

موتورهای جستجوی فارسی ارزیابی امکانات جستجو، بازیابی اطلاعات، میزان جامعیت و مانعیت، و تعیین همپوشانی میان آنها سیدمنصور رجبی | یعقوب نوروزی

چکیده

هدف: شناسایی و ارزیابی امکانات جستجو در موتورهای جستجوی فارسی، تعیین کارآمدترین موتور جستجو به لحاظ بازیابی اطلاعات مرتبط، شناسایی موتورهای جستجوی کارآمد از حیث محاسبه میزان جامعیت و مانعیت، و تعیین میزان همپوشانی میان آنها.

روش/ رویکرد پژوهش: پژوهش از نوع کاربردی است و به روش پیمایشی-تحلیلی انجام شده است. امکانات موجود در ۵ موتور جستجوی عمومی و در دسترس فارسی براساس سیاهه واریسی پژوهشگر ساخته مورد مقایسه قرار گرفت. ضریب بازیافت نتایج مرتبط منتخب براساس ۵۰ صفحه بازیابی شده برای ۷ کلیدواژه و عبارت به صورت مجزا محاسبه شد. جامعیت و مانعیت در موتورهای جستجو در صورت وجود همپوشانی و ناهمپوشانی محاسبه شد. میزان همپوشانی موجود بین موتورها نیز در مسیر محاسبه مانعیت به دست آمد.

یافته‌ها: موتورهای جستجوی فارسی از امکانات جستجوی مناسبی برخوردار هستند و می‌توان تا حدودی از عملکرد آنها اطمینان حاصل کرد. در رابطه با بازیابی نتایج مرتبط و میزان جامعیت، بین موتورهای جستجو فاصله و تفاوت معناداری وجود دارد. موتورهای دارای امکانات جستجوی بیشتر از نظر بازیافت اطلاعات مرتبط‌تر و همپوشانی بیشتر کارآمدتر معرفی شده‌اند.

نتیجه‌گیری: توجه به ظرفیت‌های خالی شناسایی شده در موتورهای جستجوی فارسی می‌تواند در تقویت نقاط قوت و بهبود نقاط ضعف آنها مفید واقع شود.

کلیدواژه‌ها

موتور جستجوی فارسی، کارآمدی، بازیابی اطلاعات، همپوشانی، جامعیت، مانعیت، ارزیابی

موتورهای جستجوی فارسی

ارزیابی امکانات جستجو، بازیابی اطلاعات، میزان جامعیت و مانعیت، و تعیین

همپوشانی میان آنها

سیدمنصوررجبی^۱

یعقوب نوروزی^۲

دریافت: ۹۲/۰۳/۱۳

پذیرش: ۹۲/۰۶/۲۵

مقدمه

متداولترین ابزارهای بازیابی اطلاعات در بستر شبکه جهانی اینترنت، فهرست‌های موضوعی، موتورهای جستجو، و ابرموتورهای جستجو هستند که هر یک برای انتخاب، سازماندهی، و بازیابی منابع اطلاعاتی مورد درخواست کاربران از فنون خاص خود بهره می‌گیرند. آمارها نشان می‌دهند که بسیاری از افراد، سفر در دنیای وب را با موتورهای جستجو آغاز می‌کنند. مراجعه به موتورهای جستجوگر آنچنان عمومیت یافته که جستجو و کار با آنها دومین فعالیت عمده کاربران در دنیای وب (بعد از ارسال و دریافت نامه‌های الکترونیکی) محسوب می‌شود. بسیاری از کاربران با استفاده از ابزارهای جستجو، اولین تجارب وب‌گردی خود را کسب کرده‌اند (بار - ایلان^۳، ۲۰۱۱).

بر اساس پژوهش‌های پی.ای.دابلیو^۴ (۲۰۰۹) ۸۸ درصد کاربران اینترنتی آمریکایی برای یافتن اطلاعات از موتورهای جستجو استفاده می‌کنند. فرایند پیشرفت در موتورهای جستجو ابتدا با چالش‌هایی روبه‌رو شد، اما سپس با سرعت و عملکرد بالا در دستیابی به ساختار مناسب خود حرکت کرد. با نگاهی به تاریخچه موتورهای جستجوی وبی باید به موتور جستجوی آرچی^۵ اشاره کرد که در سال ۱۹۹۰ توسط آلن امتیج^۶ دانشجوی دانشگاه مک‌گیل^۷ کانادا طراحی و معرفی شد. در سال ۱۹۹۱ جستجویی با عنوان گوفر^۸ توسط مارک مک‌کاهیل^۹ در دانشگاه مینه‌سوتای^{۱۰} آمریکا طراحی شد. آرچی فایل‌های رایانه‌ای را فهرست می‌کرد، در حالی که گوفر، فرایند نمایه‌سازی اسناد متنی ساده را در دستور کار خود

۱. کارشناسی ارشد علم اطلاعات و دانش‌شناسی، دانشگاه آزاد اسلامی، واحد همدان (نویسنده مسئول)
mansoor.r65@gmail.com
۲. استادیار گروه علم اطلاعات و دانش‌شناسی، دانشگاه قم
ynorouzi@gmail.com
3. Bar - Ilan
4. PEW (Pew Internet and Life)
5. Archie
6. A. Emtage
7. University of McGill University
8. Gopher
9. M. McCahill
10. University of Minnesota

داشت. بر مبنای همین رویه، در سطح جهانی، موتورهای جستجوی فراوانی با قابلیت‌ها و محدودیت‌های گوناگون به وجود آمده‌اند. بسیاری از آنها مانند گوگل و یاهو قابلیت جستجوی مدارک به زبان‌های متنوع را دارند. موتورهایی نیز به صورت تخصصی به لحاظ موضوعی و زبانی در بعضی کشورها توسعه یافته‌اند. در ایران نیز موتورهای جستجوی فارسی متنوعی به وجود آمده‌اند که به نوبه خود در این حوزه فعالیت مفیدی داشته‌اند و در یک محدوده زمانی متوقف شده‌اند. در حال حاضر، موتورهای جستجوی عمومی فعال و قدرتمندی با قابلیت‌های بالا وجود دارند که به جستجوی اطلاعات کاربران و خدمت‌رسانی می‌پردازند مانند جاماسپ^۱، ریسمون^۲، گوگل^۳، کاوشگر^۴، و پارسیک^۵.

موتورهای جستجوی فارسی بر اساس ساختار و چارچوب موتورهای استاندارد جهانی طراحی شده‌اند؛ هر چند که در محدوده کوچک‌تری خدمت‌رسانی می‌کنند و عملکرد و بازیابی مختصرتری دارند. در واقع، بازیابی اطلاعات زیربنای موتورهای جستجوی امروزی است. این فرایند زمانی آغاز می‌شود که یک کاربر وارد مرحله جستجو می‌شود؛ فرایندی که مجموعه‌ای از اطلاعات ذخیره‌شده با هدف پاسخ به اطلاعات مورد درخواست کاربران را فراهم آورده است. چنین نظامی باید بتواند در رتبه‌بندی، اسناد مرتبط با نیاز کاربر را در رتبه بالاتر قرار دهد و سپس اسناد کمتر مرتبط را به دنبال آنها بیاورد (جوچیمز^۶، ۲۰۰۲).

در حوزه موتورهای جستجو، پژوهش‌های نسبتاً زیادی انجام شده است، ولی در زمینه ارزیابی همه مؤلفه‌های کاربردی هر موتور جستجو، کمتر تلاشی صورت گرفته است. مهم‌ترین کارهای انجام‌شده را می‌توان به شکل زیر دسته‌بندی کرد:

بهاری موفق (۱۳۸۸) به مطالعه تطبیقی نتایج جستجوی کلیدواژه‌های تخصصی پزشکی در ۱۹ ابرموتوری که رایگان قابل دسترس و سازگار با اهداف پژوهش بودند پرداخت و اطلاعات با سیاهه و آرسی گردآوری شد. یافته‌های وی نشان داد که تمامی ۱۹ ابرموتور مورد بررسی از عملگر AND، جستجوی عبارتی، تنظیم تعداد نتایج در هر صفحه، حفظ سؤال جستجو شده، و راهنما (help) برخوردار هستند، اما امکانات جستجوی کلمات فرامتنی در سایت‌ها و نمایش حجم صفحات جستجو شده در هیچ ابرموتوری مشاهده نشد. همچنین، امکاناتی از قبیل کوتاه‌سازی، جستجوی کلیدواژه‌ها در عنوان و نشانی سایت، و نمایش مختصری از متن، در ۱۸ ابرموتور جستجو پشتیبانی می‌شد.

قاضی میرسعید، حقانی، و اکبری (۱۳۸۶) به مطالعه مقایسه‌ای موتورهای کاوش و ابرموتورهای کاوش در بازیابی اطلاعات فیزیوتراپی و تعیین میزان همپوشانی ۷ موتور کاوش و ۷ ابرموتور کاوش که پر استفاده‌ترین موتورها و ابرموتورهای کاوش اینترنت بر اساس معرفی سایت searchenginewatch.com بودند پرداخت. زیررده فیزیوتراپی با استفاده از سرعنوان‌های

1. Jamasp (<http://www.Jamasp.com>)
2. Rismoon (<http://www.Rismoon.com>)
3. Googler (<http://www.Googler.ir>)
4. Kavoshgar (<http://www.Kavoshgar.com>)
5. Parseek (<http://www.Parseek.com>)
6. Information Retrieval
7. Joachims

موضوعی پزشکی مش^۱ انتخاب شد. یافته‌ها نشان داد که به ترتیب موتورهای کاوش گوگل، آلتاویستا، و وبسایت all the web بیشترین نتایج بازیابی را در میان موتورهای کاوش داشتند. در میان ابرموتورهای کاوش، بیشترین نتایج مربوط به ابرموتور ایکس کویک^۲ بود. کمیجانی (۱۳۸۲) به مقایسه کارایی موتورهای کاوش عمومی و تخصصی وب در بازیابی اطلاعات کشاورزی پرداخت. اختلاف معناداری میان موتورهای کاوش در بازیابی مدارک مرتبط و معتبر وجود داشت، ولی اختلاف معناداری میان آنها در بازیابی اطلاعات روزآمد شده و تعداد صفحات غیرفعال مشاهده نشد. موتورهای کاوش گوگل و یاهو در همه معیارها در میان سه موتور کاوش رتبه نخست داشتند.

لیتون^۳ (۱۹۹۵) در پژوهشی از ۸ سؤال مرجع از یک کتابخانه دانشگاهی به عنوان پرسش‌های جستجو استفاده کرد و از طریق به کارگیری ملاک ارزشیابی مانعیت، مانعیت کامل، مانعیت حداکثری ۱۰ موردی، و تکراری بودن نتایج به مقایسه موتورهای اینفوسیک، لیکوس، وب کراولر، ورلد واید وب، و ارم پرداخت. یافته‌های وی نشان داد که لیکوس و اینفوسیک نسبت به بقیه عملکرد بهتری داشته‌اند. در پژوهش دیگری، لیتون و سربو استاوا^۴ (۱۹۹۷) میزان مانعیت در میان خدمات جستجوی شبکه جهانی (موتورهای جستجو) را محاسبه کردند. یافته‌های آنها نشان داد که موتور آلتاویستا بهتر از بقیه عمل کرده است. از نظر جامعه آماری، موتور جستجوی آلتاویستا بهتر از دیگر موتورهای جستجو در ربط مدارک مورد نیاز عمل کرده است.

عرفان منش و دیده‌گاه^۵ (۲۰۱۰) با ارزیابی تطبیقی عملکرد موتورهای جستجوی فارسی در وب نشان دادند که موتور قطره^۶ در بیشتر ویژگی‌ها از قبیل ترافیک و رفت و آمد کاربر، مدت زمان صرف شده در سایت، و تعداد صفحات نشان داده شده برای هر کاربر، بالاترین رتبه را داشته است. موتور جستجوی کاوشگر، بیشترین تعداد کاربران خارجی و بگردیم^۷ هیچ کاربر خارجی نداشته است.

کومار و پراکاش^۸ (۲۰۰۹) نیز در بررسی جامعیت و مانعیت مرتبط موتورهای گوگل و یاهو نشان دادند که مانعیت گوگل برای پرسشنامه‌های ساده چندکلمه‌ای بالاست. در مقابل، جامعیت یاهو برای پرسشنامه‌های پیچیده چندکلمه‌ای بالاست. مانعیت مرتبط گوگل برای پرسشنامه‌های ساده تک کلمه‌ای زیاد بوده، در حالی که مانعیت مرتبط یاهو برای پرسشنامه‌های پیچیده چندکلمه‌ای بالاتر بوده است. رادر، لون، و شاه^۹ (۲۰۰۸) به تعیین میزان همپوشانی نتایج بازیابی شده در صفحه نخست ۵ موتور جستجوی سایروس، آلتاویستا، بایوبوب، گوگل، و هات‌بات پرداختند. یافته‌ها نشان داد که هیچ‌گونه همپوشانی در بازیابی نتایج استخراج شده دیده نمی‌شود، اما میزان همپوشانی در چهار موتور جستجوی

1. Medical Subject Headings (MeSH)
2. Ixquick
3. Leighton
4. Srivastava
5. Erfanmanesh & Didegah
6. Ghatreh
7. Begardim
8. Kumar & Prakash
9. Rather, Lone, & Shah

آلتاویستا، گوگل، هات‌بات، و سایر بات، و سائروس به میزان ۲ درصد حاصل شد. اسفندیاری مقدم و پریخ (۲۰۰۶) با بررسی نتایج جستجو در ابرموتورهای جستجو و موتورهای تحت پوشش آنها از جنبه همپوشانی و رتبه‌بندی نشان دادند که فراموتورها از امکانات جستجوی مناسبی برخوردار هستند و می‌توان به‌منظور جستجو و بازیابی مدارک مرتبط و مناسب تا حد زیادی به آنها اعتماد کرد. به‌علاوه، بین رتبه‌بندی موتورهای ابرموتورهای جستجو تفاوت معناداری وجود دارد.

در مجموع، با توجه به پژوهش‌های انجام‌شده در حوزه موتورهای جستجو، می‌توان چنین نتیجه گرفت که ارزیابی موتورهای جستجوی خاص یک زبان کمتر مورد توجه قرار گرفته است به‌همین دلیل، در پژوهش حاضر شناسایی و ارزیابی امکانات جستجو در موتورهای جستجوی فارسی، تعیین کارآمدترین موتور جستجو به‌لحاظ بازیابی اطلاعات مرتبط، شناسایی موتورهای جستجوی کارآمد از حیث محاسبه میزان جامعیت و مانعیت، و تعیین میزان همپوشانی میان موتورهای جستجوی فارسی مد نظر قرار گرفت.

روش‌شناسی

پژوهش حاضر که از نوع کاربردی است به‌روش پیمایشی - تحلیلی انجام شده است. ۵ موتور جستجوی عمومی فارسی فعال جاماسپ، ریسمن، گوگلر، کاوشگر، و پارسیک به‌عنوان جامعه پژوهش در نظر گرفته شد. برای گردآوری داده‌ها از سیاهه واری پژوهشگر ساخته مبتنی بر متون و منابع استفاده شد. برای اطمینان از روایی ابزار پژوهش نظرات متخصصان حوزه گرفته و در سیاهه نهایی اعمال شد. به‌منظور سنجش موتورهای جستجو از لحاظ بازیابی اطلاعات مرتبط از ۷ کلیدواژه و عبارت با نظر افراد صاحب‌نظر انتخاب شد. همچنین، برای تعیین جامعیت و مانعیت، چهار عنوان مقاله در موتورهای جستجوی منتخب فارسی مورد ارزیابی قرار گرفت. در شناسایی کارآمدترین موتور جستجوی فارسی به‌لحاظ بازیابی اطلاعات مرتبط کلیدواژه‌هایی که از موقعیتی عمومی‌تر نیز برخوردار بود، انتخاب شد. محاسبه این کارآمدی براساس ارزیابی ۵۰ نتیجه بازیابی شده برای هر کلیدواژه انجام پذیرفت. محاسبه درصد کارآمدی هر موتور جستجو، براساس فرمول محاسبه درصد کارآمدی ابرموتورهای جستجو برگرفته از اسفندیاری مقدم (۱۳۸۴) طراحی شده است.

$$E = \frac{\sum_{s=1}^S \sum_{k=1}^K X_{sk}}{THs} \times 100$$

در این معادله، E نشانگر کارآمدی هر موتور جستجو؛ S معرف تعداد موتورهای جستجو؛ و K نیز معرف تعداد کلیدواژه‌هاست که در این پژوهش، S=۱-۷ و K=۱-۷ است. TH_s نشانگر تعداد بازیافت‌ها در موتورهای جستجو ضرب در تعداد موتورها در تعداد کلیدواژه‌هاست. تعداد بازیافت‌ها، ۵۰ بازیافت، تعداد موتورها، ۵ موتور، و تعداد کلیدواژه‌ها نیز ۷ کلیدواژه انتخاب بود. روش محاسبه کارآمدی موتورها با محاسبه موتور جستجوی جاماسب، به‌عنوان نمونه معرفی شده است.

$$E_1 = \frac{1}{TH_s} [X_1 + X_2 + X_3 + X_4 + X_5 + X_6 + X_7] \times 100$$

$$E_1 = \left[\frac{3 + 2 + 2 + 9 + 3 + 6 + 3}{1750} \right] \times 100$$

$$= \frac{154}{1750} \times 100 = 0.088 \times 100 = 8.8 \approx 9\%$$

به‌منظور محاسبه میزان جامعیت، ابتدا از طریق ارزیابی روابط موجود بین موتورهای جستجوی فارسی براساس نتایج بازیابی شده، میزان همپوشانی با استفاده از فرمول محاسبه گردید. در صورت وجود همپوشانی صفر با توجه به فرمول در نظر گرفته شده، میزان جامعیت محاسبه و از نتایج به‌دست‌آمده میانگین گرفته شد تا جامعیت نسبی حاصل شود. برای محاسبه میزان مانعیت در موتورهای جستجوی فارسی، معیارهایی تعیین و برای هر معیار، نمره‌ای خاص در نظر گرفته شد. پس از شناسایی مقادیر نمرات، محاسبه میزان مانعیت از طریق فرمول برگرفته از شافی و رادر (۲۰۰۵) امکان‌پذیر گردید. سرانجام ۱۰ رکورد نخست هر یک از موتورهای جستجو برای مقالات تعیین شده مورد بررسی و مقایسه قرار گرفت.

یافته‌ها

۱. امکان‌ات موتورهای جستجوی فارسی در جستجوی اطلاعات

جدول ۱. معیارهای ارزیابی عملکردهای جستجو در موتورهای جستجوی فارسی

سایر عملگرهای جستجو			عملگرهای بولی			موتورهای جستجوی فارسی	
کوتاه‌سازی	عبارتی	(NEAR) همجواری	پراتنز ترکیبی	NOT نه	OR یا		AND و
-	+	-	+	+	+	+	پارسیک
-	+	+	-	+	+	+	جاماسب

1. Shafi & Rather

-	-	-	-	+	-	+	رسمون
+	-	+	+	+	+	+	کوشگر
+	+	+	+	+	+	-	گوگلر
۲ (%۴۰)	۳ (%۶۰)	۳ (%۶۰)	۳ (%۶۰)	۵ (%۱۰۰)	۴ (%۸۰)	۴ (%۸۰)	فراوانی/ فراوانی نسبی (درصد) نسبی (درصد)

جدول ۱ نشان می‌دهد که عملگر بولی NOT در هر ۵ موتور جستجو به طور کامل (۱۰۰ درصد) فعال است. عملگرهای بولی AND و OR در ۴ موتور جستجوی فارسی فعال بودند. جستجو به روش ترکیبی، همجواری، و عبارتی در ۳ موتور جستجو مورد توجه قرار گرفته‌اند. در نهایت، فقط ۲ موتور جستجوی فارسی، توانایی جستجو به روش کوتاه‌سازی را داشتند. معیارهای دامنه جستجو در موتورهای جستجوی فارسی در جدول ۲ ارائه شده است.

جدول ۲. معیارهای دامنه جستجو در موتورهای جستجوی فارسی

فرمت اسناد	نوع مدرک				قلمرو جستجو	مکان	زمان	زبان	موتورهای جستجوی فارسی
	اخبار	وب	فیلم	تصویر					
-	+	+	-	+	-	-	-	+	پارسیک
-	-	+	+	+	+	-	-	+	جاماسب
-	+	+	-	-	-	-	-	-	رسمون
+	-	+	+	+	+	-	-	+	کوشگر
+	-	+	-	-	-	-	-	+	گوگلر
۲ (%۴۰)	۲ (%۴۰)	۵ (%۱۰۰)	۲ (%۴۰)	۳ (%۶۰)	۲ (۴۰%) ۴	۰ (%)	۰ (%)	۴ (۸۰%)	فراوانی/ فراوانی نسبی (درصد)

همان‌گونه که در جدول ۲ مشاهده می‌شود، در رابطه با دامنه جستجو در موتورهای جستجوی فارسی، محدودسازی به لحاظ زبانی، زمانی، مکانی، نوع مدارک موجود، و فرمت اسناد مورد بررسی قرار گرفتند. نتایج به دست آمده حاکی از آن بود که هر ۵ موتور جستجوی فارسی (۱۰۰ درصد) از امکان جستجوی تحت وب برخوردار هستند. از میان موتورهای مورد بررسی، ۴ موتور جستجو (۸۰ درصد) از امکان محدودسازی به زبان صفحات جستجو برخوردار بودند. محدودسازی به لحاظ جستجو در تصاویر، ۳ موتور جستجو (۶۰ درصد)

این قابلیت را به صورت پیش فرض مورد توجه قرار داده‌اند. در میان موتورهای جستجوی فارسی، ۲ موتور جستجو (۴۰ درصد) از امکان محدودسازی جستجو به لحاظ قلمرو، فیلم، اخبار، و فرمت اسناد برخوردار بوده‌اند. در نهایت، هیچ کدام از موتورهای جستجو، معیار محدودسازی جستجو به لحاظ زمانی و مکانی را مورد توجه قرار نداده‌اند. شاخص‌های ارزیابی نمایش اطلاعات مورد جستجو در موتورهای جستجوی فارسی در جدول ۳ ارائه شده است.

جدول ۳. شاخص‌های ارزیابی نمایش اطلاعات مورد جستجو

موتورهای جستجوی فارسی	حفظ و نمایش عبارت مورد جستجو	فراوانی باز یافت هادر هر صفحه	امکان تنظیم تعداد صفحات بازبایی شده	امکان مشاهده پیش نمایش در هر صفحه	امکان فیلترسازی نتایج برای جستجوی دقیق تر	سرعت بازبایی اطلاعات در هر موتور جستجو
پارسیک	+	۱۰	-	+	-	
جاماسب	+	۱۰	-	-	+	
ریسمون	+	۱۰	-	-	+	
کاوشرگر	-	۸	-	-	-	
گیگلر	+	۱۰	-	-	+	
فراوانی / فراوانی نسبی (درصد)	۴ (۸۰٪)	—	۰ (۰٪)	۱ (۲۰٪)	۲ (۴۰٪)	۳ (۶۰٪)

داده‌های جدول ۳ نشان می‌دهد که ۴ موتور جستجو (۸۰ درصد) از قابلیت حفظ و نمایش عبارت مورد جستجو در هر صفحه برخوردار هستند. به دنبال آن، ۳ موتور جستجو (۶۰ درصد) به نمایش سرعت بازبایی اطلاعات در هر صفحه موتور جستجو توجه داشته‌اند. امکان فیلترسازی نتایج به منظور جستجوی دقیق تر برای کاربران را، فقط ۲ موتور جستجو (۴۰ درصد) فراهم ساختند. مشاهده پیش نمایش در هر صفحه، امکانی است که فقط یک موتور جستجو (۲۰ درصد) به این شاخص توجه ویژه داشته است. در نهایت، هیچ یک از موتورهای جستجوی فارسی از امکان تنظیم تعداد صفحات بازبایی شده برخوردار نبودند. لازم به ذکر است که موتور جستجوی پارسیک از قابلیت نمایش تعداد باز یافت‌ها در جستجوی وب برخوردار نبود. فقط در جستجوی منابع علمی موتور جستجوی پارسیک،

تعداد بازیافت‌ها نمایش داده شد که پژوهشگر با انتخاب اصطلاح علمی و به کار بردن آن در موتورهای این مشکل را مرتفع ساخت. فراوانی بازیافت‌ها در موتورهای جستجوی فارسی در هر صفحه در جدول ۴ ارائه شده است.

جدول ۴. فراوانی بازیافت‌ها در موتورهای جستجوی فارسی در هر صفحه

تعداد بازیافت‌ها	۱۰	۸
فراوانی	۴	۱

با توجه به داده‌های به دست آمده در جدول ۴، تعداد صفحات بازیابی شده در موتورهای جستجوی فارسی در هر بازیافت به صورت پیش فرض تعیین شده است و کاربر امکان تنظیم تعداد صفحات را ندارد. ۴ موتور جستجو، ۱۰ صفحه را از مجموع صفحات موجود به صورت پیش فرض نمایش می‌دهند؛ و فقط ۱ موتور جستجو، نمایشی ۸ صفحه‌ای را پیش فرض کار خود قرار داده است.

دیگر معیارهای ارزیابی موتورهای جستجوی فارسی در جدول ۵ ارائه شده است.

جدول ۵. معیارهای ارزیابی ساختاری موتورهای جستجوی فارسی

موتورهای جستجوی فارسی	راهنمای استفاده از موتور جستجو	تصحیح‌املای کلیدواژه کلیدواژه	ارائه نتایج پیشنهادی مرتبط	قابلیت توپرشدن (Bold)	امکان جستجوی پیشرفته
پارسیک	+	+	-	+	-
جاماسب	-	+	+	+	-
رسمون	+	-	-	+	-
کلوشگر	-	+	-	+	-
گوگلر	-	+	-	+	-
فراوانی/فراوانی نسبی (درصد)	۲ (۴۰٪)	۴ (۸۰٪)	۱ (۲۰٪)	۵ (۱۰۰٪)	۰ (۰٪)

جدول ۵ نشان می‌دهد که ۵ موتور جستجوی فارسی (۱۰۰درصد) از قابلیت توپیر شدن (بلد) عبارت و یا کلیدواژه مورد جستجو برخوردار بودند. در ۴ موتور جستجو (۸۰ درصد) امکان اصلاح کلیدواژه‌ها با املائی غلط وجود داشت. در محیط موتورهای جستجوی ۲ موتور (۴۰ درصد) امکان راهنمای استفاده از موتور جستجو فراهم شده بود. یک موتور جستجو (۲۰ درصد) از توانایی ارائه نتایج پیشنهادی مرتبط برخوردار بود و هیچ‌یک از موتورهای جستجوی مورد بررسی به امکان جستجوی پیشرفته توجهی نداشته‌اند.

۲. کارآمدی موتورهای جستجو از نظر بازیابی نتایج مرتبط

به‌منظور سنجش کارآمدی، از ۷ کلیدواژه منتخب عمومی شامل بشار اسد، ادبیات فارسی، مجله ورزشی، فیلم سینمایی، نظام بانکی، فصل بهار، و روابط عمومی استفاده شد. مجموع بازیافت عینی موتورهای جستجوی فارسی در رابطه با هر عبارت جستجو از صفحات ۱ تا ۵۰ در جدول ۶ ارائه شده است.

جدول ۶. مجموع بازیافت عینی^۱ موتورهای جستجوی فارسی در رابطه با هر عبارت جستجو از صفحات ۱-۵۰

کلیدواژه‌ها							موتورهای جستجوی فارسی
بشار اسد	ادبیات فارسی	مجله ورزشی	فیلم سینمایی	نظام بانکی	فصل بهار	روابط عمومی	
تعداد بازیافت	تعداد بازیافت	تعداد بازیافت	تعداد بازیافت	تعداد بازیافت	تعداد بازیافت	تعداد بازیافت	
۷	۹	۲	۷	۳	۷	۹	پارسیک
۳۳	۲۹	۲	۳۹	۳	۱۶	۳۲	جاماسپ
۴۲	۳۸	۴	۳۲	۳۹	۲۹	۴۹	ریسمون
۳۷	۴۵	۲۶	۴۵	۴۸	۴۶	۴۷	کاوشگر
۲۸	۴۶	۱۴	۳۶	۲۳	۴۸	۴۶	گوگلر

با توجه به جدول ۶، موتور جستجوی کاوشگر با ۲۹۴ بازیافت مرتبط با عین عبارات مورد جستجو به‌عنوان کارآمدترین موتور جستجو معرفی شد. موتور جستجو بعدی گوگلر است که ۲۵۲ بازیافت مرتبط را داشته است. ریسمون با بازیابی ۲۳۳ موردی،

۱. تعداد منابع بازیابی شده مرتبط، توسط موتورها (۵۰ بازیافت نخست) ۱ تا ۱۰ را بازیافت عینی می‌گویند.

سومین موتور جستجوی کارآمد معرفی شد. جاماسب ۱۵۴ عنوان بازیافت مرتبط ارائه کرده است و در نهایت نیز موتور جستجوی پارسیک با بازیافتی ۴۴ موردی در رتبه آخر قرار دارد. لازم به اشاره است که تعداد بازیافت‌های معرفی شده از میان ۱۷۵۰ بازیافت موتورهای جستجو بوده است.

جدول ۷. فراوانی هر کلیدواژه به تفکیک در هر موتور جستجو از مجموع ۵۰ صفحه مورد بررسی

موتورهای جستجوی فارسی	روابط عمومی	فصل بهار	نظام بانکی	فیلم سینمایی	مجله ورزشی	ادبیات فارسی	بشار اسد
پارسیک	۹ (٪۱۸)	۷ (٪۱۴)	۳ (٪۶)	۷ (٪۱۴)	۲ (٪۴)	۹ (٪۱۸)	۷ (٪۱۴)
جاماسب	۳۲ (٪۶۴)	۱۶ (٪۳۲)	۳ (٪۶)	۳۹ (٪۷۸)	۲ (٪۴)	۲۹ (٪۵۸)	۳۳ (٪۶۶)
ریسمون	۴۹ (٪۹۸)	۲۹ (٪۵۸)	۳۹ (٪۷۸)	۳۲ (٪۶۴)	۴ (٪۸)	۳۸ (٪۷۶)	۴۲ (٪۸۴)
کلوشگر	۴۷ (٪۹۴)	۴۶ (٪۹۲)	۴۸ (٪۹۶)	۴۵ (٪۹۰)	۲۶ (٪۵۲)	۴۵ (٪۹۰)	۳۷ (٪۷۴)
گوگلر	۴۷ (٪۹۴)	۴۸ (٪۹۶)	۳۳ (٪۶۶)	۳۶ (٪۷۲)	۱۴ (٪۲۸)	۴۶ (٪۹۲)	۳۸ (٪۷۶)

جدول ۸. کارآمدی موتورهای جستجوی فارسی در بازیابی نتایج مرتبط از ۵۰ صفحه مورد بررسی

موتورهای جستجوی فارسی	جاماسب	ریسمون	گوگلر	کلوشگر	پارسیک
درصد کارآمدی (E)	$100 \times (154/1750) = 9\%$	$100 \times (233/1750) = 13\%$	$100 \times (252/1750) = 14\%$	$100 \times (294/1750) = 17\%$	$100 \times (44/1750) = 3\%$

کارآمدی موتورهای جستجوی فارسی در بازیابی نتایج مرتبط از ۵۰ صفحه مورد بررسی براساس فراوانی‌های به دست آمده از جدول ۷ در جدول ۸ نشان داده شده است.

۳. جامعیت و مانعیت موتورهای جستجوی فارسی در بازیابی اطلاعات

به منظور محاسبه میزان جامعیت موتورهای جستجو، به شکل زیر عمل شد:

$$\text{تعداد بازیافت‌های کل موتور جستجوی اول برای مقاله اول} = \frac{\text{تعداد بازیافت‌های کل موتور جستجوی اول برای مقاله اول} + \text{مقدار همپوشانی در موتور دوم} + \text{مقدار همپوشانی در موتور سوم} + \text{مقدار همپوشانی در موتور چهارم} + \text{مقدار همپوشانی در موتور پنجم}}{\text{تعداد بازیافت‌های کل موتور جستجوی اول برای مقاله اول}}$$

$$= \frac{42}{42 + 0 + 1 + 1 + 0} = 0/95$$

تعداد بازیافت‌های کل موتور جستجوی اول برای مقاله اول = ۴۲

تعداد بازیافت‌های کل موتور جستجوی دوم برای مقاله اول = ۲۸

تعداد بازیافت‌های کل موتور جستجوی سوم برای مقاله اول = ۱۲۲۰۰

تعداد بازیافت‌های کل موتور جستجوی چهارم برای مقاله اول = ۶۴

تعداد بازیافت‌های کل موتور جستجوی پنجم برای مقاله اول = ۱۰

جدول ۹. جامعیت موتورهای جستجوی فارسی

پارسیک	کاوشگر	گوگلر	ریسمون	جامسپ	موتورهای جستجوی فارسی
-	۰/۹۸	۰/۹۹	-	۰/۹۷	مانعیت+همپوشانی

جدول ۹ وضعیت جامعیت موتورهای جستجوی فارسی برای ۴ عنوان مقاله مورد جستجو را پس از سنجش رابطه‌ها در صورت وجود همپوشانی نشان می‌دهد. برای محاسبه جامعیت در صورتی که همپوشانی صفر باشد به طریق زیر عمل شد:

$$\text{تعداد بازیافت‌های کل موتور جستجوی اول برای مقاله اول} = \frac{\text{تعداد بازیافت‌های کل موتور جستجوی اول برای مقاله اول}}{\text{تعداد بازیافت‌های کل موتور جستجوی اول برای مقاله اول} + \text{موتور دوم} + \text{موتور سوم} + \text{موتور چهارم} + \text{موتور پنجم}}$$

$$= \frac{42}{42 + 28 + 12200 + 64 + 10} = *$$

جدول ۱۰. جامعیت موتورهای جستجوی فارسی

مقاله	پاریسک	کاوشگر	گوگلر	ریسمون	جاماسپ	موتورهای جستجوی فارسی
۱م	۰	۰/۰۰۵	۰/۹۸	۰/۰۰۲	۰/۰۰۳	جامعیت + همپوشانی صفر
۲م	۰/۰۰۱	۰/۰۱۲	۰/۹۷	۰/۰۰۷	۰/۰۰۸	جامعیت + همپوشانی صفر
۳م	۰/۰۰۳	۰/۲۱	۰/۶۲	۰	۰/۱۲	جامعیت + همپوشانی صفر
۴م	۰	۰/۰۰۲	۰/۹۹	۰	۰/۰۰۱	جامعیت + همپوشانی صفر

جدول ۱۰ نتایج محاسبه جامعیت در موتورهای جستجوی فارسی را برای ۴ عنوان مقاله مورد جستجو پس از سنجش رابطه‌ها در صورت وجود همپوشانی صفر نشان می‌دهد.

جدول ۱۱. میانگین جامعیت‌های دارای همپوشانی و همپوشانی صفر

پاریسک	کاوشگر	گوگلر	ریسمون	جاماسپ	موتورهای جستجوی فارسی
۰/۰۰۳	۰/۵۱	۰/۹۴	۰/۰۰۱	۰/۳۷	میانگین داده‌ها/جامعیت نسبی

میانگین جامعیت‌های دارای همپوشانی و همپوشانی صفر پس از محاسبه براساس داده‌های به‌دست‌آمده از جدول‌های ۹ و ۱۰ در جدول ۱۱ ارائه شده است. مانعیت، به‌عنوان دومین عامل مورد ارزیابی در فرایند بازیابی اطلاعات مورد سنجش قرار گرفت. به‌منظور محاسبه این عامل، ۶ معیار در نظر گرفته شد و به هریک نمره‌های زیر تعلق گرفت:

۱. اگر صفحه بازیابی شده تمام متن مقاله را ارائه داد (نمره ۳)؛
۲. اگر صفحه بازیابی شده چکیده مقاله را نشان داد (نمره ۲)؛
۳. اگر فقط صفحه مربوط به کتاب یا پایگاه داده که اطلاعات مقاله در آن ذکر شده است، بازیابی شد (نمره ۱)؛
۴. اگر صفحه بازیابی شده، غیر از این ۳ مورد را ارائه داد (نمره صفر)؛
۵. اگر چند صفحه یکسان را با نشانی‌های مختلف بازیابی کرد (نمره صفر)؛ و
۶. در صورتی که صفحه مورد نظر با خطای سرور مواجه شد (نمره صفر).

$$\text{مجموع نمرات کسب شده برای هر موتور جستجو و هر مقاله} = \frac{\text{محداسبه مانعیت در هر موتور جستجو}}{\text{تعداد صفحات مورد بررسی در هر موتور جستجو}}$$

جدول ۱۲. مجموع مانعیت‌های حاصل شده برای هر موتور جستجو از ۴ مقاله

پارسیک	کاوشگر	گوگلر	ریسمون	جاماسپ	موتورهای جستجوی فارسی
۰	۳/۷	۰/۴	۰/۵	۰	مجموع مانعیت‌ها

جدول ۱۲، مجموع مانعیت‌های حاصل شده برای هر موتور جستجو از ۴ عنوان مقاله را نشان می‌دهد.

جدول ۱۳. میانگین نتایج به دست آمده مانعیت در موتورهای جستجو

پارسیک	کاوشگر	گوگلر	ریسمون	جاماسپ	موتورهای جستجوی فارسی
۰	۰/۹۲	۰/۱	۰/۱۲	۰	مانعیت کل

جدول ۱۳، میانگین نتایج به دست آمده مانعیت را در موتورهای جستجوی فارسی نشان می‌دهد.

۴. میزان همپوشانی موتورهای جستجوی فارسی

جدول ۱۴. تعداد صفحات دارای همپوشانی بین موتورهای جستجوی فارسی

پارسیک	کاوشگر	گوگلر	ریسمون	جاماسپ	موتورهای جستجوی فارسی
۰	۴	۴	۰	۶	مقدار همپوشانی

تعداد صفحات دارای همپوشانی بین موتورهای جستجوی فارسی در جدول ۱۴ ارائه شده است.

جدول ۱۵. درصد وجود همپوشانی بین موتورهای جستجوی فارسی

موتورهای جستجوی فارسی	جاماسپ	ریسمون	گوگلر	کاوشگر	پارسیک
درصد همپوشانی	۶۰٪	۰	۴۰٪	۴۰٪	۰

همچنین درصد وجود همپوشانی بین موتورهای جستجوی فارسی مؤلفه دیگری بود که به تفکیک برای هر موتور جستجو براساس داده‌های به دست آمده در جدول ۱۵ نشان داده شده است.

بحث و نتیجه گیری

موتورهای جستجوی فارسی ابزارهای مناسبی برای بازیابی اطلاعات برای کاربران فارسی زبان هستند. محدودیت و نقصان در همه موتورهای جستجو وجود دارد و موتورهای جستجوی فارسی نیز از این قاعده مستثنی نیستند. نتایج نشان می‌دهد که کاربران فقط در مسیر جستجوی اطلاعات مورد نیاز نباید خود را به امکانات پیش فرض در هر موتور جستجو محدود کنند. بلکه باید از تخصص کافی برای انتخاب واژگان هدف و کلیدی برخوردار باشند. امکانات جستجو برای موتورهای جستجو به طور نسبی و جامع طراحی شده‌اند. با این حال، فقط تکیه بر این امکانات، کاربر را به نتیجه منظم و دقیق هدایت نمی‌کند، بلکه کسب مهارت و درگیری با واژگان و عبارات به گونه‌ای که از روابط بین واژگان و عبارات اطلاعات کافی داشته باشد، به کاربر در رسیدن به نتیجه درست یاری می‌رساند. موتورهای جستجوی فارسی به علت وجود نقصانها و ظرفیت‌های خالی موجود، نسبت به موتورهای جستجوی محبوب بین‌المللی عملکرد پایین تری دارند. این نقصانها در میان موتورهای جستجو متفاوت است. به لحاظ توانمندی موتورهای جستجو در استفاده از امکانات کاوش، شناسایی کدهای موجود در موتورهای جستجویی که به صورت پیش فرض امکانات خود را طراحی کرده‌اند و پس از هر فرایند جستجو، کد مربوط به نوع جستجو که محدود به قلمرو، قالب مدارک، و یا گونه‌های دیگر در کنار عبارت مورد کاوش درج می‌شوند در بازیابی نتایج مؤثرند.

نتایج به دست آمده نشان داد که تمامی موتورهای جستجوی فارسی، از عملگر NOT، جستجو در وب، و قابلیت توپر شدن عبارت یا کلیدواژه مورد جستجو برخوردار هستند. نکته مهم این است که معیار جستجو به لحاظ زمانی، مکانی، و امکانات جستجوی پیشرفته در هیچ موتور جستجویی مشاهده نشد که این خود معرف وجود نقصان در ساختار عملکردی موتور جستجو است. موتورهای جستجوی فارسی، عملگر AND، OR، جستجوی زبانی، حفظ و نمایش عبارت مورد جستجو، و اصلاح کلیدواژه‌ها با املائی غلط را در حد بالایی (۴ موتور جستجو) پشتیبانی می‌کنند. پشتیبانی این مؤلفه‌ها تطابق نسبی با مؤلفه‌های موجود

در ساختار موتورهای جستجوی محبوب و شناخته شده در سطح جهان را نشان می دهد. بررسی تطبیقی امکانات جستجوی ۵ موتور جستجوی مورد توجه در پژوهش حاضر، نشان می دهد که موتورهای جستجوی فارسی در آینده ای نزدیک ابزاری مناسب برای جستجوی کاربران، اعم از خبره و مبتدی، خواهند شد. از نظر امکانات جستجو، نتایج پژوهش همسویی نسبی با امکانات مورد سنجش در پژوهش های بهاری موفق (۱۳۸۸)، اسفندیاری مقدم و پریخ^۱ (۲۰۰۶) و همچنین پژوهش عرفان منش و دیدگاه^۲ (۲۰۱۰) داشت.

از نظر کارآمدی موتورهای جستجوی فارسی، موتور جستجوی کاوشگر با ۱۷ درصد کارآمدترین موتور جستجو معرفی شد و به دنبال آن، گوگل با ۱۴ درصد، ریسمن با ۱۳ درصد، جاماسپ با ۹ درصد و در نهایت نیز پارسیک با ۳ درصد بازیافت مرتبط برای ۷ کلیدواژه انتخابی قرار داشتند. البته نتایج حاصل شده با نتایجی که از پژوهش های کمیجانی (۱۳۸۲)، پترسون^۳ (۱۹۹۷)، اسفندیاری مقدم و پریخ^۴ (۲۰۰۶) تطابق نسبی و البته در مواردی نیز فاصله معناداری وجود دارد.

در زمینه جامعیت با وجود همپوشانی بین روابط، موتور جستجوی گوگل با ۰/۹۹، کاوشگر با ۰/۹۸، و جاماسپ با ۰/۷۱ به ترتیب بیشترین مقدار جامعیت با وجود همپوشانی بین روابط را به خود اختصاص دادند. این نتایج نشان دهنده توجه به معیار دقت در فرایند بازیابی اطلاعات برای این موتورهای جستجو است. موتورهای جستجوی ریسمن و پارسیک فاقد همپوشانی بین روابط بودند که معرف بی توجهی این دو موتور جستجو به معیار دقت در فرایند بازیابی اطلاعات است. توجه به دقت در بازیابی اطلاعات به هنگام جستجو در هر موتور جستجو از مهم ترین شاخص های سنجش کیفی است و این معیار در دو موتور جستجوی ریسمن و پارسیک مشاهده نشد.

در مسیر محاسبه میزان جامعیت و مانعیت در ۵ موتور جستجوی فارسی مشخص شد که بین میزان جامعیت و مانعیت موجود در موتورهای جستجوی فارسی و موتورهای مورد سنجش توسط لیتون^۴ (۱۹۹۵)، لیتون و سریواستاوا^۵ (۱۹۹۷)، بیتیریم، تونتا، و سور^۶ (۲۰۰۲)، و حلیم و کائور^۷ (۲۰۰۶) همسویی وجود دارد.

نتایج از نظر میزان همپوشانی موتورهای جستجو، با نتایج پژوهش قاضی میرسعید، حقانی، و اکبری (۱۳۸۶)، اسپینک^۸ و همکاران (۲۰۰۶)، رادر، لون، و شاه^۹ (۲۰۰۸) همسویی دارد. هر چند که ساختار موتورهای جستجو به لحاظ توانمندی پایگاه داده و تارهای تینده شده در هسته آن متغیر است، با این حال عامل سنجش یکسان است و موتورهای مورد سنجش در این پژوهش نیز تقریباً هم گروه و هم سطح هستند. ۲۸ درصد همپوشانی بین موتورهای مورد ارزیابی نشان از عملکرد پایین موتورهای جستجوی دارد.

بنابراین، با توجه به یافته ها پیشنهاد های زیر قابل اجراست:

• در ساختار موتورهای جستجو تمامی معیارهای جستجو در این پژوهش جهت

1. Isfandiyari-Moghaddam & Parirokh
2. Erfanmanesh & Didegah
3. Peterson
4. Leighton
5. Leighton & Srivastava
6. Bitirim, Tonta, & Sever
7. Halim & Kaur
8. Spink
9. Rather, Lone, & Shah

ارتقای سطح کیفی فرایند جستجو توسط کاربران در نظر گرفته شوند؛

- برای کاربران شرایطی مهیا شود که به هنگام جستجوی کلیدواژه‌ها یا عبارات، نتایج بازیابی شده مرتبط با عنوان یا عناوین جستجو باشد؛
- به معیار جامعیت به عنوان عامل توانمندی یک موتور جستجو توجه شود. آنچه کاربران را در عملیات جستجو راضی نگه می‌دارد، وجود دقت در نتایج بازیابی شده در هر موتور جستجو است. طراحان موتورهای جستجو باید مانعیت موجود به هنگام جستجو را به حداقل ممکن برسانند و رضایت کاربران را در نظر بگیرند؛ کاربران تا زمانی مخاطب یک موتور جستجو هستند که نتایج بازیابی شده، نیاز آنها را برطرف سازد.
- به وجود همپوشانی موتورهای جستجو به هنگام جستجوی موارد ثبت شده در سایت‌های معتبر توجه شود، به گونه‌ای که برای مقاله مورد جستجو بتوان پایگاه‌های یکسانی را مشاهده کرد که مقاله مورد نظر را نمایه کرده‌اند.

مآخذ

- اسفندیاری مقدم، علیرضا (۱۳۸۴). بررسی نتایج جستجو در ابرموتورهای کاوش و موتورهای تحت پوشش آنها از جنبه همپوشانی و رتبه‌بندی. پایان‌نامه کارشناسی ارشد، دانشگاه فردوسی، مشهد.
- بهارى موفق، زهره (۱۳۸۸). بررسی تطبیقی نتایج جستجوی کلیدواژه‌های تخصصی پزشکی در موتورهای ابرموتورهای کاوش وب. پایان‌نامه کارشناسی ارشد، دانشگاه آزاد اسلامی، همدان.
- قاضی میرسعید، جواد؛ حقانی، حمید؛ و اکبری، علیرضا (۱۳۸۶). مطالعه مقایسه‌ای موتورهای کاوش و ابرموتورهای کاوش در بازیابی اطلاعات فیزیوتراپی از شبکه جهانی وب و تعیین میزان همپوشانی میان آنها. مدیریت اطلاعات سلامت، ۴ (۱)، ۱۱-۲۱.
- کمبجانی، احمد (۱۳۸۲). مقایسه کارایی موتورهای کاوش عمومی و تخصصی وب در بازیابی اطلاعات کشاورزی. فصلنامه اطلاع‌شناسی، ۱ (۱)، ۱۲۵-۱۴۴.
- Bar-Ilan, J., & Levene, M. (2011). A method to assess search engine results. *Online Information Review*, 35 (6), 854-868.
- Bitirim, Y., Tonta, Y., & Sever, H. (2002). Information retrieval effectiveness of Turkish search engines. In *Advances in information systems (93-103). Second International Conference, ADVIS 2002, Izmir (Turkey), 23-25. Springer-Verlag, Berlin Heidelberg: Springer.*
- Erfanmanesh, M. A. & Didegah, F. (2010). Evaluating function of Persian search engines on the web using correspondence analysis. *International Journal of Information Science and Management (IJSM)*, 8 (2), 77-87, Retrieved October 14, 2015, from <http://ijism.ricest.ac.ir/index.php/ijism/article/view/51/54>
- Halim, H. & Kaur, K. (2006). Malaysian web search engines: A critical analysis. *Malaysian Journal of Library & Information Science*, 11(1), 103-122.

- Isfandyari Moghaddam, A. & Parirokh, M. (2006). A comparative study on overlapping of search results in metasearch engines and their common underlying search engines. *Library Review*, 55 (5), 301-306.
- Joachims, T. (2002). Optimizing search engines using clickthrough data. In *Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining* (KDD '02). ACM, New York, NY, USA, 133-142. DOI=10.1145/775047.775067 <http://doi.acm.org/10.1145/775047.775067>
- Kumar, B. S., & Prakash, J. N. (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.
- Leighton, H. V. (1995). Performance of four World Wide Web (WWW) index services: Infoseek, Lycos, WebCrawler, and WWWWorm. Retrieved October 14, 2015, from <http://course1.winona.edu/vleighton/webind.htm>
- Leighton, H. V. & Srivastava, J. (1997). Precision among world wide web search services (search engines): Alta Vista, Excite, Hotbot, Infoseek, Lycos, Retrieved October 14, 2015, from <http://course1.winona.edu/VLeighton/webind2/webind2.htm#REFLEI1>
- Peterson, R. E. (1997). Eight Internet search engines compared. *First Monday*, 2 (2).
- PEW (March 2000 - September 2009). Pew Internet & American Life Project tracking surveys, Pew Internet & American Life Project, Retrieved October 14, 2015, from <http://www.pewinternet.org/>
- Rather, R. A., Lone, F. A., & Shah, G. J. (2008). Overlap in Web search results: A study of five search engines. *Library Philosophy and Practice (e-journal)*, 226. Retrieved October 14, 2015, from <http://digitalcommons.unl.edu/cgi/viewcontent.cgi?article=1228&context=libphilprac>
- Shafi, S. M., & Rather, R. A. (2005). Precision and recall of five search engines for retrieval of scholarly information in the field of biotechnology. *Webology*, 2 (2).
- Spink, A., Jansen, B. J., Kathuria, V., & Koshman, S. (2006). Overlap among major web search engines. *Internet Research*, 16 (4), 419-426.

استناد به این مقاله:

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