

Syntax and Semantics: Research Trends and Directions

Hoda Homavandi 

Ph.D. Graduate of Knowledge and Information Sciences, University of Tehran, Tehran, Iran

Yaghub Norouzi * 

Professor, Department of Knowledge & Information Science, University of Qom, Qom, Iran

Bent-ol Hoda Khabbazan 

PhD Candidate in Knowledge & Information Science, University of Qom, Qom, Iran

Abstract

In an environment where the development of artificial intelligence and facilitating the interaction between man and machine is increasingly important, language challenges are becoming more and more important points in this direction. The important role of language in transferring and retrieving information cannot be examined independently of its syntax and semantics. Explaining and describing the researches of this field from quantitative and qualitative dimensions and understanding the tendencies of researchers can be an important step in understanding the importance of syntax and semantics in communication structures in modern information search and retrieval environments. Based on this, in this descriptive-analytical study, with the aim of qualitative and quantitative analysis, studies in the field of syntax and semantics have been conducted in information retrieval. In order to conduct this research, Aviard's systematic review method was used. The analysis of studies in this field showed that these studies focus on the three main axes of syntax and semantics in relation to information organization, information retrieval, and information search. Overall, the analysis of the selected studies shows that addressing the discussion of syntax and semantics in relation to the organization of information, as well as explaining, and analyzing the theoretical aspects of syntax and semantics in information retrieval is one of the things that requires more research,

* Corresponding Author: ynorouzi@gmail.com

How to Cite: Homavandi, H., Norouzi, Y., & Khabbazan, B.H. (2024). Syntax and Semantic: Research Trends and Directions. *Journal of Knowledge Retrieval and Semantic Systems*, 11(39), 205-250. DOI: 10.22054/jks.2023.75172.1599

especially interdisciplinary studies. Researchers have the fields of linguistics, information science, and computer science. The results also can have guidelines for specialists in various fields who somehow deal with storing and retrieving information. In addition, the identified axes and the methods and approaches used and the identified research gaps open the way for researchers to conduct future studies

Keywords: Syntax and Semantics, Information Retrieval, Information Search, Research Trends

1. Introduction

Recently, the development of artificial intelligence and human-computer interaction has highlighted the increasing importance of language challenges in information retrieval. The crucial role of language in disseminating, accessing, and retrieving information cannot be studied independently of syntax and semantics. Explaining and describing research in this field from both quantitative and qualitative perspectives, and understanding researchers' trends, is an important step in comprehending the significance of syntax and semantics in communication structures within modern information search and retrieval environments. Consequently, in this descriptive and analytical study, we conducted qualitative and quantitative analyses of studies in the field of syntax and semantics in information retrieval.

2. Literature Review

In recent years, there has been a lot of interdisciplinary research focusing on investigating the impact of language on the interaction between users and the web environment. These studies have discussed the language from various perspectives and have explored information retrieval across different types of information media, including web databases, search engines, commercial websites, and libraries. Tapsai (2019), Norouzi and Hamavandi (2018), Hammo (2009), Lazarinis (2008), Ofoghi, Yearwood & Ghosh (2006) have focused on different languages such as Persian, English, Arabic, and Greek. The findings show that the syntax and morphology, as well as the semantics of searched terms and phrases, have a significant impact on the retrieval of results. In addition, search tools tend to rely more on the general form of words instead of focusing on the real needs of users in order to improve the search process.

Due to the huge amount of information on the World Wide Web and the challenges related to information retrieval, researchers and software developers have turned to the Semantic Web to keep up with the changes. The Semantic Web has provided a large amount of structured and machine-understandable information on a wide range of topics (Guha, McCool & Miller, 2003). Semantic models perform well in identifying and recognizing synonyms, similar words, and semantic frameworks. Therefore, one of the most important challenges in the field of information storage and retrieval is to bridge the gap between the language used by information seekers and information providers (Rezaee Sharifabadi et al., 2010).

The current study aims to systematically review previous research findings on syntax and semantics in information storage and retrieval across different contexts. Each context represents different dimensions of knowledge representation systems, from traditional to semantics. Upon reviewing the research, it was found that no systematic review has been conducted with a focus on syntax and semantics in the field of information retrieval.

3. Methodology

In this qualitative research using Aveyard's systematic review method, we aim to address the following questions:

- 1) What is the statistical status of studies in the field of syntax and semantics in storing and retrieving information?
- 2) What are the main subject areas that researchers have focused on in studies related to syntax and semantics in storing and retrieving information?
- 3) What research methods and approaches have researchers employed in this field?
- 4) What are the research gaps and areas that require further study in this field?

To gather relevant sources from information databases, we selected search keywords based on the research questions. Then, we used search strategies and various operators to combine the keywords and phrases, ensuring a comprehensive and effective search in Persian databases such as Magiran, Irandoc, SID, NoorMagz, ISC, and Civilica, as well as databases including Scopus, Emerald, ProQuest, and Google Scholar. There was no time limitation for the search. We recorded accepted sources such as articles and theses that were

relevant and valid. By removing irrelevant and duplicate sources, we selected 12 Persian sources and 42 English sources. After categorizing, the studies were analyzed according to the type of source, research method, and tool. The results of the analysis of the studies were presented in the form of tables and graphs.

4. Results

The selected studies were categorized into three groups (their characteristics were described in detail): *information retrieval*, *information organization*, and *information search* based on the analysis of keywords and subjects raised in the sources. The results of the study revealed that among the 54 reviewed studies, Iranian researchers had conducted the most research in the field of syntax and semantics in information retrieval, with 12 studies. The United States followed with 5 studies, and China and Vietnam tied for third place with 4 studies each. The majority of the studies focused on syntax and semantics in information retrieval.

5. Discussion

Analysis of 54 selected studies has shown that these studies were conducted over a period of 26 years. The oldest study was included in the review back to 1997, while the most recent one is from 2022. This shows the dynamic nature of the field under investigation and demonstrates how it is constantly changing and being influenced by the advancement of web technologies. Furthermore, a thematic analysis of the research, based on the studies' keywords, reveals that "Ontology," as a tool of the semantic web, is closely linked to the semantic and syntax aspects of language in information retrieval.

Moreover, in a total of 54 studies, the majority were experimental (19), followed by applied (15) and analytical (9). Additionally, there were 6 studies that combined applied and analytical methods. Content analysis and comparative analysis each had 2 instances, while case studies were the least frequent with only 1 case. These studies have utilized tools such as ontology, search engines, and techniques including natural language processing, annotation, tagging, and indexing.

The discussion about exploring syntax and semantics in relation to information retrieval across different languages is believed to make a significant contribution to the development of future research


literature in this field. This is because users' native language plays a central role in forming search terms for information retrieval, based on subjective meanings, context, and content. Considering this point can effectively enhance information retrieval systems.


6. Conclusion


Although many studies have addressed various aspects of syntax and semantics in information retrieval, more research is needed to investigate syntax and semantics in information organization. It is also important to delve into and analyze their theoretical aspects in information retrieval, especially through interdisciplinary studies.

Moreover, the interconnectedness of various areas of study demonstrates the close relationship between syntax and semantics and linguistic issues in nearly every field that involves organizing, storing, and retrieving information. These areas include the study of syntax and semantics in relation to environmental sensors, indexing, identification and summarization of texts, plagiarism detection, natural language processing, repositories, improvement of query users and data retrieval in repositories and search engines, metadata enrichment, and image retrieval. The results of this research, such as its main themes, identified methods and approaches, and research gaps, can offer valuable insights for future studies.

نحو و معنا: روندها و گرایش‌های پژوهشی

هدی همبندی  دانش‌آموخته دکتری علم اطلاعات و دانش‌شناسی، دانشگاه تهران، تهران، ایران

یعقوب نوروزی * استاد، گروه علم اطلاعات و دانش‌شناسی، دانشگاه قم، قم، ایران

بنت‌الهدی خبازان  دانشجوی دکتری علم اطلاعات و دانش‌شناسی، دانشگاه قم، قم، ایران

چکیده

در فضایی که توسعه هوش مصنوعی و تسهیل تعامل بین انسان و ماشین به‌طور فزاینده مورد توجه است، چالش‌های زبانی به‌عنوان نقاطی مهم در این مسیر بیش‌ازپیش خودنمایی می‌کنند. نقش مهم زبان در انتقال و بازیابی اطلاعات را نمی‌توان فارغ از نحو و معنای آن بررسی نمود. تبیین و تشریح پژوهش‌های این حوزه از ابعاد کمی و کیفی و درک گرایش‌های پژوهشگران می‌تواند گام مهمی در درک اهمیت نحو و معنا در ساختارهای ارتباطی در محیط‌های نوین جستجو و بازیابی اطلاعات باشد. بر همین اساس این مطالعه که از نوع توصیفی - تحلیلی است، با هدف تحلیل کیفی و کمی مطالعات حوزه نحو و معنا در بازیابی اطلاعات انجام شده است. به‌منظور انجام این پژوهش از روش مرور نظام‌مند آویارد استفاده شد. تحلیل مطالعات این حوزه نشان داد که این مطالعات بر سه محور اصلی نحو و معنا در ارتباط با سازمان‌دهی اطلاعات، بازیابی اطلاعات و جستجوی اطلاعات تمرکز دارند. در مجموع تحلیل مطالعات منتخب نشان می‌دهد پرداختن به بحث نحو و معنا در ارتباط با سازمان‌دهی اطلاعات و همچنین تبیین و واکاوی وجوه نظری نحو و معنا در بازیابی اطلاعات از مواردی است که نیاز به پژوهش بیشتر، خصوصاً مطالعات بین‌رشته‌ای بیشتر از سوی پژوهشگران حوزه‌های زبان‌شناسی، علم اطلاعات و علوم رایانه دارد. نتایج پژوهش حاضر می‌تواند رهنمودهایی برای متخصصان حوزه‌های مختلفی داشته باشد که به‌نوعی با ذخیره و بازیابی اطلاعات سروکار دارند. به‌علاوه محورهای شناسایی‌شده و روش‌ها و رویکردهای به‌کار رفته و شکاف‌های پژوهشی شناسایی‌شده راهگشای پژوهشگران در انجام مطالعات آتی است.

کلیدواژه‌ها: نحو و معنا، بازیابی اطلاعات، جستجوی اطلاعات، روندهای پژوهش

مقدمه

برقراری ارتباط از اساسی‌ترین نیازهای موجودات زنده از جمله انسان است. خصوصاً در عصر کنونی که قابلیت برقراری ارتباطات حتی فراتر از مرزها و محدودیت‌های زمانی و مکانی از ضروریات جوامع به شمار می‌رود. زبان به ما امکان برقراری ارتباط می‌دهد و تقریباً همه موجودات زنده به‌نوعی دارای این توانایی هستند. زبان وسیله‌ای است که برای تفاهم و انتقال پیام بین آدمیان به وجود آمده است؛ مانند دیگر موجودات زنده و حتی جمادات که هر کدام برای تعامل، زبان ویژه خود را دارند، زبان به ما این امکان را می‌دهد که بتوانیم اندیشه‌ها و مفاهیم پیچیده را منتقل کنیم (نوروزی، ۱۳۹۸). زبان طبیعی یک پدیده چندوجهی است که در یک کانال ارتباطی برای انتقال مفهوم بین افراد یک جامعه به کار می‌رود. دوسوسور^۱ زبان را متشکل از دو وجه «صورت» و «معنا» می‌داند که صورت به‌واسطه نظام آوایی یا نوشتاری تجلی عینی می‌یابد و معنا در این کانال ارتباطی جنبه انتزاعی دارد (نقل در قیومی، ۱۳۹۸). در زبان‌شناسی نوین به‌طور کلی لایه‌های مختلفی برای نظام زبان در نظر گرفته می‌شود و یکی از این لایه‌ها معنا است که با درک مفهوم واژه‌ها و جملات سروکار دارد (رفیعی، ۱۳۹۴). معنا را می‌توان خون همواره در جریان رگ‌های زبان دانست. کنار هم قرار گرفتن زنجیره آواها و واج‌ها و سایر عناصر سازنده زبان همه به‌منظور برآورده ساختن یک هدف گویشوران صورت می‌گیرد که آن انتقال معنا است. درواقع می‌توان گفت رابطه دانش‌های نحو و معنی‌شناسی به‌گونه‌ای است که اگر یکی از این دو به شکل درست به کار گرفته نشود دیگری نیز به تأثیر از آن مخدوش و ناقص می‌شود و مقصود متکلم، به‌درستی به مخاطب انتقال نمی‌یابد. همسو با دیدگاه ویکری و ویکری^۲ (۱۹۸۹) اگر به مجموعه جملات موجود در یک متن بیندیشیم، می‌توان دریافت که چه ساختار ارتباطی پیچیده‌ای دارد. درون جمله، واژه‌ها در طبقات نحوی جای می‌گیرند. در ورای این، جملات نیز به شیوه‌ای با یکدیگر ارتباط می‌یابند که بتوانند بحث موجود در یک متن را پیش ببرند. در نتیجه، با مجموعه‌ای از مفاهیم آشکار و پنهان در درون متن مواجه هستیم که روند ذخیره، سازمان‌دهی و بازیابی آن را به‌شدت تحت‌الشعاع خود قرار می‌دهد.

1. de Saussure

2. Vickery & Vickery

از همین روست که در سال‌های اخیر، بسیاری از پژوهش‌های بین‌رشته‌ای به بررسی تأثیر مسائل زبانی در تعامل کاربران و محیط وب معطوف شده‌اند؛ پژوهش‌هایی که هر یک از دیدگاهی بحث زبان در کاوش و بازیابی اطلاعات را در انواع رسانه‌های اطلاعاتی اعم از پایگاه‌های تحت وب، موتورهای کاوش، وب‌سایت‌های تجاری، کتابخانه‌ای و ... مورد کنکاش قرار داده و مشکلات ناشی از آن را بررسی نموده‌اند. یافته‌های پژوهش تاپسای^۱ (۲۰۱۹)، نوروزی و هم‌اوندی (۲۰۱۸)، هامو^۲ (۲۰۰۹)، لازارینیس^۳ (۲۰۰۸)، افقی و همکاران^۴ (۲۰۰۶) با تمرکز بر زبان‌های مختلف از جمله فارسی، انگلیسی، عربی، یونانی و... نشان می‌دهد که نحو یا ریخت‌شناسی واژه‌ها و عبارت‌های جستجو شده و معنای هر یک از آن‌ها بر بازیابی نتایج اثر دارد و ابزارهای جستجو به جای تمرکز بر نیاز واقعی کاربران در جهت بهبود فرایند کاوش، بیشتر بر شکل کلی واژه‌ها تکیه می‌کنند. با وجود مقبولیت و استفاده زیاد از فضای وب، موتورهای کاوش و حتی ابزارهای هوش مصنوعی برای یافتن انواع اطلاعات توسط طیف گسترده‌ای از کاربران و فراهم‌شدن ابزارهای متنوع با امکانات متفاوت برای جستجو در محیط وب، ویژگی‌های زبانی اعم از نحو و معنا، همچنان به‌عنوان عاملی مهم در فرایند سازمان‌دهی و ذخیره و بازیابی اطلاعات مطرح هستند و در موارد بسیاری به آن‌ها پرداخته شده است.

با توجه به آنچه ذکر شد، علی‌رغم استفاده زیاد کاربران از موتورهای کاوش و برخورداری موتورهای جستجو از امکانات مناسب و محبوبیت آن‌ها برای جستجوی وب توسط کاربران، این ابزارها معایب و مشکلات عدیده‌ای نیز دارند. موتورهای کاوش به دلیل جستجو بر اساس (تطبیق صوری و نه نحوی و معنایی) کلیدواژه‌ها در بسیاری از مواقع، نتایج مرتبط با نیاز کاربر را ارائه نمی‌دهند و او را در میان حجم وسیعی از اطلاعات غیرمرتبط و اضافی رها می‌کنند، به طوری که کاربر باید زمان زیادی را به بازمینی و مرور در مدارک بازیابی شده صرف کند تا منبع موردنیاز خود را بیابد و در بیشتر مواقع، عدم دسترسی به اطلاعات موردنیاز وجود دارد (دری، ۱۳۹۳). با وجود مقدار زیاد اطلاعات در شبکه جهانی وب و مشکلات موجود در رابطه با جستجوی اطلاعات،

1. Tapsai, Ch.

2. Hammo, B.H.

3. Lazarinis, F.

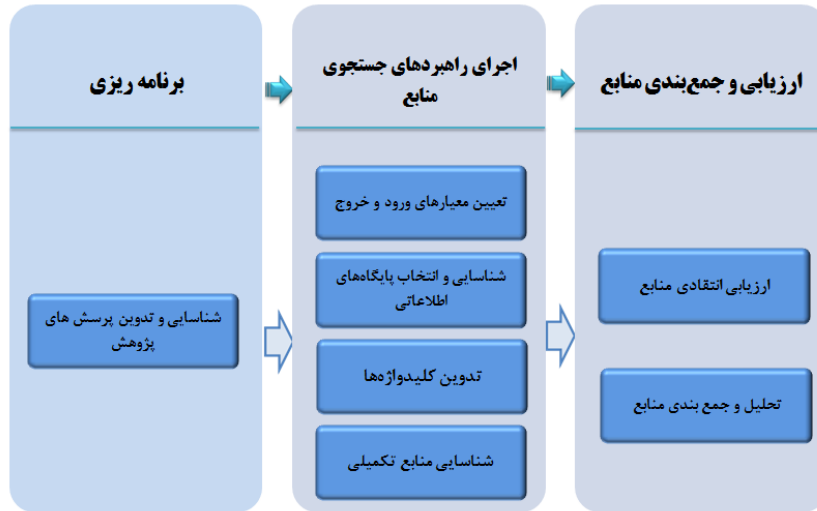
4. Ofoghi et al.

پژوهشگران و تولیدکنندگان نرم‌افزارها، برای همگام شدن با تغییرات جدید، به سمت شکل جدیدی از فناوری به نام وب معنایی روی آوردند. ظهور وب معنایی حجم زیادی از اطلاعات ساختارمند و قابل فهم برای ماشین را در مورد طیف وسیعی از اشیا، در دسترس قرار داده است (Guha et al., 2003). پرداختن به پژوهش‌هایی که بر چالش نزدیک کردن زبان سامانه‌های جستجو و بازیابی اطلاعات به زبان کاربران و غنی‌سازی معنایی آن تمرکز دارند و در بافت‌های موضوعی و زبانی مختلف بحث نحو و معنا را در بازیابی اطلاعات مورد مطالعه قرار دادند، باعث درک بهتر وضعیت کنونی این حوزه خواهد بود. علاوه بر این داشتن دیدگاهی جامع از این حوزه موضوعی دورنمای شفاف‌تری از مسیر پیش‌روی پژوهشگران، کاربران و ساینز ذینفعان ارائه می‌کند.

در این راستا، پژوهش حاضر بر آن است تا با مروری نظام‌مند، یافته‌های پژوهش‌های انجام‌شده در ایران و جهان را در خصوص نحو و معنا در ذخیره و بازیابی اطلاعات در بافت‌های مختلف که هر یک نشان‌دهنده ابعاد گوناگون این مقوله در نظام‌های بازنمون دانش از سنتی تا معنایی هستند، ارائه دهد. همچنین مطالعه مروری نظام‌مند این حوزه و تحلیل و بررسی مطالعات انجام‌شده به منظور شناسایی خلأها و کاستی‌های پژوهش در این زمینه و آشنایی پژوهشگران حوزه ذخیره و بازیابی اطلاعات به‌ویژه در داخل کشور از مواردی است که ضرورت انجام این مطالعه را روشن می‌سازد. همچنین با مرور متون پژوهشی، می‌توان گفت، تاکنون پژوهشی به روش مرور نظام‌مند و با تمرکز بر نحو و معنا در حوزه بازیابی اطلاعات انجام نشده است.

روش‌شناسی پژوهش

پژوهش حاضر از نوع کیفی است که به روش مرور نظام‌مند انجام شده است و بر مبنای الگوی آویارد (۱۳۹۰) یافته‌های پژوهشی مرتبط با متغیرهای نحو و معنا را مورد بررسی قرار داده است. در نمودار ۱ گام‌های اجرای پژوهش حاضر ارائه شده است:



نمودار ۱. روش اجرای مرور نظام‌مند در پژوهش حاضر

با انجام مرور نظام‌مند منابع، علاوه بر ارائه دیدگاهی کلی از حوزه مورد مطالعه و روند مطالعات انجام‌شده به شناسایی ابعاد مختلف موضوع مورد پژوهش، شکاف‌ها، زمینه‌ها و ظرفیت‌های انجام پژوهش‌های جدید پرداخته می‌شود. در ادامه گام‌های انجام این پژوهش به تفکیک ارائه می‌شود.

تعیین و شناسایی پرسش‌های پژوهش، گام بسیار مهمی در روش شناسی پژوهش مرور نظام‌مند است. بنابراین پژوهش حاضر به دنبال پاسخ به پرسش‌های ذیل است:

۱. وضعیت آماری مطالعات حوزه نحو و معنا در ذخیره و بازیابی اطلاعات در ایران و جهان به چه صورتی است؟
۲. محورهای موضوعی که پژوهشگران در پژوهش‌های مرتبط با نحو و معنا در ذخیره و بازیابی اطلاعات به آن پرداخته‌اند، کدام‌اند؟
۳. پژوهشگران در پژوهش‌های حوزه نحو و معنا در ذخیره و بازیابی اطلاعات از چه روش‌ها و رویکردهای پژوهشی استفاده کرده‌اند؟
۴. شکاف‌های پژوهشی و زمینه‌های نیازمند مطالعات بیشتر در این حوزه کدام‌اند؟

به‌منظور اطمینان از اینکه منابع شناسایی‌شده با پرسش‌های پژوهش مرتبط هستند، معیارهای ورود و خروج که در جدول ۱ آمده تعریف شد.

جدول ۱. معیارهای پذیرش و عدم پذیرش (ورود و خروج) منابع بازیابی شده

عنوان	معیار پذیرش	معیار عدم پذیرش
موضوع پژوهش	مطالعات ارائه‌دهنده ارتباط نحو و معنا با سازمان‌دهی، بازیابی و جستجوی اطلاعات	مواردی غیر از موضوعات یادشده
زبان پژوهش	فارسی / انگلیسی	غیر از زبان فارسی و انگلیسی
نوع مطالعه	مقالات مجلات / فراهمایی‌ها / پایان‌نامه‌ها	مواردی مانند فصول کتاب، نقد و نظرها و پوستر

در گام بعدی ابتدا به‌منظور استخراج منابع مرتبط از پایگاه‌های اطلاعاتی، کلیدواژه‌های متناسب با سؤال پژوهش که در جدول ۲ آمده است، تدوین شد. پس از آن کلیدواژه‌های مدنظر و عبارت حاصل از ترکیب آن‌ها با استفاده از راهبردهای جستجو و عملگرهای متنوع جهت حفظ جامعیت و مانعیت جستجو، در پایگاه‌های فارسی بانک اطلاعات نشریات کشور (مگیران^۱)، پژوهشگاه علوم و فناوری اطلاعات ایران (ایرانداک^۲)، پایگاه مرکز اطلاعات علمی جهاد دانشگاهی (اس‌آی‌دی^۳)، پایگاه مجلات تخصصی نور (نورمگز^۴)، پایگاه استنادی علوم جهان اسلام (ای‌اس‌سی^۵) و مرجع دانش (سیویلیکا^۶) و پایگاه‌های اطلاعاتی انگلیسی شامل اسکوپوس^۷، امرالد^۸، پروکوئست^۹ و گوگل اسکالر^{۱۰}

1. Magiran

2. Irandoc

3. Sid

4. Noormags

5. Isc

6. Civilica

7. Scopus

8. Emerald

9. ProQuest

10. Google Scholar

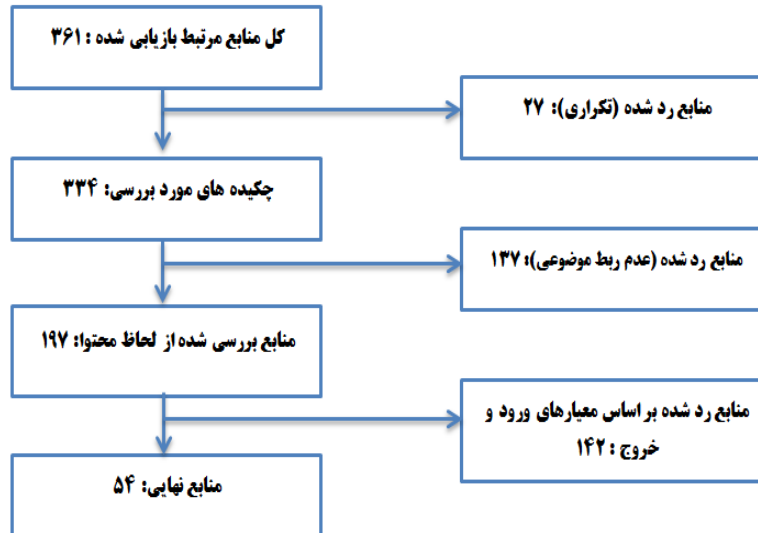
مورد جستجو قرار گرفتند و مقالاتی که از نظر ارتباط موضوعی و اعتبار^۱ مورد قبول بودند به منظور انجام مراحل بعدی ذخیره و اطلاعات آن‌ها ثبت شد. لازم به ذکر است که در این پژوهش به منظور حفظ جامعیت و بررسی سیر تاریخی پژوهش‌ها از اعمال محدودیت زمانی به هنگام جستجو اجتناب شد.

جدول ۲. کلیدواژه‌های فارسی و انگلیسی مورد استفاده جهت جستجوی منابع

کلیدواژه‌های فارسی	کلیدواژه‌های انگلیسی
نحو	Syntax
معنا	Semantic
بازیابی اطلاعات	Information retrieval
سازمان‌دهی اطلاعات	Information organization
جستجوی اطلاعات	Information search

سپس منابعی که در عنوان و چکیده به کلیدواژه‌های مورد جستجو اشاره داشتند با توجه به پرسش‌ها و هدف پژوهش و معیارهای ورود، مورد بررسی قرار گرفتند. در ادامه منابع غیرمرتبط و تکراری حذف شدند. در نهایت ۱۲ منبع فارسی و ۴۲ منبع انگلیسی انتخاب شد (نمودار ۲).

^۱ در مورد اعتبار مقالات موارد معدودی که فاقد داوری بوده و یا یک محتوا به چند صورت منتشر شده و مواردی از این قبیل در مطالعه وارد نشدند.



نمودار ۲. فرایند بررسی و انتخاب منابع

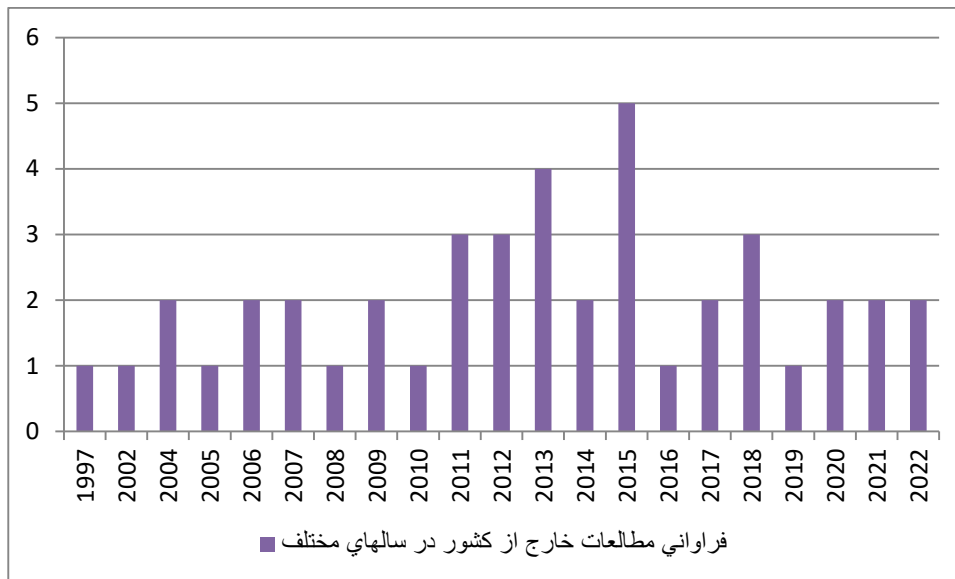
جهت تجزیه و تحلیل منابع، ابتدا در فایل اکسل جدولی با اختصاص کدی به هر منبع تهیه گردید. قابل ذکر است که تحلیل کلیدواژه‌ها و موضوعات مطرح شده در منابع مبنای گروه‌بندی مطالعات نحو و معنا در چهار گروه بازیابی اطلاعات، سازمان‌دهی اطلاعات و جستجوی اطلاعات قرار گرفت. بعد از گروه‌بندی، مطالعات از جهت نوع منبع، روش و ابزار پژوهش مورد بررسی قرار گرفتند و تحلیل شدند.

یافته‌ها

وضعیت آماری مطالعات حوزه نحو و معنا در ایران و جهان

از بین ۵۴ مطالعه مورد بررسی، ۴۲ عنوان در خارج از کشور و ۱۲ عنوان در داخل کشور انجام و منتشر شده است. در بین مطالعات مورد بررسی، ۱۰ عنوان مربوط به پایان‌نامه‌ها (فارسی و انگلیسی) و ۴۴ عنوان متعلق به مقالات پژوهشی است. در این پژوهش بیشترین مطالعات حوزه نحو و معنا در خارج از کشور بین سال‌های ۲۰۱۱ تا ۲۰۱۵ میلادی انجام شده و روند انجام پژوهش‌ها در سال‌های مختلف ادامه داشته و متوقف نشده است (نمودار ۳).

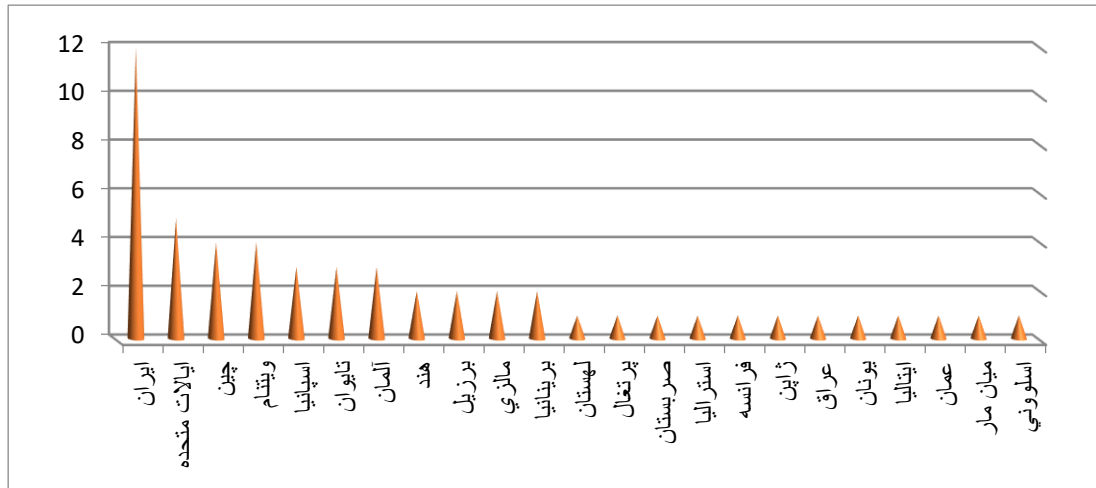
در ایران نیز مطالعات در این حوزه از سال ۱۳۹۰ آغاز شده است و به‌جز سال‌های ۱۳۹۳ و ۱۳۹۹ در بقیه سال‌ها مطالعات در هر سال یک یا دو مورد انجام شده است.



نمودار ۳. فراوانی مطالعات خارج از کشور در مهر و موم‌های مختلف

از نظر مشارکت پژوهشگران کشورهای مختلف در مطالعات حوزه نحو و معنا با احتساب برون‌دادها و فعالیت‌های پژوهشی ارائه‌شده به زبان فارسی، از بین ۵۴ پژوهش بررسی‌شده کشور ایران با ۱۲ مطالعه رتبه نخست را دارد^۱ و پس از آن ایالات متحده با ۵ مطالعه و چین و ویتنام با ۴ مطالعه در جایگاه دوم و سوم قرار دارند (نمودار ۴).

^۱ به نظر می‌رسد اگرچه شروع پژوهش‌های داخلی مرتبط با نحو و معنا در بازیابی اطلاعات در ایران در مقایسه با پژوهش‌های خارج از کشور با تأخیر بوده است، اما به لحاظ تعداد و کمیت (به دلایلی از جمله ویژگی‌های خاص زبان فارسی در مواجهه با محیط‌های جدید اطلاعاتی) از فراوانی قابل توجهی برخوردار است.



نمودار ۴. فراوانی مشارکت محققان کشورهای مختلف

تحلیل موضوعی مطالعات مربوط به نحو و معنا در بازیابی اطلاعات

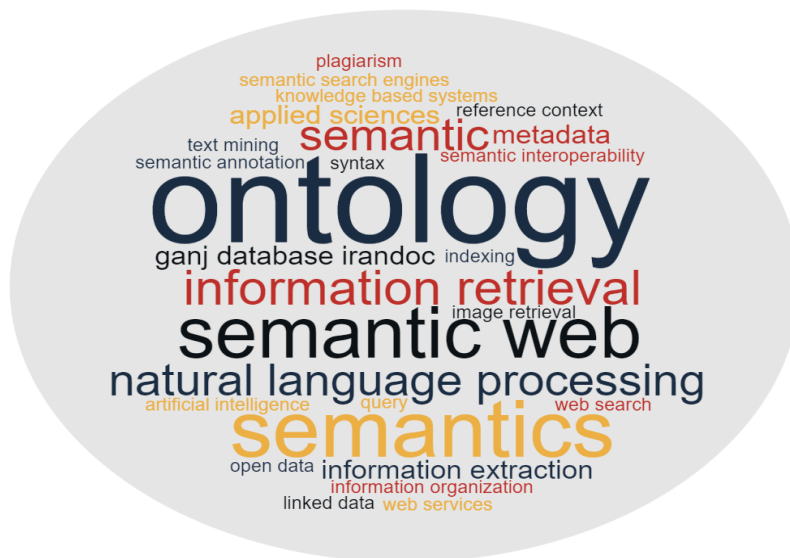
بررسی و تحلیل مطالعات نشان می‌دهد نحو و معنا در حوزه‌های مختلف و متعدد مورد توجه بوده‌اند. بررسی و تحلیل موضوعی کلیدواژه‌های مطالعات، این پراکندگی موضوعی را بهتر نمایان می‌کند. شکل ۱ فراوانی کلیدواژه‌های مطالعات مورد بررسی را نشان می‌دهد. از مجموع کلیدواژه‌های مقالات، ۹۸ کلیدواژه با فراوانی بیش از یک مرتبه و بین ۲ تا ۱۴ مرتبه تکرار شده‌اند. بیشترین فراوانی متعلق است به هستی‌شناسی/ها با ۱۴ تکرار، بازیابی اطلاعات با ۱۱ تکرار، معنا^۱ با ۹ تکرار، وب معنایی با ۸ تکرار، پردازش زبان طبیعی با ۵ تکرار و علوم کاربردی، پایگاه گنج^۲، استخراج اطلاعات و فراداده هر کدام با ۳ تکرار. در همین راستا شاهد پراکندگی موضوعی قابل توجهی در مطالعات مربوط به نحو و معنا هستیم و در تعداد قابل توجهی از منابع مورد بررسی، پژوهشگران از زوایای مختلف به آن پرداخته‌اند. از این رو دسته‌بندی اهداف و محورهای موضوعی بر مبنای جنبه‌هایی است که پژوهش‌ها تمرکز بیشتری بر روی آن‌ها در مطالعات مربوط به نحو و معنا داشته‌اند.

^۱. Semantic/semantics

^۲. در بین واژگان عبارت (ganj database irandoc) هم به چشم می‌خورد که مربوط به مطالعاتی است که از این پایگاه به‌عنوان بستر جهت جستجو و بازیابی اطلاعات استفاده کرده‌اند.

براساس آنچه ذکر شد، می‌توان مطالعات را در سه حوزه موضوعی اصلی ذیل تقسیم‌بندی نمود:

- نحو و معنا در ارتباط با سازمان‌دهی اطلاعات
- نحو و معنا در ارتباط با بازیابی اطلاعات
- نحو و معنا در ارتباط با جستجوی اطلاعات



شکل ۱. نقشه فراوانی موضوعات مبتنی بر کلیدواژه‌ها در منابع مورد بررسی^۱

بررسی فراوانی مطالعات در ارتباط با محورهای موضوعی نشان می‌دهد که بیشترین مطالعات به ترتیب در رابطه با نحو و معنا در بازیابی اطلاعات (۳۳ مورد) انجام شده است و پس از آن نحو و معنا در جستجوی اطلاعات (۱۴ مورد) و نحو و معنا در سازمان‌دهی اطلاعات (۷ مورد) بیشترین توجه پژوهشگران را به خود معطوف کرده‌اند. در ادامه به تحلیل مطالعات هریک از این محورهای موضوعی پرداخته می‌شود.

الف- نحو و معنا در سازمان‌دهی اطلاعات

^۱ نقشه فراوانی موضوعات مبتنی بر کلیدواژه‌ها با استفاده از بستر آنلاین <https://www.wordclouds.com>

ساخته شده است.

مطالعات این بخش که بین سال‌های ۱۹۹۷ تا ۲۰۱۸ انجام شده‌اند (جدول ۳)، به نقش نحو و معنا در سازمان‌دهی اطلاعات پرداخته‌اند. قدیمی‌ترین مطالعه موردبررسی نیز در این مقوله موضوعی جای دارد که در سال ۱۹۹۷ صورت گرفته است (۲). از جمله محورهای این مطالعات مباحث تئوری و استدلال منطقی و ریاضی در تحلیل متن در نظام‌های سازمان‌دهی اطلاعات (۱)، استفاده از عناصر معنایی از جمله نمایه‌سازی معنایی مفاهیم در تکمیل عناصر نحوی در سازمان‌دهی منابع (۲) و (۴) و (۶)، کاربرد هستی‌شناسی برای پوشش شکاف بین نحو و معنا و غنی‌سازی داده‌ها (۳) و (۵) و (۷) است.

جدول ۳. مطالعات مربوط به نحو و معنا در سازمان‌دهی اطلاعات

نویسنده	هدف	سال	قالب ^۱	کشور	روش	ابزار/فن
نصیری و خندان ۱	تحلیل دلالت‌های منطق ریاضی برتراند راسل ^۲ از جمله تحلیل متن و تحلیل بیان‌های مختلف زبان‌شناختی، از قبیل اسم خاص، طبقات موضوعی، روابط، واژگان منطقی، گزاره‌ها و جملات در طراحی نظام‌های سازمان‌دهی برای پوشش کاستی‌های زبان طبیعی در مقابل زبان‌های منطقی	۱۳۹۵	م	ایران	تحلیلی - کاربردی	متون / منابع کتابخانه‌ای

^۱. قالب انتشار منابع بررسی شده شامل مقاله مجله (م)، مقاله کنفرانسی (ک) و پایان‌نامه (پ) که در جدول با حروف اختصاری نشان داده شده است.

^۲. Bertrand Russell

ابزار/فن	روش	کشور	قالب ^۱	سال	هدف	نویسنده	
منطق توصیفی ^۲	کاربردی	فرانسه	ک	۱۹۹۷	غنی‌سازی معنایی اسناد با استفاده از اضافه نمودن عناصر معنایی به ساختار نحوی SGML	پولت و همکاران ^۱	۲
هستی‌شناسی	کاربردی - تحلیلی	ایالات متحده	ک	۲۰۰۹	استفاده از ارتباطات معنایی در هستی‌شناسی‌های مختلف برای پر کردن شکاف بین نحو و معنا و ناهمگونی ساختار نحو و معنا در داده‌های میان‌افزارهای مربوط به حسگرهای ساحلی	دوربها و همکاران ^۳	۳
نمایه‌سازی معنایی مفاهیم	کاربردی	بریتانیا	پ	۲۰۱۰	کاربرد فن نمایه‌سازی معنایی مفاهیم برای بازنمون مؤثر و خوشه‌بندی مجموعه‌ای از اسناد بدون ساختار یا ساختار یافته به منظور غلبه بر برخی مشکلات مانند عدم	بارسی ^۴	۴

^۱. Pouillet et al.
^۲. Description logic
^۳. Durbha et al.
^۴. Barresi, S.

ابزار/فن	روش	کشور	قالب ^۱	سال	هدف	نویسنده	
					توجه به مترادف‌ها و همچنین پوشش تمایز درجه اهمیت معنایی اصطلاحات		
هستی‌شناسی / تحلیل رسمی مفاهیم ^۲	تحلیلی	ژاپن	ک	۲۰۱۱	کاربرد معیارهای تشابه معنایی در هستی‌شناسی برای غلبه بر محدودیت‌های معیارهای تشخیص مشابهت نحوی	اکمل و باترس ^۱	۵
چارچوب توصیف منبع	کاربردی	تایوان	م	۲۰۱۵	رویکرد مبتنی بر چهارچوب توصیف منبع (RDF) برای تبدیل فراداده از نگاشت عناصر واژگانی معادل به نگاشت معنایی با روابط متنی مختلف	چن ^۳	۶
هستی‌شناسی با معماری OntoGenesis ^۵	کاربردی - تحلیلی	برزیل	م	۲۰۱۸	غنی‌سازی معنایی داده‌ها بر اساس هستی‌شناسی‌های دامنه برای حرکت از خدمات مبتنی بر داده‌ها براساس نحو	اولیویرا و همکاران ^۴	۷

۱. Akmal & Batres

۲. Formal Content Analysis

۳. Chen

۴. Oliveira et al.

۵. نوعی معماری برای غنی‌سازی خودکار معنایی داده‌ها

ابزار/فن	روش	کشور	قالب ^۱	سال	هدف	نویسنده
					به سمت خدمات مبتنی بر داده‌های معنایی	

ب- نحو و معنا در بازیابی اطلاعات

مطالعات مربوط به این محور موضوعی با احتساب مطالعات داخلی بین سال‌های ۲۰۰۲ تا ۲۰۲۲ صورت گرفته‌اند و بیشترین تعداد مطالعات موردبررسی را به خود اختصاص داده‌اند (جدول ۴). این پژوهش‌ها به رابطه نحو و معنا در بازیابی اطلاعات پرداخته‌اند. تمرکز مطالعات این بخش، بر محورهایی است که در ادامه به آن اشاره می‌شود. توجه به معنا در شیوه‌های مرسوم وزن‌دهی به واژگان (۱)، استفاده از رویکردهای مبتنی بر نحو و معنا و هستی‌شناسی در خلاصه‌سازی و شباهت‌یابی و بهبود بازیابی متون (۲)، (۴)، (۱۰)، (۱۹)، (۲۰)، (۳۱)، استفاده از فنون مبتنی بر معنا و هستی‌شناسی به منظور افزایش دقت بازیابی و بازیابی معنایی تصاویر (۳)، (۲۸)، توجه به نحو، بافت و معنا در بهبود ترجمه ماشینی (۵)، کاربرد بافتار معنایی و معنا برای غلبه بر دشواری‌های زبان‌ها از جمله هم‌نویسه‌ها، مترادفات و... در بازیابی اطلاعات (۶)، (۷)، (۹)، (۱۵) کاربرد هستی‌شناسی‌ها در بهبود بازیابی اطلاعات، انجام جستجوی معنایی و افزایش ربط نتایج (۸)، (۲۵)، (۳۳)، استفاده از نحو و معنا در پردازش زبان طبیعی و درک بهتر زبان طبیعی در بازیابی و مدیریت اطلاعات (۱۱)، (۱۲)، (۱۸)، کاربرد معنا به منظور بهبود سیستم پرسش و پاسخ و افزایش دقت بازیابی، افزایش تطبیق‌پذیری نتایج و رضایت کاربران (۱۳)، (۲۲)، (۲۳)، (۲۷)، استفاده از نحو و معنا در طراحی موتورهای کاوش (۱۴)، (۲۶)، ارزیابی نظام‌های بازیابی اطلاعات معنایی مبتنی بر ابزارهای معنایی از جمله هستی‌شناسی (۲۱)، کاربرد هستی‌شناسی و روابط معنایی برای گسترش پرس‌وجو کاربر در نظام‌های بازیابی (۱۶)، به کارگیری ویژگی‌های نحوی و معنایی برای تشخیص سرقت ادبی (۱۷)، (۲۹)، استفاده از حاشیه‌نویسی نحوی و معنایی

برای پوشش شکاف معنایی در بازیابی اطلاعات (۳۰)، به کارگیری نحو برای استنتاج روابط معنایی و ابهام‌زدایی از اطلاعات (۳۲).

جدول ۴. مطالعات مربوط به نحو و معنا در بازیابی اطلاعات

ابزار	روش	کشور	قالب	سال	هدف	نویسنده	
روش وزن دهی آماری به نام TF- IDF	تجربی	ایران	پ	۱۳۹۰	روش جدید با رویکرد توجه به ویژگی‌های اصلی زبان برای وزن دهی به کلمات که با مبنا قرار دادن روش وزن دهی به نام TF-IDF و تغییر (از دو جنبه معنا با استفاده از تئوری زبانی نظریه مرکزیت و گرامر) پارامتر TF که یک پارامتر اندازه‌گیری فرکانس در سطح یک متن است، می‌پردازد.	کامیار	۱
فرهنگ واژگان/ تحلیل معانی پنهان ^۱	تجربی	ایران	پ	۱۳۹۱	پیشنهاد یک روش جدید برای خلاصه‌سازی خودکار متون فارسی با رویکردی مبتنی بر معنا	محمدیان جدول قدم	۲

^۱. Latent semantic analysis

ابزار	روش	کشور	قالب	سال	هدف	نویسنده	
اصطلاح‌نامه ریختاری	مروری	ایران	م	۱۳۹۱	ارائه اصطلاح‌نامه ریختاری برای پر کردن شکاف‌های معنی شناختی با ترکیب فنون مبتنی بر متن و فنون مبتنی بر محتوا و نیز نزدیک کردن دیدگاه‌های نمایه‌سازان و متخصصان رایانه، برای بهبود دقت بازیابی تصاویر	صدوقی و همکاران	۳
رویکرد مبتنی بر معنا و نحو	تجربی	ایران	پ	۱۳۹۲	پیشنهاد یک روش شباهت‌یابی جملات، طراحی‌شده برای زبان فارسی بر پایه معنا و نحو برای بهبود دقت بازیابی نظام‌های خلاصه‌ساز چند سنده	نوجوان آقدرق	۴
فرهنگ واژگان/پرسشنامه	تجربی	ایران	م	۱۳۹۴	ابهام‌زدایی واژگانی صفات چندمعنایی در ترجمه ماشینی با تأکید و توجه به بافت و معنا	ذوالفقار کندری و موسوی میانگاه	۵
مجموعه آزمایشی مدارک	تحلیل محتوا	ایران	پ	۱۳۹۴	مقایسه قدرت پنج نوع بافتار (زمینه معنایی) (context)	هوشیار	۶

ابزار	روش	کشور	قالب	سال	هدف	نویسنده	
					متنی شامل بافتار استناد متنی، ارجاع، عنوان ارجاع، عنوان و متن مقاله در ابهام‌زدایی معنایی از هم‌نویسه‌ها		
بررسی منابع	مروری - تحلیلی	ایران	م	۱۳۹۸	نقش و اهمیت زبان با تأکید بر زبان مادری در شکل‌گیری واژگان جست‌وجو در بازیابی اطلاعات مبتنی بر معانی ذهنی افراد، بافت و محتوای مورد جست‌وجو	نوروزی	۷
قواعد نظام معناشناسی رودلف کارناب و هستی‌شناسی	فلسفی - تحلیل مفهومی	ایران	م	۱۴۰۰	شناسایی و تبیین نظریه‌های معنایی و نحوی رودلف کارناب ^۱ و اهمیت آن در توسعه نظری و عملی هستی‌شناسی‌ها جهت بازیابی مؤثر اطلاعات	نصیری و همکاران	۸
سیاهه واریسی	کاربردی -تحلیل محتوا	ایران	م	۱۴۰۱	بررسی تأثیر مشکلات خط فارسی در ابعاد صرفی، معنایی و رسم‌الخط بر بازیابی مدارک	آخشیک و همکاران	۹

1. Rudolf Carnap

نویسنده	هدف	سال	قالب	کشور	روش	ابزار
۱۰ گامالو و همکاران ^۱	استفاده از اطلاعات نحوی برای استخراج اطلاعات معنایی در پیکره متنی زبان پرتغالی	۲۰۰۲	م	پرتغال	تجربی	پیکره متنی زبان پرتغالی
۱۱ هومندا ^۲	استفاده از نحو و معنا در پردازش زبان طبیعی	۲۰۰۴	ک	لهستان	تجربی	پارادایم محاسبات / پردازش دانه‌ای یا گرانول
۱۲ مینکر ^۳	استفاده از اطلاعات نحوی به منظور ابهام‌زدایی و بهبود رمزگشایی و درک بهتر زبان طبیعی در بازیابی اطلاعات (مورد مطالعه: اطلاعات حمل و نقل با قطار/ریلی)	۲۰۰۴	م	آلمان	مطالعه موردی	
۱۳ افقی و همکاران	کاربرد معنا به منظور بهبود سیستم پرسش و پاسخ و افزایش دقت بازیابی در نظام‌های بازیابی اطلاعات	۲۰۰۶	ک	استرالیا	تجربی	فریم نت ^۴ و آکواپنت ^۵

۱. Gamallo et al.

۲. Homenda, W.

۳. Minker, W.

۴. پیکره FrameNet یک پایگاه داده واژگانی انگلیسی است که هم توسط انسان و هم توسط ماشین قابل خواندن

است، بر اساس مثال‌هایی از نحوه استفاده از کلمات در متون واقعی.

۵. پیکره AQUAINT متشکل از داده‌های متنی خبری به زبان انگلیسی است.

ابزار	روش	کشور	قالب	سال	هدف	نویسنده	
پایگاه داده علائق کاربر ^۲ / یادگیری ماشین	کاربردی	چین	ک	۲۰۰۶	پیشنهاد مدل موتور کاوش تعاملی هوشمند با استفاده از اطلاعات معنایی موجود در داده‌های جستجوی کاربر	منگ و هانگ ^۱	۱۴
موتورهای جستجو	کاربردی	ایالات متحده	پ	۲۰۰۷	بررسی مشکل نحو و معنای ترکیبات اسمی با استفاده از الگوریتم‌های جدید، با استفاده از وب به‌عنوان یک مجموعه و موتورهای جستجو به‌عنوان رابط برای آن مجموعه	ناکوو ^۳	۱۵
هستی‌شناسی	کاربردی	ویتنام	م	۲۰۰۸	کاربرد هستی‌شناسی و روابط معنایی بین کلمات برای گسترش پرس‌وجوی کاربر	نگوین و فن ^۴	۱۶
موتور کاوش	تجربی	تایوان	ک	۲۰۱۲	استفاده از ویژگی‌های نحوی و معنایی برای تشخیص سرتقت ادبی	لین و همکاران ^۵	۱۷

^۱. Meng & Huang

^۲. user information preference c

^۳. Nakov, P. I.

^۴. Nguyen & Phan

^۵. Lin et al.

ابزار	روش	کشور	قالب	سال	هدف	نویسنده	
زبان نشانه‌گذاری گسترش پذیر (ایکس‌ام‌ال)، چهارچوب توصیف منبع (آردی‌اف و آردی‌اف‌ای)، زبان هستی‌شناسی وب (ادبیوال) ^۲	تجربی	اسپانیا	م	۲۰۱۲	ارائه ابزار کاربرمحور حاشیه‌نویسی معنایی محتوای وب به زبان طبیعی برای توسعه یک سیستم معنایی مدیریت اطلاعات وب (WCMS (semantic web content management systems)	ناوارو - گالیندو و ساموس ^۱	۱۸
هستی‌شناسی	تحلیلی	ایتالیا	م	۲۰۱۲	محیط خلاصه / چکیده‌نویسی معنایی مبتنی بر هستی‌شناسی برای منابع چندرسانه‌ای قضایی	فرسینی و سارتور ^۳	۱۹
برچسب‌گذاری	کاربردی - تحلیلی	صربستان	م	۲۰۱۳	استفاده از اطلاعات نحوی برای تعیین شباهت معنایی متون	باتانویچ و بوجیچ ^۴	۲۰

¹. Navarro-Galindo & Samos

². XML, RDF, RDFa, OWL

³. Fersini & Sartori

⁴. Batanović & Bojić

ابزار	روش	کشور	قالب	سال	هدف	نویسنده	
نظام‌های بازیابی معنایی (موتور جستجوی هستی‌شناسی، موتورهای جستجو برای ايرادده یک موتور جستجو، دایرکتوری‌ها و کاتالوگ‌های هستی‌شناسی/ دایرکتوری‌ها و کاتالوگ‌های فرا داده)	تحلیلی مقایسه‌ای	اسپانیا	م	۲۰۱۳	ارزیابی انواع مختلف نظام‌های بازیابی وب معنایی با توجه به توانایی آن‌ها در مدیریت و بازیابی اسناد معنایی و ارائه معیارهایی مهم و لازم در توسعه نظام‌های بازیابی اطلاعات معنایی	موراتو و همکاران ^۱	۲۱
استفاده از معیارهای ارتباط معنایی، فنون پردازش زبان طبیعی و مدل‌های آماري	تجربی	هند	م	۲۰۱۵	استفاده از نحو و معنا در بهبود بازیابی و افزایش تطبیق‌پذیری نتایج برای کاربران	شارما و همکاران ^۲	۲۲
(واسپارگاه Dspace)	کاربردی مطالعه – موردی	یونان	م	۲۰۱۵	ارائه روش ارزیابی برای جستجوی (پرسش و پاسخ) معنایی در مخازن دیجیتال	سولومو و کوتسومترو پولوس ^۳	۲۳

^۱. Morato et al.

^۲. Sharma et al.

^۳. Solomou & Koutsomitropoulos

ابزار	روش	کشور	قالب	سال	هدف	نویسنده	
هستی‌شناسی و پایگاه دانش	تحلیلی	ویتنام	م	۲۰۱۵	نشانه‌گذاری خودکار معنایی اخبار فوتبال برای بهبود دقت در بازیابی آن‌ها نسبت به بازیابی محدود به نحو	نگوین و کائو ^۱	۲۴
هستی‌شناسی	ارزیابی تجربی	ویتنام	م	۲۰۱۵	کاربرد هستی‌شناسی برای غنای معنایی جستجوی سستی براساس کلیدواژه در مرحله ایجاد استراتژی جستجو و رتبه‌بندی اطلاعات مرتبط	نگوین و همکاران ^۲	۲۵
هستی‌شناسی دامنه	کاربردی	عمان	م	۲۰۱۶	ارائه یک موتور جستجوی معنایی عربی کارآمد و مقیاس‌پذیر بر اساس یک نمودار هستی‌شناختی دامنه (متناسب با ریخت‌شناسی و دستور زبان مربوط به زبان عربی)	سید و المقریشی ^۳	۲۶

^۱. Nguyen & Cao

^۲. Nguyen et al.

^۳. Sayed & Al Muqrishi

ابزار	روش	کشور	قالب	سال	هدف	نویسنده	
زبان پرس‌وجوی نکسی ^۲	کاربردی - تحلیلی	مالزی	م	۲۰۱۷	یک مدل پرس‌وجوی معنایی - نحوی برای ساخت انواع مختلف پرس‌وجوهای ساخت‌یافته در بازیابی در محیط‌های مبتنی بر xml	جن و فنگ ^۱	۲۷
هستی‌شناسی	تحلیلی	مالزی	م	۲۰۱۷	کاهش فاصله معنایی با ارائه مدلی برای یکپارچه‌سازی شاخص‌های ویژگی‌های متنی و بصری از طریق هستی‌شناسی برای بهبود بازیابی معنایی تصاویر	ام‌کی و محد نوح ^۳	۲۸
پایگاه داده	تجربی	هند	م	۲۰۱۸	استفاده از ویژگی‌های نحوی و معنایی برای تشخیص سرقت ادبی	وانی و گوپتا ^۴	۲۹
واژگان کنترل‌شده / نمونه واسپارگاه‌های وب	کاربردی	بریتانیا	پ	۲۰۱۸	غنی و پایدار کردن ارزش محتوا در واسپارگاه‌های خودکار با حاشیه‌نویسی محتوا توسط متخصصان	خان ^۵	۳۰

^۱. Gan & Phang

^۲. an information retrieval (IR) query language for searching structured and semi-structured document collections

^۳. MK & Mohd Noah

^۴. Vani & Gupta

^۵. Khan, A. A.

ابزار	روش	کشور	قالب	سال	هدف	نویسنده	
					موضوعی علاوه بر حاشیه‌نویسی معنایی خودکار برای پر کردن شکاف بین نحو و معنا		
مجموعه واژگان SemEval-2012 task برای ارزیابی نظام‌های معنایی	تجربی	چین	ک	۲۰۱۹	تحلیل شباهت معنایی متون با استفاده از درخت نحوی جملات	تانگ و همکاران ^۱	۳۱
هستی‌شناسی	کاربردی	ایالات متحده	ک	۲۰۲۰	یک روش مبتنی بر نحو برای استخراج روابط معنایی از یک نوع اطلاعات و ابهام‌زدایی از خط‌مشی محرمانگی در اپلیکیشن‌ها	بکایی حسینی و همکاران ^۲	۳۲
هستی‌شناسی	کاربردی	تایوان	م	۲۰۲۱	توسعه یک مدل هستی‌شناسی برای کمک به کاربران برای جستجوی معنایی مجموعه‌های فرهنگی بر اساس مفاهیم موضوعی	سانگ و چی ^۳	۳۳

ج- نحو و معنا در جستجوی اطلاعات

- ^۱. Tang et al.
- ^۲. Bokaei Hoseini et al.
- ^۳. Sung & Chi

مطالعات مربوط به این محور موضوعی با احتساب مطالعات داخلی بین سال‌های ۲۰۰۵ تا ۲۰۲۲ صورت گرفته و به رابطه نحو و معنا در جستجوی اطلاعات پرداخته‌اند. در ادامه به موضوعات مورد تمرکز این مطالعات پرداخته می‌شود. استفاده از نظام‌های معنایی در کنار ابزارهای پردازش زبان طبیعی برای بهبود ترجمه در جستجوی اطلاعات بین‌زبانی (۱)، تحلیل مؤلفه‌های معنایی جستجوی کاربران به‌منظور بررسی رفتار اطلاعاتی آن‌ها (۲)، بهبود پرس‌وجوهای کلیدواژه‌ای در فرایند جستجوی اطلاعات با استفاده از روش‌های مبتنی بر معنا و هستی‌شناسی در کنار اطلاعات نحوی (۳)، (۴)، (۸)، (۹)، (۱۰) کاربرد نحو و معنا در توسعه سیستم‌های پرس‌وجو به زبان طبیعی و پردازش زبان طبیعی در سامانه‌های مختلف (۵)، استفاده از معنا و فناوری‌های وب معنایی از جمله هستی‌شناسی در بهبود الگوریتم‌های جستجو و تطبیق معنایی و ارتقا بر داده‌ها و جستجوی شبکه‌ای (۶)، (۷)، مقایسه موتورهای کاوش نحوی و معنایی (۱۱)، (۱۲)، به‌کارگیری روش‌های یادگیری عمیق و شبکه عصبی برای تحلیل نحوی زبان و پردازش زبان طبیعی (۱۳)، کاربرد استنتاج معنایی در کنار نحوی به‌منظور پوشش شکاف معنایی در ابزارهای برنامه‌نویسی (۱۴).

جدول ۵. مطالعات مربوط به نحو و معنا در جستجوی اطلاعات

ابزار	روش	کشور	قالب	سال	هدف	نویسنده	
ابزارهای پردازش زبان طبیعی و شبکه معنایی	تجربی	ایران	پ	۱۳۹۰	بررسی اثر ترکیب ابزارهای پردازش زبان طبیعی و شبکه معنایی در بهبود عملیات رفع ابهام ترجمه در سیستم جستجوی اطلاعات (بین‌زبانی) فارسی-انگلیسی	ابوترابی، گودرزی	۱

نویسنده	هدف	سال	قالب	کشور	روش	ابزار
۲ کریمی و همکاران	تحلیل لاگ کاربران پایگاه اطلاعات گنج (از نظر ویژگی‌های معنایی جستجو و اصلاح جستجو در میان کاربران) به منظور بررسی رفتار اطلاع‌یابی کاربران	۱۳۹۷	م	ایران	تحلیلی	سیاهه‌وارسی / پایگاه گنج
۳ استوچانوویچ ^۱	بهبود پالایش پرس‌وجو در روش‌های نحوی با استفاده از فناوری‌های معنایی و استخراج خودکار مفاهیم از پرس‌وجوی متنی	۲۰۰۵	م	آلمان	تجربی	هستی‌شناسی
۴ تریلو و همکاران ^۲	استفاده از هستی‌شناسی برای اضافه کردن معانی به کلیدواژه‌های جستجو به منظور بهبود جستجوی مبتنی بر نحو واژگان	۲۰۰۷	م	اسپانیا	تجربی	هستی‌شناسی
۵ دانگ و تویین ^۳	سیستم پرسش و پاسخ برای سیستم جستجوی کتاب‌های الکترونیکی در کتابخانه با تمرکز بر رویکردهای	۲۰۰۹	م	ویتنام	تجربی	پردازش زبان طبیعی

^۱. Stojanovic, N.

^۲. Trilloe et al.

^۳. Dang & Tuyen

ابزار	روش	کشور	قالب	سال	هدف	نویسنده	
					تجزیه و تحلیل و نمایش نحو، نمایش معنایی و تبدیل از ساختار نحوی به ساختار معنایی		
	تجربی	چین	ک	۲۰۱۱	ارائه الگوی جدید الگوریتم تطبیق معنایی بر اساس فاصله معنایی برای رفع نقایص الگوریتم‌های صرفاً مبتنی بر نحو در سامانه‌های یودی‌دی‌آی ^۲	ژو و همکاران ^۱	۶
هستی‌شناسی	کاربردی	ایالات متحده	پ	۲۰۱۱	بهبود ماهیت ابرداده‌های علمی فعلی و جستجوی منابع شبکه‌ای با ارائه یک چهارچوب هستی‌شناختی با استفاده از فناوری‌های وب معنایی با توانایی تشخیص مطابقت‌های رشته‌ای رایج و به‌علاوه معادل‌های اصطلاحی و سایر روابط معنایی	گوپتا ^۳	۷

^۱. Xu et al.

^۲. Universal Description, Discovery, and Integration (UDDI) specification defines a way to publish and discover information about Web services.

^۳. Gupta, C.

نویسنده	هدف	سال	قالب	کشور	روش	ابزار
۸ روی و زنگ ^۱	ارائه روش پردازش زبان طبیعی برای اضافه کردن معانی در تولید و ایجاد پرس‌وجوهای مبتنی بر علوم شناختی و وب معنایی	۲۰۱۳	م	ایالات متحده	تجربی	پردازش زبان طبیعی
۹ رامادا و همکاران ^۲	پردازش معنایی پرس‌وجوهای (علاوه بر نحو) کاربران برای افزایش دقت و ربط بازیابی از پایگاه‌های داده رابطه‌ای وب	۲۰۱۴	ک	برزیل	کاربردی	پایگاه داده رابطه‌ای
۱۰ یوسیتلی و همکاران ^۳	کاربرد هستی‌شناسی جستجو در امکان ایجاد پرس‌وجوهای قوی و پیچیده معنایی	۲۰۱۴	م	آلمان	کاربردی	هستی‌شناسی
۱۱ هوسان ^۴	انجام مقایسه جامع بین موتورهای جستجوی نحوی و موتورهای جستجوی معنایی	۲۰۲۰	م	عراق	پیمایش مقایسه‌ای	موتورهای جستجوی نحوی و معنایی
۱۲ فیو ^۵	استفاده از الگوریتم جستجوی مبتنی بر مفهوم بر اساس تحلیل رسمی مفاهیم	۲۰۲۱	م	میانمار	تحلیلی	پیکره وردنت/ ویکی‌پدیا/ تحلیل رسمی مفاهیم

^۱. Roy & Zeng

^۲. Ramada et al.

^۳. Uciteli et al.

^۴. Hussan, B. K.

^۵. Phyoo, S. S.

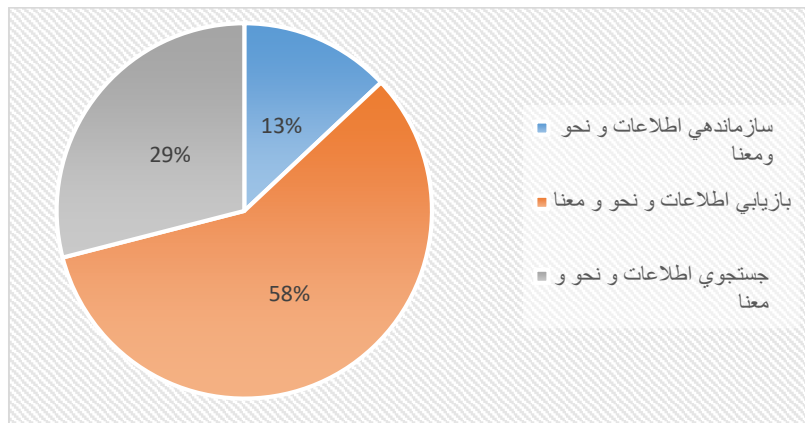
نویسنده	هدف	سال	قالب	کشور	روش	ابزار
	(FCA) منابع دانش خارجی با ایجاد سلسله‌مراتب مفاهیم برای دستیابی به نتایج جستجوی معنادار دقیق و کارآمد (مدیریت پیشینه‌ها) در بازیابی اطلاعات در حوزه مقالات علمی					
۱۳	روش‌های یادگیری عمیق و شبکه عصبی برای تحلیل نحوی زبان طبیعی (پردازش زبان طبیعی)	۲۰۲۲	م	چین	تحلیلی	پردازش زبان طبیعی / یادگیری ماشین
۱۴	ارتقای استنتاج گرامر به استنتاج معنایی زبان‌های دامنه خاص با استفاده از میانجی معنایی، به منظور تسهیل زبان خاص دامنه برای کارشناسان موضوعی بدون دانش برنامه‌نویسی	۲۰۲۲	پ	اسلونی	کاربردی	ابزار تولیدکننده کامپایلر/مفسر ^۳

^۱. Zhou, Y.

^۲. Kovačević, Ž.

^۳. compiler/interpreter generator tool

همان‌طور که اطلاعات جدول‌های ۳ تا ۵ نشان می‌دهد، بیشترین فعالیت پژوهشی در محور موضوعی نحو و معنا، در ارتباط با بازیابی اطلاعات انجام شده است که در نمودار ۵ نیز قابل مشاهده است.



نمودار ۵. فراوانی محورهای موضوعی

نمودار ۵ حاکی از آن است که تمرکز ۵۸٪ مطالعات در حوزه بازیابی اطلاعات و نحو و معنا بوده و ۲۹٪ مطالعات را حوزه جستجوی اطلاعات و نحو و معنا تشکیل داده است و تنها ۱۳٪ پژوهشگران به جنبه سازمان‌دهی اطلاعات و نحو و معنا پرداخته‌اند. بنابراین میزان مطالعات در حوزه سازمان‌دهی اطلاعات از نظر کمی در سطح پایینی ارزیابی می‌شود.

۴. روش‌ها و رویکردهایی پژوهشگران در پژوهش‌های حوزه نحو و معنا در بازیابی اطلاعات

از نظر روش‌ها و ابزارهای مورداستفاده پژوهشگران، طیف متنوعی از روش‌ها و ابزارها در مطالعات مرتبط با حوزه نحو و معنا به کار رفته است (جدول ۳-۵). بر اساس بررسی‌های به عمل آمده، از مجموع ۵۴ پژوهش، بیشتر آن‌ها از نوع تجربی (۱۹ مورد) و کاربردی (۱۵ مورد) هستند و مطالعات تحلیلی (۹)، کاربردی - تحلیلی (۶ مورد)، تحلیل محتوا و مقایسه‌ای (هر یک با ۲ مورد) مطالعه موردی (۱ مورد) از لحاظ فراوانی در رتبه‌های بعدی قرار دارند که به تناسب از ابزارهایی نظیر هستی‌شناسی، موتورهای کاوش و فونونی مثل

پردازش زبان طبیعی، حاشیه‌نویسی، برجسب‌گذاری و نمایه‌سازی استفاده کرده‌اند که در جدول‌های ۳ تا ۵ نشان داده شده است.

۵. شکاف‌های پژوهشی حوزه نحو و معنا در بازیابی اطلاعات

با توجه به پیشرفت‌های قابل‌توجه در زمینه وب معنایی، هوش مصنوعی و فناوری‌ها و ابزارهای وابسته به آن‌ها که هر یک به نوعی در ارتباط با بازیابی اطلاعات نیز هستند، این مقوله به جزء لاینفک بسیاری از حوزه‌های ذکر شده تبدیل شده است. تنوع موضوعات و حوزه‌ها در مطالعات مورد بررسی هم حاکی از این مورد است. از طرفی تمرکز مطالعات چه در داخل و چه در خارج از کشور، بیشتر در حوزه بازیابی و بعد از جستجوی اطلاعات بوده و موضوع سازمان‌دهی اطلاعات نیازمند پژوهش‌های بیشتر است (نمودار ۵). چراکه نظام‌های سازمان‌دهی اطلاعات به‌عنوان مبنا و به‌نوعی نقطه آغاز فرایند جستجو و بازیابی اطلاعات نقش مهمی در این زمینه دارند. اگرچه در ابعاد فنی نحو و معنا در بازیابی اطلاعات در داخل کشور مطالعاتی به‌صورت پیمایش یا به‌صورت تجربی انجام شده، اما با توجه به مشکلات کاربران فارسی‌زبان و ویژگی‌های نحوی و معنایی زبان فارسی، لازم است پژوهشگران رشته‌های مختلف از جمله زبان‌شناسی، علم اطلاعات، رایانه، هوش مصنوعی و پردازش زبان طبیعی توجه ویژه‌ای به ابعاد یادشده داشته باشند. همچنین مباحث زبانی خصوصاً مباحث نظری مربوط به نحو و معنا در ارتباط با زبان فارسی به‌جز یک مطالعه (نوروزی، ۱۳۹۸) مغفول مانده که با توجه به تعداد روزافزون کاربران فارسی‌زبان وب و سامانه‌های بازیابی اطلاعات نیازمند توجه بیشتر است. هرچند پژوهش‌هایی در مسائل نحوی و معنایی زبان فارسی در بحث بازیابی صورت گرفته است؛ اما فقدان مطالعات متمرکز بر ارائه مبناهای علمی و عملی برای مدنظر داشتن ویژگی‌های نحوی و معنایی این زبان (در سامانه‌های سازمان‌دهی، جستجو و بازیابی اطلاعات) احساس می‌شود.

بحث و نتیجه‌گیری

در مقاله حاضر تجزیه و تحلیل پژوهش‌های مربوط به نحو و معنا با استفاده از رویکرد مرور نظام‌مند مورد تأکید و بررسی قرار گرفت. در این بررسی، شناسایی خلأها و اولویت‌های

پژوهشی در زمینه‌های علمی طی دهه‌های گذشته (از دهه ۱۹۹۰ میلادی تا اکنون)، فرصتی را برای متخصصین و علاقه‌مندان ایجاد می‌کند که با تحلیل و بررسی‌های صورت گرفته، به موضوع نحو و معنا اشراف بیشتر و کامل‌تری پیدا کنند.

بررسی پژوهش‌ها در حوزه نحو و معنا در این پژوهش نشان داد که در مجموع ۵۴ عنوان پژوهش در داخل و خارج از کشور انجام گرفته است که نتایج آن حاکی از سهم قابل توجه پژوهشگران ایرانی در این حوزه است. علاوه بر این، بر اساس آنچه در یافته‌های پژوهش ملاحظه شد بررسی پژوهش‌ها نشان داد که از نظر زمانی، مطالعات در بازه زمانی ۲۶ ساله صورت گرفته‌اند به طوری که قدیمی‌ترین مطالعه مورد بررسی مربوط به سال ۱۹۹۷ و به‌روزترین آن مربوط به سال ۲۰۲۲ است، اما در ایران اولین مطالعه در این حوزه در سال ۲۰۱۱ و با فاصله زمانی قابل توجهی صورت گرفته است (جداول ۳-۵). این امر نشان‌دهنده پویایی حوزه مورد بررسی و تأثیرپذیری و تأثیرگذاری متقابل آن در ارتباط با فناوری‌های وب معنایی دارد.

همچنین تحلیل موضوعی پژوهش‌های مورد بررسی بر اساس کلیدواژه‌های مطالعات، حاکی از این است که «هستی‌شناسی»، به‌عنوان یکی از ابزارهای وب معنایی پیوند تنگاتنگی با ابعاد معنایی و نحوی زبان در بازیابی اطلاعات دارد. از سوی دیگر کاربرد هستی‌شناسی‌ها به‌عنوان مبنای بسیاری از سامانه‌های بازیابی معنایی اطلاعات، همسو با پیشرفت‌های فناوریانه در وب معنایی در حوزه‌های گوناگون دانش، مورد توجه روزافزون قرار گرفته است. در بررسی‌های انجام گرفته، مبحث بازیابی اطلاعات در مقایسه با سازمان‌دهی و جستجوی اطلاعات، بیشتر در کانون توجه است (نمودار ۵). بیشترین روش‌ها و رویکردهای مورد استفاده در پژوهش‌های بررسی شده، از نوع تجربی و کاربردی هستند و انواع مطالعات تحلیلی و مقایسه‌ای از لحاظ فراوانی در رتبه‌های بعدی قرار دارند و به‌تناسب، ابزارهایی نظیر هستی‌شناسی، موتورهای کاوش و فونونی مثل پردازش زبان طبیعی، حاشیه‌نویسی و برجسب‌گذاری و نمایه‌سازی در آن‌ها بکار رفته است (جداول ۳-۵).




نکته قابل توجه در پژوهش‌های مورد بررسی، استفاده پژوهشگران داخلی از پایگاه گنج (مربوط به پایان‌نامه‌ها) پژوهشگاه علوم و فناوری اطلاعات ایران به‌عنوان بستر جهت

۱. بازه زمانی مذکور شده منطبق بر تاریخ انتشار منابع جستجو شده و تحلیل شده در این پژوهش است.

بررسی و پیمایش ابعاد نحوی و معنایی زبان فارسی در ارتباط با بازیابی اطلاعات است. همچنین پراکندگی مطالعات در رشته‌های مختلف دانش نشان‌دهنده پیوند ناگسستنی نحو و معنا و مسائل زبانی با تقریباً هر حیطه‌ای است که فعالیت ذینفعان آن در ارتباط با سازمان‌دهی، ذخیره و بازیابی اطلاعات است. از جمله این حوزه‌ها می‌توان به مطالعه نحو و معنا در ارتباط با حسگرهای محیطی، نمایه‌سازی منابع، مشابهت‌یابی و چکیده‌سازی متون، تشخیص سرقت ادبی، پردازش زبان طبیعی، واسپارگاه‌ها، بهبود پرس‌وجوی کاربران و بازیابی اطلاعات در واسپارگاه‌ها و موتورهای کاوش و دیگر سامانه‌ها، غنی‌سازی فراداده‌ها و بازیابی تصاویر اشاره کرد (جداول ۳-۵). از نظر زبانی نیز زبان‌های غیر انگلیسی از جمله فارسی، عربی و پرتهالی از مواردی هستند که پژوهشگران، نحو و معنا آن‌ها را در ارتباط با بازیابی اطلاعات در بافت‌های مختلف مدنظر داشته‌اند. به نظر می‌رسد بحث بررسی نحو و معنا در ارتباط با بازیابی اطلاعات در زبان‌های مختلف، سهم بسزایی در شکل‌گیری ادبیات پژوهشی آینده این حوزه نیز دارد، چراکه زبان مادری نقش محوری در شکل‌گیری واژگان جستجو در بازیابی اطلاعات مبتنی بر معانی ذهنی افراد، بافت و محتوای مورد جستجو دارد که توجه به آن می‌تواند در بهبود سامانه‌های بازیابی اطلاعات نقش مؤثری داشته باشد.

در مجموع تحلیل مطالعات منتخب نشان می‌دهد پرداختن به بحث نحو و معنا در ارتباط با سازمان‌دهی اطلاعات و همچنین تبیین، واکاوی و تقویت وجوه نظری نحو و معنا در بازیابی اطلاعات از مواردی است که نیاز به پژوهش بیشتر، خصوصاً مطالعات بین‌رشته‌ای بیشتر از سوی پژوهشگران حوزه‌های زبان‌شناسی، علم اطلاعات و علوم رایانه - به‌عنوان حوزه‌هایی از دانش که ذخیره، سازمان‌دهی و بازیابی اطلاعات از دغدغه‌های اصلی آن‌ها است - دارد. نتایج پژوهش حاضر می‌تواند رهنمودهایی برای متخصصان حوزه‌های مختلفی داشته باشد که به‌نوعی با ذخیره و بازیابی اطلاعات سروکار دارند. علاوه بر این محورهای شناسایی‌شده و روش‌ها و رویکردهای به‌کاررفته و شکاف‌های پژوهشی شناسایی‌شده نیز می‌تواند راهگشای پژوهشگران در انجام مطالعات آتی باشد.

ORCID

Hoda Homavandi  <https://orcid.org/0000-0002-2722-3904>
Yaghub Norouzi  <https://orcid.org/0000-0003-3030-7647>
Bent-ol Hoda Khabbazan  <https://orcid.org/0009-0000-0297-8393>

منابع

- آخشیکی، سمیه سادات، نگهداری، خدیجه و امامی، علیرضا. (۱۴۰۱). خط فارسی در گنج: واکاوی تأثیر مسائل صرف، معنا و رسم‌الخط بر بازیابی اطلاعات در پایگاه اطلاعات علمی ایران. علوم و فنون مدیریت اطلاعات، ۸(۱)، ۱۹۳-۲۲۰.
- آویارد، هلن. (۲۰۰۷). چگونه یک تحقیق مروری انجام دهیم؟: راهنمای کاربردی نگارش مقاله و پایان‌نامه در علوم اجتماعی و پزشکی. ترجمه پوریا صرامی فروشانی و فردین علی پور گراوند (۱۳۹۰). تهران: جامعه شناسان.
- ابوترابی گودرزی، حامد. (۱۳۹۰). بررسی و امکان‌سنجی بهبود سیستم‌های جستجوی اطلاعات فارسی-انگلیسی [پایان‌نامه کارشناسی ارشد، دانشگاه شاهد، تهران].
- دری، راحله. (۱۳۹۳). مقایسه و ارزیابی موتورهای جستجوی معنایی. پژوهشنامه پردازش و مدیریت اطلاعات، ۳۰(۲)، ۴۶۷-۴۸۷.
- رضایی شریف‌آبادی، سعید، خسروی، عبدالرسول، حاجی زین‌العابدینی، محسن. (۱۳۸۶). امکان‌سنجی مهار مستندات موضوعی پزشکی در پایگاه‌های اطلاعاتی فارسی موجود بر روی وب. مطالعات تربیتی و روان‌شناسی، ۸(۳)، ۱۸۳-۲۰۱.
- رفیعی، امیررضا. (۱۳۹۴). استخراج روابط هم‌معنایی و تضاد [پایان‌نامه کارشناسی ارشد، مرکز زبان‌ها و زبان‌شناسی، دانشگاه صنعتی شریف، تهران].
- صدوقی، فرحناز، ولی‌نژادی، علی، حسن‌زاده، حافظ محمد، بورقی، حمید و پاسیار، پریسا. (۱۴۰۱). چالش‌های بازیابی معنایی تصاویر و کاربرست نوین اصطلاح‌نامه‌ها. پژوهشنامه پردازش و مدیریت اطلاعات، ۲۷(۳)، ۶۴۱-۶۶۶.
- قیومی، مسعود. (۱۳۹۸). تعیین خودکار معانی واژه‌های فارسی با استفاده از تعبیه معنایی واژه. پژوهشنامه پردازش و مدیریت اطلاعات، ۳۵(۱)، ۲۵-۵۰.
- کامیار، حسین. (۱۳۹۰). روش جدید وزن دهی معنایی به کلمات در کاربردهای پردازش متن [پایان‌نامه کارشناسی ارشد، دانشگاه فردوسی مشهد، مشهد].

کریمی، المیرا، بابایی، محمود و حسینی بهشتی، ملوک السادات. (۱۳۹۷). تحلیل رفتار اصلاح پرس و جوی کاربر براساس ویژگی‌های معنایی: تحلیل لاگ کاربران پایگاه اطلاعات گنج (ایرانداک). *تعامل انسان و اطلاعات*، ۵ (۳)، ۱-۱۴.

محمدیان جدول قدم، فراز. (۱۳۹۱). *خلاصه‌سازی خودکار متون فارسی به روش معنایی* [پایان‌نامه کارشناسی ارشد، دانشگاه اصفهان، اصفهان].

ذوالفقار کندری، زهره و موسوی میانگاه، طیبه. (۱۳۹۴). ابهام زدایی واژگانی صفات مبهم در ترجمه ماشینی: بررسی پیکره بنیاد. *پژوهشنامه پردازش و مدیریت اطلاعات*، ۳۰ (۳)، ۷۱۹-۷۳۵.

نصیری، طاهر و خندان، محمد. (۱۳۹۵). منطق ریاضی برتراند راسل و دلالت‌های آن در سازمان‌دهی اطلاعات و سازه‌های موضوعی آن. *مطالعات ملی کتابداری و سازماندهی اطلاعات*، ۲۷ (۱)، ۷-۲۴.

نصیری، طاهر، ریاحی نیا، نصرت، نشاط، نرگس، شقاقی، مهدی و رسولی پور، رسول. (۱۴۰۰). معناشناسی کارناپ و کاربست آن در بازیابی آنتولوژیک اطلاعات. *تعامل انسان و اطلاعات*، ۸ (۳)، ۱-۸.

نوجوان آقدرق، باقر. (۱۳۹۲). *طراحی سامانه‌ای برای شناسایی جملات مشابه برای استفاده در خلاصه‌سازی چند سنده* [پایان‌نامه کارشناسی ارشد، موسسه آموزش عالی نبی اکرم].
نوروزی، یعقوب. (۱۳۹۸). بافت و معنا در بازیابی اطلاعات: با تأکید بر زبان مادری. *پژوهشنامه پردازش و مدیریت اطلاعات*، ۳۵ (۱)، ۱-۲۴.

ویکری، برایان و ویکری، الینا. (۱۹۸۹). *علم اطلاع‌رسانی در نظر و عمل*. ترجمه عبدالحسین فرج پهلوی (۱۳۸۰). مشهد: دانشگاه فردوسی مشهد.

هوشیار، مژگان. (۱۳۹۴). *مقایسه قدرت انواع بافتار متن در ابهام‌زدایی معنایی از هم‌نویسه‌های انگلیسی* [پایان‌نامه کارشناسی ارشد، دانشگاه شیراز].

References

- Akmal, S., & Batres, R. (2011, December). Ontology-based semantic similarity measures for product-service system design. In *2011 IEEE/SICE International Symposium on System Integration (SII)* (pp. 932-937). IEEE
- Barresi, S. (2010). *Semantic based indexing technique for optimisation and intelligent document representation: application to structured and unstructured document clustering* [Doctoral dissertation, University of Salford].

- Batanović, V., & Bojić, D. (2014). Evaluation and classification of syntax usage in determining short-text semantic similarity. *Telfor Journal*, 6(1), 64-68.
- Bokaei Hosseini, M., Slavin, R., Breaux, T., Wang, X., & Niu, J. (2020). Disambiguating requirements through syntax-driven semantic analysis of information types. In *Requirements Engineering: Foundation for Software Quality: 26th International Working Conference, REFSQ 2020, Pisa, Italy, March 24–27, 2020, Proceedings 26* (pp. 97-115). Springer International Publishing.
- Chen, Y. N. (2015). A RDF-based approach to metadata crosswalk for semantic interoperability at the data element level. *Library hi tech*, 33(2), 175-194
- Dang, N., & Tuyen, D. T. T. (2009). Natural language question answering model applied to document retrieval system. *World Academy of Science, Engineering and Technology*, 51, 36-39.
- Durbha, S. S., King, R. L., Amanchi, S. K., Bheemireddy, S., & Younan, N. H. (2009, December). Information Services and Middleware for the Coastal Sensor Web. In *2009 IEEE International Conference on Data Mining Workshops*. 656-661. IEEE
- Fersini, E., & Sartori, F. (2012). Semantic storyboard of judicial debates: a novel multimedia summarization environment. *Program*, 46(2), 199-219.
- Gamallo, P., Gasperin, C., Agustini, A., Lopes, G. P., & de Lima, V. L. S. (2005). Using Syntax-based methods for extracting semantic information. *Linguistica Computazionale*, Pisa-Roma, 22, 201-229.
- Gan, K. H., & Phang, K. K. (2017). A semantic-syntax model for XML query construction. *International Journal of Web Information Systems*, 13(2), 155–172.
- Guha, R., McCool, R., & Miller, E. (2003, May). Semantic search. In *Proceedings of the 12th international conference on World Wide Web* (pp. 700-709).
- Gupta, C. (2011). Semantic search for grid resources and scientific metadata using an ontological framework. State University: Binghamton.
- Hammo, B.H. (2009). Towards enhancing retrieval effectiveness of search engines for diacritized Arabic documents. *Information Retrieval*, 12(3), 300 -323.
- Homenda, W. (2004). Information structuring in natural language communication: Syntax versus semantic. In *Artificial Intelligence and Soft Computing-ICAISC 2004: 7th International Conference*, Zakopane, Poland, June 7-11, 2004. Proceedings 7,872-878. Springer Berlin Heidelberg.
- Hussan, B. K. (2020). Comparative study of semantic and keyword based search engines. *Adv. Sci. Technol. Eng. Syst. J*, 5 (1), 106 -111.
- Khan, A. A. (2018). *Exploiting Linked Open Data (LoD) and Crowdsourcing-based semantic annotation & tagging in web repositories to improve and sustain relevance in search results* [Doctoral dissertation, University of Southampton].

- Kovačević, Ž. (2022). *Automatic Compiler/Interpreter Generation from Programs for Domain-Specific Languages Using Semantic Inference* [Doctoral dissertation, Univerza v Mariboru (Slovenia)].
- Lazarinis, F. (2008). Improving concept-based web image retrieval by mixing semantically similar Greek queries. *Program: electronic library and information systems*, 42(1), 56-67
- Lin, W. Y., Peng, N., Yen, C. C., & Lin, S. D. (2012, July). Online Plagiarized Detection Through Exploiting Lexical, Syntax, and Semantic Information. In *Proceedings of the ACL 2012 System Demonstrations*, 145-150.
- Luis Morato, J., Sánchez-Cuadrado, S., Dimou, C., Yadav, D., & Palacios, V. (2013). Evaluation of semantic retrieval systems on the semantic web. *Library hi tech*, 31(4), 638-656.
- Meng, D., & Huang, X. (2006, October). *An Interactive Intelligent Search Engine Model Research Based on User Information Preference*. In 9th Joint International Conference on Information Sciences (JCIS-06). Atlantis Press.
- Minker, W. (2004). Introducing Syntax Information in a Stochastically-Based Semantic Case Grammar Parser. *International Journal of Speech Technology*, 7, 45-54. DOI:10.1023/B:IJST.0000004807.22285.36
- MK, Y. I. A., & Mohad Noah, S. A. (2017). Semantic text-based image retrieval with multi-modality ontology and DBpedia. *The Electronic Library*, 35(6), 1191-1214
- Nakov, P. I. (2007). *Using the web as an implicit training set: Application to noun compound syntax and semantics*. University of California: Berkeley
- Navarro-Galindo, J. L., & Samos, J. (2012). The FLERSA tool: adding semantics to a web content management system. *International Journal of Web Information Systems*, 8 (1), 1-10.
- Nguyen, H. M., Nguyen, H. Q., Tran, K. N., & Vo, X. V. (2015). GeTFIRST: ontology-based keyword search towards semantic disambiguation. *International Journal of Web Information Systems*, 11 (4), 442-467.
- Nguyen, T., & Phan, T. (2008). A hybrid solution of ontology-based query expansion. *International Journal of Web Information Systems*, 4 (2), 215-227.
- Nguyen, Q. M., & Cao, T. D. (2015). A novel approach for automatic extraction of semantic data about football transfer in sport news. *International Journal of Pervasive Computing and Communications*, 11(2),233-252.
- Norouzi, Y., & Homavandi, H. (2018). Image search and retrieval problems in web search engines: A case study of Persian language writing style challenge. *Online Information Review*, 42 (6),752-767.
- Ofoghi, B., Yearwood, J., & Ghosh, R. (2006, January). A semantic approach to boost passage retrieval effectiveness for question

- answering. In *Proceedings of the 29th Australasian Computer Science Conference*- 48, 95-101.
- Oliveira, B. C., Huf, A., Salvadori, I. L., & Siqueira, F. (2019). Ontogenesis: an architecture for automatic semantic enhancement of data services. *International Journal of Web Information Systems*, 15(1), 2-27.
- Phyo, S. S. (2022). Content analysis-based documentation and exploration of research articles. *Data Technologies and Applications*, 56(1), 24-43.
- Poulet, L., Pinon, J. M., & Calabretto, S. (1997, July). Semantic structuring of documents. In *Proceedings of the Third Basque International Workshop on Information Technology-BIWIT'97-Data Management Systems* (pp. 118-124). IEEE.
- Ramada, M., Silva, J. C., & DE SÁ Leitao-Junior, P. (2014). A method for semantic analysis of keyword-based queries to access information in web databases. In *Proceedings of the IADIS International Conference WWW/Internet*, 35-42.
- Roy, S. D., & Zeng, W. (2014). Cognitive canonicalization of natural language queries using semantic strata. *ACM Transactions on Speech and Language Processing (TSLP)*, 10(4), 1-30.
- Sayed, A., & Al Muqrishi, A. (2016). CASONTO: An efficient and scalable Arabic semantic search engine based on a domain specific ontology and question answering. *International Journal of Web Information Systems*, 12(2), 242-262.
- Sharma, S., Lather, J. S., & Dave, M. (2015). A Novel Framework for Semantic Discovery of Web Services using Integrated Semantic Model. *INFOCOMMUNICATIONS JOURNAL*, 7(2), 10-18.
- Solomou, G., & Koutsomitropoulos, D. (2015). Towards an evaluation of semantic searching in digital repositories: a DSpace case-study. *Program*, 49 (1), 63-90.
- Stojanovic, N. (2005). On the conceptualisation of the query refinement task. *Library Management*, 26(4/5), 281-294.
- Sung, H. Y., & Chi, Y. L. (2021). Applications of Semantic Web in integrating open data and bibliographic records: a development example of an infomediary of Taiwanese indigenous people. *The Electronic Library*, 39(2), 337-353.
- Tang, Z., Zhu, L., Le, Y., Li, K., & Cai, C. (2019, August). *Fusing Syntax and Word Embedding Knowledge for Measuring Semantic Similarity*. In 2019 IEEE 21st International Conference on High Performance Computing and Communications; IEEE 17th International Conference on Smart City; IEEE 5th International Conference on Data Science and Systems (HPCC/SmartCity/DSS). 359-366. IEEE
- Tapsai, Ch. (2018). *Information Processing and Retrieval from CSV File by Natural Language*. IEEE 3rd International Conference on Communication and Information Systems (ICCIS), Singapore. 212-216. DOI: 10.1109/ICOMIS.2018.8644947.
- Trillo, R., Gracia, J., Espinoza, M., & Mena, E. (2007). Discovering the Semantics of User Keywords. *J. Univers. Comput. Sci.*, 13(12), 1908-1935.

- Uciteli, A., Goller, C., Burek, P., Siemoleit, S., Faria, B., Galanzina, H., Weiland, T., Drechsler-Hake, D., Bartussek, W. & Herre, H. (2014). Search Ontology, a new approach towards Semantic Search. in FoRESEE: Future Search Engines 44. annual meeting of the GI Stuttgart - GI Edition Proceedings LNI, Kollen:Lecture Notes in Informatics, P-232, 667-672.
- Vani, K., & Gupta, D. (2018). Integrating syntax-semantic-based text analysis with structural and citation information for scientific plagiarism detection. *Journal of the Association for Information Science and Technology*, 69(11), 1330-1345.
- Xu, D., Jiang, S., & Qi, Y. (2011). A New Matching Method Based on Improved Semantic Distance. In *2011 International Conference on Internet Technology and Applications* (pp. 1-4). IEEE.
- Zhou, Y. (2022). Natural language processing with improved deep learning neural networks. *Scientific Programming*, 2022 (1), 6028693.

References [In Persian]

- Akhshik, S. S., Negahdari, K., & Emami, A. (2022). Persian Writing in GANJ: Investigating the Impact of Morphology, Semantics, and Writing Style on Iran's Treasure of Scientific and Technical Information. *Sciences and Techniques of Information Management*, 8(1), 193-220. DOI: 10.22091/stim.2021.6418.1505 [In Persian]
- Aveyard, H. (2007). *Doing a Literature Review in Health and Social Care*. Translated by Pouria Sarami Forooshani & Fardin Alipour Gravand (2011). Tehran: Jameshenasan. [In Persian]
- Abutorabi Gudarzi, H. (2012). *Persian-English Cross-Language Information Retrieval: survey and improvement feasibility* [Masters Thesis, Shahed University, Tehran] [In Persian]
- Dorri, R. (2015). Comparison and Evaluation of Semantic Search Engines. *Iranian Journal of Information Processing and Management*, 30(2), 467-490. DOI: 10.35050/JIPM010.2015.044 [In Persian]
- Rezai-Sharifabadi, S., Khosravi, A., Haji Zainalabedini, M. (2007). The Feasibility of Subject Authority Control of Persian Medical Databases available on the Web. *Educational and Psychological Studies*, 8(3), 183-201. [In Persian]
- Rafiee, A.R. (2015). *Synonym and Antonym Extraction* [Masters Thesis, Languages and Linguistics Center, Sharif University of Technology, Tehran] [In Persian]
- Sadoughi, F., Valinejadi, A., Hassanzadeh, H. M., Bouraghi, H., & Pasyar, P. (2012). Image Semantic Retrieval Challenges and Thesauri Modern Applications. *Iranian Journal of Information Processing and Management*, 27(3), 641-666. [In Persian]
- Ghayoomi, M. (2019). Identifying Persian Words Senses Automatically by Utilizing the Word Embedding Method. *Iranian Journal of Information Processing and Management*, 35(1), 25-50. DOI: 10.35050/JIPM010.2019.001 [In Persian]
- Kamyar, H. (2011). *Novel Semantic Term Weighting Approach in Text*

- Processing Applications* [Masters Thesis, Ferdowsi University of Mashhad, Mashhad] [In Persian]
- Karimi, E., Babae, M., Hosseini Beheshti, M. (2018). Analysis of User query refinement behavior based on semantic features: user log analysis of Ganj database (IranDoc). *Human Information Interaction*, 5 (3), 1-14. [In Persian]
- Mohammadian Jadval Ghadam, F. (2012). *Provide a method for automatic text Summarization Farsi* [Masters Thesis, University of Isfahan, Isfahan] [In Persian]
- Zolfaghar Kondori, Z., & Mosavi Miangah, T. (2015). Lexical Disambiguation of Polysemeous Adjectives in MT: A Corpus Based Study. *Iranian Journal of Information Processing and Management*, 30(3), 719-735. DOI: 10.35050/JIPM010.2015.030 [In Persian]
- Nassiri, T., & Khandan, M. (2016). Bertrand Russell's Mathematical Logic and Its Implications in Information Organization and Subject Headings. *Librarianship and Information Organization Studies*, 27(1), 7-24. [In Persian]
- Nasiri, T., Riahinia, N., Neshat, N., Shaghaghi, M., Rasoli Poor, R. (2021). Carnapean Modal Semantics and its Implications for Ontology-Based Information Retrieval. *Human Information Interaction*, 8 (3), 1-8. URL: <http://hii.khu.ac.ir/article-1-3014-fa.html> [In Persian]
- Nojavan Aghdaragh, B. (2013). *Similar sentences detection system for Multi - document summarization* [Masters Thesis, University College of Nabi Akram] [In Persian]
- Norouzi, Y. (2019). Context and Meaning in Information Retrieval: Emphasis on Mother Tongue. *Iranian Journal of Information Processing and Management*, 35(1), 1-24. DOI: 10.35050/JIPM010.2019.063 [In Persian]
- Vickery, A., & Vickery, B. C. (1989). *Information Science in Theory and Practice*. Translated by Abdulhossein Farajpahlou (1380). Mashhad: Ferdowsi University of Mashhad. [In Persian]
- Hooshyar, M. (2016). *The comparison of powers of different kinds of text contexts in sense disambiguation of English homographs* [Masters Thesis, Shiraz University, Shiraz] [In Persian]

استناد به این مقاله: هماوندی، هدی، نوروزی، یعقوب و خبازان، بنت‌الهدی. (۱۴۰۳). نحو و معنا: روندها و

گرایش‌های پژوهشی. فصلنامه بازیابی دانش و نظام‌های معنایی، ۱۱ (۳۹)، ۲۵۰-۲۰۵. DOI:

10.22054/jks.2023.75172.1599



Journal of Knowledge Retrieval and Semantic Systems is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.