

Analyzing the Central Role of Language in the Information Retrieval Process: From Semantics to Enhancing User Interaction and System Performance

naeimeh zarif ghasemian¹

Abstract

The present study aims to examine the central role of language in the information retrieval process and to analyze the effects of syntactic structures, semantic features, and linguistic differences between users and search systems on the accuracy and relevance of retrieved results. It seeks to elucidate existing linguistic barriers in transforming information needs into search logic and to propose solutions for enhancing convergence between user language and the language of retrieval systems. This research adopts an analytical-descriptive retrieval systems approach. The data were obtained through the analysis of 200 scientific and specialized documents selected from reputable information databases. The documents were chosen based on criteria of scientific quality, topical relevance, and diversity of subject domains, and were subsequently examined using qualitative and quantitative analyses. The results grounded in syntactic, lexical, and semantic indicators indicate that clear and unambiguous syntactic structures in user queries play a decisive role in improving the precision and recall of search systems. Semantic analysis further revealed that the use of accurate, domain-specific, and contextually appropriate terminology significantly enhances retrieval quality and reduces the incidence of irrelevant documents. The study also demonstrates that a substantial portion of search inefficiency arises from the gap between users' natural

¹ PhD student in Information Science and Knowledge; Faculty of Educational Sciences and Psychology, Ferdowsi University; Library expert, Faculty of Literature, Ferdowsi University; Mashhad; Iran

language and the structured language of retrieval systems. Employing language translation tools, standardized vocabularies, and expert professional feedback can effectively mitigate this gap. From a theoretical perspective, the findings suggest that linguistic analysis in information retrieval can benefit from insights drawn from the philosophy of language and theories of meaning in use. The study confirms that language is not merely a tool for expressing information needs but a fundamental element in their formation, interpretation, and processing. Simultaneous attention to the syntactic, semantic, and contextual dimensions of queries, along with improvements in translation mechanisms between natural language and search languages, can lead to substantial enhancements in the accuracy and efficiency of information retrieval systems. These findings provide a foundation for the development of semantic-based systems, improved human-machine interaction, and the design of search engines that are more closely aligned with users' real informational needs.

Keywords: Role of Language; Syntactic Structure; Precision and Recall; Search Language; Information Retrieval; Relevance; Meaning



واکاوی نقش محوری زبان در فرآیند بازیابی اطلاعات: از معناشناسی تا بهبود تعامل کاربر و عملکرد نظام‌های اطلاعاتی

نعمه ظریف قاسمیان^۱

چکیده

پژوهش حاضر با هدف واکاوی نقش محوری زبان در فرآیند بازیابی اطلاعات انجام شده است و تلاش دارد تأثیر ساختارهای نحوی، ویژگی‌های معناشناختی و تفاوت‌های زبانی میان کاربر و سامانه‌های جستجو را بر دقت و مرتبط بودن نتایج بازیابی تحلیل کند. این مطالعه در پی تبیین موانع زبانی موجود در تبدیل نیاز اطلاعاتی به منطق جستجو و ارائه راهکارهایی برای ارتقای همگرایی میان زبان کاربران و زبان سیستم‌های بازیابی است. این پژوهش از رویکرد تحلیلی-توصیفی بهره برد و داده‌های آن از تحلیل ۲۰۰ سند علمی و تخصصی انتخاب شده از پایگاه‌های معتبر اطلاعاتی به دست آمد. اسناد بر اساس معیارهای کیفیت علمی، ارتباط موضوعی و تنوع حوزه محتوایی گزینش و سپس با استفاده از شاخص‌های نحوی، واژگانی و معناشناختی مورد تحلیل کیفی و کمی قرار گرفتند. نتایج نشان داد ساختارهای نحوی روشن و فاقد ابهام در پرسش‌های کاربران، نقش تعیین‌کننده‌ای در افزایش دقت و بازیافت سامانه‌های جستجو دارند. تحلیل معنایی نیز آشکار ساخت که انتخاب واژگان دقیق، تخصصی و متناسب با زمینه موضوعی موجب ارتقای کیفیت بازیابی و کاهش اسناد نامرتبط می‌شود. پژوهش همچنین نشان داد که بخش مهمی از ناکارآمدی در جستجو ناشی از شکاف میان زبان طبیعی کاربران و زبان ساخت‌یافته سیستم‌هاست و بهره‌گیری از ابزارهای ترجمه زبانی، واژگان استاندارد و بازخورد حرفه‌ای کارشناسان می‌تواند این شکاف را کاهش دهد. نتایج از منظر نظری نیز نشان داد که تحلیل زبان در بازیابی اطلاعات می‌تواند از دیدگاه‌های فلسفه زبان و نظریه کاربرد معنا بهره‌مند شود. پژوهش حاضر اثبات کرد که زبان نه تنها ابزار بیان نیاز اطلاعاتی، بلکه عنصر بنیادین در شکل‌گیری، تفسیر و پردازش آن است. توجه همزمان به ابعاد نحوی، معناشناختی و زمینه‌ای پرسش‌ها و بهبود سازوکارهای ترجمه میان زبان طبیعی و زبان جستجو می‌تواند موجب ارتقای چشمگیر دقت و کارایی نظام‌های بازیابی شود. این نتایج زمینه‌ساز توسعه سیستم‌های معنایه، بهبود تعامل انسان و ماشین و طراحی موتورهای جستجویی سازگارتر با نیازهای واقعی کاربران خواهد بود.

^۱ دانشجوی دکتری علم اطلاعات و دانش‌شناسی دانشگاه فردوسی مشهد، مدرس دانشگاه فردوسی مشهد و متخصص کتابخانه دانشکده ادبیات و علوم انسانی دانشگاه فردوسی مشهد. Na.zarif@mail.um.ac.ir

کلیدواژه‌ها: نقش زبان، ساختار نحوی، دقت و بازیافت، زبان جستجو، بازیابی اطلاعات،
ارتباط، معنا

زودایند ویراستاری نشده (فصلنامه بازیابی دانش و نظام‌های معنایی)

مقدمه

توانایی انسان در سازماندهی، تفسیر و بهره‌گیری از دانش، یکی از بنیادی‌ترین وجوه تمایز او از سایر موجودات به‌شمار می‌آید. این توانایی نه صرفاً محصول انباشت داده‌ها، بلکه حاصل فرایندی پیچیده از دریافت، پردازش و معناپردازی اطلاعات است؛ فرایندی که در بستر تعاملات زبانی و شناختی شکل می‌گیرد. در عصر دیجیتال، که بخش عمده‌ای از دانش انسانی از طریق سامانه‌های اطلاعاتی تولید و بازیابی می‌شود، کیفیت این فرایند بیش از هر زمان دیگری به چگونگی تعامل میان کاربر، زبان و نظام‌های بازیابی اطلاعات وابسته است.

در این میان، زبان نقشی محوری ایفا می‌کند. زبان نه تنها ابزار بیان نیاز اطلاعاتی، بلکه سازوکار اصلی مفهوم‌سازی، صورت‌بندی مسئله و ارزیابی نتایج جست‌وجو است. پژوهش‌های کلاسیک در حوزه بازیابی اطلاعات نشان داده‌اند که نیاز اطلاعاتی کاربر پیش از آنکه به یک پرسش صریح تبدیل شود، حالتی ذهنی، ناپایدار و وابسته به زمینه دارد (Belkin, 1980). این نیاز ذهنی تنها از طریق زبان به سامانه منتقل می‌شود و کیفیت این انتقال، تأثیر مستقیمی بر کارآمدی فرایند بازیابی دارد. از این منظر، تعامل انسان با نظام‌های اطلاعاتی اساساً تعاملی زبان‌محور است؛ تعاملی که موفقیت آن به هم‌خوانی میان بازنمایی زبانی نیاز کاربر و ساختارهای زبانی و مفهومی سامانه وابسته است (Marchionini, 2008).

با وجود این، زبان طبیعی ذاتاً واجد ابهام، چندمعنایی و وابستگی به بافت است. معنا نه صرفاً در سطح واژگان، بلکه در بستر کاربرد آن‌ها و در پیوند با دانش پیشین، هدف جست‌وجو و موقعیت کاربر شکل می‌گیرد. این امر سبب می‌شود که ترجمه زبان طبیعی کاربر به زبان جست‌وجوی ساخت‌یافته، همواره با چالش‌هایی اساسی همراه باشد. پژوهش‌های حوزه معناشناسی و بازیابی تعاملی نشان می‌دهند که ناهماهنگی میان زبان کاربر و زبان نظام‌های نمایه‌سازی و بازیابی، یکی از عوامل اصلی افت دقت و رضایت کاربران است (Ingwersen & Järvelin, 2005; Baeza-Yates & Ribeiro-Neto, 2011).

در سال‌های اخیر، با گسترش پایگاه‌های اطلاعاتی پیچیده، تنوع کاربران و رشد فناوری‌های مبتنی بر پردازش زبان طبیعی، توجه پژوهشگران بیش از پیش به نقش معنا، بافت و تعامل زبانی در بازیابی اطلاعات معطوف شده است (Manning et al., 2008; Croft, 2018; Metzler, & Strohman, 2018). با این حال، بخش قابل توجهی از پژوهش‌های موجود تا به امروز بر جنبه‌های فنی الگوریتم‌های بازیابی تمرکز داشته‌اند، یا زبان را صرفاً به عنوان ورودی صورتی سامانه تلقی کرده‌اند. در نتیجه، تحلیل یکپارچه زبان به عنوان پدیده‌ای نحوی، معنایی و زمینه‌مند در کل چرخه بازیابی اطلاعات، همچنان با کمبود مطالعات نظری منسجم مواجه است.

بر این اساس، خلأ اصلی در ادبیات پژوهش را می‌توان در فقدان نگاهی جامع دانست که زبان را نه تنها در سطح پرسش، بلکه در تمامی مراحل شکل‌گیری نیاز اطلاعاتی، تعامل کاربر با سامانه و تفسیر نتایج بازیابی مورد واکاوی قرار دهد. این پژوهش با هدف پاسخ‌گویی به این خلأ، به تبیین نقش محوری زبان در فرایند بازیابی اطلاعات می‌پردازد و تأثیر ویژگی‌های نحوی، معنایی و زمینه‌ای زبان را بر کیفیت جست‌وجو و تعامل کاربر تحلیل می‌کند. در این راستا، ضمن بازخوانی مبانی نظری زبان و ارتباط در حوزه علم اطلاعات، چالش‌های ناشی از فاصله میان زبان طبیعی کاربر و زبان نظام‌های بازیابی بررسی شده و راهکارهایی مفهومی برای افزایش همگرایی میان این دو سطح زبانی پیشنهاد می‌شود.

زبان چیست؟

زبان از بنیادی‌ترین سازوکارهای ارتباط و شناخت انسانی است و همواره در کانون توجه زبان‌شناسان، فیلسوفان و نظریه‌پردازان معنا قرار داشته است. در تعریف‌های کلاسیک، زبان به مثابه نظامی قراردادی از نشانه‌های آوایی و نمادین در نظر گرفته می‌شود که اعتبار و کارکرد آن‌ها در چارچوب توافقی‌های اجتماعی یک جامعه زبانی شکل می‌گیرد. بر اساس این دیدگاه، زبان ابزار اصلی انتقال اندیشه، تجربه و معنا میان افراد است (حجازی، ۱۳۷۹).

نجدفی (۱۳۵۸) نیز با تأکید بر کارکرد ارتباطی زبان، آن را نظامی می‌داند که تجربه انسانی را به واحدهای معنادار تقسیم کرده و از طریق ساختارهایی چون واج، تکواژ و جمله صورت‌بندی می‌کند.

در نظریه‌های زبان‌شناسی معاصر، برداشت‌های متنوع‌تری از ماهیت زبان مطرح شده است. در رویکردهای تجربه‌گرا، زبان عمدتاً به‌عنوان مجموعه‌ای از رفتارهای قابل مشاهده میان گوینده و شنونده تحلیل می‌شود؛ در مقابل، رویکرد زایشی - شناختی، به‌ویژه در آثار چامسکی، بر جنبه‌های ذهنی و ساختارهای درونی زبان تأکید دارد. چامسکی با تفکیک میان «توانش زبانی» و «کنش زبانی»، زبان را نه صرفاً مجموعه‌ای از رفتارهای زبانی، بلکه بازتاب دانشی ضمنی و نظام‌مند می‌داند که امکان تولید و درک بی‌نهایت جمله را برای گویشور فراهم می‌سازد (Chomsky, 1965, Chapman, 2000). این تمایز، نقش سازوکارهای شناختی در پردازش معنا و تفسیر زبان را برجسته کرده و زمینه‌ای نظری برای تحلیل زبان در تعامل انسان و نظام‌های اطلاعاتی فراهم می‌آورد.

در سنت ساختارگرایی، فردینان دو سوسور زبان را نهادی اجتماعی و نظامی از روابط می‌داند که معنا در آن نه به‌صورت فردی، بلکه در شبکه‌ای از تمایزها و تقابلهای شکل می‌گیرد. از نظر سوسور، زبان پدیده‌ای جمعی است که خارج از اراده فرد وجود دارد و کارکرد اصلی آن، امکان‌پذیر ساختن ارتباط در بستر اجتماعی است (Saussure, 1916). این نگاه، توجه را از واژه‌های منفرد به ساختار کلی زبان و نقش نظام‌مند آن در تولید معنا معطوف می‌کند.

در ادامه این جریان، زبان‌شناسانی چون ادوارد ساپیر و بنجامین لی ورف بر نقش فعال زبان در شکل‌دهی به تجربه و ادراک انسانی تأکید کرده‌اند. فرضیه نسبیت زبانی، هرچند در شکل قوی آن مورد مناقشه قرار گرفته، نشان می‌دهد که ساختارهای زبانی می‌توانند بر شیوه‌های دسته‌بندی تجربه، توجه و تفسیر واقعیت اثرگذار باشند (استاینبرگ، ۱۳۸۶). این دیدگاه،

اهمیت زبان را نه فقط به عنوان بازتاب جهان، بلکه به مثابه عاملی مؤثر در سازماندهی شناخت انسانی برجسته می‌سازد.

در حوزه فلسفه زبان، به‌ویژه در اندیشه متأخر ویتگنشتاین، معنا نه ویژگی ذاتی واژگان، بلکه حاصل کاربرد آن‌ها در بافت‌های اجتماعی و «بازی‌های زبانی» تلقی می‌شود. از این منظر، فهم معنا مستلزم توجه به زمینه‌های کاربردی زبان است و هرگونه جداسازی معنا از بافت استعمال، به ساده‌سازی مفرط می‌انجامد (واله، ۱۳۸۲). این رویکرد، نقش زمینه و کاربرد را به عنوان مؤلفه‌ای اساسی در تفسیر زبان برجسته می‌کند؛ مؤلفه‌ای که در تعامل کاربر با نظام‌های بازیابی اطلاعات اهمیتی بنیادین دارد.

در میان زبان‌شناسان ایرانی، باطنی زبان را نظامی چندبعدی می‌داند که هم واجد ساختارهای صوری و نحوی است و هم پیوندی باگسستگی با معنا و جهان بیرونی دارد. به‌زعم او، تحلیل زبان بدون توجه هم‌زمان به این دو بُعد، به‌درکی ناقص از کارکردهای ارتباطی و شناختی آن می‌انجامد (باطنی، ۱۳۷۸؛ ۱۳۸۵؛ ۱۳۸۶). این نگاه، امکان بررسی زبان به عنوان واسطه‌ای میان ذهن، معنا و واقعیت را فراهم می‌سازد.

با وجود تنوع این رویکردها، اجماعی کلی بر این نکته وجود دارد که زبان پدیده‌ای پیچیده، پویا و میان‌رشته‌ای است و نمی‌توان آن را به یک سطح تحلیلی واحد فروگاست. آنچه در تمامی این دیدگاه‌ها مشترک است، تأکید بر نقش بنیادین زبان در صورت‌بندی، انتقال و تفسیر معناست؛ نقشی که در حوزه بازیابی اطلاعات اهمیتی مضاعف می‌یابد. در این حوزه، تمامی مراحل از شکل‌گیری نیاز اطلاعاتی در ذهن کاربر تا بیان پرسش، تعامل با سامانه و تفسیر نتایج، بر بستری زبانی استوار است. از این رو، فهم دقیق ماهیت زبان و سازوکارهای معنایی و زمینه‌ای آن، پیش‌شرط تحلیل چالش‌هایی است که در ترجمه زبان طبیعی کاربران به زبان قابل پردازش نظام‌های اطلاعاتی پدید می‌آید.

زبان و ارتباط

مفهوم «ارتباط» ریشه در واژه لاتین *communicare* به معنای «سهیم شدن» یا «به اشتراک نهادن» دارد و از همان آغاز، ناظر بر فرایند انتقال معنا میان افراد بوده است. با وجود کاربرد گسترده این مفهوم در حوزه‌های مختلف، ارائه تعریفی جامع از ارتباط دشوار است، زیرا این پدیده همواره در بستر تعاملات انسانی و در چارچوب‌های فرهنگی، شناختی و اجتماعی متنوع شکل می‌گیرد.

یکی از تأثیرگذارترین الگوهای نظری در این حوزه، مدل ارتباطی شانون و ویور (۱۹۴۹) است که ارتباط را فرایندی متشکل از منبع، پیام، کانال، گیرنده و مقصد می‌داند. هرچند این مدل در اصل برای ارتباطات فنی طراحی شده بود، اما به دلیل شفاف‌سازی عناصر اصلی انتقال پیام، تأثیر عمیقی بر مطالعات ارتباط انسانی گذاشت. در این چارچوب، ارتباط مستلزم رمزگذاری پیام توسط فرستنده و رمزگشایی آن توسط گیرنده است؛ فرایندی که بدون وجود نظامی مشترک از نشانه‌ها، تحقق نمی‌یابد.

در اغلب اشکال ارتباط انسانی، زبان طبیعی نقش محوری در رمزگذاری و رمزگشایی پیام ایفا می‌کند. زبان امکان می‌دهد معنا در قالب نشانه‌هایی مشترک سازمان یابد و از ذهن گوینده به ذهن شنونده منتقل شود. در مدل‌های ارتباطی بعدی، از جمله الگوی شرام (۱۹۵۴)، بر نقش تجربه، زمینه و بازخورد تأکید بیشتری شده است. این الگو نشان می‌دهد که فهم پیام تنها زمانی ممکن است که میدان‌های تجربی فرستنده و گیرنده همپوشانی داشته باشد.

عنصر «نویز» یا اختلال نیز بخش جدایی‌ناپذیر از فرایند ارتباط است و می‌تواند در سطوح مختلفی بروز یابد: نویز فیزیکی، زبانی، فرهنگی یا شناختی. در سطح زبانی، ابهام واژگانی،

چندمعنایی و تفاوت در پیش فرض‌های معنایی می‌تواند باعث تحریف پیام شود و فهم مشترک را مختل سازد (Shannon & Weaver, 1949).

زبان طبیعی در فرایند ارتباط، هم از منظر ساختاری و هم از منظر معنایی نقش آفرین است. از دیدگاه نحوی، پیام باید مطابق قواعد ترکیب زبان سامان یابد تا واحدی معتبر و قابل تفسیر ایجاد کند؛ از منظر معنایی، این واحدها باید بتوانند به جهان خارج ارجاع دهند و معنای موردنظر گوینده را منتقل سازند. پیوند میان این دو سطح است که زبان را به نظامی مولد و پویا تبدیل می‌کند و امکان تولید بی‌نهایت پیام معنادار را فراهم می‌آورد (Carpintero, 1996).

با توجه به این ملاحظات، روشن می‌شود که زبان صرفاً ابزار انتقال اطلاعات نیست، بلکه چارچوب مفهومی‌ای فراهم می‌کند که در آن، خود فرایند ارتباط شکل می‌گیرد. این امر در حوزه بازیابی اطلاعات اهمیتی مضاعف دارد؛ زیرا پرسش کاربر به‌عنوان نخستین کنش ارتباطی، باید در قالب زبانی قابل تفسیر به نظام اطلاعاتی منتقل شود. هرگونه ناهماهنگی میان زبان کاربر و زبان نظام، یا بروز نویز معنایی در مراحل رمزگذاری و رمزگشایی، می‌تواند به سوءتعبیر نیاز اطلاعاتی و کاهش دقت و کارایی بازیابی منجر شود. از این رو، تحلیل رابطه زبان و ارتباط، بنیانی نظری برای فهم چالش‌های تعامل میان کاربر و نظام‌های بازیابی اطلاعات فراهم می‌آورد.

زبان در بازیابی اطلاعات

تحولات فناورانه نیمه دوم قرن بیستم، شیوه‌های تولید، سازمان‌دهی و دسترسی به دانش را به‌طور بنیادین دگرگون کرد و زمینه را برای شکل‌گیری نظام‌های رایانه‌ای بازیابی اطلاعات فراهم آورد. در این بستر نوین، بازیابی اطلاعات به یکی از رایج‌ترین اشکال تعامل انسان با دانش ذخیره‌شده بدل شد؛ تعاملی که اگرچه به معنای دقیق کلمه واجد ویژگی‌های ارتباط

میان فردی نیست، اما در بنیان خود بر مؤلفه‌ای مشترک با ارتباط انسانی استوار است: زبان (Salton, 1968; Ingwersen, 1992).

مطالعه بازیابی اطلاعات، از این منظر، مستلزم توجه هم‌زمان به سازوکارهای فنی سامانه‌های رایانه‌ای و ابعاد زبانی و شناختی تعامل کاربر با این سامانه‌هاست. معنا و کارآمدی فرایند بازیابی نه صرفاً در الگوریتم‌ها یا ساختارهای فنی، بلکه در نقطه تلاقی بازنمایی زبانی نیاز اطلاعاتی، پردازش نظام‌مند داده‌ها و تفسیر شناختی نتایج از سوی کاربر شکل می‌گیرد. همین پیوند سه‌گانه‌اش که بازیابی اطلاعات را به حوزه‌ای ذاتاً میان‌رشته‌ای تبدیل می‌کند (Belkin, 1980; Ingwersen & Järvelin, 2005).

کاربران معمولاً نیازهای اطلاعاتی خود را در قالب زبان طبیعی صورت‌بندی می‌کنند؛ حالتی ذهنی که اغلب با نوعی عدم قطعیت یا احساس کمبود دانش همراه است. بلکین این وضعیت را "وضعیت غیرعادی دانش" می‌نامد و بر این نکته تأکید می‌کند که بیان نیاز اطلاعاتی همواره پیش از آن که دقیق و صریح باشد، مبهم و تدریجی است (Belkin, 1980). با این حال، سامانه‌های بازیابی اطلاعات توان پردازش مستقیم این صورت‌بندی‌های طبیعی و مبهم را ندارند و ناگزیر به استفاده از بازنمایی‌های صوری و ساخت‌یافته‌اند. از این رو، کاربر باید نیاز اطلاعاتی خود را به شکلی بازنمایی کند که برای سامانه قابل پردازش باشد (Belkin, Oddy, & Brooks, 1982).

این بازنمایی واسطه، که معمولاً از آن با عنوان «زبان جست‌وجو» یا «معادله جست‌وجو» یاد می‌شود، نقشی کلیدی در فرایند بازیابی ایفا می‌کند. زبان جست‌وجو مجموعه‌ای از نمادها، اصطلاحات و قواعد نحوی است که امکان تبدیل بیان طبیعی نیاز اطلاعاتی به ساخت‌هایی قابل تفسیر برای سامانه را فراهم می‌آورد (Jansen, Booth, & Spink, 2008). فرایند گذار از زبان طبیعی به زبان جست‌وجو معمولاً شامل دو گام اساسی است: نخست، تحلیل

¹ Anomalous State of Knowledge

مفهومی پرسش و شناسایی مفاهیم محوری؛ و دوم، ترجمه این مفاهیم به ساخت‌هایی که با منطق پردازشی سامانه هم‌خوانی دارند. این ساخت‌ها سپس با نمایه‌های اسنادی – که خود حاصل فرایندهای زبانی نمایه‌سازی هستند – مقایسه می‌شوند تا مجموعه‌ای از اسناد با بیشترین میزان ارتباط بازیابی گردد (Lancaster, 1992).

از این منظر، کارایی بازیابی اطلاعات تا حد زیادی به توانایی کاربر در صورت‌بندی زبانی مناسب نیاز اطلاعاتی وابسته است. زبان جست‌وجو نه تنها باید از نظر نحوی درست ساخت باشد، بلکه باید بتواند روابط معنایی موردنظر کاربر را نیز، هرچند به صورت محدود، بازنمایی کند. در نظام‌های کلاسیک، این بازنمایی عمدتاً از طریق عملگرهای منطقی و تطابق اصطلاحی انجام می‌شد؛ اما در رویکردهای مدرن‌تر، از جمله جست‌وجوی معنایی و مدل‌های برداری، تلاش می‌شود لایه‌های عمیق‌تری از معنا و شباهت مفهومی در فرایند بازیابی لحاظ شود (Manning, Raghavan, & Schütze, 2008; Zhao et al., 2022).

از منظر معناشناختی، معادله جست‌وجو را می‌توان نوعی کنش زبانی دانست؛ کنشی که در آن کاربر با زبان، سامانه را به بازیابی مجموعه‌ای خاص از اسناد فرا می‌خواند. در این چارچوب، معنا صرفاً ویژگی اسناد بازیابی شده نیست، بلکه در خود فرایند صورت‌بندی پرسش و نحوه زبان‌بندی آن نیز تحقق می‌یابد. همان‌گونه که ون ریسبرگن تأکید می‌کند، نظریه بازیابی اطلاعات ناگزیر مبتنی بر نوعی درک نظام‌مند از «معنا» است؛ درکی که بدون توجه به مبانی زبان‌شناختی و معناشناختی ناقص خواهد بود (van Rijsbergen, 1986).

با گسترش سامانه‌های بازیابی اطلاعات و افزایش پیچیدگی ساختارهای زبانی مورد استفاده در آن‌ها، ضرورت دقت در هر دو سوی این تعامل بیش از پیش آشکار شده است. از یک سو، توصیف محتوای اطلاعات از طریق نمایه‌سازی و از سوی دیگر، توصیف نیازهای اطلاعاتی کاربران، هر دو کنش‌هایی اساساً زبانی‌اند. کیفیت بازیابی، در نهایت، به میزان هم‌گرایی این

دو سطح وابسته است. از این رو، تحلیل زبان در بازیابی اطلاعات نشان می‌دهد که این فرایند را نمی‌توان صرفاً عملیاتی فناورانه دانست، بلکه باید آن را پدیده‌ای دانست که در بنیان خود بر سازوکارهای نحوی، معنایی و شناختی زبان استوار است.

کلمات و معانی

پرداختن به نقش زبان در فرایند بازیابی اطلاعات ناگزیر ما را به قلمرو معناشناسی رهنمون می‌سازد؛ زیرا هر کنش جست‌وجو بر پایه نسبت میان واژگان و معانی آن‌ها شکل می‌گیرد. آنچه در سامانه‌های بازیابی پردازش می‌شود، صرفاً رشته‌هایی از نشانه‌های زبانی نیست، بلکه بازنمایی‌هایی از معناست که باید با نیاز اطلاعاتی کاربر هم‌خوانی داشته باشد. از این رو، درک ماهیت معنا و چگونگی شکل‌گیری آن، پیش‌شرط تحلیل دقیق فرایندهای جست‌وجو و بازیابی به‌شمار می‌آید.

بحث درباره معنا در فلسفه زبان پیشینه‌ای طولانی دارد، اما در فلسفه معاصر، آراء لودویگ ویتگنشتاین جایگاهی محوری یافته‌اند. ویتگنشتاین در پژوهش‌های فلسفی بر این نکته تأکید می‌کند که بسیاری از ابهام‌ها و بن‌بست‌های نظری، نه از پیچیدگی واقعیت، بلکه از سوءفهم کارکرد زبان ناشی می‌شوند و تصریح می‌کند که "فلسفه نبردی است با افسون زبان" (Wittgenstein, 1953, p. 47). از منظر او، معنا نه جوهری ثابت و مستقل، بلکه حاصل کاربرد واژه‌ها در زمینه‌های معین و در چارچوب آن چیزی است که «بازی‌های زبانی» نامیده می‌شود. فهم معنا، در این نگاه، مستلزم توانایی مشارکت در این بازی‌ها و به‌کارگیری واژگان در موقعیت‌های واقعی است، نه صرفاً دانستن تعاریف قاموسی.

دیدگاه‌های معناشناختی متأخر نیز این رویکرد پویا به معنا را تقویت کرده‌اند. پاتنم با طرح ایده‌هایی همچون «تقسیم کار زبانی»، نشان می‌دهد که معنا نه صرفاً امری ذهنی است و نه کاملاً درون‌زبان، بلکه در پیوند با عوامل اجتماعی، بافتی و جهان بیرونی شکل می‌گیرد. بر

این اساس، معنا را نمی‌توان به مجموعه‌ای از ویژگی‌های ثابت فروکاست؛ بلکه باید آن را پدیده‌ای دانست که در تعامل میان زبان، کاربران و زمینه‌های کاربردی پدید می‌آید. مرور تحلیلی این جریان‌ها در آثار Lycan (2000) و Blackburn (1984) به روشنی نشان می‌دهد که معنا همواره امری وابسته به کاربرد و زمینه است.

این برداشت کاربردمحور از معنا، پیامدهای مهمی برای نظریه و عمل بازیابی اطلاعات دارد. در رویکردهای کلاسیک بازیابی اطلاعات، غالباً فرض می‌شود که کاربر از نیاز اطلاعاتی نسبتاً مشخصی برخوردار است و سپس این نیاز را به شکلی زبانی برای تعامل با سامانه بیان می‌کند. اما اگر، چنان‌که رویکردهای زبانی و فلسفی جدید نشان می‌دهند، اندیشیدن خود در بستر زبان صورت گیرد، آنگاه زبان جست‌وجو تنها ابزار بیان یک نیاز پیشینی نخواهد بود، بلکه چارچوبی است که خود به شکل‌گیری و تحدید نیاز اطلاعاتی کمک می‌کند (Frohmann, 1994; Hjørland, 1997). در این حالت، محدودیت‌ها و امکانات زبان جست‌وجو مستقیماً بر نحوه تصور کاربر از مسئله و دامنه جست‌وجوی او اثر می‌گذارند.

از این منظر، فرایند جست‌وجو را می‌توان نوعی کنش زبانی دانست که شباهت‌هایی با گفت‌وگو دارد: کاربر پرسشی مطرح می‌کند و سامانه با مجموعه‌ای از اسناد پاسخ می‌دهد. برای کارآمدی این تعامل، اصولی مشابه قواعد همکاری گرایس باید برقرار باشد؛ به این معنا که پرسش کاربر از نظر وضوح، ارتباط و تناسب با زمینه بهینه باشد و پاسخ سامانه نیز با مقصود کاربر هم‌خوانی داشته باشد. با این حال، برخلاف گفت‌وگوی انسانی، در اغلب سامانه‌های بازیابی اطلاعات امکان اصلاح فوری سوء تفاهم‌ها و بازخورد تعاملی محدود است؛ امری که به گفته بلر (Blair, 1990) می‌تواند مانع شکل‌گیری معنای دقیق و مشترک میان کاربر و سامانه شود، به‌ویژه در جست‌وجوهای پیچیده و چندبعدی.

اهمیت معنا در کاربرد در بازیابی اطلاعات از این نکته نیز ناشی می‌شود که زبان طبیعی اساساً برای تمایز دقیق و نظام‌مند میان اسناد طراحی نشده است. در زبان روزمره، ابهام معنایی،

چندمعنایی و تنوع واژگانی امری طبیعی است؛ کاربران ممکن است برای یک مفهوم واحد از واژگان متفاوتی استفاده کنند یا یک واژه را در زمینه‌های گوناگون با معانی متفاوت به کار برند (Furnas et al., 1987). در مقابل، سامانه‌های بازیابی اطلاعات ناگزیرند زبان را به شکلی ساخت یافته و استاندارد شده بازنمایی کنند تا امکان پردازش ماشینی فراهم شود. این تفاوت بنیادین، اغلب به پدیده «ناسازگاری واژگانی» می‌انجامد؛ وضعیتی که در آن واژگان به کاررفته در پرسش کاربر با واژگان موجود در اسناد مرتبط هم‌پوشانی ندارند و در نتیجه، بازیابی مناسب یا دشواری مواجه می‌شود (Zhao & Callan, 2010).

در چنین شرایطی، نقش میانجی‌گرانه جست‌وجوگران حرفه‌ای، به‌ویژه کتابداران و متخصصان اطلاعات، برجسته می‌شود. این افراد می‌کوشند فاصله میان زبان طبیعی و سیال کاربران و زبان ساخت یافته و محدود سامانه‌های بازیابی را کاهش دهند. این فرایند را می‌توان نوعی ترجمه میان دو منطق زبانی دانست: زبانی که در زندگی روزمره معنا را در بستر کاربرد تولید می‌کند و زبانی که در سامانه‌های اطلاعاتی باهدف تفکیک، دسته‌بندی و بازیابی دقیق طراحی شده است.

در مجموع، نظریه‌های معناشناختی – از رویکرد کاربر‌محور ویتگنشتاین تا دیدگاه‌های زمینه‌محور پاتنم – بر این نکته تأکید دارند که معنا پدیده‌ای ایستا و مستقل از تجربه نیست. این فهم پویا از معنا برای بازیابی اطلاعات اهمیتی بنیادین دارد؛ زیرا بدون توجه به نقش کاربرد، زمینه و کنش‌های زبانی کاربران، نه توصیف اسناد و نه فرایند جست‌وجو قادر نخواهند بود بازنمایی دقیقی از نیازهای اطلاعاتی فراهم آورند.

مروری بر پژوهش‌ها

پژوهش‌های مرتبط با زبان و بازیابی اطلاعات طی بیش از شش دهه گذشته نشان داده‌اند که کارآمدی نظام‌های اطلاعاتی به‌طور بنیادین به چگونگی بازنمایی، تفسیر و پردازش زبان

طبیعی کاربران وابسته است. از نخستین مطالعات کلاسیک در دهه‌های ۱۹۶۰ و ۱۹۷۰ تا رویکردهای معاصر مبتنی بر معناشناسی محاسباتی و یادگیری عمیق، زبان همواره به عنوان عنصر محوری در تعامل میان کاربر و سامانه‌های بازیابی اطلاعات مورد توجه بوده است.

در پژوهش‌های اولیه، تمرکز اصلی بر سنجش اثربخشی بازیابی و تعریف مفهوم «ارتباط» قرار داشت. کوپر (۱۹۷۳) با نقد معیارهای سنتی مانند دقت و بازیافت، بر ضرورت در نظر گرفتن ارزش ذهنی نتایج برای کاربر تأکید کرد و نشان داد که ارزیابی بازیابی اطلاعات بدون توجه به تجربه و تفسیر کاربر ناقص است. هم‌زمان، ویلسون (۱۹۷۳) و ون ریسبرگن (۱۹۷۹) با طرح مفهوم «ارتباط موقعیتی»، نقش زمینه و شرایط کاربر را در تشخیص مرتبط بودن اسناد برجسته ساختند و مبنای نظری لازم برای نگاه زمینه‌محور به بازیابی اطلاعات را فراهم آوردند.

از دهه ۱۹۸۰ به بعد، توجه به ابعاد زبانی و معناشناختی بازیابی اطلاعات افزایش یافت. بلکین (۱۹۸۰) با معرفی مفهوم "حالت غیرعادی دانش" نشان داد که نیاز اطلاعاتی کاربر اغلب مبهم و ناتمام است و تنها از طریق زبان و تعامل تدریجی با سامانه شکل می‌گیرد. این دیدگاه در پژوهش بلکین، اودی و بروکس (۱۹۸۲) گسترش یافت و بر دشواری ترجمه زبان طبیعی کاربر به زبان قابل پردازش سامانه تأکید شد. در همین راستا، ون ریسبرگن (۱۹۸۶) بازیابی اطلاعات را مستلزم درکی نظام‌مند از معنا دانست و پیوند میان نظریه بازیابی و مبانی زبان‌شناختی را اجتناب‌ناپذیر تلقی کرد.

در حوزه فلسفه زبان و تأثیر آن بر بازیابی اطلاعات، پژوهش‌های فروممن (۱۹۹۰) و بلیر (۱۹۹۰، ۱۹۹۲) نقش تعیین‌کننده‌ای ایفا کرده‌اند. این مطالعات با بهره‌گیری از آرای ویتگنشتاین، آستین، سرل و گریس، نشان دادند که بازنمایی معنا در نظام‌های اطلاعاتی بدون توجه به کاربرد زبان، کنش گفتاری و زمینه ارتباطی با محدودیت‌های جدی مواجه است.

بلیر به‌ویژه تأکید می‌کند که پرسش‌های کاربران نوعی کنش زبانی‌اند و ناهماهنگی میان زبان کاربر و زبان سامانه ریشه بسیاری از ناکامی‌های بازیابی است. در ادامه، هارتر (۱۹۹۲) با اتکا به نظریه ارتباط اسپربر و ویلسون، جست‌وجو را شکلی از کنش ارتباطی دانست که در آن استنباط معنایی نقشی کلیدی دارد.

با ورود به دهه ۱۹۹۰ و اوایل دهه ۲۰۰۰، رویکردهای زمینه‌محور و تعاملی گسترش یافتند. راب و ون وون ریسبرگ (۱۹۹۶) بازیابی اطلاعات را فرایندی گفت‌وگو‌محور تلقی کردند که در آن سامانه باید قادر باشد زمینه جاری جست‌وجو را بر اساس بازخورد کاربر تولید و به‌روزرسانی کند. هیولند (۱۹۹۸) نیز با مقایسه دو رویکرد تصویری معنا و بازی‌های زبانی و یتگنشتاین، نشان داد که توجه به بافت اجتماعی و کاربردی زبان، چارچوب مناسب‌تری برای طراحی نظام‌های بازیابی اطلاعات فراهم می‌کند. در همین مسیر، کورلند و لی (۲۰۰۴) با تلفیق مدل‌های زبانی سند و خوشه اسناد، اهمیت ساختار کلی مجموعه اسناد را در بهبود دقت و بازیافت نشان دادند.

در دهه‌های اخیر، تمرکز پژوهش‌ها به تدریج به سوی معناشناسی محاسباتی، رانش معنایی و یادگیری ماشین سوق یافته است. تورنلی و گیب (۲۰۰۷، ۲۰۰۹) با تحلیل تغییرات معنایی زبان در بستر اجتماعی، نشان دادند که مدل‌های سنتی بازیابی اطلاعات در مواجهه با پویایی معنا کارایی محدودی دارند. لیوما و همکاران (۲۰۱۶) با افزودن ویژگی‌های معنایی-کیفی اسناد مانند واقعیت‌نمایی و عینیت به مدل‌های زبانی، امکان بهبود مرتبط بودن نتایج را نشان دادند. در ادامه، کارپوخین و همکاران (۲۰۲۰) با معرفی روش‌های بازیابی متراکم مبتنی بر بازنمایی‌های برداری، تحولی اساسی در سامانه‌های پرسش و پاسخ ایجاد کردند که بر فاصله گرفتن از تطابق صرف واژگانی و حرکت به سوی تطابق معنایی دلالت دارد. پژوهش فن و همکاران (۲۰۲۱) نیز با رویکردی زبان‌شناختی نشان داد که سنجش مرتبط بودن مستلزم در نظر گرفتن هم‌زمان ابعاد مفهومی و معنایی پرسش و سند است.

در سطح پژوهش‌های داخلی، مطالعات انجام‌شده نشان می‌دهد که بازیابی اطلاعات فارسی با چالش‌های زبانی خاصی مواجه است. مرتضایی (۱۳۹۲) به مشکلات ناشی از نبود استانداردهای اصطلاح‌گزینی و ناهمسانی واژگانی اشاره می‌کند. برنجیان (۱۳۹۴) و نوذری (۱۳۹۴) با تمرکز بر هم‌معنایی و شبکه‌های معنایی، نشان داده‌اند که گسترش معنا محور جست‌وجو می‌تواند به بهبود نتایج بازیابی کمک کند. نوروزی (۱۳۹۸) نقش زبان مادری در شکل‌گیری مفاهیم ذهنی و بیان نیاز اطلاعاتی را برجسته می‌سازد. پژوهش‌های باقری و همکاران (۱۳۹۸، ۱۳۹۹) و باقری و نوروزی (۱۳۹۹) نیز با ارائه الگوها و شاخص‌های کاربرد فناوری معنایی در کتابخانه‌های دیجیتال، بر ضرورت حرکت به سوی بازیابی معنا محور تأکید دارند. در تازه‌ترین پژوهش‌ها، هماوندی و همکاران (۱۴۰۳) و منیری، شلوسر و کوورکو (۲۰۲۴) با رویکردهای تحلیلی و یادگیری عمیق، کمبود منابع استاندارد و پیچیدگی صرفی زبان فارسی را از موانع اصلی بازیابی مؤثر دانسته‌اند.

در مجموع، مرور پژوهش‌ها نشان می‌دهد که در سطح بین‌المللی، پیوند میان زبان، معنا، زمینه و فناوری‌های بازیابی اطلاعات به‌طور مستمر تقویت شده و مسیر توسعه نظام‌های نوین بازیابی به سوی رویکردهای معناشناختی و تعاملی هموار شده است. با این حال، در ادبیات علمی ایران، علی‌رغم افزایش مطالعات فنی، هنوز پژوهش‌هایی که به‌صورت یکپارچه به مبانی زبانی، فلسفی و شناختی بازیابی اطلاعات بپردازند، محدودند. این خلأ به‌ویژه در تبیین رابطه میان زبان طبیعی کاربر، بافت کاربردی و زبان ساخت‌یافته سامانه‌های بازیابی مشهود است. پژوهش حاضر در چنین بستری شکل گرفته و می‌کوشد با رویکردی تحلیلی و میان‌رشته‌ای، نقش محوری زبان را در تمامی مراحل بازیابی اطلاعات واکاوی کرده و راهکارهایی برای همگرایی بیشتر میان زبان کاربر و زبان نظام‌های اطلاعاتی ارائه دهد.

بیان مساله:

در جهان امروز، دسترسی مؤثر، دقیق و به‌هنگام به اطلاعات به یکی از پیش‌شرط‌های اساسی فعالیت‌های علمی، آموزشی، سازمانی و تصمیم‌گیری بدل شده است. گسترش شتابان داده‌ها در محیط‌های دیجیتال، ضرورت بهره‌گیری از نظام‌های ذخیره و بازیابی اطلاعات را بیش از پیش آشکار ساخته و تعامل انسان با این نظام‌ها را به یکی از رایج‌ترین اشکال مواجهه با دانش تبدیل کرده است. در این تعامل، زبان طبیعی کاربران نقش واسطه‌ای بنیادین ایفا می‌کند، زیرا نیازهای اطلاعاتی، پرسش‌ها و ارزیابی نتایج همگی در بستری زبانی شکل می‌گیرند.

با این حال، یکی از چالش‌های دیرپای حوزه بازیابی اطلاعات، ناهمخوانی میان زبان طبیعی کاربران و زبان ساخت‌یافته نظام‌های بازیابی اطلاعات است، چالشی که به‌طور مستقیم بر دقت، بازیافت و مرتبط بودن نتایج تأثیر می‌گذارد (Blair, 1990; Hjørland, 1998). کاربران معمولاً نیازهای اطلاعاتی خود را با بهره‌گیری از زبان روزمره، وابسته به زمینه و سرشار از ابهام‌های معنایی بیان می‌کنند، در حالی که بسیاری از نظام‌های بازیابی بر پایه بازنمایی‌های واژگانی، آماری یا ساختارهای صورتی طراحی شده‌اند و ظرفیت محدودی برای تفسیر قصد زبانی، روابط مفهومی و ظرایف معنایی دارند (Furnas et al., 1987; Manning et al., 2008).

از منظر زبان‌شناسی و فلسفه زبان، زبان صرفاً ابزاری خنثی برای انتقال اطلاعات نیست، بلکه چارچوبی است که معنا در آن ساخته، تفسیر و تثبیت می‌شود. نظریه‌های کنش‌گفتاری نشان می‌دهند که معنا نه فقط در محتوای واژگان، بلکه در نحوه به‌کارگیری آن‌ها در بافت‌های خاص ارتباطی تحقق می‌یابد (Austin, 1962; Searle, 1969). همچنین، رویکردهای فلسفی متأثر از ویتگنشتاین و گریس بر این نکته تأکید دارند که فهم معنا همواره وابسته به زمینه، قصد گوینده و قواعد نانوشته تعامل زبانی است (Wittgenstein, 1953; Grice, 1969).

(1975). این دیدگاه‌ها آشکار می‌سازند که هرگونه ساده‌سازی معنا به تطابق صرف واژگان، ناگزیر به کاهش کیفیت ارتباط و سوءتفسیر منجر می‌شود.

در عمل، بسیاری از نظام‌های بازیابی اطلاعات—از موتورهای جست‌وجو گرفته تا پایگاه‌های داده تخصصی—همچنان بر مدل‌های نمایه‌سازی واژگانی یا الگوریتم‌های آماری کلاسیک و نیمه‌مدرن تکیه دارند. اگرچه این رویکردها در مقیاس‌پذیری و سرعت موفق بوده‌اند، اما در مواجهه با ابهام معنایی، تنوع واژگانی، چندمعنایی و تفاوت‌های زمینه‌ای زبان طبیعی با محدودیت‌های جدی روبه‌رو هستند. در نتیجه، معناهای موردنظر کاربران اغلب به‌طور ناقص یا تحریف‌شده در زبان نظام بازنمایی می‌شود و فاصله‌ای میان «آنچه کاربر می‌جوید» و «آنچه سامانه بازیابی می‌کند» شکل می‌گیرد.

این مسئله در بافت زبان فارسی اهمیت مضاعفی پیدا می‌کند. پیچیدگی‌های صرفی و نحوی زبان فارسی، تنوع شکل‌های واژگانی، نبود استانداردهای یکنواخت در اصطلاح‌گزینی و کمبود منابع معنایی غنی، همگی بر دشواری همگرایی میان زبان کاربر و زبان سامانه می‌افزایند. با وجود پیشرفت‌های اخیر در پردازش زبان طبیعی و معناشناسی محاسباتی، بخش قابل توجهی از پژوهش‌های انجام‌شده در ایران همچنان بر جنبه‌های فنی و الگوریتمی متمرکز بوده و ابعاد زبانی، فلسفی و شناختی بازیابی اطلاعات کمتر مورد توجه قرار گرفته است. این رویکرد یک‌سویه، مانع از درک جامع سازوکارهای تولید معنا در تعامل کاربر-سامانه شده و به تداوم شکاف زبانی در فرآیند جست‌وجو انجامیده است.

با توجه به این زمینه، پژوهش حاضر به بررسی این پرسش کلیدی می‌پردازد که ویژگی‌های زبانی و معناشناختی زبان طبیعی چگونه بر دقت، کیفیت و مرتبط بودن بازیابی اطلاعات تأثیر می‌گذارند و چه راهکارهایی می‌توان برای کاهش شکاف میان زبان کاربر و زبان نظام‌های بازیابی ارائه کرد. در راستای پاسخ به این پرسش، این مطالعه هدف دارد چالش‌ها و محدودیت‌های استفاده از زبان طبیعی کاربران در فرایند بازیابی اطلاعات را شناسایی کند،

سازوکار تعامل میان زبان کاربران و زبان سامانه‌های بازیابی اطلاعات را تحلیل نماید، نقاط ناسازگاری و شکاف‌های معناشناختی را مشخص سازد و در نهایت چارچوبی مفهومی و عملی برای افزایش همگرایی میان زبان کاربران و زبان سامانه ارائه دهد.

برای تحقق این اهداف، پژوهش به بررسی چهار پرسش اصلی پرداخته است:

۱. چگونه ویژگی‌های نحوی و ساختاری زبان طبیعی کاربران بر دقت و مرتبط بودن اطلاعات بازیابی شده تأثیر می‌گذارد؟

۲. نقش زمینه و کاربرد معنایی واژگان در تطبیق پرسش‌های کاربران با اسناد موجود در سامانه‌های بازیابی اطلاعات چیست؟

۳. چه چالش‌هایی در فرآیند ترجمه زبان طبیعی کاربران به زبان جستجوی سامانه‌ها وجود دارد و این چالش‌ها چه تأثیری بر کیفیت بازیابی اطلاعات دارند؟

۴. چه راهکارهایی می‌تواند به افزایش همگرایی میان زبان کاربران و زبان سامانه منجر شود و چگونه می‌تواند کیفیت و کارایی بازیابی اطلاعات را بهبود دهد؟

روش پژوهش:

این مطالعه از رویکرد تحلیلی-توصیفی بهره برد و هدف آن بررسی تأثیر ابعاد زبانی، شامل ویژگی‌های نحوی و معناشناختی پرسش‌های کاربران، بر دقت و مرتبط بودن نتایج بازیابی اطلاعات بود. برای تحقق این هدف، مجموعه‌ای از اسناد علمی و تخصصی به‌عنوان داده پژوهش انتخاب شد و محتوای آن‌ها با استفاده از شاخص‌های کمی و کیفی تحلیل گردید.

۱. جامعه و نمونه پژوهش

جامعه پژوهش شامل اسناد دیجیتالی منتشرشده در پایگاه‌های معتبر علمی و تخصصی بود که قابلیت بازیابی بر اساس کلیدواژه‌ها و پرسش‌های متنی کاربران را دارند. برای دستیابی به

نمونه‌ای متنوع و نماینده از موضوعات مرتبط با نیازهای اطلاعاتی کاربران، ۲۰۰ سند از میان این جامعه انتخاب شد. معیارهای انتخاب شامل:

- ارتباط موضوعی با محورهای پژوهش،
- تنوع حوزه‌های اطلاعاتی،
- کیفیت علمی منابع.

اسناد منتخب پس از دسته‌بندی موضوعی، به عنوان مبنای تحلیل نحوی، واژگانی و معنایی در بخش‌های مختلف پژوهش مورد استفاده قرار گرفتند.

۲. پرسش‌های نمونه و اهداف آنها

برای بررسی نقش زبان در فرآیند بازیابی اطلاعات، پرسش‌های نمونه‌ای طراحی شد که هر یک جنبه‌ای از تعامل میان کاربر و سامانه را روشن می‌کردند.

جدول ۱. پرسش‌های نمونه‌ای طراحی شد برای هر یک از جنبه‌های تعامل میان کاربر و سامانه

شماره پرسش	پرسش نمونه	هدف پرسش
۱	تأثیر ساختار نحوی عبارت جستجو بر دقت بازیابی چیست؟	بررسی اثر ویژگی‌های نحوی زبان طبیعی کاربران بر دقت بازیابی اطلاعات
۲	نقش زمینه و کاربرد معنایی واژگان در مرتبط بودن نتایج بازیابی چیست؟	تحلیل اهمیت معانی ضمنی و زمینه در تطبیق پرسش‌ها با اسناد
۳	چه مشکلاتی در ترجمه پرسش‌های کاربران به زبان جستجوی سیستم وجود دارد؟	شناسایی ناسازگاری‌ها و خطاهای ترجمه زبان طبیعی کاربران به زبان سیستم
۴	راهکارهای مؤثر برای بهبود تطبیق زبان کاربر و زبان سیستم چیست؟	ارائه راهکارهای عملی برای افزایش دقت و مرتبط بودن نتایج بازیابی

این چهار پرسش چارچوب اصلی تحلیل پژوهش را شکل دادند و مبنایی برای طراحی مراحل جمع‌آوری و تحلیل داده‌ها فراهم ساختند.

فرآیند گردآوری داده‌ها در این پژوهش به صورت مرحله‌ای و بر اساس الگویی نظام‌مند انجام گرفت. در گام نخست، مجموعه‌ای از پرسش‌های اطلاعاتی با الهام از نیازهای واقعی کاربران و مبانی نظری زبان طبیعی تدوین شد تا تنوع لازم از نظر ساختارهای نحوی و معناشناختی فراهم شود. سپس هر پرسش به معادله جستجو تبدیل گردید و این تبدیل با استخراج واژگان کلیدی، شناسایی اصطلاحات وابسته و تنظیم روابط منطقی با بهره‌گیری از عملگرهای بولی انجام شد تا شکاف میان زبان کاربر و زبان سامانه کاهش یابد. در مرحله بعد، معادلات طراحی شده در پایگاه‌های اطلاعاتی منتخب اجرا شد و تمامی نتایج در شرایط زمانی و تنظیمات ثابت ثبت و مستندسازی شد. برای هر پرسش، فهرست کامل اسناد بازیابی شده گردآوری و در قالب جداول مقایسه‌ای سامان‌دهی گردید و سپس دو ارزیاب مستقل میزان انطباق مفهومی و موضوعی این اسناد با پرسش‌های اولیه را بررسی کردند تا اعتبار نتایج ارزیابی تضمین شود. در ادامه نیز عملکرد نظام‌های بازیابی با استفاده از شاخص‌های دقت، بازیافت و معیار اف سنجیده شد؛ به گونه‌ای که دقت نشان‌دهنده سهم اسناد مرتبط از میان خروجی‌های جستجو، بازیافت بیانگر توان نظام در شناسایی همه اسناد مرتبط بالقوه، و معیار اف ارائه‌دهنده تصویری ترکیبی و متوازن از کارایی کلی سیستم بود.

برای تحلیل داده‌ها، از مجموعه‌ای هماهنگ از شاخص‌های کمی و کیفی بهره گرفته شد تا عملکرد نظام‌های بازیابی اطلاعات از زوایای مختلف ارزیابی شود. در بخش کمی، شاخص‌های دقت، بازیافت و معیار اف به عنوان سنجه‌های اصلی کارآمدی سیستم محاسبه گردید؛ دقت میزان سهم اسناد مرتبط در خروجی جستجو را نشان می‌داد، بازیافت توان سیستم در شناسایی تمامی اسناد مرتبط بالقوه را می‌سنجید، و معیار اف با فراهم کردن میانگینی هم‌ساز از این دو شاخص تصویری متوازن از کارایی کلی ارائه می‌کرد. داده‌های

حاصل در قالب جداول مقایسه‌ای تنظیم شد تا تفاوت عملکرد سیستم در مواجهه با ساختارهای گوناگون نحوی و معنایی پرسش‌ها به روشنی قابل بررسی باشد، و تمامی محاسبات نیز برای افزایش اعتبار نتایج با استفاده از نرم‌افزارهای آماری استاندارد انجام شد. در بخش کیفی، محتوای پرسش‌ها و اسناد بازیابی شده با رویکردی تفسیری تحلیل گردید تا منابع احتمالی ناهماهنگی میان زبان طبیعی کاربران و زبان کنترل شده پایگاه‌ها شناسایی شود؛ تحلیلی که با تمرکز بر معانی واژگان، بافت کاربرد آن‌ها، نحوه تبدیل مفاهیم کاربر به معادل‌های جستجویی، و مسائلی چون ابهام واژگانی، چندمعنایی اصطلاحات و ناسازگاری‌های نحوی انجام گرفت. نتایج این بررسی‌ها، عوامل مؤثر بر کاهش دقت و بازیافت را روشن ساخت و زمینه را برای ارائه پیشنهادهایی جهت بهبود فرایند تطبیق زبانی در سامانه‌های بازیابی اطلاعات فراهم کرد.

یافته‌های پژوهش

پرسش ۱: تأثیر ساختار نحوی عبارت جستجو بر دقت بازیابی

برای ارزیابی نقش ساختار نحوی در عملکرد نظام بازیابی اطلاعات، مجموعه‌ای از پرسش‌ها با الگوهای نحوی متفاوت طراحی و در سامانه جستجو اجرا شد. هدف از این مرحله آن بود که مشخص شود تا چه حد رعایت انسجام نحوی در نحوه بیان پرسش می‌تواند بر دقت و بازیافت نتایج تأثیر بگذارد. شاخص‌های دقت، بازیافت و معیار ترکیبی اف_۱ به عنوان مبنای تحلیل کمی مورد استفاده قرار گرفتند.

جدول ۲. شاخص‌های ارزیابی عملکرد بازیابی برای پرسش‌های دارای ساختار نحوی صحیح

مقدار	معیار
۰.۷۸	دقت
۰.۷۱	بازیابی

۰.۷۴	معیار اف
------	----------

نتایج نشان دادند که پرسش‌هایی با ساختار نحوی روشن و منسجم، میزان انطباق بیشتری با واژگان و نمایه‌های مورد استفاده در پایگاه اطلاعاتی دارند. دقت ۰.۷۸ بیانگر آن است که بخش قابل توجهی از اسناد بازیابی شده، از نظر محتوایی با نیاز اطلاعاتی پرسش‌ها همخوانی داشته‌اند. همچنین مقدار ۰.۷۱ برای بازیافت نشان می‌دهد که ساختار نحوی مناسب به نظام بازیابی کمک می‌کند تا بخش بیشتری از اسناد مرتبط را شناسایی کند. مقدار معیار اف^۱ نیز که ۰.۷۴ محاسبه شده، ارزیابی نسبی مثبتی از کارایی کلی جستجو در این شرایط ارائه می‌کند.

برای بررسی آماری تفاوت میان پرسش‌های با ساختار نحوی صحیح و پرسش‌های با ساختار نحوی ضعیف، آزمون t مستقل بر روی دقت و بازیافت انجام شد. نتایج نشان داد که تفاوت میان گروه‌ها از نظر دقت ($p < 0.01$) و بازیافت ($p < 0.05$) معنادار است، که بر اهمیت ساختار نحوی در هم‌ترازی میان زبان کاربران و زبان نمایه‌سازی اسناد تأکید دارد.

این یافته‌ها نشان می‌دهند که سازمان‌دهی صحیح عناصر نحوی در پرسش‌ها می‌تواند به‌طور مستقیم کیفیت بازیابی اطلاعات را بهبود بخشد، به‌ویژه در سیستم‌های مبتنی بر پردازش خودکار زبان طبیعی که وابستگی بالایی به هم‌خوانی بین پرسش کاربران و نمایه‌های متنی دارند. بنابراین، توجه به ویژگی‌های نحوی در طراحی رابط‌های جستجو و الگوریتم‌های بازیابی، می‌تواند نقش مؤثری در افزایش دقت و کارایی سیستم ایفا کند.

پرسش ۲: نقش انتخاب واژگان دقیق در دقت بازیابی اطلاعات

برای بررسی تأثیر انتخاب واژگان بر نتایج جستجو، پرسش‌ها در دو گروه طراحی شدند:

- (۱) پرسش‌هایی مبتنی بر واژگان عمومی
- (۲) پرسش‌هایی شامل واژگان دقیق و مرتبط با حوزه موضوعی.

^۱ F-Measure

هدف از این مقایسه آن بود که مشخص شود سطح تخصصی بودن واژگان چگونه بر دقت و بازیافت نظام بازیابی اطلاعات تأثیر می‌گذارد.

جدول ۳. مقایسه عملکرد سیستم در دو نوع واژگان جستجو

نوع پرسش جستجو	دقت	بازیافت	توضیح
واژگان عمومی	۵۵٪	۶۰٪	گسترده‌گی دامنه معنایی و نبود محدودسازی موضوعی باعث بازیابی اسناد غیرمرتبط شده است.
واژگان دقیق/تخصصی	۸۵٪	۷۸٪	استفاده از اصطلاحات دقیق و مرتبط با حوزه موضوعی منجر به کاهش نویز و افزایش دقت نتایج شده است.

یافته‌ها نشان می‌دهند که پرسش‌هایی که از واژگان دقیق و تخصصی استفاده می‌کنند، هماهنگی بیشتری با نمایه‌های مورد استفاده در پایگاه اطلاعاتی دارند. دقت ۸۵ درصدی بیانگر این است که بخش عمده اسناد بازیابی شده در این حالت با موضوع اصلی پرسش مطابقت داشته‌اند. همچنین افزایش بازیافت به ۷۸ درصد نشان می‌دهد که انتخاب واژگان تخصصی نه تنها میزان اسناد غیرمرتبط را کاهش داده بلکه به شناسایی بخش بیشتری از اسناد مرتبط نیز کمک کرده است.

برای اعتبارسنجی آماری تفاوت عملکرد میان دو گروه، آزمون t مستقل بر روی دقت و بازیافت انجام شد. نتایج نشان داد که تفاوت‌ها از نظر دقت ($p < 0.01$) و بازیافت ($p < 0.05$) معنادار هستند. این یافته‌ها تأکید می‌کنند که انتخاب هوشمندانه واژگان، به‌ویژه واژگان تخصصی، نقش کلیدی در هم‌ترازی پرسش کاربران با نمایه‌های پایگاه اطلاعاتی و بهبود عملکرد سیستم‌های بازیابی اطلاعات دارد.

این نتایج نشان می‌دهد که واژگان عمومی، به دلیل دامنه معنایی گسترده و ابهام بیشتر، احتمال بازیابی اسناد ناهمخوان را افزایش می‌دهند و در نتیجه دقت جستجو کاهش می‌یابد. در مقابل، واژگان دقیق‌تر - که قابلیت انطباق بالاتری با ساختار نمایه‌سازی دارند - می‌توانند کیفیت کلی بازیابی را به شکل قابل توجهی بهبود بخشند. بنابراین انتخاب هوشمندانه واژگان در مرحله طرح پرسش نقش مهمی در ارتقای سطح عملکرد نظام‌های بازیابی اطلاعات دارد.

پرسش ۳: تأثیر زمینه و معنا بر بازیابی اطلاعات

به منظور بررسی نقش زمینه معنایی در جستجو، پرسش‌ها در دو حالت طراحی و اجرا شدند:

۱) حالت بدون توجه به زمینه، که در آن تنها واژه‌های سطحی پرسش مبنای جستجو قرار گرفت؛

۲) حالت مبتنی بر زمینه و معنا، که طی آن معنای مفهومی واژگان، روابط معنایی و نقش آن‌ها در بافت جمله در تحلیل پرسش لحاظ شد.

جدول ۴. مقایسه عملکرد جستجو در حالت بدون زمینه و با زمینه معنایی

نوع پرسش جستجو	دقت	بازیابی	توضیح
بدون توجه به زمینه	۶۰٪	۶۵٪	نادیده گرفتن نقش معنایی واژگان موجب شد بخشی از اسناد مرتبط به دلیل عدم انطباق زمینه‌ای بازیابی نشوند.
با توجه به زمینه و معنا	۸۸٪	۸۲٪	لحاظ کردن روابط معنایی و بافت مفهومی باعث افزایش همزمان شناسایی اسناد مرتبط و کاهش اسناد غیرمرتبط شد.

نتایج نشان می‌دهد توجه به زمینه معنایی در مرحله طرح پرسش تأثیر چشمگیری بر کیفیت بازیابی اطلاعات دارد. دقت از ۶۰ به ۸۸ درصد افزایش یافته است که بیانگر کاهش محسوس اسناد غیرمرتبط است. همچنین بازیافت از ۶۵ به ۸۲ درصد رسیده که نشان می‌دهد بخش

بیشتری از اسناد مرتبط نیز کشف شده‌اند. این بهبود همزمان نشان‌دهنده آن است که درک صحیح از بافت موضوعی و روابط معنایی (از جمله هم‌معنایی، چندمعنایی و وابستگی‌های معنایی) نقش اساسی در هم‌ترازی پرسش‌های کاربران با ساختار نمایه‌سازی پایگاه داده ایفا می‌کند.

برای اعتبارسنجی آماری، آزمون t مستقل بر روی دقت و بازیافت دو گروه انجام شد و نتایج نشان داد که تفاوت‌ها از نظر دقت ($p < 0.01$) و بازیافت ($p < 0.05$) معنادار هستند. این یافته‌ها تأکید می‌کنند که درک صحیح از بافت موضوعی، هم‌معنایی، چندمعنایی و وابستگی‌های معنایی واژگان نقش اساسی در هم‌ترازی پرسش‌های کاربران با نمایه‌های پایگاه داده دارد.

به طور کلی، این یافته‌ها تأکید می‌کنند که مدل‌های جستجوی مبتنی بر معنا و پرسش‌هایی که با توجه به زمینه بازنویسی می‌شوند، نسبت به جستجوی صرفاً واژه‌محور عملکرد بسیار بهتری دارند و می‌توانند به شکل قابل توجهی کیفیت بازیابی اطلاعات را ارتقا دهند. بنابراین، استفاده از مدل‌های جستجوی مبتنی بر معنا و بازنویسی پرسش‌ها با توجه به زمینه مفهومی می‌تواند عملکرد سیستم‌های بازیابی اطلاعات را به طور قابل توجهی بهبود بخشد.

پرسش ۴: نقش بازخورد حرفه‌ای در بهبود بازیابی اطلاعات

به منظور سنجش تأثیر مداخله حرفه‌ای بر کیفیت جستجو، پرسش‌های کاربران در دو مرحله، پیش و پس از دریافت راهنمایی تخصصی، ارزیابی شد. در مرحله نخست، پرسش‌ها بدون هیچ‌گونه اصلاح یا مداخله ثبت و اجرا گردید. در مرحله دوم، کاربران پس از دریافت بازخورد حرفه‌ای—که شامل روشن‌سازی نیاز اطلاعاتی، پیشنهاد واژگان دقیق‌تر، و بازآرایی ساختاری پرسش بود—مجدد اقدام به جستجو کردند.

جدول ۵. مقایسه عملکرد کاربران قبل و بعد از دریافت بازخورد حرفه‌ای

مرحله جستجو	دقت	بازیابی	توضیح
قبل از بازخورد حرفه‌ای	۷۰٪	۶۸٪	وجود ابهام در پرسش‌ها و استفاده از واژگان غیرهدفمند موجب کاهش کیفیت نتایج شده بود.
بعد از بازخورد حرفه‌ای	۹۲٪	۸۷٪	بازنویسی هدفمند پرسش‌ها و تعریف روشن‌تر دامنه جستجو منجر به افزایش قابل توجه دقت و فراخوانی اسناد شد.

نتایج نشان می‌دهد که دریافت بازخورد حرفه‌ای نقش مؤثری در بهبود عملکرد جستجو دارد. ارتقای دقت از ۷۰ به ۹۲ درصد بیانگر آن است که اصلاح واژگان و رفع ابهام‌های معنایی توانسته حجم اسناد غیرمرتبط را به میزان چشمگیری کاهش دهد. همچنین، افزایش بازیافت از ۶۸ به ۸۷ درصد نشان می‌دهد که بازخورد حرفه‌ای نه تنها دامنه جستجو را محدود نکرده، بلکه با هدایت کاربر به استفاده از واژگان و ساختارهای مناسب، امکان بازیابی طیف گسترده‌تری از اسناد مرتبط را فراهم کرده است.

برای بررسی معناداری تفاوت‌ها، آزمون t زوجی بر روی دقت و بازیافت قبل و بعد از بازخورد انجام شد. نتایج نشان داد که افزایش دقت و بازیافت پس از دریافت بازخورد از نظر آماری معنادار است ($p < 0.01$). این یافته‌ها تأکید می‌کنند که بازخورد حرفه‌ای می‌تواند با هدایت کاربران به انتخاب واژگان دقیق‌تر و ساختار پرسش مناسب، همگرایی میان زبان کاربران و زبان نمایه‌سازی اسناد را افزایش دهد و کیفیت بازیابی اطلاعات را به طور قابل توجهی بهبود بخشد.

به طور کلی، یافته‌های این مطالعه نشان می‌دهد که نقش کتابداران و کارشناسان بازیابی اطلاعات فراتر از ارائه راهنمایی ساده است. مداخله حرفه‌ای به عنوان عاملی میانجی، پرسش‌های کاربران را به زبانی تبدیل می‌کند که هم با منطق سیستم بازیابی اطلاعات سازگار باشد و هم نیاز اطلاعاتی دقیق کاربر را بازتاب دهد. این فرآیند، شکاف موجود میان زبان

طبیعی کاربران و زبان نمایه‌سازی سیستم را کاهش می‌دهد و بهبود همگرایی میان پرسش و نمایه‌ها را تسهیل می‌کند. در نتیجه، مداخله تخصصی و هدایت کاربر در انتخاب واژگان و ساختار پرسش، نقش مهمی در افزایش دقت و بازیافت و ارتقای کارایی کلی نظام‌های بازیابی اطلاعات ایفا می‌کند.

نتیجه‌گیری

پژوهش حاضر با تمرکز بر نقش بنیادین زبان در فرآیند بازیابی اطلاعات نشان داد که کیفیت و کارآمدی جستجو، به طور مستقیم با نحوه صورت‌بندی پرسش‌ها در زبان طبیعی کاربران مرتبط است. نتایج تحلیل داده‌ها آشکار ساخت که ساختار نحوی روشن و منسجم، همراه با انتخاب واژگان دقیق و تخصصی، می‌تواند میزان تطابق میان نیاز اطلاعاتی کاربر و نمایه‌های سامانه را به شکل چشمگیری افزایش دهد.

در پرسش‌هایی که عناصر نحوی به‌درستی سازمان یافته بودند، هر دو شاخص دقت و بازیافت بهبود یافته‌اند؛ این امر نشان می‌دهد که حتی چینش واژگان نیز بر تفسیر نیت کاربر و فرآیند پردازش پرسش تأثیرگذار است. یافته‌ها با پژوهش‌های پیشین در حوزه پردازش زبان طبیعی و تحلیل نحوی هم‌راستا بوده و اهمیت توجه به وجه ساختاری زبان در طراحی پرسش‌ها و سامانه‌های جستجو را تأکید می‌کند.

از منظر معناشناختی، پژوهش نشان داد که واژگان دقیق و تخصصی نقش کلیدی در فهم صحیح پرسش‌ها و تطبیق آن‌ها با محتوای اسناد دارند. استفاده از واژگان کلی، مبهم یا چندمعنایی باعث افزایش بازیابی‌های نامرتب و کاهش دقت می‌شود، در حالی که واژگان زمینه‌مند و معنادار، همزمان دقت و بازیافت را بهبود می‌بخشند. این یافته‌ها نشان می‌دهد که تحلیل روابط معنایی، هم‌معنایی، چندمعنایی و زمینه کاربرد واژگان باید به طور جدی در طراحی نظام‌های جستجو لحاظ شود.

تحلیل نحوه ترجمه پرسش‌های زبان طبیعی به زبان جستجو نیز نشان داد که بخش قابل توجهی از ناکارآمدی‌ها ناشی از شکاف میان زبان کاربران و زبان سامانه است. پیچیدگی‌های نحوی یا واژگان نامناسب مانع از بازشناسی صحیح نیاز کاربر می‌شوند. در نتیجه، توسعه ابزارهای ترجمه، استانداردسازی واژگان و ارائه راهنمایی‌های هوشمند و تعاملی می‌تواند این شکاف را کاهش دهد و فرآیند تبدیل نیاز اطلاعاتی به پرسش‌های جستجویی مؤثر را تسهیل کند.

نقش بازخورد حرفه‌ای در پژوهش نیز برجسته شد؛ بازنویسی پرسش‌ها با مداخله کارشناسان، افزایش قابل توجهی در دقت و بازیافت ایجاد کرد. این امر نشان می‌دهد که حضور متخصصان اطلاعات و کتابداران، نه تنها نقش پشتیبانی دارد، بلکه به عنوان پلی میان زبان طبیعی کاربران و زبان سازمان‌یافته سامانه‌های جستجو عمل می‌کند و به کاهش ابهام‌ها و هدایت کاربران به واژگان و ساختارهای استاندارد کمک می‌کند.

در سطح نظری، یافته‌ها بیانگر آن است که تحلیل زبان در بازیابی اطلاعات را نمی‌توان به جنبه‌های صرفاً فنی محدود کرد. رویکردهای فلسفه زبان، به ویژه نظریه معنا در بستر کاربرد، می‌توانند درک عمیق‌تری از تعامل کاربر و سامانه فراهم کنند. اگر جستجو را نوعی گفت‌وگو میان کاربر و سامانه در نظر بگیریم، معنا نه در واژگان ثابت، بلکه در کاربرد آن‌ها و هدف کاربر شکل می‌گیرد. این دیدگاه می‌تواند راهگشای طراحی سامانه‌های هوشمندتر و تحلیل دقیق‌تر پرسش‌ها و اسناد باشد.

به طور کلی، یافته‌ها نشان می‌دهند که زبان نه تنها ابزار بیان نیاز اطلاعاتی، بلکه بخشی از خود فرآیند شکل‌گیری آن نیاز است. توجه همزمان به جنبه‌های نحوی، معناشناختی، زمینه کاربرد واژگان و سازوکارهای ترجمه میان زبان طبیعی و زبان جستجو، می‌تواند زمینه‌ساز توسعه سامانه‌هایی شود که نزدیک‌تر به نیاز واقعی کاربران، دقیق‌تر و کارآمدتر عمل کنند. این پژوهش همچنین می‌تواند به عنوان پایه‌ای برای مطالعات آینده در حوزه‌های تعامل انسان

و ماشین، تحلیل گفتمان در بازیابی اطلاعات، موتورهای جستجوی معناپایه و تأثیر بافت‌های اجتماعی و فرهنگی بر رفتار اطلاعاتی کاربران مورد استفاده قرار گیرد.

منابع

- استاینبرگ، دنی. (۱۳۸۶). ترجمه ارسلان گل‌فام. درآمدی بر روانشناس زبان. تهران: نشر سمت.
- باطنی، محمدرضا. (۱۳۷۸). زبان و تفکر. تهران: نشر فرهنگ معاصر.
- باطنی، محمدرضا. (۱۳۸۶). چهار گفتار درباره زبان. تهران: نشر آگاه.
- باطنی، محمدرضا. (۱۳۸۵). پیرامون زبان و زبان شناسی. تهران: نشر فرهنگ معاصر.
- باقری، ت. (۱۳۹۹). بررسی و شناسایی فناوری‌های معنایی و میزان اهمیت آن‌ها در بازیابی اطلاعات کتابخانه دیجیتال. مجله علم‌سنجی کاسپین دانشگاه علوم پزشکی بابل، ۱۷(۱)، ۱۶-۲۵. <https://doi.org/10.22088/cjs.7.1.52.25>
- باقری، ت.، نوروزی، ی.، اسفندیاری، ع.، و زارعی، ع. (۱۳۹۸). ارائه الگویی مفهومی برای کاربرد فناوری معنایی در بازیابی اطلاعات کتابخانه‌های دیجیتال. مطالعات ملی کتابداری و سازماندهی اطلاعات، ۳۰(۲)، ۱۲۹-۱۵۱.
- <https://doi.org/10.30484/nastinfo.2019.2145.1820>
- باقری، ت.، و نوروزی، ی. (۱۳۹۹). مقایسه بین وضع موجود و مطلوب کاربرد فناوری معنایی در بازیابی اطلاعات در نرم‌افزارهای کتابخانه‌های دیجیتال ایران. کتابداری و اطلاع‌رسانی، ۲۳(۴)، ۷۵-۷۸. <https://doi.org/10.30481/lis.2022.345678.1234.48>
- برنجیان، ش. ر. (۱۳۹۴). جستجوی ماشینی کلمات هم معنا در بازیابی اطلاعات. مقاله ارائه‌شده در اولین همایش ملی پیشرفت‌ها و چالش‌ها در علوم، مهندسی و فناوری، شیراز، ایران.
- <https://civilica.com/doc/560306>
- چامسکی، نوام (۱۳۷۷). ترجمه کوروش صفوی. زبان و ذهن. تهران: نشر هرمس.
- چپمن، شیوان. (۱۳۸۴). ترجمه حسین صافی. از فلسفه به زبانشناسی. تهران: نشر گام نو.
- صفوی، کوروش. (۱۳۸۰). گفتارهایی در زبان شناسی. تهران: نشر هرمس.
- فهمی حجازی، محمود. (۱۳۷۹). زبان شناسی عربی. تهران: نشر سمت.
- مرتضائی، ل. (۱۳۹۲). مسائل زبان فارسی خط و ذخیره در بازیابی و سازی اطلاعات. اطلاع‌رسانی و کتابداری، فصل‌نامه مرکز مدارک و علمی ایران، ۷۱(۱-۲).
- نجفی، ابوالحسن. (۱۳۵۸). مبانی زبان شناسی و کاربرد آن در زبان فارسی. تهران: نشر دانشگاه آزاد.

نوذری، س. (۱۳۹۴). کاربرد شبکه معنایی در بازیابی شعر: نوذری سودابه. پژوهشنامه پردازش و مدیریت اطلاعات، ۳۱(۱)، ۱۲۹-۱۵۱. <https://doi.org/10.35050/JIPM010.2015.006>

نوروزی، ی. (۱۳۹۸). بافت و معنا در بازیابی اطلاعات (با تأکید بر زبان مادری). فصلنامه پژوهشگاه علوم و فناوری اطلاعات ایران، ۳۵(۱)، ۱-۲۴.

واله، حسین (۱۳۸۲). متافیزیک و فلسفه زبان. تهران: نشر هرمس.

هماونلی، م، نوروزی، ی.، و خبازان، ب. (۱۴۰۳). نحو و معنا: روندها و گرایش‌های پژوهشی. فصلنامه نظام‌های معنایی و بازیابی دانش، ۱۱(۲۳۹)، ۴۵-۵۴. <https://doi.org/10.22054/jks.2023.75172.1599>

References

- Austin, J. L. (1962). *How to do things with words* (J. O. Urmson, Ed.). Oxford University Press
- Blackburn, S. (1984). *Spreading the word: Groundings in the philosophy of language*. Oxford University Press.
- Blair, D. C. (1990). *Language and representation in information retrieval*. Amsterdam: Elsevier Science Publishers.
- Blair, D. C. (1992). *Information retrieval and the philosophy of language*. *The Computer Journal*, 35(3), 200-207. <https://doi.org/10.1093/comjnl/35.3.200>
- Breuer, T., Frihat, S., Fuhr, N., Lewandowski, D., Schaef, P., & Schenkel, R. (2025). Large language models for information retrieval: Challenges and chances. *Datenbank-Spektrum*, 25, 71-8
- Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.
- Chapman, S. (2000). *Linguistics and philosophy: An introduction*. New York: Routledge.
- Cooper, W. S. (1973). On selecting a measure of retrieval effectiveness. *Journal of the American Society for Information Science*, 24(2), 87-100. <https://doi.org/10.1002/asi.4630240204>
- Fan, Y., Guo, J., Ma, X., Zhang, R., Lan, Y., & Cheng, X. (2021, March 1). A linguistic study on relevance modeling in information retrieval. arXiv. <https://doi.org/10.48550/arXiv.2103.00956>
- Frohmann, B. (1990). Rules of indexing: A critique of mentalism in information retrieval theory. *Journal of Documentation*, 46(2), 81-101. <https://doi.org/10.1108/eb026855>
- Furnas, G. W., Landauer, T. K., Gomez, L. M., & Dumais, S. T. (1987). The vocabulary problem in human-system communication. *Communications of the ACM*, 30(11), 964-971. <https://doi.org/10.1145/32206.32212>
- García-Carpintero, M. (1996). *Las palabras, las ideas y las cosas: Una presentación de la filosofía del lenguaje*. Ariel.
- Grice, H. P. (1975). *Logic and conversation*. In P. Cole & J. L. Morgan (Eds.), *Syntax and semantics* (Vol. 3, pp. 41-58). Academic Press.
- Harter, S. P. (1992). Psychological relevance and information science. *Journal of the American Society for Information Science*, 43(9), 602-615.
- Hjørland, B. (1998). Theory and metatheory of information science: A new interpretation. *Journal of Documentation*, 54(5), 606-621. <https://doi.org/10.1108/EUM0000000007183>.

- Karpukhin, D., Oguz, B., Min, S., Lewis, P., Wu, L., Edunov, S., Chen, D., & Yih, W.-T. (2020). Dense passage retrieval for open-domain question answering. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)* (pp. 6769–6781). Association for Computational Linguistics.
<https://doi.org/10.18653/v1/2020.emnlp-main.550>
- Kimbrough, S. O. (1990). On representation schemes for promising electronically. *Decision Support Systems*, 6(2), 99–122.
- Kurland, O., & Lee, L. (2004). Corpus structure, language models, and ad hoc information retrieval. In *Proceedings of the 27th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval* (pp. 186–193). ACM.
- Lancaster, F. W. (1992). *Vocabulary control for information retrieval*. Information Resources Press.
- Lioma, C., Larsen, B., Lu, W., & Huang, Y. (2016). A study of factuality, objectivity and relevance: Three desiderata in large-scale information retrieval? In *Proceedings of the 3rd IEEE/ACM International Conference on Big Data Computing, Applications and Technologies (BDCAT 2016)* (pp. 107–117). ACM
- Lycan, W. G. (2000). *Philosophy of language: A contemporary introduction*. Routledge.
- Manning, C. D., Raghavan, P., & Schütze, H. (2008). *Introduction to information retrieval*. Cambridge University Press.
- manning, C. D., Raghavan, P., & Schütze, H. (2008). *Introduction to information retrieval*. Cambridge University Press. balyan.ir+1
- Marchionini, G. (2008). Human–information interaction research and development. *Library & Information Science Research*, 30(3), 165–174.
http://ils.unc.edu/~march/Marchionini_Inf_interact_LISR_2008
- Moniri, S., Schlosser, T., & Kowerko, D. (2024). Investigating the Challenges and Opportunities in Persian Language Information Retrieval through Standardized Data Collections and Deep Learning. *Computers*, 13(8), 212.
<https://doi.org/10.3390/computers13080212> MDPI+1
- Mothe, J. (2022). Analytics Methods to Understand Information Retrieval Effectiveness — A Survey. *Mathematics*, 10(12), 2135. MDPI
- Mothe, J. (2022). Analytics methods to understand information retrieval effectiveness—A survey. *Mathematics*, 10(12), 2135.
<https://doi.org/10.3390/math10122135>

- Reddy, S., Patro, S., Dixit, T., & Shen, X. (2025). Large Language Models for Information Retrieval: Challenges and Chances. *Datenbank-Spektrum*, 25, 71–81. SpringerLink
- Ruthven, I., & van Rijsbergen, C. J. (1996). Context generation and information retrieval. In J. Stewman (Ed.), *Proceedings of the Florida Artificial Intelligence Research Symposium (FLAIRS 1996)* (pp. 380–384).
- Saussure, F. de. (1916). *Cours de linguistique générale* (C. Bally & A. Sechehaye, Eds.). Payot.
- Schramm, W. (1954). *The Process and Effects of Mass Communication*. University of Illinois Press.
- Searle, J. R. (1969). *Speech acts: An essay in the philosophy of language*. Cambridge University Press.
- Shannon, C. E., & Weaver, W. (1949). *The Mathematical Theory of Communication*. Urbana, IL: The University of Illinois Press, 1-117.
- Thornley, C., & F. Gibb. 2007. A dialectical approach to information retrieval. *Journal of documentation*, 63 (5): 755-764.
- Thornley, C., & Gibb, F. (2009). Meaning in philosophy and meaning in information retrieval. *Journal of Documentation*, 65(1), 133–150. <https://doi.org/10.1108/00220410910926158>.
- van Rijsbergen, C. J. (1979). *Information retrieval*, 2nd ed., London: Butterworths
- van Rijsbergen, C. J. (1986). A new theoretical framework for information retrieval. In *Proceedings of the 9th Annual International ACM SIGIRConference on Research and Development in Information Retrieval* (pp. 194–200). ACM.
- Wilson, P. (1973). Situational relevance. *Information Storage & Retrieval*, 9(5), 457–471.
- Wittgenstein, L. (1953). *Philosophical investigations* (G. E. M. Anscombe, Trans.). New York: Macmillan
- Zhao, W. X., Liu, J., Ren, R., & Wen, J.-R. (2022). Dense Text Retrieval based on Pretrained Language Models: A Survey. arXiv. arXiv
- Zhao, W. X., Liu, J., Ren, R., & Wen, J.-R. (2022, November 27). Dense text retrieval based on pretrained language models: A survey. arXiv. <https://doi.org/10.48550/arXiv.2211.14876>