

Measuring the Relevance of Retrieved Images in Search Engines based on Persian Language Writing Styles

Yaghub Norouzi*

PhD in Knowledge and Information Science; Associate Professor;
Qom University; Qom, Iran Email: ynorouzi@gmail.com

Hoda Homavandi

PhD in Knowledge and Information Science; University of Tehran;
Tehran, Iran Email: H.homavandi@gmail.com

Bent-ol hoda Khabbazan

MA in Knowledge and Information Science;
Head of Hazrat-e Masoumeh University Library; Qom, Iran;
Email: Khabbazan.lib@gmail.com

Iranian Journal of
**Information
Processing and
Management**

Received: 12, Aug. 2021

Accepted: 11, Apr. 2022

Abstract: The aim of this study was to investigate the image retrieval from selected search engines according to the written and semantic features of Persian language and determine their relevance using recall and precision formulas and to identify the most efficient search engine in retrieving images in Persian. It was done using direct observation technique. After reviewing related researches, a list of search keywords was formed in the form of a checklist based on the written and semantic features of Persian language. Each of these keywords was searched in the studied search engines, including "Google" and "Bing" and the semantic search engine "DakDakGo" in the Persian language, and the retrieved results were recorded. Then, the recall and precision of search results in each search engine were calculated and the relevance of images based on these features in each of the studied search engines was investigated. A variety of descriptive statistical techniques were applied to analyze the data along with Kolmogorov-Smirnov, Shapiro-Wilk, Kruskal-Wallis and Friedman tests. Findings demonstrated that Google, Bing and Duckduckgo search engines do not pay enough attention to the written and semantic features of Persian language and many of these features are ignored while searching and retrieving images. In the present study, Google search engine had a higher recall and precision than the other two search engines, and despite the claim of semantic search engines to provide better and more relevant information than other search engines, Duckduckgo search engine did not show good performance in retrieving images related to the written and semantic part of Persian language. There is also a significant difference

Iranian Research Institute
for Information Science and Technology
(IranDoc)

ISSN 2251-8223

eISSN 2251-8231

Indexed by SCOPUS, ISC, & LISTA

Vol. 38 | No. 2 | pp. 639-670

Winter 2023

<https://doi.org/10.35050/JIPM010.2022.037>



* Corresponding Author

between the recall and the precision of the three studied search engines.

Keywords: Image Retrieval, Relevance, Recall, Precision, Search Engines, Persian Language

سنجش میزان ربط تصاویر بازیابی شده در موتورهای کاوش مبتنی بر ویژگی‌های نگارشی زبان فارسی

یعقوب نوروزی

دکتری علم اطلاعات و دانش‌شناسی؛ دانشیار؛ گروه
علم اطلاعات و دانش‌شناسی، دانشگاه قم؛ قم، ایران؛
ynorouzi@gmail.com پدیده‌آور رابط

هدی همآوندی

دکتری علم اطلاعات و دانش‌شناسی؛ دانشگاه تهران؛
تهران، ایران H.homavandi@gmail.com

بنت‌الهدی خبازان

کارشناسی ارشد علم اطلاعات و دانش‌شناسی؛ رئیس
کتابخانه دانشگاه حضرت معصومه (س)؛ قم، ایران؛
Khabazan.lib@gmail.com



مقاله برای اصلاح به مدت ۸ روز نزد پدیدآوران بوده است.

پدیش: ۱۴۰۱/۰۱/۲۲

دو یافت: ۱۴۰۰/۰۵/۲۱

نشریه علمی | رتبه بین‌المللی
پژوهشگاه علوم و فناوری اطلاعات ایران
(ایرانداک)

شاپا (چاپی) ۸۲۲۳-۲۲۵۱

شاپا (الکترونیکی) ۸۲۳۱-۲۲۵۱

نمایه در SCOPUS، ISC، LISTA، و

jipm.irandoc.ac.ir

دوره ۳۸ | شماره ۲ | صص ۶۳۹-۶۷۰

زمستان ۱۴۰۱

<https://doi.org/10.35050/JIPM010.2022.037>



چکیده: این پژوهش با هدف بررسی تصاویر بازیابی شده از موتورهای کاوش برگزیده طبق ویژگی‌های نوشتاری و معنایی زبان فارسی و تعیین میزان ربط آن‌ها با استفاده از فرمول‌های جامعیت و دقت و نیز شناسایی کارآمدترین موتور کاوش در بازیابی تصاویر به زبان فارسی و به روش پیمایشی-تحلیلی و با استفاده از شیوه مشاهده مستقیم انجام گرفت. پس از مرور پژوهش‌های مرتبط، کلیدواژه‌های کاوش در قالب یک سیاهه بر پایه ویژگی‌های نوشتاری و معنایی زبان فارسی شکل گرفت. هر یک از این کلیدواژه‌ها در موتورهای کاوش مورد مطالعه شامل دو موتور کاوش معمولی «گوگل» و «بینگ» و موتور کاوش معنایی «داک داک گو» که در زمره موتورهای کاوش پُر استفاده هستند و قابلیت جست‌وجوی تصاویر به زبان فارسی را نیز فراهم نموده‌اند، جست‌وجو و تعداد نتایج بازیابی شده مرتبط و غیرمرتبط ثبت گردید. سپس، جامعیت و دقت نتایج جست‌وجو در هر موتور کاوش محاسبه و به بررسی میزان ربط تصاویر برپایه این ویژگی‌ها در هر یک از موتورهای کاوش مورد مطالعه پرداخته شد. برای تجزیه و تحلیل داده‌ها از انواع فنون آماری توصیفی به همراه آزمون «کلموگروف-اسمیرنوف»، «شاپیرو-ویلک»، «کروسکال-والیس» و «فریدمن» استفاده شد. یافته‌های پژوهش نشان داد که موتورهای کاوش مورد مطالعه

نسبت به ویژگی‌های نوشتاری و معنایی زبان فارسی توجه کافی ندارند و بسیاری از این ویژگی‌ها را در هنگام جست‌وجو و بازیابی تصاویر نادیده می‌گیرند. در پژوهش حاضر معلوم شد که موتور کاوش «گوگل» از جامعیت و دقت بالاتری نسبت به دو موتور کاوش دیگر برخوردار است و به‌رغم ادعای موتورهای کاوش معنایی مبنی بر فراهم‌آوری اطلاعات باکیفیت‌تر و مرتبط‌تر نسبت به دیگر موتورهای جست‌وجو، موتور کاوش «داک‌داک‌گو» عملکرد مطلوبی در زمینه بازیابی تصاویر مرتبط با بخش نوشتاری و معنایی زبان فارسی از خود نشان نداد. همچنین، بین میزان جامعیت و دقت در سه موتور کاوش مورد پژوهش اختلاف معناداری وجود دارد.

کلیدواژه‌ها: بازیابی تصاویر، ربط، جامعیت، مانعیت، موتورهای کاوش، زبان فارسی

۱. مقدمه

در عصر حاضر شاهد رشد روزافزون تولید علم و فناوری هستیم و بر این اساس تولید انواع اطلاعات در قالب‌های مختلف در حال افزایش است و هر روز بر حجم آن افزوده می‌شود. در میان انواع اطلاعات موجود در وب، تصاویر از مقبول‌ترین داده‌ها هستند و از اهمیت ویژه‌ای برخوردارند؛ چرا که برای برخی از کاربران ارزش اطلاعاتی یک تصویر می‌تواند بهتر از هزاران کلمه باشد (خادمی‌زاده، عصاره و مبینی ۱۳۹۷). از سوی دیگر، امروزه اشاعه تصاویر به دلیل توسعه فناوری‌های دیجیتال و افزایش حجم عظیم اطلاعات دیداری رشد بی‌سابقه‌ای داشته و روزبه‌روز بر گسترش کاربردهای داده‌های تصویری افزوده می‌شود. این روند از سویی کاربران اطلاعات را با سردرگمی و مشکل در بازیابی اطلاعات مناسب مواجه می‌سازد (جلالی دیزجی، خالقی و طاهری ۱۳۹۵) و از دیگر سو، نیاز به ارزیابی مستمر ابزارهای دسترسی به اطلاعات وب را به‌منظور اصلاح و بهبود آن در تناسب با نیازها و ملزومات روزآمد ضروری می‌نماید. افزون بر این، دسترسی و استفاده از اینترنت به‌عنوان یک منبع اطلاعاتی روزمره و پرمراجعه نیاز به ارزیابی کیفیت آن را مضعف می‌کند؛ چرا که استانداردها و کنترل‌هایی که در انتشار و سازماندهی در فضاهای سنتی اطلاعاتی به کار می‌روند، در اینترنت مورد استفاده قرار نمی‌گیرند (Kumar & Prakash 2009). همچنین، حجم عظیم تصاویر ذخیره‌شده و نارسایی نظام‌های کنونی بازیابی اطلاعات غیرممتنی در وب باعث شده است که کاربران در جست‌وجوی تصاویر دچار مشکل شوند. بنابراین، این پرسش مطرح می‌شود که افراد چگونه می‌توانند تصاویر مورد نظر خود را از میان حجم انبوه تصاویر موجود در وب به‌گونه‌ای کارا و مؤثر بازیابی

کنند (خادمی‌زاده، عصاره و مبینی ۱۳۹۷). به نظر می‌رسد که کاربران برای به‌دست آوردن تصاویر مورد نیاز باید از ابزارهای جست‌وجوی کارآمد و قابل اعتمادی استفاده کنند تا بتوانند در میان حجم عظیم مجموعه‌های تصاویر آنچه را که می‌خواهند بازیابی کنند. یکی از ابزارهای قدرتمندی که در این عرصه شروع به فعالیت کرده‌اند، موتورهای جست‌وجو هستند. موتورهای جست‌وجو در حال حاضر نه تنها وسیله‌ای برای دسترسی به اطلاعات متنی هستند، بلکه برای جلب بیشتر رضایت کاربران، سازماندهی و بازیابی تصاویر را نیز جزء اهداف خود قرار داده‌اند (نوروزی و سوری ۱۳۹۳). از آنجا که حدود ۹۰ درصد زمانی که کاربران صرف جست‌وجوی محتوای وب می‌کنند در محیط موتورهای کاوش وب است، می‌توان گفت موتورهای کاوش محبوب‌ترین ابزار جست‌وجوی محتوای وب نزد کاربران هستند؛ به گونه‌ای که عملکردها و امکانات این ابزارها در کاوش انواع اطلاعات، مورد توجه محققان و حتی کاربران است. با وجود انواع موتورهای جست‌وجو که بر پایه الگوریتم‌ها و روش‌های بازیابی طراحی شده‌اند، بسیاری از کاربران در مورد این که کدام یک از این ابزارها با توجه به ویژگی‌ها و نیازهای خاص آن‌ها دارای عملکرد بهتری است، اطلاع دقیقی ندارند (Bute, Hussaini & Adamu 2017). بنابراین، انتخاب موتور جست‌وجوی مناسب که جوابگوی نیازهای اطلاعاتی آن‌ها باشد و دسترسی به موقع به اطلاعات صحیح و مرتبط را فراهم آورد، از اهمیت بالایی برخوردار است. بیشتر کاربران اینترنت نیازهای اطلاعاتی خود را از طریق موتورهای جست‌وجوی متداول و شناخته‌شده‌ای مانند «گوگل»، «یاهو» و «بینگ» برآورده می‌کنند که از جمله موتورهای جست‌وجوی پرکاربرد و محبوب وب محسوب می‌شوند. این سیستم‌ها به‌رغم برخورداری از امکانات مناسب و محبوبیت آن‌ها برای جست‌وجوی وب توسط کاربران، معایب و مشکلاتی نیز دارند. این موتورهای جست‌وجو در بسیاری از مواقع، نتایج مرتبط با نیاز کاربر را ارائه نمی‌دهند؛ به طوری که کاربر می‌بایست زمان زیادی را به بازیابی و مرور در مدارک بازیابی‌شده صرف کند تا منبع مورد نیاز خود را بیابد (دُری ۱۳۹۳). با وجود مقدار زیاد اطلاعات در شبکه جهانی وب و مشکلات موجود در رابطه با جست‌وجوی اطلاعات، محققان و تولیدکنندگان نرم‌افزارها برای همگام شدن با تغییرات جدید، به سمت شکل جدیدی از فناوری به نام وب معنایی روی آورده‌اند. از جمله برنامه‌های کاربردی توسعه داده‌شده به منظور استفاده از فناوری معنایی، موتورهای جست‌وجوی معنایی هستند. فناوری وب معنایی توانایی تمرکز بر فراداده به جای نحو را

دارد، که باعث شده موتورهای جست‌وجوی معنایی به‌جای نحو کلمات کلیدی، معنای آن‌ها را جست‌وجو کنند (Hussan 2020). در واقع، این نوع موتورها می‌توانند خواسته کاربر را تجزیه و تحلیل کرده و با استفاده از استدلال منطقی، نتایج دقیق‌تری را بازبایی کنند.

با وجود این، «نوتس» معتقد است که ردیابی سریع اطلاعات مورد نیاز در اینترنت به‌صورت نیازی پیچیده در آمده است (Notess 1997). یکی از دلایل عمده این پیچیدگی می‌تواند گوناگونی کاربران و زبان‌ها و فرهنگ‌های آنان باشد؛ چرا که یافته‌های جهانی در وب‌سایت پایش جهانی وضعیت اینترنت^۱ در سال ۲۰۲۰، نشان می‌دهد که ۷۴/۱ درصد از کاربران اینترنت را افراد غیرانگلیسی‌زبان تشکیل می‌دهند و همین ویژگی، وب را یک مکان اطلاعاتی چندزبانه و چندفرهنگی می‌سازد. بدین ترتیب، نیازها و خواسته‌های کاربران غیرانگلیسی‌زبان در طراحی هر سیستم اطلاعاتی تحت وب و به‌ویژه در سامانه‌های بازبایی آن و موتورهای کاوش باید در نظر گرفته شود (هماوندی ۱۳۹۲). بنابراین، با توجه به طیف گسترده کاربران از نظر ملیت و زبان، یکی از مسائل چالش‌برانگیز در مورد وب و فرایند نمایه‌سازی در آن، بحث زبان است. پردازش، خلاصه‌سازی و بازبایی اطلاعات از متون و تحلیل آن، همه از دسته مواردی هستند که به‌دلیل محدودیت‌های خاص زبانی، استفاده از آن‌ها به‌صورت کامل توسط کاربران با زبان‌های مختلف انجام نمی‌پذیرد. بنابراین، با وجود این که موتورهای جست‌وجوی متعددی برای تسهیل جست‌وجو در محیط وب وجود دارند، اما به نظر می‌رسد که توجه آن‌ها به زبان‌های غیرانگلیسی در مقایسه با انگلیسی کافی نیست و کاربران ایرانی نیز از این قاعده مستثنا نیستند؛ چرا که بر پایه آخرین آمار اعلام‌شده از ضریب نفوذ اینترنت در کشور از سوی «سازمان تنظیم مقررات و ارتباطات رادیویی» بالغ بر ۹۴ درصد جمعیت کشور به اینترنت متصل می‌شوند. به بیان دیگر، از هر ۱۰۰ نفر در کشور ۹۴ نفر از اینترنت استفاده می‌کنند. این آمار نشان می‌دهد که تا پایان خرداد ماه ۱۳۹۹ بالغ بر ۷۸ میلیون و ۸۶ هزار و ۶۶۳ نفر در ایران از اینترنت استفاده می‌کردند و در واقع، مشترک اینترنت بودند (خبرگزاری مهر ۱۳۹۹/۵/۴). از همین روست که عدم توجه به ویژگی‌های تأثیرگذار زبان فارسی از جمله ویژگی‌های نگارشی و معنایی در مراحل ذخیره‌سازی، پردازش، جست‌وجو، و بازبایی

1. Internet World Stats: Number of Internet Users by Language, 2020

اطلاعات از پایگاه‌های اطلاعاتی به‌ویژه موتورهای جست‌وجو، موجب ایجاد چالش‌ها و موانعی پیش روی کاربران فارسی‌زبان در دستیابی به اطلاعات مورد نیازشان شده است؛ چرا که دادوگرفت در بازیابی، زمانی موفقیت‌آمیز است که مدارک بازیابی‌شده مناسب درخواست‌کننده باشد. از این رو، ربط را می‌توان محک توفیق بازیابی و مهم‌ترین مفهوم در بازیابی اطلاعات دانست (پائو ۱۹۸۹). چالش‌های زبانی در این دادوگرفت در موارد متعددی منجر به بازیابی نتایج غیرمرتبط با حوزه موضوعی مورد نظر می‌گردد. بنابراین، ممکن است نتایج بازیابی‌شده به‌زبانی دیگر از ربط کافی با پرسش جست‌وجو برخوردار نباشد. این است که پرداختن به پژوهش در حوزه ارزیابی میزان ربط تصاویری که از این نظام‌ها در اختیار کاربر قرار می‌گیرد، مهم بوده و گامی در جهت بهبود کیفیت نظام‌های بازیابی اطلاعات است که می‌تواند به طراحی و ایجاد نظام اطلاعاتی بهتر بیانجامد (ابوالقاسم مسلمان، مؤمنی و زین‌العابدینی ۱۳۹۴)؛ چرا که ربط، یکی از معیارهایی است که برای ارزیابی عملکرد نظام‌های بازیابی اطلاعات به‌ویژه موتورهای جست‌وجو و کارایی کاوش‌های انجام‌شده به کار می‌رود. در حقیقت، هسته کلیه فعالیت‌های طراحی نظام‌های بازیابی اطلاعات و ارزیابی این نظام‌ها مفهوم ربط است. به‌طوری که می‌توان گفت میزان ربط نتایج بازیابی‌شده یا به‌عبارت دیگر، بررسی جامعیت و مانعیت نتایج از شاخصه‌های مهمی است که بازیابی مؤثر و مطلوب را در پی خواهد داشت اهمیت پرداختن به این موضوع در مطالعه نظام‌های دیداری و بازیابی تصاویر دو چندان می‌شود، زیرا تصاویر نه‌تنها به‌عنوان منبع اطلاعاتی و ابزار ارتباطی به کار می‌روند، بلکه با توسعه روش‌های دقیق تصویرسازی، درک انسان از اطلاعات را ارتقا داده‌اند (Menard 2007)؛ به‌گونه‌ای که امروزه، انگیزه‌های زیادی برای استفاده از تصاویر در اهداف تجاری و پژوهش‌های علمی تا انگیزه‌های دیگری مثل سرگرمی به‌وجود آمده است. این علاقه و انگیزه کاربران به داده‌های تصویری در برخی از آمارهای اینترنتی به‌خوبی قابل مشاهده است. به نقل از «کلارکسون»، ۹۳ درصد از پربازدیدترین پست‌های «فیس‌بوک» را تصاویر تشکیل می‌دهند و ۶۵ درصد از کاربران اینترنت علاقه‌مند به ارسال تصاویر در پست الکترونیک خود هستند (Clarkson 2014). روشن است که این آمارها با توسعه بسترهای نرم‌افزاری و شبکه‌های اجتماعی افزایش پیدا کرده‌اند. هم‌زمان، حجم تصاویر موجود در وب و اندازه پایگاه‌های داده‌ای تصویری به‌سرعت در حال بزرگ شدن بوده و یافتن تصاویر مرتبط برای کاربران دشوار شده است (بهمن‌آبادی ۱۳۸۱). ارزیابی کارایی یک

نظام بازیابی اطلاعات صرف نظر از این که با چه نوع پایگاه اطلاعاتی روبه‌رو باشیم، بدون بررسی جامعیت و مانعیت ناتمام خواهد ماند و به‌درستی نمی‌توان درباره کارایی آن قضاوت کرد (عباسی دشتکی، چشمه سهرابی، ۱۳۹۸). بر همین پایه ارزشیابی امری بسیار ضروری در حل مشکلات نظام‌های اطلاع‌رسانی است و در جریان شناسایی نقاط ضعف، می‌توان توصیه‌هایی برای کارکردهای خاص مثل نمایه‌سازی، مهار واژگان و مراحل جست‌وجو به‌منظور بهبود بازیابی ارائه داد (پائو ۱۹۸۹). با توجه به اهمیت و کاربرد وسیع موتورهای کاوش، یکی از موارد قابل توجه در ارتباط با این ابزارها شمار اندک مطالعات انجام‌شده با موضوع ارزیابی کارایی آن‌ها در تأمین نیازهای اطلاعاتی کاربران با در نظر گرفتن ویژگی‌های زبان‌های مختلف غیرانگلیسی است (Gross 2014). افزون بر این، انجام پژوهش‌هایی از این دست با کشف و بازنمون چالش‌های عمده زبانی در بازیابی تصاویر در طراحی ابزارهای جست‌وجوی بومی نیز کمک‌کننده است. تاکنون پژوهش‌های مختلفی به بررسی جامعیت و دقت موتورهای کاوش عمومی در بازیابی اطلاعات به زبان فارسی پرداخته‌اند، اما مطالعه‌ای که در بحث بازیابی تصاویر، ویژگی‌های اصلی نوشتاری و معنایی زبان فارسی را افزون بر موتورهای کاوش معمولی در ارتباط با موتورهای کاوش معنایی مورد آزمون قرار داده باشد، انجام نشده است. بنابراین، با توجه به مسائل مطرح‌شده، پژوهش حاضر تلاش دارد میزان ربط تصاویر بازیابی‌شده به زبان فارسی را با استفاده از سنج‌های جامعیت و دقت، از موتورهای کاوش معمولی «گوگل» و «بینگ» و موتور کاوش معنایی «داک‌داک‌گو» که بر پایه رتبه‌بندی‌ها جزء پُراستفاده‌ترین موتورهای کاوش با قابلیت پشتیبانی زبان فارسی و همچنین، قابلیت جست‌وجو و بازیابی تصاویر هستند، ارزیابی کرده و عملکرد موتورهای کاوش مورد پژوهش را بررسی و کارآمدترین موتور کاوش در بازیابی تصاویر به زبان فارسی را (بر پایه ویژگی‌های نوشتاری و معنایی) معرفی نماید.

۲. پرسش‌های پژوهش

۱. میزان جامعیت و دقت در تصاویر بازیابی‌شده توسط موتورهای کاوش مورد مطالعه در این پژوهش چقدر است؟
۲. آیا بین موتورهای کاوش مورد مطالعه در میزان جامعیت و دقت تفاوت معناداری وجود دارد؟

۱. لیست پراستفاده‌ترین موتورهای کاوش از سایت‌های

<https://netmarketshare.com> و <https://clever-solution.com/blog/top-search-engines-in-the-world> استخراج شده‌اند.

۳. کدام‌یک از موتورهای کاوش مورد مطالعه عملکرد بهتری در بازیابی تصاویر به زبان فارسی دارند؟

۳. پیشینه پژوهش

با توجه به تبدیل شدن موتورهای جست‌وجوی اینترنت به ابزاری پُراستفاده در جست‌وجو و بازیابی اطلاعات در میان افراد مختلف، تحقیقات متعددی در این زمینه انجام شده که بر نمایه‌سازی، بازیابی اطلاعات از این ابزارهای کاوش و بررسی ربط آن‌ها و نیازهای زبانی کاربران کشورهای مختلف دنیا تأکید دارند. در ادامه، مرتبط‌ترین پژوهش‌های داخلی و خارجی انجام‌شده با پژوهش حاضر در جداول ۱ و ۲ ارائه می‌گردد.

جدول ۱. پیشینه‌های داخلی مربوط به پژوهش حاضر

نویسنده و سال	عنوان	هدف	روش	نتیجه
نوروزی، خویدکی و صدرآبادی (۱۴۰۰)	مرور نظام‌مند پژوهش‌های حیطه ربط در ایران	شناسایی شکاف در پژوهش‌های کنونی و ارائه چارچوب/زمینه به‌منظور فعالیت‌های پژوهشی جدید	مطالعه نظام‌مند	<ul style="list-style-type: none"> ◇ تمرکز بیشتر مطالعات بر ربط عینی و موتورهای جست‌وجو ◇ تعلق جدیدترین مطالعات به قضاوت ربط ◇ تعلق کمترین مطالعات به بازخورد ربط
خادمی‌زاده، عصاره و مبینی (۱۳۹۷)	بررسی تأثیر شیوه‌های نمایه‌سازی بر به‌سازی بازیابی تصاویر با استفاده از موتور جست‌وجوی گوگل	مقایسه تأثیر استفاده از دو روش نمایه‌سازی متن‌محور و فولکسونومی یا نمایه‌سازی مردمی تصاویر بر میزان بازیابی موفق تصاویر	پژوهش مبتنی بر فناوری و شبه‌تجربی	تأثیر مثبت استفاده از روش نمایه‌سازی مردمی تصاویر در مقایسه با روش متن‌محور در بازیابی موفق تصاویر به‌وسیله موتور جست‌وجوی گوگل
یقظین و جوکار (۱۳۹۶)	کارآمدی فیلد زبان در موتورهای کاوش بین‌المللی برای بازیابی نتایج در زبان فارسی و عربی	کارآمدی فیلد محدودیت زبانی در سه موتور کاوش گوگل، بینگ و یاهو	پیمایشی	کارآمدی بیشتر موتورهای جست‌وجوی بینگ و یاهو در بازیابی نتایج مرتبط با زبان فارسی و عربی
جلالی، دیزجی، خالقی و طاهری (۱۳۹۵)	میزان جامعیت و مانعیت بازیابی تصاویر در موتورهای کاوش Bing و Google Image	تعیین میزان جامعیت و مانعیت دو موتور کاوش بینگ و گوگل در بازیابی محتوای محصور تصاویر	شبه‌آزمایشی	برتری موتور کاوش گوگل در مقایسه با بینگ از نظر میزان جامعیت و مانعیت بیشتر موتور کاوش بینگ

نویسنده و سال	عنوان	هدف	روش	نتیجه
رجبی و نوروزی (۱۳۹۴)	موتورهای جست‌وجوی فارسی؛ ارزیابی امکانات جست‌وجوی بازایی اطلاعات، میزان جامعیت و مانعیت و تعیین همپوشانی میان آن‌ها	بررسی و ارزیابی امکانات جست‌وجو در موتورهای جست‌وجوی فارسی، شناسایی کارآمدترین موتور جست‌وجو به لحاظ بازایی اطلاعات مرتبط، شناسایی موتورهای جست‌وجوی کارآمد به لحاظ بازایی اطلاعات از حیث محاسبه میزان جامعیت و مانعیت در هر موتور جست‌وجو و تعیین میزان همپوشانی موجود بین آن‌ها	پیمایشی - تحلیلی	<ul style="list-style-type: none"> ◇ برخورداری موتورهای جست‌وجوی فارسی از امکانات جست‌وجوی مناسب و عملکرد قابل اطمینان آن‌ها ◇ وجود تفاوت معنادار بین موتورهای جست‌وجو از نظر بازایی نتایج مرتبط و میزان جامعیت ◇ کارآمدی بیشتر موتورهای دارای امکانات جست‌وجوی بیشتر از نظر باز یافت اطلاعات مرتبط تر و همپوشانی بیشتر
ابوالقاسم مسلمان، مؤمنی و حاجی زین العابدینی (۱۳۹۴)	سنجش میزان ربط تصاویر باز یاب شده در موتورهای جست‌وجوی یاهو، گوگل، پیک سرچ ^۱ و فلیکر ^۲ و ارائه الگوی پیشنهادی بهینه	سنجش میزان ربط تصاویر باز یاب شده در موتورهای جست‌وجوی یاهو، گوگل، پیک سرچ و فلیکر به زبان فارسی، بر پایه میزان جامعیت و مانعیت نتایج بازایی شده و معرفی بهترین موتور جست‌وجو و ارائه الگوی بهینه	پیمایشی	<ul style="list-style-type: none"> ◇ عملکرد بهتر موتور جست‌وجوی یاهو نسبت به سایر موتورهای جست‌وجو با اختصاص بیشترین میزان مانعیت ◇ کسب نتایج بهتر با استفاده از نظام مبتنی بر محتوا و به کارگیری مجموعه‌ای از ویژگی‌های متن و رنگ
نوروزی و سوری (۱۳۹۳)	ارزایی موتورهای جست‌وجو در بازایی تصاویر بر پایه نمایه‌سازی مبتنی بر متن و محتوا	شناسایی توانمندی‌ها و تفاوت‌های موتورهای جست‌وجوی مورد مطالعه در بازایی تصاویر بر پایه نمایه‌سازی مبتنی بر متن و محتوا	تحقیقات ارزیابانه	<ul style="list-style-type: none"> ◇ کارآمدی بالای موتور جست‌وجوی گوگل در بازایی تصاویر نسبت به سایر موتورهای جست‌وجو
حریری و کیلی مفرد (۱۳۹۲)	مقایسه دقت موتورهای کاوش عمومی و تخصصی پزشکی در بازایی تصاویر پزشکی	تعیین میزان دقت موتورهای کاوش عمومی و تخصصی پزشکی در بازایی تصاویر پزشکی و مقایسه آن‌ها	پیمایشی / مقایسه‌ای	<ul style="list-style-type: none"> ◇ دقت بیشتر موتور کاوش گوگل و موتور کاوش پزشکی اومنی‌مدیکال ◇ سرچ در بازایی تصاویر مرتبط و کاملاً مرتبط پزشکی ◇ تفاوت در الگوریتم نمایه‌سازی منابع موجود در وب

نویسنده و سال	عنوان	هدف	روش	نتیجه
کوکبی و همکاران (۱۳۹۲)	تأملاتی بر نمایه‌سازی تصاویر: یک تصویر ارزشی برابر با هزار واژه	بررسی شیوه‌های مختلف نمایه‌سازی (نمایه‌سازی مفهوم‌محور، نمایه‌سازی محتوا‌محور، نمایه‌سازی مشارکتی یا فاکسونومی)	اسنادی	<ul style="list-style-type: none"> ◇ توصیفات متنی، بهترین روش برای نمایه‌سازی تصاویر ◇ عملکرد مناسب موتورهای جست‌وجوی موجود در بازیابی منابع متنی و عدم موفقیت آن‌ها در بازیابی تصاویر
احمدی و چشمه‌سهرابی (۱۳۹۱)	مقایسه میزان جامعیت و مانعیت در ابزارهای کاوش اطلاعات فازی و غیرفازی	مقایسه میزان جامعیت و مانعیت در ابزارهای جست‌وجوی فازی و غیرفازی و شناسایی قابلیت‌های موجود در سیستم‌های بازیابی اطلاعات فازی	پیمایشی	<ul style="list-style-type: none"> ◇ مانعیت بیشتر نظام‌های فازی در مقایسه با نظام‌های غیرفازی ◇ برتری نظام‌های جست‌وجوی فازی از نظر جامعیت

همان‌طور که در جدول ۱، مشاهده می‌شود، مرور پیشینه‌های مورد مطالعه که روش بیشتر آن‌ها پیمایشی است، در مجموع، حاکی از آن است که طی سال‌های گذشته مطالعه بر روی تصاویر به‌عنوان یکی از انواع مهم اطلاعات در محیط وب روندی صعودی داشته است. البته، این مسئله را می‌توان ناشی از افزایش توجهات به تصاویر به‌عنوان منابع اطلاعاتی دانست که هم‌اکنون خود دارای هویت اطلاعاتی مستقلی شده‌اند و از همین رو، انواع پایگاه‌ها و موتورهای جست‌وجو در این زمینه راه‌اندازی شده است؛ ضمن این که سازمان‌ها و به‌ویژه سازمان‌های تصویر‌محور به یکباره با انباشت تصاویری مواجه شده‌اند که برنامه‌ای برای سازماندهی آن‌ها ارائه نکرده‌اند. بنابراین، لازم است شیوه‌های نمایه‌سازی و تعیین میزان جامعیت و دقت بازیابی اطلاعات به‌ویژه تصاویر در ابزارهای ذخیره و بازیابی اطلاعات به‌خصوص موتورهای کاوش بیشتر مورد توجه و ارزیابی قرار گیرد. در رصد پژوهش‌های انجام‌شده مشاهده می‌شود که میزان توجه به موتورهای کاوش «گوگل» و «بینگ» بیشتر بوده است. این مسئله یادآور می‌شود که توجه به رسانه‌ها و بسترهای تخصصی حوزه ذخیره و بازیابی تصاویر به‌نوعی مغفول مانده است. بنابراین، رویکردها باید از تأکید صرف بر بسترهای عام ارائه‌دهنده تصاویر که در موارد زیادی تصویر را به‌عنوان هویت مستقلی از متن در نظر نمی‌گیرند، افزایش یابد. البته، این پژوهش منوط به داخل کشور نیست و در خارج از ایران نیز مطالعات گوناگونی در این زمینه انجام شده است که برخی از آن‌ها در جدول ۲، ارائه شده است.

جدول ۲. پیشینه‌های خارجی مربوط به موضوع پژوهش

نویسنده و سال	عنوان	هدف	روش	نتیجه
CheshmehSohrabi, and Adnani Sadati (2021)	ارزیابی عملکرد موتورهای جستجوی وب در ارزیابی تصویر: یک مطالعه تجربی	تعیین میزان بازیابی تصاویر مربوط به نیاز کاربران از طریق سنجش ربط موتورهای جستجوی عمومی و تخصصی تصویر	پیمایشی	برتری موتور جستجوی عمومی تصویر نسبت به موتورهای تخصصی تصویر بر پایه بازیابی تصاویر مربوط به درخواست‌های کاربران و از نظر تعداد تصاویر غیر تکراری بازیابی شده
Hussan (2020)	بررسی تطبیقی موتورهای جستجوی معنایی و مبتنی بر کلمات کلیدی	بررسی و مقایسه جامع برای موتورهای جستجوی مبتنی بر نحو و موتورهای جستجوی مبتنی بر معنایی و فناوری‌های آن‌ها	پیمایشی	استفاده موتورهای جستجوی مبتنی بر نحو از فناوری‌های معمولی و موتورهای جستجوی معنایی از فناوری‌های وب معنایی جستجوی نتایج دقیق در درخواست جستجوی کاربر مزیت متقابل موتورهای جستجوی معنایی
Norouzi & Homavandi (2018)	بررسی مشکلات جستجو و بازیابی تصاویر در موتورهای کاوش وب (مطالعه موردی چالش‌های سبک نگارش زبان فارسی)	تعیین مشکلات جستجو و بازیابی تصاویر در موتورهای کاوش برگزیده مبتنی بر ویژگی‌های نگارشی زبان فارسی	پیمایشی - تحلیلی	عدم توجه کافی موتورهای کاوش گوگل، بینگ و یاهو، نسبت به ویژگی‌های نوشتاری و معنایی زبان فارسی انطباق بیشتر قابلیت‌های گوگل در بخش معنایی و نوشتاری زبان فارسی
Negi and Kumar (2014)	تحلیل تطبیقی موتورهای جستجوی مبتنی بر کلمات کلیدی و معنایی	بررسی عملکرد جستجوی معنایی دو موتور جستجوی معمولی گوگل و یاهو و سه موتور جستجوی معنایی هاکیا، داک‌داک‌گو و بینگ و سنجش عملکرد آن‌ها در پردازش پرس‌وجوهای زبان طبیعی	پیمایشی	عملکرد بالاتر جستجوی معنایی موتورهای جستجوی معنایی نسبت به موتورهای جستجوی عادی

نویسنده و سال	عنوان	هدف	روش	نتیجه
Lopez and Ribeiro (2011)	ارزیابی و مقایسه موتورهای کاوش وب در بازیابی اطلاعات سلامت	ارزیابی تطبیقی موتورهای کاوش عمومی و تخصصی حوزه سلامت در بازیابی اطلاعات میزان ربط اطلاعات بازیابی‌شده از آن‌ها	پیمایشی - تحلیلی	دقت بیشتر موتورهای کاوش عمومی از موتورهای کاوش تخصصی حوزه سلامت ارائه بهترین نتایج توسط موتور جست‌وجوی گوگل
Hammo (2009)	افزایش کارایی موتورهای جست‌وجو برای پیشینه‌های نشانه‌گذاری‌شده به زبان عربی	ارائه یک قالب کاری برای افزایش کارایی موتورهای کاوش برای متون عربی دارا و فاقد اعراب‌گذاری از طریق روش‌های گسترش سؤال	پیمایشی	تأثیر گسترش سؤال بر بهبود جست‌وجو و بازیابی متون عربی و افزایش کارایی موتورهای کاوش با استفاده از ابزارهای پیشرفته پردازش زبان طبیعی

همان‌گونه که در جداول ۱ و ۲، مشاهده می‌شود، پژوهشگران حوزه‌های گوناگون از جمله علوم رایانه، هوش مصنوعی و متخصصان علم اطلاعات به موازات رشد و گسترش موتورهای کاوش به‌عنوان ابزارهای دسترسی به انبوه اطلاعات در وب به بررسی و مطالعه این ابزارهای قدرتمند ذخیره و بازیابی اطلاعات پرداخته‌اند. «ساراسویک» نیز معتقد است که پیشرفت‌های علمی در بازیابی اطلاعات بیش از آنکه وابسته به توسعه فناوری باشد، به شناخت بهتر ماهیت ربط و به‌کارگیری این ادراک در فرایند بازیابی اطلاعات بستگی دارد (Saracevic 2007). با وجود این، هنوز هم در پژوهش‌های انجام‌شده، بسیاری از جنبه‌های مرتبط با این حوزه بررسی و ارزیابی نشده‌اند. آنچه می‌توان از جمع‌بندی پیشینه‌ها به‌دست آورد، حاکی از آن است که پژوهش‌ها بیشتر به‌صورت پیمایشی و مطالعات ارزیابی و مقایسه‌ای بر روی چند موتور کاوش انجام گرفته‌اند که عمدتاً از موتورهای کاوش عمومی و مبتنی بر کلمات کلیدی هستند و در آن‌ها بیشتر به ارزیابی عملکردهای آن‌ها، شیوه‌های نمایه‌سازی و یا به میزان جامعیت و دقت در بازیابی اطلاعات پرداخته شده است. در این میان پژوهش‌های اندکی نظیر پژوهش‌های صورت گرفته توسط Hussan (2020) و Negi & Kumar (2014)، به‌ویژه در سال‌های اخیر روی موتورهای کاوش معنایی صورت گرفته است. بنابراین، این مسئله نشان از مغفول ماندن این موضوع با توجه به پیشرفت فناوری معنایی در سال‌های اخیر دارد و از همین رو، نیازمند توجه بیشتر است. همچنین، ویژگی‌ها و مشکلات زبانی نیز به‌عنوان عاملی مهم در بحث ذخیره و بازیابی

اطلاعات مطرح هستند که در برخی آثار از جمله (Hammo, Norouzi & Homavandi (2018) (2009)، «یقظین و جوکار» (۱۳۹۶)، «ابوالقاسم مسلمان، مؤمنی و حاجی زین العابدینی» (۱۳۹۴)، مورد توجه قرار گرفته‌اند. در این پژوهش‌ها قابلیت‌ها و کاستی‌های موتورهای کاوش در پوشش زبان‌های غیرانگلیسی و با هدف شناخت و ارائه راهکارهایی جهت اصلاح چالش‌های زبانی مورد توجه قرار گرفته است؛ هرچند در مورد زبان فارسی به دلیل ماهیت و خصوصیات منحصر به فرد آن در بخش نوشتاری و معنایی و ظریفی که این زبان دارد، جای تحقیق و تأمل بسیار دارد. بر پایه نتایج برخی پژوهش‌های یادشده، اگرچه طی سال‌های اخیر در موتورهای کاوش اصلاحاتی جهت غلبه بر موانع و چالش‌های زبانی در فرایند جست‌وجوی اطلاعات، از جمله ارائه مترادف کلمات جست‌وجو و املائی صحیح آن به کاربر انجام شده است، اما آنچه در عمل مشاهده می‌شود، به‌ویژه در مورد زبان‌های غیرانگلیسی از جمله فارسی، همچنان مشکلات متعدد حل نشده و کشف نشده‌ای فرا روی کاربران وجود دارد. بر همین پایه، تاکنون پژوهشی که به‌طور خاص با تمرکز بر بررسی میزان پوشش و دقت موتورهای کاوش معمولی و معنایی در بازیابی تصاویر که در آن افزون بر ویژگی‌های ریخت‌شناسی و نوشتاری، مسائل معنایی زبان فارسی هم مد نظر باشد، انجام نشده است. از این رو، پژوهش حاضر در پی آن است که تا حد ممکن این هدف را تحقق بخشد و با بررسی جنبه‌های مختلف نوشتاری و معنایی زبان فارسی در بازیابی تصاویر از موتورهای کاوش معمولی و معنایی، به سنجش میزان پوشش و دقت بازیابی این موتورهای کاوش بر پایه فرمول‌های ربط (جامعیت و دقت) بپردازد.

۴. روش پژوهش

پژوهش حاضر از نوع کاربردی است و در آن برای پاسخگویی به سؤالات پژوهش از روش پیمایشی-تحلیلی استفاده شده است که با استفاده از مشاهده مستقیم انجام گرفت. جامعه آماری پژوهش دو موتور کاوش معمولی «گوگل» و «بینگ» و موتور کاوش معنایی «داک‌داک‌گو» که در زمره موتورهای کاوش پُر استفاده هستند و افزون بر این، قابلیت بازیابی تصاویر به زبان فارسی را دارند، انتخاب شده‌اند. پس از بررسی منابع مرتبط و پیشینه‌های فارسی پژوهش و با الگوبرداری از آن‌ها نسبت به تهیه سیاهه محقق ساخته اقدام شد. در واقع، کلیدواژه‌های موجود در سیاهه این پژوهش پیونددهنده ویژگی‌های نگارشی و معنایی زبان فارسی و توانایی موتورهای کاوش در پاسخگویی به این خصوصیات است؛ به

این ترتیب که برای هر یک از ویژگی‌های نوشتاری و معنایی زبان فارسی واژه‌ای انتخاب شد تا به‌عنوان کلیدواژه کاوش، مبنای قرار گیرد. سپس، به جهت ماهیت بین رشته‌ای موضوع پژوهش، روایی سیاهه مذکور که شامل کلیدواژه‌های موجود در جداول ۳ و ۴ است، با مشورت چهار نفر از اساتید علم اطلاعات و دانش‌شناسی و زبان و ادبیات فارسی مورد تأیید قرار گرفت. به‌منظور گردآوری داده‌ها، در تاریخ معین^۱ (نیمه اول اردیبهشت ماه سال ۱۴۰۰) هر یک از متغیرها (واژه‌های موجود در سیاهه) توسط پژوهشگران^۲ به تفکیک، وارد بخش جست‌وجوی تصاویر موتورهای کاوش مورد پژوهش شد. در بخش مسائل نوشتاری صورت‌های مختلف متصور برای هر واژه، به تفکیک در هر سه موتور جست‌وجو، درج و تعداد کل نتایج بازایی شده (مرتبط و نامرتبط) در میان ۵۰ نتیجه نخست^۳ ثبت شد. در بخش مسائل معنایی نیز به‌همین ترتیب عمل شد؛ با این تفاوت که در مورد کلمات فاقد صورت‌های مختلف نوشتاری و دارای معانی گوناگون، مانند واژه «شیر» و «شور» تعداد نتایجی که حاوی معانی مختلف کلیدواژه جست‌وجو شده بودند، شمارش و ثبت گردید. از آنجا که مفاهیم نهفته در تصاویر سریع‌تر از متون ادراک می‌شود، در این پژوهش از قضاوت ربط دوگانی استفاده شده است؛ بر این مبنای که تصاویر «مربوط» و یا «نامربوط» هستند. بنابراین، در نتایج بازایی شده تصاویری که با متغیرها (واژه‌های موجود در سیاهه) همخوانی نداشته و در واقع، نامرتبط محسوب می‌شدند، از تصاویر مرتبط جدا شدند و سرانجام، میانگین جامعیت و مانعیت در هر موتور جست‌وجو محاسبه شد. در این پژوهش منظور از جامعیت یا بازیافت، توانایی یک نظام در بازایی تمام مدارک مرتبط در پایگاه است (Prasher 1989). بنابراین، برای محاسبه مقدار آن، نه تنها تعداد مدارک

۱. لازم است تاریخ جست‌وجو و بازایی نتایج به‌دلیل تغییرات مداوم محتوای موتورهای کاوش مد نظر قرار گیرد.

۲. برپایه دیدگاه گروهی از متخصصان بازایی اطلاعات، از جمله کوپر (Cooper) که ربط را به معنای نتیجه منطقی تعریف می‌کند، می‌توان مفهوم ربط را به‌منزله یک امر اکیداً فردی و شخصی کنار گذاشت (پائو ۱۹۸۹) و نتایج حاصل از پژوهش‌های ربط را که توسط متخصصان انجام می‌شوند، تعمیم‌پذیر دانست.

۳. به‌علت تعداد زیاد تصاویر بازایی شده در هر یک از موتورهای کاوش برگزیده و عدم امکان بررسی تمامی یافته‌ها، پژوهشگران ملزم به انتخاب تعداد محدودی از مدارک برای داوری ربط در هر موتور کاوش بودند. با توجه به پیشینه‌های پژوهش که در بسیاری از موارد ۲۰ نتیجه نخست را مورد ارزیابی قرار داده‌اند، و به جهت جامع‌تر شدن پژوهش، ۵۰ نتیجه (تصویر) نخست بازایی شده برای هر واژه در هر موتور کاوش در نظر گرفته و اساس کار ثبت یافته‌ها شد.

مرتبط بازیابی شده، بلکه تعداد مدارک مرتبط بازیابی نشده را نیز باید داشته باشیم. با این توصیف به نظر می‌رسد که محاسبه جامعیت یک پایگاه تقریباً غیرممکن باشد. چون نمی‌توان به تعداد مدارک مرتبطی دست یافت که به دلایلی در بازیابی ملاحظه نمی‌شود (Clarke and Willett 1997)، روشی برای محاسبه جامعیت نسبی در محیط وب به کار برده‌اند. آنان جامعیت نسبی را به این صورت تعریف می‌کنند:

$$\text{تعداد کل مدارک مرتبط بازیابی شده توسط یک پایگاه} \\ \text{جامعیت نسبی} = \frac{\text{تعداد کل مدارک مرتبط بازیابی شده توسط کلیه پایگاه‌های مورد مطالعه}}{\text{تعداد کل مدارک مرتبط بازیابی شده توسط یک پایگاه}} \times 100$$

از سوی دیگر، به اعتقاد «لنکستر» میزان مانعیت عبارت است از نسبت تعداد اسناد بازیابی شده مربوط به تعداد کل اسناد بازیابی شده (Lancaster 1979). به گفته دیگر، مانعیت نسبت اسناد بازیابی شده‌ای است که به‌واقع مربوط هستند و میزان آن را به شکل فرمول زیر به درصد محاسبه می‌کنند:

$$\text{تعداد مدارک مرتبط بازیابی شده} \\ \text{دقت} = \frac{\text{تعداد کل مدارک بازیابی شده}}{\text{تعداد مدارک مرتبط بازیابی شده}} \times 100$$

در این پژوهش، هنگام بررسی موتورهای کاوش، برای توصیف و تحلیل داده‌های به‌دست‌آمده، انواع فنون آماری توصیفی و استنباطی به کار گرفته شد. به‌منظور مقایسه بین موتورهای کاوش و توصیف و تحلیل داده‌ها، از آزمون‌های «کلموگروف-اسمیرنوف»، «شاپیرو-ویلک»، «کروسکال-والیس» و «فریدمن» استفاده گردید. برای این منظور از نرم‌افزار آماری «اس‌پی‌اس‌اس» ۲۵ و برای تبیین، نمایش و تفسیر داده‌ها از نرم‌افزار «اورد» و «اکسل» استفاده شد.

۵. تجزیه و تحلیل یافته‌ها

برای پاسخگویی به پرسش‌های پژوهش، پس از وارد کردن واژه‌های موجود در سیاهه محقق ساخته در قسمت جست‌وجوی تصاویر و بررسی پنجاه تصویر نخست بازیابی شده، داده‌ها ثبت و پس از تجزیه و تحلیل آماری نتایج، میزان جامعیت و دقت هر یک از موتورهای کاوش در بخش نوشتاری و معنایی محاسبه و در قالب جدول ۳، و بر حسب درصد ارائه شد.

پرسش اول: میزان جامعیت و دقت در تصاویر بازیابی‌شده توسط موتورهای کاوش مورد مطالعه در این پژوهش چقدر است؟

در پاسخ به این پرسش، با توجه به یافته‌های به‌دست‌آمده در جدول ۳، به تفکیک هر یک از موتورهای کاوش میزان جامعیت و دقت نتایج بازیابی‌شده گزارش شده است.

جدول ۳. توزیع فراوانی و میزان جامعیت و دقت در تصاویر بازیابی‌شده توسط موتورهای کاوش مورد مطالعه

گوگل			بینگ			داک‌داک‌گو			متغیر	مسئله نوشتاری و معنایی
تعداد	تعداد	تعداد	تعداد	تعداد	تعداد	تعداد	تعداد	تعداد		
۱۴	۲۸	۲۸	۱۷	۳۴	۳۴	۱۹	۳۸	۳۸	ضبط واژگان لاتین	تفاوت در معنای واژگان
۲۰	۳۷/۰۴	۴۰	۱۵	۲۸/۷۸	۳۰	۱۹	۳۵/۱۹	۳۸	آنفلوآنزا	
۱۸	۲۹/۵۱	۳۶	۱۹	۳۱/۱۵	۳۸	۲۴	۳۹/۳۴	۴۸	آنفلوآنزا	
۲۱	۶۰	۴۲	۸	۲۲/۸۶	۱۶	۶	۱۷/۱۴	۱۲	تیتانیم	واژگان دخیل و معادل آنها
۱۱	۳۷/۹۳	۲۲	۱۲	۴۱/۳۸	۲۴	۶	۲۰/۶۹	۱۲	تیتانیوم	
۳۶	۵۳/۷۳	۷۲	۱۳	۱۹/۴۰	۲۶	۱۸	۲۶/۸۷	۳۶	کامپیوتر	
۲۸	۶۲/۲۲	۵۶	۷	۱۵/۵۶	۱۴	۱۰	۲۲/۲۲	۲۰	رایانه	اشتقاق واژگان
۴۳	۶۹/۳۵	۸۶	۷	۱۱/۲۹	۱۴	۱۲	۱۹/۳۵	۲۴	پستیچی	
۳۹	۷۹/۵۹	۷۸	۵	۱۰/۲۰	۱۰	۵	۱۰/۲۰	۱۰	پست‌چی	
۳۲	۷۲/۷۳	۶۴	۰	۰	۰	۱۲	۲۷/۲۷	۲۴	کتاب‌خانه (وسیله)	مرکب
۳۴	۸۰/۹۵	۶۸	۱	۲/۳۸	۲	۷	۱۶/۶۷	۱۴	کتابخانه	
۹	۲۷/۲۷	۱۸	۵	۱۵/۱۵	۱۰	۱۹	۵۷/۵۸	۳۸	کتاب‌خانه (مکان)	
۱۵	۱۸/۲۹	۳۰	۲۴	۲۹/۲۷	۴۸	۴۳	۵۲/۴۴	۸۶	کتابخانه	

داک داک گو			بینک			گوگل			متغیر	مسئله نوشتاری و معنایی
ردیف	ک.م.ن	ت.م.ن	ردیف	ک.م.ن	ت.م.ن	ردیف	ک.م.ن	ت.م.ن		
۸	۱۲/۹۰	۴	۶	۹/۶۸	۳	۴۸	۷۷/۴۲	۲۴	دانشسرا	مشق مرکب
۰	۰	۰	۰	۰	۰	۲۲	۱۰۰	۱۱	دانش سرا	
۲	۹/۰۹	۱	۶	۲۷/۲۷	۳	۴۲	۶۵/۶۳	۲۱	فن آوری	
۸	۱۲/۵۰	۴	۱۴	۲۱/۸۸	۷	۱۴	۶۳/۶۴	۷	فناوری	
۵۲	۲۹/۲۱	۲۶	۵۲	۲۹/۲۱	۲۶	۷۴	۴۱/۵۷	۳۷	گلها	علائم جمع
۵۲	۳۴/۲۱	۲۶	۱۸	۱۱/۸۴	۹	۸۲	۵۳/۹۵	۴۱	گل ها	
۰	۰	۰	۶	۳۰	۳	۴۴	۲۶/۸۳	۲۲	معلمین	
۲۰	۵۰	۱۰	۴	۱۰	۲	۴۴	۴۵/۸۳	۲۲	معلمان	
۱۴	۴۳/۷۵	۷	۱۶	۵۰	۸	۳۸	۶۱/۲۹	۱۹	معلمها	
۴	۲۸/۵۷	۲	۴	۲۸/۵۷	۲	۳۲	۳۰/۷۷	۱۶	معلم ها	
۳۲	۱۹/۷۵	۱۶	۶۴	۳۹/۵۱	۳۲	۲۲	۶۴/۷۱	۱۱	روحانیون	
۴۴	۳۰/۵۶	۲۲	۵۶	۳۸/۸۹	۲۸	۱۴	۳۰/۴۳	۷	روحانیان	
۰	۰	۰	۰	۰	۰	۷۶	۵۸/۴۶	۳۸	روحانیها	
۰	۰	۰	۴	۶۶/۶۷	۲	۸۰	۶۳/۴۹	۴۰	روحانی ها	
۷۴	۴۵/۱۲	۳۷	۴۶	۲۸/۰۵	۲۳	۷۰	۸۱/۴۰	۳۵	درختان	
۳۴	۳۵/۴۲	۱۷	۱۸	۱۸/۷۵	۹	۶۶	۴۰/۷۴	۳۳	درخت ها	
۱۲	۱۹/۳۵	۶	۱۲	۱۹/۳۵	۶	۴۴	۳۰/۵۶	۲۲	درختها	
۲۴	۱۸/۴۶	۱۲	۳۰	۲۳/۰۸	۱۵	۲	۱۰۰	۱	سزیجات	
۲۰	۱۵/۸۷	۱۰	۲۶	۲۰/۶۳	۱۳	۲	۳۳/۳۳	۱	سزی ها	
۶	۶/۹۸	۳	۱۰	۱۱/۶۳	۵	۱۴	۷۰	۷	سزیها	
۳۴	۳۲/۶۹	۱۷	۳۸	۳۶/۵۴	۱۹	۱۶	۴۰	۸	حیوانات	
۴	۱۱/۷۶	۲	۸	۲۳/۵۳	۴	۲	۶/۲۵	۱	حیوانها	
۲۰	۴۳/۴۸	۱۰	۱۲	۲۶/۰۹	۶	۶	۴۲/۸۶	۳	حیوان ها	

داک‌داک‌گو			بینگ			گوگل			متغیر	مسئله نوشتاری و معنایی
ردیف	تاریخ و بهیود	نتایج و بهیود	ردیف	تاریخ و بهیود	نتایج و بهیود	ردیف	تاریخ و بهیود	نتایج و بهیود		
۶۲	۳۴/۰۷	۳۱	۲۰	۱۰/۹۹	۱۰	۱۰۰	۵۴/۹۵	۵۰	مدارس	انواع جمع‌ها (جمع‌های فارسی و مکسر)
۲۸	۲۴/۱۴	۱۴	۱۶	۱۳/۷۹	۸	۷۲	۶۲/۰۷	۳۶	مدرسه‌ها	
۲۸	۱۹/۷۲	۱۴	۴۰	۲۸/۱۷	۲۰	۷۴	۵۲/۱۱	۳۷	کتب	
۶	۷/۵۰	۳	۴	۵	۲	۷۰	۸۷/۵۰	۳۵	کتابها	
۸	۱۰	۴	۴	۵	۲	۶۸	۸۵	۳۴	کتاب‌ها	
۳۲	۳۶/۳۶	۱۶	۲۸	۳۱/۸۲	۱۴	۲۸	۳۱/۸۲	۱۴	نذورات	
۴	۵/۴۱	۲	۶	۸/۱۱	۳	۶۴	۸۶/۴۹	۳۲	نذرها	طریقه نگارش الف مقصوره
۵۸	۳۹/۷۳	۲۹	۴۴	۳۰/۱۴	۲۲	۴۴	۳۰/۱۴	۲۲	کسری (طاق)	
۴	۱۸/۱۸	۲	.	.	.	۱۸	۸۱/۸۲	۹	کسرا	
۲۲	۴۰/۷۴	۱۱	۱۲	۲۲/۲۲	۶	۲۰	۳۷/۰۴	۱۰	کسری (شخص)	
۴۸	۳۳/۸۰	۲۴	۵۸	۴۰/۸۵	۲۹	۳۶	۲۵/۳۵	۱۸	کسرا	
۵۲	۳۲/۵۰	۲۶	۲۸	۱۷/۵۰	۱۴	۸۰	۵۰	۴۰	مصلى	
۲۴	۱۷/۶۵	۱۲	۲۴	۱۷/۶۵	۱۲	۸۸	۶۴/۷۱	۴۴	مصلا	واژه‌هایی با شکل نوشتاری یکسان و تلفظ متفاوت
۶۸	۴۳/۰۴	۳۴	۴۴	۲۷/۸۵	۲۲	۴۶	۲۹/۱۱	۲۳	مسکن (خانه)	
.	۲	۱۰۰	۱	مسکن (قرص)	
۶۶	۲۹/۷۳	۳۳	۶۲	۲۷/۹۳	۳۱	۹۴	۴۲/۳۴	۴۷	دایره‌المعارف	استفاده از تاي منقوط
۶۴	۲۹/۰۹	۳۲	۶۲	۲۸/۱۸	۳۱	۹۴	۴۲/۷۳	۴۷	دایره‌المعارف	
۷۸	۳۴/۵۱	۳۹	۵۸	۲۵/۶۶	۲۹	۹۰	۳۹/۸۲	۴۵	تیروئید	نحوه نگارش همزه میانی و پایانی کلمات با کرسی واو، دندان، الف و بدون کرسی
۳۲	۲۱/۰۵	۱۶	۲۴	۱۵/۷۹	۱۲	۹۶	۶۳/۱۶	۴۸	تیروئید	
۵۶	۳۸/۸۹	۲۸	۱۶	۱۱/۱۱	۸	۷۲	۵۰	۳۶	امضاء	
۹۲	۴۲/۲۰	۴۶	۳۰	۱۳/۷۶	۱۵	۹۶	۴۴/۰۴	۴۸	امضا	مؤذن
۴۰	۳۵/۰۹	۲۰	۳۴	۲۹/۸۲	۱۷	۴۰	۳۵/۰۹	۲۰	مؤذن	
۶	۱۵/۸۹	۳	۶	۱۵/۷۹	۳	۲۶	۶۸/۴۲	۱۳	مؤذن	

داک‌داک‌گو			بینگ			گوگل			متغیر	مسئله نوشتاری و معنایی
ردیف	شماره	تاریخ برپا	ردیف	شماره	تاریخ برپا	ردیف	شماره	تاریخ برپا		
۰	۰	۰	۲۶	۲۹/۵۵	۱۳	۶۲	۷۰/۴۵	۳۱	شیء	نوشته نوشتاری
۰	۰	۰	۰	۰	۰	۱۲	۱۰۰	۶	شیء	
۸	۱۲/۹۰	۴	۸	۱۲/۹۰	۴	۴۶	۷۴/۱۹	۲۳	مأمور	
۱۴	۱۳/۴۶	۷	۶	۵/۷۷	۳	۸۴	۸۰/۷۷	۴۲	مامور	
۴۰	۲۸/۹۹	۲۰	۲۸	۲۰/۲۹	۱۴	۷۰	۵۰/۷۲	۳۵	زمرّد	استفاده و عدم استفاده از تشدید
۴۰	۲۸/۱۷	۲۰	۳۰	۲۱/۱۳	۱۵	۷۲	۵۰/۷۰	۳۶	زمرّد	
۴۰	۲۷/۰۳	۲۰	۳۶	۲۴/۳۲	۱۸	۷۲	۴۸/۶۵	۳۶	اعضا بدن	کسرء اضافه و بدل‌های آن
۲۰	۱۹/۲۳	۱۰	۲۴	۲۳/۰۸	۱۲	۶۰	۵۷/۶۹	۳۰	اعضاء بدن	
۲۰	۱۶/۶۷	۱۰	۲۰	۱۶/۶۷	۱۰	۸۰	۶۶/۶۷	۴۰	اعضای بدن	
۰	۰	۰	۰	۰	۰	۲	۱۰۰	۱	واژه‌های دواملائی آذوقه	
۴	۲۵	۲	۱۰	۶۲/۵۰	۵	۲	۱۲/۵۰	۱	(واژه‌هایی با واج آذوقه یا آوای مشترک	
۸۴	۴۲	۴۲	۲۰	۱۰	۱۰	۹۶	۴۸	۴۸	و شکل نوشتاری تهران	
۲۸	۲۶/۴۲	۱۴	۱۰	۹/۴۳	۵	۶۸	۶۴/۱۵	۳۴	طهران	متفاوت)
۳۲	۱۷/۳۹	۱۶	۵۲	۲۸/۲۶	۲۶	۱۰۰	۵۴/۳۵	۵۰	پاییز	جابه‌جایی ی و همزه در کلمات فارسی
۱۸	۱۲	۹	۳۲	۲۱/۳۳	۱۶	۱۰۰	۶۶/۶۷	۵۰	پائیز	
۷۰	۳۸/۸۹	۳۵	۱۰	۵/۵۶	۵	۱۰۰	۵۵/۵۶	۵۰	آئینه	
۷۴	۸۰/۴۳	۳۷	۰	۰	۰	۱۸	۱۹/۵۷	۹	آئینه	
۳۲	۲۱/۳۳	۱۶	۱۸	۱۲	۹	۱۰۰	۶۶/۶۷	۵۰	خانه بازی	نحوء نگارش و غیرملفوظ و میانجی
۲	۶/۲۵	۱	۱۰	۳۱/۲۵	۵	۲۰	۶۲/۵۰	۱۰	خانه ی بازی	
۳۶	۲۶/۰۹	۱۸	۸	۵/۸۰	۴	۹۴	۶۸/۱۲	۴۷	خانه بازی	
۲۰	۲۱/۷۴	۱۰	۴	۴/۳۵	۲	۶۸	۷۳/۹۱	۳۴	خونه	استفاده از زبان محاوره
۱۰۰	۵۱/۵۵	۵۰	۲۴	۱۲/۳۷	۱۲	۷۰	۳۶/۰۸	۳۵	خانه	

داک‌داک‌گو			بینگ			گوگل			متغیر	مسئله نوشتاری و معنایی
ردیف	تاریخ	تعداد	ردیف	تاریخ	تعداد	ردیف	تاریخ	تعداد		
۲۴	۲۷/۹۱	۱۲	۲۲	۲۵/۵۸	۱۱	۴۰	۴۶/۵۱	۲۰	پشاهنگ	کاربرد و حذف مد در کلمات فارسی
۱۴	۱۸/۴۲	۷	۱۶	۲۱/۰۵	۸	۴۶	۶۰/۵۳	۲۳	پشاهنگ	
۱۴	۳۱/۸۲	۷	۱۰	۲۲/۷۳	۵	۲۰	۴۵/۴۵	۱۰	خوراسان	نحوه نگارش واو معدوله
۳۶	۲۶/۸۷	۱۸	۸	۵/۹۷	۴	۹۰	۶۷/۱۶	۴۵	خراسان	
۲۲	۱۸/۶۴	۱۱	۱۶	۱۳/۵۶	۸	۸۰	۶۷/۸۰	۴۰	خانه‌ایی	نحوه نگارش یای وحدت نکره بعد از های مخفی
۱۴	۲۱/۲۱	۷	۱۶	۲۴/۲۴	۸	۳۶	۵۴/۵۵	۱۸	خانه بی	
۸۴	۴۱/۵۸	۴۲	۳۲	۱۵/۸۴	۱۶	۸۶	۴۲/۵۷	۴۳	خانه	
۷۲	۴۱/۳۸	۳۶	۴۲	۲۴/۱۴	۲۱	۶۰	۳۴/۴۸	۳۰	اسحاق	نحوه نگارش الف مستوره
۶۴	۴۷/۰۶	۳۲	۴۸	۳۵/۲۹	۲۴	۲۴	۱۷/۶۵	۱۲	اسحق	
۶۰	۳۶/۵۹	۳۰	۴	۲/۴۴	۲	۱۰۰	۶۰/۹۸	۵۰	شیر (حیوان)	همنامی یا واژه‌های یکسان با معانی متفاوت (واژگان مشترک شیر (لبنی))
۲	۱۰۰	۱	۰	۰	۰	۰	۰	۰	شیر آب	
۴	۱۰۰	۲	۰	۰	۰	۰	۰	۰	شیر (لبنی)	
۱۶	۱۶/۶۷	۸	۲	۲/۰۸	۱	۷۸	۸۱/۲۵	۳۹	شور (طعم)	لفظی / هم‌آوا و هم‌نویسه
۰	۰	۰	۴	۱۰۰	۲	۰	۰	۰	شور و شوق	
۰	۰	۰	۰	۰	۰	۲	۱۰۰	۱	کره (لبنی)	واژگان هم‌نویسه با معانی متفاوت (واژگان مشترک کره زمین)
۰	۰	۰	۰	۰	۰	۲۰	۱۰۰	۱۰	کره زمین	
۰	۰	۰	۰	۰	۰	۳۰	۱۰۰	۱۵	کره (جغرافیا)	نوشتاری / هم‌نویسه و با آوای متفاوت
۰	۰	۰	۰	۰	۰	۸	۱۰۰	۴	کره (کشور)	
۴۴	۲۹/۳۳	۲۲	۳۴	۲۲/۶۷	۱۷	۷۲	۴۸	۳۶	مرد (جنسیت)	
۰	۰	۰	۲	۱۰۰	۱	۰	۰	۰	مرد (مردن)	
۳۴	۲۴/۶۴	۱۷	۴۸	۳۴/۷۸	۲۴	۵۶	۴۰/۵۸	۲۸	دریای خزر	هم‌معنایی و مترادف
۸	۹/۵۲	۴	۲۸	۳۳/۳۳	۱۴	۴۸	۵۸/۱۴	۲۴	دریای کاسپین	
۲۴	۲۰/۶۹	۱۲	۲۴	۲۰/۶۹	۱۲	۶۸	۵۸/۶۲	۳۴	دریای مازندران	
۲۸/۴۶	۲۵/۱۹	۱۴/۲۳	۱۹/۸۸	۲۰/۳۸	۱۰	۵۰/۶۹	۵۴/۴۳	۲۵/۳۵		میانگین کل

به‌طور کلی، میزان جامعیت موتور کاوش «گوگل» ۵۴/۴۲ درصد و دقت آن ۵۰/۶۹ درصد در بازیابی تصاویر گزارش شد. بنابراین، می‌توان نتیجه گرفت که میزان بازیافت (جامعیت) موتور کاوش «گوگل» نسبت به دقت آن بیشتر است و این مطلب نشان‌دهنده پایین بودن دقت این موتور کاوش در بازیابی تصاویر به زبان فارسی است.

مطابق با یافته‌های به‌دست آمده، به‌طور کلی، میزان جامعیت موتور کاوش «بینگ» ۲۰/۳۸ درصد و دقت آن ۱۹/۸۸ درصد در بازیابی تصاویر گزارش شد. بنابراین، می‌توان نتیجه گرفت که اگرچه میزان جامعیت و دقت این موتور کاوش مناسب ارزیابی نشد، اما به‌طور کلی، میزان بازیافت (جامعیت) موتور کاوش «بینگ» نسبت به دقت آن بیشتر است و این مطلب نشان‌دهنده پایین بودن دقت این موتور کاوش در بازیابی تصاویر به زبان فارسی است. به نظر می‌رسد، که این وضعیت به احتمال زیاد به دلیل عدم پوشش گسترده صفحات سایر زبان‌ها نظیر زبان فارسی در پایگاه داده‌های «بینگ» یا ناشی از ضعف الگوریتم‌های به‌کاررفته توسط ربات خزنده این موتور کاوش در جست‌وجو و بازیابی پیوندهای مرتبط و نمایه‌سازی آن‌ها و یا فاصله روزآمدسازی پایگاه داده‌ها در پایگاه خود باشد. این در حالی است که موتور کاوش «گوگل»، تصاویر مرتبط بیشتری را از همان واژه‌ها بازیابی می‌کند و احتمالاً در سایر امکانات مربوط به ذخیره و بازیابی اطلاعات نیز کامل‌تر است.

در مورد موتور کاوش «داک‌داک‌گو» نیز به‌طور کلی، میزان جامعیت این موتور کاوش ۲۵/۱۸ درصد و دقت آن ۲۸/۴۶ درصد در بازیابی تصاویر گزارش شد. با وجود پایین بودن میزان جامعیت و دقت این موتور کاوش، به‌طور کلی، می‌توان نتیجه گرفت که میزان دقت موتور کاوش «داک‌داک‌گو» نسبت به بازیافت آن در بازیابی تصاویر به زبان فارسی بیشتر است.

پرسش دوم: آیا بین موتورهای کاوش مورد مطالعه در میزان جامعیت و دقت تفاوت معناداری وجود دارد؟

ابتدا به‌منظور مقایسه میانگین سه موتور کاوش مورد پژوهش، نرمال بودن توزیع متغیرهای جامعیت و مانعیت بررسی شدند.

جدول ۴. نرمال بودن متغیر جامعیت و دقت در موتورهای کاوش مورد پژوهش

نوع متغیر	کولموگروف اسمیرنوف				شاپیرو-ویلک	
	سطح معناداری	درجه آزادی	آماره آزمون	سطح معناداری	درجه آزادی	آماره آزمون
جامعیت	۰	۳۱۲	۰/۱۰۴	۰	۳۱۲	۰/۹۲۶
دقت	۰	۳۱۲	۰/۱۲۵	۰	۳۱۲	۰/۹۰۹

بر پایه جدول ۴، با توجه به این که مقدار احتمال دو آزمون «کولموگروف-اسمیرنوف» و «شاپیرو-ویلک» کمتر از ۰/۰۵ محاسبه شده، بنابراین فرض نرمال بودن متغیرها رد می‌شود و برای بررسی متغیرها باید از روش‌های ناپارامتری استفاده کرد.

جدول ۵. نتایج آزمون «کراسکال وایس» مربوط به جامعیت و دقت موتورهای کاوش مورد پژوهش در بخش نوشتاری و معنایی

کروسکال-وایس	درجه آزادی	سطح معناداری	
۱۱۱/۸۵۲	۲	۰	جامعیت
۵۳/۴۵۷	۲	۰	دقت

با توجه نرمال نبودن توزیع متغیر جامعیت و دقت، برای مقایسه آنها بین موتورهای کاوش از آزمون ناپارامتری «کروسکال-وایس» استفاده شد. بر پایه جدول ۵، با توجه به این که مقدار احتمال آزمون در متغیر جامعیت (۰/۰۰) کوچک‌تر از ۰/۰۵ محاسبه شده، بنابراین بین موتورهای کاوش مورد مطالعه در میزان جامعیت آنها تفاوت معناداری وجود دارد. همچنین، با توجه به این که مقدار احتمال آزمون در متغیر دقت (۰/۰۰) کوچک‌تر از ۰/۰۵ محاسبه شده، بنابراین بین موتورهای کاوش مورد مطالعه در میزان دقت آنها نیز تفاوت معناداری وجود دارد.

پرسش سوم: کدام یک از موتورهای کاوش مورد مطالعه عملکرد بهتری در بازیابی تصاویر به زبان فارسی دارند؟

با استفاده از آزمون «فیدمن» و گرفتن رتبه میانگین و میانگین جامعیت و دقت واژه‌ها مشخص شد که کدام موتور کاوش در بازیابی تصاویر به زبان فارسی بر پایه سیاهه ارزیابی پژوهش از جامعیت و دقت بالاتری برخوردار است. در مجموع، بررسی نتایج از نظر جامعیت در دو بخش نوشتاری و معنایی در جدول ۴، و از نظر دقت نیز در جدول ۵، ارائه شده است.

جدول ۴. نتایج آزمون «فریدمن» مربوط به جامعیت موتورهای کاوش، در بازیابی کلیدواژه‌های مورد پژوهش در بخش نوشتاری و معنایی

	گوگل	داک‌داک‌گو	بینگ
بخش نوشتاری	۳۴	۳	۰
بخش معنایی	۲	۱	۲

با استفاده از آزمون «فریدمن» و گرفتن رتبه میانگین و میانگین جامعیت و دقت واژه‌ها مشخص شد که کدام موتور کاوش در بازیابی تصاویر به زبان فارسی بر پایه سیاهه ارزیابی پژوهش، از جامعیت و دقت بالاتری برخوردار است. در مجموع، بررسی نتایج از نظر جامعیت در دو بخش نوشتاری و معنایی نشان می‌دهد که موتور کاوش «گوگل» با ۳۴ مورد برتری در بخش نوشتاری و ۲ مورد برتری در بخش معنایی در رتبه اول و موتور کاوش «داک‌داک‌گو» با ۳ مورد برتری در بخش نوشتاری و ۱ مورد برتری در بخش معنایی در رتبه دوم و موتور کاوش «بینگ» تنها با ۲ مورد برتری در بخش معنایی در رتبه سوم قرار دارند.

جدول ۵. نتایج آزمون «فریدمن» مربوط به دقت موتورهای کاوش در بازیابی کلیدواژه‌های مورد پژوهش در بخش نوشتاری و معنایی

	گوگل	داک‌داک‌گو	بینگ
بخش نوشتاری	۲۷	۹	۱
بخش معنایی	۴	۱	۰

همچنین از نظر دقت نیز موتور کاوش «گوگل» با ۲۷ مورد برتری در بخش نوشتاری و ۴ مورد برتری در بخش معنایی در رتبه اول و موتور کاوش «داک‌داک‌گو» با ۹ مورد برتری در بخش نوشتاری و ۱ مورد برتری در بخش معنایی در رتبه دوم و موتور کاوش «بینگ» تنها با ۱ مورد برتری در بخش نوشتاری در رتبه سوم قرار دارند.

با توجه به نتایج به دست آمده می‌توان اظهار کرد که موتور کاوش «گوگل» در تعداد بیشتری از ویژگی‌های نوشتاری و معنایی نسبت به «داک‌داک‌گو» و «بینگ» برتری به دست آورد و موتور کاوش «داک‌داک‌گو» و «بینگ» به ترتیب، در رتبه‌های بعدی قرار دارند. بنابراین، موتور کاوش «گوگل» نسبت به سایر موتورهای کاوش در پژوهش حاضر، عملکرد بهتری در بازیابی تصاویر به زبان فارسی دارد.

۶. نتیجه‌گیری

با توجه به یافته‌های پژوهش، به‌طور کلی، موتورهای کاوش مورد پژوهش، در جست‌وجو و بازیابی تصاویر به زبان فارسی عملکرد موفق و قابل قبولی ندارند و ویژگی‌های نوشتاری و معنایی زبان فارسی را نادیده می‌گیرند؛ به گونه‌ای که هر یک از ویژگی‌های نوشتاری و معنایی زبان فارسی که در این پژوهش مورد مطالعه قرار گرفت، کاربران فارسی زبان را در جست‌وجو و بازیابی اطلاعات از موتورهای کاوش مورد پژوهش با چالش‌هایی مواجه می‌سازند. در بخش ویژگی‌های نوشتاری، نتایج پژوهش حاضر نشان داد که مشکلاتی نظیر پیوسته‌نویسی و جدانویسی واژگان، واژگان دخیل و معادل آن‌ها، استفاده از زبان محاوره و کلمات عربی موجود در زبان فارسی مانند انواع جمع‌ها و انواع شیوه‌های نگارش همزه، مهم‌ترین مسائل و مشکلات کاربران فارسی زبان در استفاده از محیط وب، به‌ویژه در جست‌وجو و بازیابی تصاویر در موتورهای کاوش هستند. در بخش معنایی نیز با توجه به نتایج پژوهش حاضر می‌توان اظهار کرد که ویژگی‌های معنایی زبان فارسی در موتورهای کاوش معمولی و معنایی مورد پژوهش مهجور مانده، در حالی که ابعاد معنایی زبان فارسی نقش مهمی در جست‌وجو و بازیابی اطلاعات مورد نظر کاربران از موتورهای کاوش دارد و عدم توجه به این ویژگی می‌تواند سبب شکل‌گیری جست‌وجوی ناموفق در موتورهای کاوش و عدم دسترسی کاربر به اطلاعات مورد نیاز خود و ایجاد مشکلاتی از جمله بازیافت نتایج نامربوط و ریزش کاذب و در نتیجه، عدم رضایت او از عملکرد موتورهای کاوش گردد. در مجموع، با توجه به یافته‌های پژوهش حاضر و سایر پژوهش‌های مرتبط نظیر پژوهش «هماوندی» (۱۳۹۲)، در بخش معنایی، ویژگی‌هایی نظیر همانمی، مترادف و هم‌نویسی، مشکلات جدی کاربران در استفاده از موتورهای کاوش هستند.

آنچه ذکر شد، بخشی از مهم‌ترین مسائل و ویژگی‌های زبان فارسی است که کم‌توجهی به آن‌ها توسط کاربران و طراحان موتورهای کاوش می‌تواند منجر به بروز مشکلاتی در جست‌وجو و بازیابی انواع اطلاعات به‌ویژه تصاویر شود. به‌طور کلی، نتایج این پژوهش نشان داد که میزان جامعیت موتور کاوش «گوگل» نسبت به دقت آن بیشتر است و این مطلب نشان‌دهنده پایین بودن دقت این موتور کاوش در بازیابی تصاویر به زبان فارسی است. این یافته‌ها با پژوهش «حریری و وکیلی مفرد» (۱۳۹۲) که موتور کاوش «گوگل» با دقت ۹۰ درصد دارای بیشترین دقت در بازیابی تصاویر مرتبط و کاملاً

مرتبط در بازیابی تصاویر پزشکی بود، همسویی ندارد، اما با نتایج پژوهش «احمدی و چشمه‌سهرابی» (۱۳۹۱) که گزارش کردند در میان نظام‌های جست‌وجوی فازی، «گوگل» بیشترین جامعیت را در بین سایرین دارد، و نتایج پژوهش «دیزجی» و همکاران (۱۳۹۵) مبنی بر این که موتور کاوش گوگل برای کاربرانی که به دنبال جامعیت بیشتری هستند گزینه مناسبی است و عملکرد بهتری نسبت به موتور کاوش «بینگ» دارد، همسویی دارد. همچنین در مورد موتور کاوش «بینگ» یافته‌های این پژوهش با بخشی از یافته‌های «جلالی دیزجی» و همکاران (۱۳۹۵) که گزارش کردند موتور کاوش «بینگ» برای کاربرانی که به دنبال دقت بیشتری در بازیابی هستند ارجح است، همسویی ندارد. در موتور کاوش «داک‌داک‌گو» نیز یافته‌ها حاکی از آن است که میزان دقت این موتور کاوش نسبت به جامعیت آن در بازیابی تصاویر به زبان فارسی بیشتر است. این یافته‌ها با نتایج پژوهش Negi & Kumar (2014) مبنی بر این که موتور کاوش «گوگل» و «هاکیا» عملکرد بهتری در پردازش پرس‌وجوهای زبان طبیعی ارائه دادند، در حالی که، «بینگ» نتایج مرتبط بیشتری را بازیابی کرد و موتور کاوش «داک‌داک‌گو» عملکرد قابل قبولی از خود نشان نداد، همسویی دارد. نتایج پژوهش حاضر در آزمون «کراسکال‌والیس» نیز نشان داد که بین موتورهای کاوش مورد مطالعه در میزان جامعیت و دقت آن‌ها تفاوت معناداری وجود دارد که با بخشی از نتایج پژوهش «جلالی دیزجی، خالقی و طاهری» (۱۳۹۵) مبنی بر این که بین میزان جامعیت دو موتور کاوش «گوگل» و «بینگ» تفاوت معناداری وجود دارد، همسوست.

در مجموع، با توجه به یافته‌های این پژوهش می‌توان گفت که موتورهای کاوش معمولی و حتی موتورهای کاوش معنایی که مدعی فراهم‌آوری اطلاعات باکیفیت‌تر و مرتبط‌تری نسبت به دیگر موتورهای کاوش هستند، به ویژگی‌های نوشتاری و معنایی زبان فارسی توجه کافی ندارند؛ به طوری که از بین سه موتور کاوش مورد پژوهش، تنها موتور کاوش «گوگل» از نظر ربط (میزان جامعیت و دقت) تصاویر بازیابی شده، عملکرد بهتری نسبت به دیگر موتورهای کاوش از خود نشان داد، اما باز هم مشکلاتی در بازیابی تصاویر به زبان فارسی در این موتور کاوش دیده می‌شود. یافته‌های ذکر شده با نتایج حاصل از پژوهش Norouzi and Homavandi (2018) که نشان داد در بخش معنایی و نوشتاری، قابلیت‌های «گوگل» در انطباق با زبان فارسی نسبت به همتایانش بیشتر است، همسویی دارد و با یافته‌های حاصل از پژوهش «یقطین و جوکار» (۱۳۹۶) مبنی بر این که موتور کاوش «گوگل» به لحاظ بازیابی تعداد نتایج مرتبط با زبان هدف، عملکرد ضعیف‌تری نسبت

به موتور کاوش «بینگ» و «یاهو» داشته است، و با یافته‌های پژوهش «ابوالقاسم مسلمان، مؤمنی و حاجی زین‌العابدینی» (۱۳۹۴) مبنی بر این که موتور کاوش «یاهو» بیشترین مقدار دقت را نسبت به موتورهای کاوش «گوگل»، «پیک‌سرچ» و «فلیکر» به خود اختصاص داده و عملکرد بهتری نشان داده است و همچنین، با نتایج یافته‌های CheshmehSohrabi and Adnani Sadati (2021) در خصوص این که از نظر جامعیت و دقت بین موتورهای کاوش «یاهو» و «گوگل» و «داک‌داک‌گو» رقابت تنگاتنگی وجود دارد، اما موتور کاوش «یاهو» از سایر موتورهای کاوش برتر است، همسویی ندارد.

همچنین، با توجه به یافته‌های پژوهش حاضر، بخشی از این مسائل را می‌توان به دلیل عدم پوشش گسترده صفحات سایر زبان‌ها نظیر زبان فارسی در پایگاه داده‌های موتورهای کاوش یا ناشی از ضعف الگوریتم‌های به کاررفته توسط ربات خزنده موتورهای کاوش در جست‌وجو و بازیابی پیوندهای مرتبط و به مرحله‌نمایه‌سازی مربوط دانست که باید طی آن ویژگی‌های زبان فارسی را مد نظر قرار داد و با اقداماتی نظیر ایجاد یکپارچگی و هماهنگی در شیوه‌های نگارش و کاربرد واژگان در محیط وب، استفاده از اصطلاحنامه‌ها و آگاه نمودن کاربران و جست‌وجوگران از بروز چنین مشکلاتی جلوگیری نمود. همچنین، بررسی و شناسایی نیازهای اطلاعاتی دیداری کاربران، کاهش شکاف معنایی بین ویژگی‌های سطح پایین تصاویر و محتوای معنایی سطح بالای تصاویر، استانداردسازی توصیف تصاویر و ارائه فنون مناسب نمایه‌سازی و بازیابی تصاویر با استفاده از فنون دو روش محتوا محور و مفهوم محور تصاویر از جمله مواردی که لازم است بیشتر مورد توجه متخصصان نمایه‌سازی و بازیابی تصاویر قرار گیرد. احتیاج فارسی‌زبانان، به‌ویژه به موتورهای کاوش بومی که مبتنی بر ویژگی‌های زبانی خودشان طراحی شده و در تعامل با کاربران فارسی‌زبان به ظرایف و باید و نبایدهای زبانی آن‌ها بیشتر توجه شده، بیش از پیش احساس می‌گردد. از سوی دیگر، موتورهای کاوش «گوگل» و «بینگ» و «داک‌داک‌گو» که جزء موتورهای کاوش محبوب و پُر استفاده هستند، نیز می‌بایست جهت فراهم‌ساختن نیازهای زبانی کاربران و اصلاح تعامل با کاربران تلاش بیشتری نمایند؛ چرا که طبق آمارهای جهانی ۷۴/۱ درصد از کاربران اینترنت را افراد غیرانگلیسی‌زبان تشکیل می‌دهند. از جمله این زبان‌ها می‌توان به زبان فارسی اشاره کرد که در تعدادی از کشورها به‌جز ایران از این زبان استفاده می‌کنند و نادیده انگاشتن و یا کم‌توجهی به شاخصه‌ها و ویژگی‌های زبانی کاربران از سوی موتورهای کاوش، موجب بروز مسائلی

در امر جست‌وجو و بازیابی اطلاعات می‌شود که در نهایت، از دست رفتن اطلاعات مفید و یا بازیابی اطلاعات ناخواسته را به همراه خواهد داشت. همچنین، نمایه‌سازی تکه‌های بزرگ داده از جمله مجموعه تصاویر موجود در وب، یک چالش است و موتورهای کاوش برای پیاده‌سازی آن‌ها به فناوری وب معنایی نیاز دارند. در حال حاضر، موتورهای کاوش، به‌ویژه موتورهای کاوش معنایی از فناوری‌های معنایی استفاده می‌کنند، اما نه به‌طور کامل؛ و این موضوع هنوز به‌صورت یک مشکل اساسی باقی مانده است. این مسئله به‌ویژه در مورد زبان‌های غیرانگلیسی بیشتر مشهود است. موتورهای کاوش معنایی بر خلاف ادعای خود، توجه زیادی به فراهم‌آوری امکانات و قابلیت‌های جست‌وجو و بازیابی اطلاعات برای کاربران ندارند. در حالی که موتور کاوش «گوگل» این امر مهم را به نحو احسن انجام می‌دهد. بنابراین، موتورهای کاوش معنایی باید مسیری طولانی طی کنند تا به یک موتور کاوش معنایی واقعی تبدیل شوند.

با توجه به یافته‌های پژوهش، پیشنهادهای زیر می‌تواند جهت رفع چالش‌های پیش روی کاربران فارسی‌زبان در استفاده از موتورهای کاوش راهگشا باشد:

◇ یافته‌های پژوهش نشان داد که موتورهای کاوش مورد بررسی توجه زیادی به فراهم‌آوری امکانات و قابلیت‌های جست‌وجو و بازیابی اطلاعات ندارند. بنابراین، از طریق افزودن برخی گزینه‌های جست‌وجو، بازیابی، و نمایش، مانند محدودگرها، ارائه پیشنهاد برای جست‌وجو (مانند آنچه برای زبان انگلیسی تا حدودی فراهم شده)، راهنمای کمکی و غیره در موتورهای کاوش انتخابی می‌توان کیفیت نتایج جست‌وجو را بهبود بخشید.

◇ با توجه به لزوم ایجاد یکپارچگی و هماهنگی در شیوه‌های نگارش در محیط وب وجود یک نرم‌افزار واژه‌پرداز منطبق با ویژگی‌های زبانی یا دست کم استفاده از افزونه‌های ویراستار فارسی در محیط واژه‌پردازهای رایج می‌تواند راه‌حلی مناسب برای ایجاد یکدستی متون از ابتدای خلق آن‌ها در محیط رایانه‌ای باشد.

◇ با توجه به این که بخش عمده مشکلات شناسایی شده در این پژوهش از ویژگی‌های نوشتاری و تنوع سبک نوشتار در محیط وب ناشی می‌شود، لزوم تدوین شیوه‌نامه‌ای برای سبک نوشتار در محیط وب، برای کاهش چندگانگی نوشتاری و کمک به موتورهای کاوش در بازیابی تصاویر مورد نظر کاربران احساس می‌شود.

◇ تدوین شیوه‌نامه آموزشی برای جست‌وجو و بازیابی اطلاعات مبنی بر ویژگی‌های

زبان فارسی و در نظر گرفتن مشکلات شناسایی‌شده در این پژوهش و پژوهش‌های مشابه می‌تواند به کاربران فارسی‌زبان در استفاده و بازیابی موفق‌تر از موتورهای کاوش یاری رساند.

◇ برای آگاه‌سازی برنامه‌نویسان در محیط اینترنت، تدوین یک شیوه‌نامه آموزشی برای سازماندهی و نمایه‌سازی اطلاعات، با تأکید بر مشکلات عمده شناسایی‌شده در این پژوهش می‌تواند راهگشا باشد و باعث سازماندهی و نمایه‌سازی کارآمدتر شود.

◇ با توجه به تبیین شیوه‌نامه‌ای از سوی «سازمان استاندارد جهانی» برای سازماندهی هرچه بهتر اطلاعات، به‌ویژه تصاویر به شکل معنایی، در سال‌های اخیر توجه به سازماندهی معنایی اطلاعات معطوف شده، اما هنوز به نتیجه قابل قبولی منتهی نشده است. با توجه به این که در عصر فرادیداری هستیم و اطلاعات به شکل تصویری می‌تواند قابل فهم‌تر ارائه گردد، متولیان موتورهای کاوش، به‌ویژه موتورهای کاوش معنایی می‌بایست نسبت به سازماندهی معنایی تصاویر به‌خصوص در زبان‌های غیرانگلیسی توجه بیشتری داشته باشند.

◇ در مواردی مانند واژگان دخیل و معادل آن‌ها که در موتورهای کاوش مورد پژوهش توجه لازم به آن صورت نگرفته است، استفاده از واژگان معادل و تلاش برای رواج آن از طریق رسانه‌های مکتوب و غیرمکتوب، به‌ویژه رسانه‌های رسمی از ابتدای ورود می‌تواند در پذیرش و کاربرد گسترده آن توسط افراد مؤثر بوده و سرانجام منجر به جلوگیری از رواج چندین شکل متفاوت از یک واژه شود.

◇ در مورد مسئله تنوع صورت‌های نوشتاری واژگان، موتورهای کاوش می‌توانند از طریق فراهم ساختن نمایه‌های مناسب و استفاده از اصطلاحنامه، کاربران را از وجود صورت‌های مختلف نوشتاری یک واژه آگاه سازند و در صورت لزوم به آن‌ها ارجاع دهند. در واقع، موتورهای کاوش می‌توانند سازوکاری جهت آگاهی‌دادن به کاربر در هنگام جست‌وجو، یعنی نمایش مترادفات، صورت‌های املائی گوناگون و ریشه‌یابی کلمات در مورد مشکلات نوشتاری و معنایی عمده که بر بازیابی کارآمد کلیدواژه تأثیر دارند، با درج کلیدواژه توسط کاربر در قسمت جست‌وجو (همان‌طور که در مورد زبان انگلیسی این امکان تاحدودی فراهم شده) ایجاد کنند.

◇ ایجاد سامانه‌ای مختص بخش مشکلات معنایی زبان و قسمت جست‌وجوی تصاویر، بدین صورت که با درج کلیدواژه توسط کاربر، فهرستی از شکل‌های حاوی معانی

گونگون واژه برای کاربر نمایش داده شود، ضروری است. به‌عنوان مثال، با درج واژه «کره» در کادر جست‌وجوی تصویر موتور کاوش، لیستی از تصاویر حاوی معانی مختلف این واژه (کشور کره، کره جغرافیایی، کره خوراکی) ظاهر شود و ضمن آگاه‌ساختن کاربر، امکان انجام جست‌وجوی مناسب و بدون از دست دادن نتایج را برای وی فراهم آورد.

◇ در انتها پیشنهاد می‌شود با توجه به نتایج پژوهش حاضر و پژوهش‌های مشابه، سیاست‌های کلانی از سوی نهادهای مسئول در کشور نظیر «فرهنگستان زبان و ادب فارسی»، «وزارت فرهنگ و ارشاد اسلامی»، و «کتابخانه ملی» جهت رعایت یکپارچگی و هماهنگی در متون تولیدشده به‌صورت چاپی و الکترونیکی برنامه‌ریزی و اجرایی شود، تا ضمن حراست از زبان غنی فارسی، از بروز برخی چالش‌های یادشده جلوگیری شده و به کاربران فارسی‌زبان در بازیابی اطلاعات خود به شایستگی کمک گردد.

فهرست منابع

- ابوالقاسم مسلمان، طاهره، عصمت مؤمنی، و محسن حاجی زین‌العابدینی. ۱۳۹۴. سنجش میزان ربط تصاویر بازیابی‌شده در موتورهای جست‌وجوی یاهو، گوگل، پیک‌سرچ و فلیکر و ارائه الگوی پیشنهادی بهینه. *دانش‌شناسی* ۸ (۲۸): ۱-۱۲.
- احمدی، فاطمه، و مظفر چشمه‌سهرابی. ۱۳۹۱. مقایسه میزان جامعیت و مانعیت در ابزارهای کاوش اطلاعات فازی و غیرفازی. *نظام‌ها و خدمات اطلاعاتی* ۱ (۳): ۷۵-۸۸.
- بهمن‌آبادی، علیرضا. ۱۳۸۱. *مبانی تاریخچه و فلسفه علم اطلاع‌رسانی (گزیده مقالات)*. تهران: سازمان اسناد و کتابخانه ملی جمهوری اسلامی ایران.
- پائو، میراندا لی. ۱۹۸۹. *مفاهیم بازیابی اطلاعات*. ترجمه اسدالله آزاد و رحمت‌الله فتاحی. ۱۳۸۷. مشهد: انتشارات دانشگاه فردوسی.
- جلالی دیزجی، علی، رضا خالقی، و مهدی طاهری. ۱۳۹۵. میزان جامعیت و مانعیت بازیابی تصاویر در موتورهای کاوش Bing و Google Image. *نظام‌ها و خدمات اطلاعاتی* ۵ (۳۴): ۳۳-۴۲.
- حرری، عباس. ۱۳۷۲. کامپیوتر و رسم الخط فارسی. *تحقیقات اطلاع‌رسانی و کتابخانه‌های عمومی*. ۳ (۱): ۶-۱۱. <http://publiij.ir/article-1-488-fa.html> (دسترسی در ۱۴۰۰/۲/۱۵)
- حریری، نجلا، و حسین و کیلی مفرد. ۱۳۹۲. مقایسه دقت موتورهای کاوش عمومی و تخصصی پزشکی در بازیابی تصاویر پزشکی. *مدیریت اطلاعات سلامت* ۱۰ (۶): ۸۳۰-۸۳۹.

خادمی‌زاده، شهناز، فریده عصاره، و خدیجه مبینی. ۱۳۹۷. بررسی مقایسه‌ای تأثیر نمایه‌سازی متن‌محور و فولکسونومی تصاویر بر میزان بازیابی تصاویر در موتور کاوش گوگل. *تعامل انسان و اطلاعات* ۵ (۳): ۱۵-۲۷.

خبرگزاری مهر (۱۴ مرداد ۱۳۹۹). آخرین آمار از ضریب نفوذ اینترنت در ایران، بازیابی شده در ۱۳۹۹/۵/۳۰، از <https://www.mehrnews.com/news/4990207/%D8%A2%D8%AE%D8%B1%DB%8C%D9%86%D8%A2%D9%85%D8%A7%D8%B1%D8%B6%D8%B1%DB%8C%D8%A8%D9%86%D9%81%D9%88%D8%B0%D8%A7%DB%8C%D9%86%D8%AA%D8%B1%D9%86%D8%AA%D8%AF%D8%B1%D8%A7%DB%8C%D8%B1%D8%A7%D9%86>

دری، راحله. ۱۳۹۳. مقایسه و ارزیابی موتورهای جست‌وجوی معنایی. *پژوهشنامه پردازش و مدیریت اطلاعات* ۳۰ (۲): ۴۶۷-۴۸۷.

رجبی، سید منصور و یعقوب نوروزی. ۱۳۹۴. موتورهای جست‌وجوی فارسی: ارزیابی امکانات جست‌وجو، بازیابی اطلاعات، میزان جامعیت و مانعیت، و تعیین همپوشانی میان آن‌ها. *فصلنامه مطالعات ملی کتابداری و سازماندهی اطلاعات* ۲۶ (۳): ۱۳۳-۱۵۰.

عباسی دشتکی، ندا، و مظفر چشمه‌سهرابی. ۱۳۹۸. عملکرد موتورهای جست‌وجوی گوگل، یاهو و بینگ در بازیابی اطلاعات فارسی. *مطالعات ملی کتابداری و سازماندهی اطلاعات* ۳۰ (۲): ۹۶-۱۱۱.

کوکبی، مرتضی، صالح رحیمی، فریده عصاره، و علیرضا نوروزی. ۱۳۹۲. تأملاتی بر نمایه‌سازی تصاویر: یک تصویر ارزشی برابر با هزار واژه. *تحقیقات اطلاع‌رسانی و کتابخانه‌های عمومی* ۱۹ (۲): ۲۵۷-۲۷۶.

نوتس، گری. ۱۹۹۷. راهبردها و شیوه‌های جست‌وجو در اینترنت. ترجمه سیمین نیازی. ۱۳۸۲. *فصلنامه مطالعات ملی کتابداری و سازماندهی اطلاعات* ۱۴ (۳): ۱۲۶-۱۳۳.

نوروزی، یعقوب، سمانه خویدکی و زهرا صدرآبادی. ۱۴۰۰. پژوهش‌های حیطه ربط در ایران: مرور نظام‌مند. *پژوهشنامه پردازش و مدیریت اطلاعات* ۳۶ (۳): ۶۲۱-۶۵۴.

نوروزی، یعقوب، و فرشته سوری. ۱۳۹۳. ارزیابی موتورهای جست‌وجو در بازیابی تصاویر بر پایه نمایه‌سازی مبتنی بر متن و محتوا. *کتابداری و اطلاع‌رسانی* ۱۷ (۱): ۱۸۳-۲۰۵.

هماوندی، هدی. ۱۳۹۲. بررسی مشکلات جست‌وجو و بازیابی تصاویر در موتورهای کاوش برگزیده مبتنی بر ویژگی‌های نگارشی زبان فارسی. پایان‌نامه کارشناسی ارشد. دانشگاه قم.

یقظین، مریم، و عبدالرسول جوکار. ۱۳۹۶. کارآمدی فیلد زبان در موتورهای کاوش بین‌المللی برای بازیابی نتایج مرتبط با زبان فارسی و عربی (مطالعه مقایسه‌ای: گوگل، بینگ، و یاهو). *مطالعات ملی کتابداری و سازماندهی اطلاعات* ۲۸ (۲): ۱۴۱-۱۷۵.

References

- Bute, M. S., A. Hussaini, & I. H. Adamu 2017 .. Evaluating search effectiveness of some selected search engines. *Global Journal of Pure and Applied Sciences* 23 (1): 139-149.
- CheshmehSohrabi, M. (Mozaffar), and E. Adnani Sadati. 2021. Performance evaluation of web search engines in image retrieval: An experimental study. *Information Development* 138 (4): 522-534. Retrieved from https://doi.org/10.1177/02666666_9211010211 (May 25, 2021)
- Clarke, S., and P. Willett .1997. Estimating the recall performance of search service performance. *Aslib Proceedings* 49 (7): 184 -189.
- Clarkson, Kari. 2014. 12 fun state that proof power of images. Proceeding of the SIGCHI conference of computer and security. Toronto, Canada, 2: 456-459.
- Gross, A. M. 2014. Search engine behavior and satisfaction of Arab students from a user perspective. *International Journal of Computational Linguistics Research* 585-98 :(3) .
- Hammo, Bassam H. 2009. Towards enhancing retrieval effectiveness of search engines for diacritized Arabic documents. *Information Retrieval* 12 (3): 300-323.
- Hussan, Bzar. 2020. Comparative Study of Semantic and Keyword Based Search Engines. *Advances in Science, Technology and Engineering Systems Journal* 5: 106-111.
- Internet World Stats. (31 March 2020) Number of Internet Users by Language. Miniwatts Marketing Group, Retrieved from <https://www.internetworldstats.com/stats7.htm> (accessed Nov. 10, 2020)
- Kumar, B. S., & J. N. Prakash. 2009. Precision and relative recall of search engines: A comparative study of Google and Yahoo. *Singapore Journal of Library & Information Management* 38 (1): 124-137.
- Lancaster, Fredrick Wilfrid. 1979. *Information Retrieval Systems Characteristics, testing, and evaluation*. NewYork: Wiley.
- Lopez, CT, and C. Ribeiro. 2011. Comparative evaluation of web search engines in health information retrieval. *Online information review* 35 (6): 869-892.
- Menard, E. 2007. Image Indexing: How Can I Find a Nice Pair of Italian Shoes? *Bulletn of American Society for Information Science and Technology* 34 (1): 2-21.
- Negi, Y. S., and S. Kumar. 2014. A comparative analysis of keyword- and semantic-based search engines. *Intelligent computing, Networking, and informatics* 243: 727-736.
- Norouzi, Yaghoub, and Hoda Homavandi. 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.
- Prasher, Ram Gopal. 1989. *Index and indexing systems*. Delhi: Medallion press.
- Saracevic. T. 2007. Relevance: A Review of the Literature and a Framework for Thinking on the Notion in Information Science. Part III: Behavior and Effects on Relevance. *Journal of the American Society for Information Science and Technology* 58 (13): 2126—2144.

یعقوب نوروزی

متولد سال ۱۳۵۱، دارای مدرک تحصیلی دکتری در رشته علم اطلاعات و دانش‌شناسی از دانشگاه آزاد اسلامی واحد علوم تحقیقات تهران است. ایشان هم‌اکنون دانشیار گروه علم اطلاعات و دانش‌شناسی دانشگاه قم است. کتابخانه‌های دیجیتالی، سازماندهی اطلاعات، طراحی رابط کاربری، فناوری اطلاعات، نرم‌افزارهای کتابخانه‌ای و اطلاع‌رسانی از جمله علایق پژوهشی وی است.



هدی هماوندی

دکتری علم اطلاعات و دانش‌شناسی (گرایش بازیابی اطلاعات) از دانشگاه تهران است. ذخیره و بازیابی تصاویر، نمای‌سازی و سازماندهی اطلاعات با تمرکز بر مسائل زبان‌شناختی از علایق پژوهشی وی است.



بنت‌الهدی خبازان

متولد سال ۱۳۶۴، دارای مدرک تحصیلی کارشناسی ارشد در رشته علم اطلاعات و دانش‌شناسی از دانشگاه قم است. ایشان هم‌اکنون رئیس کتابخانه دانشگاه حضرت معصومه (س) است. ذخیره و بازیابی تصاویر، سازماندهی اطلاعات، ربط، فناوری اطلاعات، رابط کاربری و موتورهای کاوش از جمله علایق پژوهشی وی است.

