generative adversarial network search for zero-shot learning //IEEE Transactions on Evolutionary Computation, 2023.

[7] Awad N., Mallik N., Hutter F. Differential evolution for neural architecture search //arXiv preprint arXiv:2012.06400, 2020.

[8] Yang Z. et al. Cars: Continuous evolution for efficient neural architecture search //Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2020, pp. 1829-1838.

[9] From semantics to meaning: the future of search technologies. [Electronic resource] Access mode: https://seb.e.lanbook.com/tpost/ilia2ndts1-ot-semantiki-k-smislu-buduschee-poiskovi (date of application: 14.11.2024).