

مقایسه عملکرد موتورهای کاوش در هدایت کاربران به صفحات خانگی در پاسخ به پرسشهای راهبری

علی سادات موسوی^۱
اورانوس تاج‌الدینی^۲
دکتر نجلا حریری^۳

چکیده

این پژوهش به روشی نظام‌مند، عملکرد موتورهای جستجو را در رابطه با پرسشهای راهبری مقایسه و نتایج را با مطالعات صورت گرفته درباره عملکرد بازیابی موتورهای جستجو برای پرسشهای اطلاعاتی مقایسه می‌کند. برای انجام این پژوهش، پنج موتور جستجوی گوگل، یاهو، بینگ، اسک و آوال بر پایه رتبه‌بندی موتورهای کاوش در سایت الکسا در انتخاب شدند. ۱۰۰ پرسش به زبان فارسی همراه با عبارت دقیق پرسش، توصیف کوتاهی از نیاز اطلاعاتی و آدرس اینترنتی صفحه مورد نظر از یکایک دانشجویان گردآوری و سپس توسط پژوهشگر در قسمت جستجوی هر یک از موتورهای جستجوی مذکور وارد و موقعیت نتایج صحیح ثبت شد. اگر صفحه خانگی در میان ده نتیجه اول یافت نمی‌شد، جستجو متوقف می‌شد. برای محاسبه عملکرد بازیابی جهت پرسشهای راهبری، از سنج‌های بازیابی متعددی نظیر سنجه موفقیت N و رتبه معکوس میانگین استفاده شده است. یافته‌های پژوهش نشان می‌دهند سه موتور جستجوی گوگل، آوال و اسک تقریباً با ۹۶، ۹۱ و ۸۷٪ موفقیت در بازیابی، عملکرد نسبتاً خوبی در بازیابی پاسخ پرسشهای راهبری به زبان فارسی دارند، اما دو موتور جستجوی یاهو و بینگ تقریباً با ۳۵ و ۴۷٪ عدم موفقیت، عملکرد ضعیفی در بازیابی پاسخ پرسشهای راهبری به زبان فارسی دارند. اکثر نتایج به دست آمده توسط گوگل و اسک و آوال در رتبه اول صفحه نتایج قرار دارند. علاوه بر آن، همه موتورهای جستجوی مورد نظر در این پژوهش از نتیجه اول به نتیجه پنجم پیشرفت دارند. به عبارت دیگر، آنها تقریباً در بیش از

-
۱. دانشجوی دکتری علم اطلاعات و دانش‌شناسی، دانشگاه آزاد اسلامی، واحد علوم و تحقیقات تهران، ایران.
moosavi56@gmail.com
 ۲. دانشجوی دکتری علم اطلاعات و دانش‌شناسی، دانشگاه آزاد اسلامی، واحد علوم و تحقیقات تهران، ایران.
tajedini.o@gmail.com
 ۳. دانشیار گروه علم اطلاعات و دانش‌شناسی، دانشگاه آزاد اسلامی، واحد علوم و تحقیقات تهران، ایران.
nadjlahariri@gmail.com

۹۰٪ موارد قادرند صفحات درخواستی خود را بدون مرور کردن نتایج ارائه شده در پایین فهرست نتایج ارائه شده به کاربر در اولین صفحه نتایج نشان دهند.

کلیدواژه‌ها: پرسشهای راهبری، موتورهای جستجو، عملکرد بازیابی، ارزیابی موتورهای جستجو.

مقدمه

شبکه جهانی وب منبع عظیمی از اطلاعات است که به طور روز افزون بر حجم آن افزوده و امروزه به ابزاری ضروری در زندگی بسیاری از مردم تبدیل شده است. موتورهای جستجو نیز مدخل اصلی ورود به آن به شمار می‌روند (وسترولد^۱، کراییز^۲ و هایمسترا^۳، ۲۰۰۱). موتورهای جستجو «ابزارهای شناخته شده‌ای» برای دستیابی به اطلاعات، سایتهای اینترنتی و خدمات وبی هستند که بسیاری از مردم در زندگی روزانه خود از آنها استفاده می‌کنند (جانسن^۴، بوث^۵ و اسپینک^۶، ۲۰۰۷). اهمیت موتورهای جستجو برای یافتن اطلاعات مرتبط در وب انکارناپذیر است به طوری که نزدیک به ۷۰٪ جستجوگران وب از یک موتور جستجو به عنوان نقطه شروع در دستیابی به اطلاعات استفاده می‌کنند. موتورهای جستجوی مهم، روزانه میلیونها پرسش از سوی کاربران دریافت می‌کنند و هر هفته در جواب پرسشهای مذکور چندین میلیون پاسخ بازیابی و در اختیار کاربران قرار داده می‌شود (سولیوان^۷، ۲۰۰۶).

پرس و جوهای وبی معمولاً به سه دسته تقسیم می‌شوند: اطلاعاتی^۸، راهبری^۹ و تراکشی^{۱۰}. در پرسشهای اطلاعاتی کاربران قصد دارند در مورد یک موضوع خاص اطلاعاتی پیدا کنند. چنین پرسشهایی اغلب به یک مجموعه از نتایج می‌انجامد تا تنها

1. Westerveld.
2. Kraaij.
3. Hiemstra.
4. Jansen.
5. Booth.
6. Spink.
7. Sullivan.
8. Informational.
9. Navigational.
10. Transactional.

یک مدرک مناسب. پرسشهای اطلاعاتی شبیه پرسشهایی هستند که به سیستم‌های بازیابی اطلاعات مبتنی بر متن ارائه می‌گردند. برای مثال «ابولؤلؤ کیست؟» یک پرسش اطلاعاتی است که کاربر می‌خواهد در مورد یک موضوع خاص اطلاعاتی را به دست آورد. چنین پرسشهایی معمولاً صفحات وب ایستا را مورد هدف قرار می‌دهند. اصطلاح «ایستا» در این جا به این مسئله اشاره دارد که وقتی صفحه به کاربر ارائه می‌شود، تعامل بیشتری برای به دست آوردن اطلاعات درخواست شده لازم نیست (برودر^۱، ۲۰۰۲؛ نقل در مک‌فارلان^۲، ۲۰۰۷).

در یک پرسش تراکنشی، کاربران علاقه‌مند به یافتن مدرکی هستند که خدمات توصیف شده در پرسش تراکنشی را ارائه می‌دهد. نتایج پرس و جوهای تراکنشی وب‌سایتی هستند که برای رسیدن به مقصود، تعاملات بیشتری را طلب می‌کنند. تراکنش می‌تواند دانلود یک برنامه یا فایل، خرید یک محصول یا جستجوهای بعدی در یک سایت باشد. برای مثال، کاربری که برای بازیهای رایگان در وب جستجو می‌کند، تمایل دارد یا بازی را دانلود کند و یا به صورت آنلاین آن را انجام دهد. در هر دو صورت، در سایت بازیابی شده تراکنش ضروری خواهد بود. کاربران از موتور جستجو به عنوان وسیله‌ای برای انجام تراکنش در تجارت الکترونیکی (مانند سیستم پرداخت گوگل^۳) استفاده می‌کنند. آنها در مقایسه با پاسخ مستقیم به یک نیاز اطلاعاتی با موضوع خاص، بیشتر دسترسی به محتوای مجموعه‌ای از عکس‌ها، موسیقی‌ها، ویدئوها و ... را فراهم می‌کنند. آنها همچنین امکان دسترسی به سرویسهای تراکنشی مانند نقشه‌ها، مزایده‌های آنلاین، راهنماهای رانندگی و حتی دیگر موتورهای جستجو را فراهم می‌کنند. در نتیجه، محققان توجه زیادی به کیفیت نتایج موتورهای جستجو دارند (رز و لوینسون، ۲۰۰۴).

هرچند این رویکرد تا حدودی در مورد جستجوگران صدق می‌کند، یافته‌ها حکایت از آن دارد که کاربران در جستجوی وبی پرسشهای خود را به منظور دیگری هم طرح

1. Broder.
2. MacFarlane.
3. Google's payment system.

می‌کنند (جانسون، بوث و اسپینک^۱، ۲۰۰۸)؛ به منظور هدایت شدن به صفحه وب خاصی که از پیش می‌شناسند یا گمان می‌کنند وجود دارد. این وب‌سایت می‌تواند متعلق به یک شخص یا سازمان باشد. موفقیت این نوع پرسشها عامل موفقیت مهمی برای موتورهای جستجوی وب است. برای مثال «آدرس اینترنتی سایت بیمارستان علوم پزشکی تهران؟» یا «بیمارستان علوم پزشکی تهران» یک پرسش راهبری است که هدف از طرح آن یافتن صفحه وب ورودی دانشگاه علوم پزشکی تهران است. ملاک قضاوت درباره نتایج بازیابی شده آن است که آیا این صفحه، صفحه‌ای است که به دنبال آن بوده است؟ آیا این صفحه‌خانگی همان است که مدنظر وی بوده یا او آن را نامگذاری کرده است؟ (حریری^۲، ۲۰۱۰).

پرسشهای راهبری اغلب تنها با یک نتیجه پاسخ داده می‌شوند و به محض اینکه نتیجه درست پیدا شود، نیاز اطلاعاتی شخص برطرف می‌شود. با این حال، همه پرسشها برای جستجوی افراد راهبری نیستند. «روز و لوینسون^۳» (۲۰۰۴) معتقدند بیشتر پرسشهایی که برای جستجوی افراد انجام می‌گیرند، راهبری نیستند. بنا به اظهارات آنها «در جستجو برای افراد مشهوری نظیر کامرون دیاز یا بن‌افلک نتایج متفاوتی از سایتهای هواداران، سایتهای رسانه‌ها و نظایر آن بازیابی خواهد شد. به عبارت دیگر، احتمال اینکه کاربری که نام فرد مشهور را به عنوان یک پرسش جستجو می‌کند قصد دیدن یک سایت خاص را داشته باشد، کم است.» این مورد ممکن است درست باشد، اما از سوی دیگر این پرسشها را نمی‌توان به عنوان پرسشهای اطلاعاتی در نظر گرفت، زیرا نمی‌توان فرض کرد که کاربر می‌خواهد اسناد و مدارک مختلفی را در این مورد مطالعه کند. آنچه می‌توان تصور کرد آن است که این پرسشها چیزی بین پرسشهای اطلاعاتی و راهبری هستند (لواندوسکی، درچسلر و ماچ، ۲۰۱۲).

مفهوم پرسشهای راهبری مبتنی است بر جستجوی اقلام مشخص شده در مجموعه‌های کتابخانه‌ای یا به صورت کلی‌تر، مبتنی است بر تفکیک نیازهای اطلاعاتی

1. Jansen, Booth, and Spink.

2. Hariri.

3. Rose and Levinson.

مسئله - محور و نیازهای اطلاعاتی عینی. در پرسشهای راهبری، کاربر اثر موجود در یک مجموعه را می‌شناسد و می‌خواهد همان اثر را بازیابی کند. بنابراین، تنها یک نتیجه، نتیجه مرتبط است و زمانی جستجو کامل می‌شود که اثر مورد نظر پیدا شود. راهبری می‌تواند به یک سایت تراکنشی باشد یا راهبری به یک سایت اطلاعاتی. از لحاظ یک موتور جستجو، هدف بردن کاربر به وبسایت مناسب است. طبیعتاً از دید کاربر، زمانی که وی به یک مقصد خاص می‌رسد، ممکن است سایر هدفهایش را دنبال کند. بنابراین، کاربر می‌تواند جستجوی راهبری را به عنوان وسیله‌ای برای رسیدن به هدفهای بزرگتر جستجو در نظر بگیرد (بلکین^۱ و دیگران، 1982).

بیان مسئله

درحالی که شمار دقیق نسبت پرسشهای راهبری در بررسی‌های گوناگون متفاوت است، نتیجه مسلم همه مطالعات آن است که پرسشهای راهبری شمار قابل توجهی از پرسشها را به خود اختصاص داده‌اند (هاوکینگ و کراسول^۲، 2004). اصطلاح متفاوتی که برای این نوع پرسشها وجود دارد «پیدا کردن صفحه خانگی^۳» است (لواندوسکی^۴، 2008). به دلیل آنکه در بیشتر موارد صفحه خانگی، اولین صفحه‌ای است که کاربران خواهند دید، این صفحه اهمیت بالایی دارد. علاوه بر آن، صفحه خانگی زمینه آشنایی با سایت، طراح و محتوای آن را فراهم می‌سازد؛ معمولاً به صفحات دیگر سایت لینک می‌دهد و امکان دسترسی به دیگر امکانات سایت مانند جستجو را فراهم می‌سازد. یوآرال صفحه خانگی، اغلب به عنوان یوآرال تمام سایت فرض می‌شود (هاوکینگ و کراسول، 2004).

از طرف دیگر، نمایه این موتورها با هم متفاوت هستند، و تمام مدارک موجود در شبکه در نمایه یک موتور موجود نیست (اسمیت، 2001) همچنین، مدارکی که از طریق یک موتور جستجو یافت می‌شود، ممکن است توسط موتور دیگر به دست نیاید. هنگام

-
1. Belkin.
 2. Hawking and Craswell.
 3. homepage finding.
 4. Lewandowski.

جستجو در وب، نظر به اینکه کاربران نمی‌توانند مقایسه‌ای نظام‌مند از نتایج موتورهای جستجوی مختلف داشته باشند، باید صرفاً بر مشاهدات خود تکیه کنند. به علاوه، از آنجا که آنها معمولاً یک یا چند نتیجه را برای یک پرسش مورد توجه قرار می‌دهند، باید به برداشت خود از فهرست نتایج اعتماد کنند (گرانکا، جوچیمز و گی^۱، 2004؛ کین، ابراین و اسمایف^۲، 2008).

با توجه به پوشش متنوع موضوعی در نمایه موتورهای جستجو و تحقیقاتی که تا حال بر روی موتورهای جستجو در مورد توانایی پاسخگویی آنها به سؤالهای مختلف انجام شده است، می‌توان انتظار داشت این موتورها توانایی پاسخگویی به سؤالها را داشته باشند. از سوی دیگر، ارزیابی‌ها در مورد عملکرد موتورهای کاوش در بازیابی برای پرسشهای اطلاعاتی نشان می‌دهد گوگل در این زمینه پیش‌تاز است (لواندوسکی، 2008). گوگل بالاترین نسبت از شرح نتایج مرتبط را ارائه می‌دهد و کاربران آن را بهترین موتور جستجو می‌دانند (ماچایل^۳ و دیگران، 2003؛ اسچمایدمینز و بومهاردت^۴، 2005). اما در مورد سؤالهای راهبری وضعیت چگونه است؟ آیا در جستجوی سؤالهای راهبری نیز گوگل بهترین عملکرد را دارد؟

فرضیه پژوهش

برای این پژوهش، فرضیه‌ای به شکل ذیل مطرح می‌گردد:
در بازیابی پاسخ پرسشهای راهبری، موتور جستجوی گوگل نسبت به سایر موتورهای جستجوی مورد استفاده در این پژوهش، بهترین عملکرد را دارد.

پرسشهای پژوهش

برای این پژوهش چهار پرسش به شکل ذیل مطرح می‌گردد.
۱. پاسخ چه تعداد و چند درصد از پرسشهای راهبری ارائه شده به موتورهای جستجوی مورد نظر در این پژوهش، بازیابی می‌گردد؟

-
1. Granka, Joachims and Gay.
 2. Keane, O'Brien and Smyth.
 3. Machill.
 4. SchmidtMaenz and Bomhardt.

۲. کدام یک از موتورهای جستجوی مورد نظر در این پژوهش، پاسخ پرسشهای راهبری را در رتبه اول فهرست نتایج بازیابی شده ارائه می‌دهند؟
۳. موقعیت نتایج بازیابی شده توسط موتورهای جستجوی مورد نظر در این پژوهش چگونه است؟
۴. در پاسخ به پرسشهای راهبری، نسبت یافته‌ها در موتورهای جستجوی مورد نظر در این پژوهش چگونه است؟

پیشینه پژوهش

بررسیهای زیادی در مورد عملکرد بازیابی در موتورهای جستجوی وب انجام شده است. با این حال، در اکثر موارد در این بررسی‌ها از پرسشهای اطلاعاتی استفاده شده است. پرسشهای اطلاعاتی، پرسشهایی هستند که کاربر می‌خواهد حداقل چند مدرک را در مورد موضوع مورد جستجوی خود پیدا کند. نیاز اطلاعاتی معمولاً با یک نتیجه برطرف نمی‌شود و کاربر پیشاپیش نمی‌داند که کدام مدرک (یا در محیط وب، کدام صفحه) بهترین نتیجه را به دست می‌دهد.

«آزادی» (۱۳۸۴) در پژوهشی با عنوان «میزان دقت موتورهای کاوش وب در بازیابی اطلاعات کتابداری و اطلاع‌رسانی» نحوه عملکرد ۷ موتور کاوش را بررسی کرد. نتایج نشان داد اختلاف معناداری در میزان دقت موتورهای کاوش مورد بررسی در بازیابی اطلاعات کتابداری و اطلاع‌رسانی وجود ندارد. در این پژوهش، موتورهای کاوش از نظر دقت به این ترتیب رتبه‌بندی شدند: اینفوسیک^۱، هات‌بات^۲، گوگل، آلتاویستا^۳، اکسایت^۴، لایکاس^۵ و وب‌کراولر^۶.

«کميجانی» (۱۳۸۱) در پایان‌نامه خود با عنوان «مقایسه کارایی موتورهای کاوش عمومی و تخصصی وب در بازیابی اطلاعات کشاورزی» ۵ موتور کاوش وب شامل

1. Infoseek.
2. Hotbot.
3. Altavista.
4. Altavista.
5. Lycos.
6. web crawler.

اگری سرف^۱، وب‌اگری^۲، گوگل، لایکاس و یاهو را در بازیابی اطلاعات تخصصی حوزه کشاورزی با هم مقایسه کرد. این مقایسه نشان داد در معیار ارتباط و اعتبار، سه موتور کاوش گوگل، لایکاس و یاهو به ترتیب رتبه اول تا سوم را داشتند و موتورهای کاوش تخصصی کشاورزی در رتبه‌های آخر بودند.

«شیلکتینگ و نیلسن»^۳ (1966) چهار موتور جستجوی پر استفاده در وب جهانگستر شامل لایکاس، آلتاویستا، اکسایت و اینفوسیک را ارزشیابی کردند. ساده‌ترین گزارش از یافته‌ها این بود که از مجموع ۲۰۰ پیوند ارائه شده با پنج جستجو، در مجموع ۵۴ پیوند مربوط به موضوع پیدا شد. لایکاس بیشترین پیوندهای مربوط را با ۱۹ مورد پیدا کرد. اکسایت با ۱۴ مورد، اینفوسیک با ۱۲ مورد و آلتاویستا با ۹ مورد پیوند مربوط، در ردیف‌های بعدی قرار داشتند.

«گانگ وو و جی لی»^۴ (1999) هفت موتور جستجوی آلتاویستا، اینفوسیک، اکسایت، نورثرن لایت^۵، یاهو، هات‌بات، و مدیکال ورلد سرچ^۶ را مقایسه کردند تا کارایی آنها را در پاسخگویی به سؤالهای کاربران علوم بهداشتی بسنجند. آنها سه موتور جستجوی اینفوسیک، آلتاویستا و اکسایت را به عنوان موتورهای کاوش رده بالا رتبه‌بندی کردند. این موتورها در مجموع از نظر روزآمد بودن منابع اطلاعاتی و ویژگیهای پیشرفته جستجو، عملکرد خوبی داشتند. یاهو و نورثرن لایت برای جستجوی منابع دارویی و محصولات طبیعی موتورهای خوبی بودند، ولی همچنین دارای بالاترین تعداد پیوندهای غیرفعال می‌باشند. هات بات پوشش جامع و ویژگیهای جستجوی پیشرفته را دارد و موتور خوبی برای جستجوی تصاویر و فایل‌های چندرسانه‌ای است، اما مشکل پیوندهای تکراری در این موتور کاوش به چشم می‌خورد. همچنین، این مطالعه نشان

-
1. Agrisurf.
 2. Web-agri.
 3. Schlichting and Nilsen.
 4. Gang wu and Jie li.
 5. Northern Light.
 6. Mwsearch.

داد حتی با استفاده از بهترین موتورهای کاوش، فقط نیمی از موارد بازیابی شده، مرتبط خواهند بود.

«مارتینز و سانچه»^۱ در مقاله‌ای با عنوان «مقایسه‌ای بین ابزارهای جستجوی اینترنت» ده موتور کاوش را برای تحقیق انتخاب کردند. در این پژوهش، امکانات مختلف جستجو مانند عملکرد بولی، کوتاه‌سازی، استفاده از پرانتز و امکانات پیشرفته نظیر اصلاح نتیجه جستجو، جستجو در فیلدهای خاص، امکان محدود کردن موضوع و نیز زبان کنترل‌شده، مقایسه و ارزیابی شد. آنها به این نتیجه رسیدند که هیچ رابطه‌ای میان رایج و عامه‌پسند بودن ابزارهای کاوش مورد مطالعه با امکانات جستجوی اطلاعات آنها وجود ندارد. دو نمونه آشکار در راهنمای موضوعی یاهو و موتور کاوش نورثرن لایت قابل مشاهده است. اگرچه موتور کاهش نورثرن لایت از نظر درجه مقبولیت آخرین رتبه را دارد، در فهرست رتبه‌بندی امکانات جستجو اولین رتبه را از آن خود کرد. حال آنکه یاهو که از رایج‌ترین و عامه‌پسندترین ابزارهای کاوش است، از لحاظ بازیابی اطلاعات رتبه بالایی را کسب نکرده است (به نقل از آزادی، ۱۳۸۴).

«برودر» (2004) بر پایه تحلیل لاگ‌فایل‌ها و بررسی کاربران موتور جستجوی آلتاویستا، دریافت که هر یک از انواع پرسشها نماینده شمار قابل توجهی از کل جستجوهاست. پرسشهای راهبری ۲۰ تا ۲۴/۵٪، پرسشهای اطلاعاتی ۳۹ تا ۴۸٪، و پرسشهای تراکنشی ۲۲ تا ۳۶٪ از کل پرسشها را شامل می‌شوند.

«رز و لوینسون»^۲ (2004) تعریفهای مختلفی از پرسشهای راهبری و تراکنشی ارائه داده‌اند. با توجه به بررسی رز، پرسشهای اطلاعاتی ۶۱ تا ۶۳٪، پرسشهای تراکنشی ۲۱ تا ۲۷٪ و پرسشهای اطلاعاتی ۱۱ تا ۱۵٪ از کل پرسشها را شامل می‌شوند.

«لواندووسکی» (2006) در یک بررسی که از داده‌های مربوط به سه موتور جستجوی آلمانی استفاده شد، دریافت که پرسشهای راهبری ۴۰٪ کل پرسشها را تشکیل می‌دهند.

1. Martinez and Sanche.

2. Rose and Levinson.

روش‌شناسی پژوهش

در این پژوهش، از دو روش متفاوت استفاده شده است. برای پیدا کردن موتورهای کاوش عمومی وب که کاربرد بیشتری دارند، از روش اسنادی یا کتابخانه‌ای و برای سنجش میزان دقت موتورهای کاوش، از روش مقایسه‌ای استفاده شده است. برای انجام مقایسه، پنج موتور جستجوی گوگل، یاهو، بینگ، اسک و موتور جستجوی آوال بر پایه این شرط که باید نمایه شخصی خود را داشته باشند، انتخاب شدند.

Google	https://www.google.com/
Yahoo	http://search.yahoo.com/
Bing	http://www.bing.com
Ask	http://www.ask.com/
AOL	http://www.aol.com/

این ابزارهای کاوش در دسته موتورهای کاوش عمومی و جزء پر استفاده‌ترین آنها هستند. راهنماهای موضوعی، ابرموتورهای کاوش، و موتورهای کاوش کم استفاده، در این جامعه وارد نشده‌اند. برای انتخاب موتورهای کاوش از رتبه‌بندی موتورهای کاوش در سایت الکسا^۱ استفاده شد. سایت الکسا یک سایت رتبه‌بندی سایتها بر اساس میزان ترافیک مصرفی آنهاست؛ به این معنا که متناسب با میزان پهنای باند مصرفی، رتبه سایت را مشخص می‌کند. رتبه‌ها از ۱ شروع شده و تا اعدادی حدود ۱۰ میلیون ادامه دارد. در این سایت، برعکس پیج رنک، هرچه عدد پایین‌تر باشد بهتر است؛ یعنی رتبه‌هایی که به یک نزدیکتر، سایت‌هایی با میزان بازدید بیشتری هستند. البته، ذکر این نکته لازم است که لزوماً ترافیک سایت تابعی از تعداد بازدید نیست، اما برای به دست آوردن یک تقریب، معیار خوبی است (الکسا، ۲۰۱۲).

برای جمع‌آوری اطلاعات، نخست مفهوم پرسشهای راهبری برای گروهی از دانشجویان دانشگاه تهران تعریف و از آنها خواسته شد آخرین پرسشهایی از این دست را که در موتور جستجو استفاده کرده‌اند، ارائه کنند. ۱۰۰ پرسش به زبان فارسی همراه عبارت دقیق پرسش، توصیف کوتاهی از نیاز اطلاعاتی و آدرس اینترنتی صفحه مورد نظر، از یکایک دانشجویان گردآوری شد.

1. alexa.com.

تمامی پرسشها توسط پژوهشگر در قسمت جستجوی هر یک از موتورهای جستجوی مورد بررسی وارد و موقعیت نتایج صحیح ثبت شد. اگر صفحه درست میان ۱۰ نتیجه اول یافت نمی‌شد، جستجو متوقف می‌شد.

برخلاف بررسی‌هایی که در مورد پرسشهای اطلاعاتی انجام شده است، قضاوت در مورد ربط نتایج در این مطالعه مشکل نیست، زیرا واضح است که کدام نتیجه مرتبط و کدام نتیجه نامرتب است.

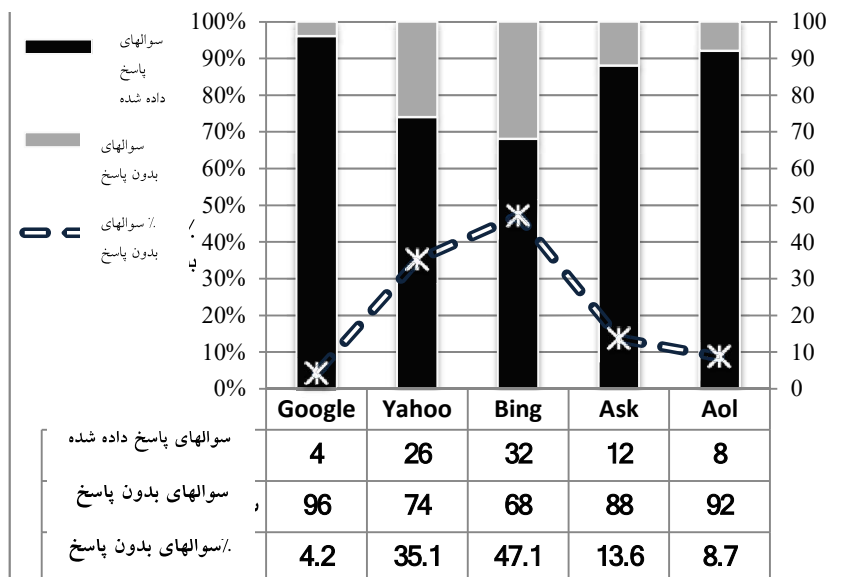
بعضی از شرکتهای بزرگ، صفحات خانگی منطقه‌ای یا مختص به یک زبان را ایجاد می‌کنند. از آنجا که پرسشهای مطرح شده در این پژوهش به زبان فارسی هستند، فقط صفحه خانگی فارسی مورد نظر است.

در این پژوهش، صفحه درخواستی، صفحه‌ای تعریف شده است که دقیقاً با آدرس اینترنتی که کاربر تعیین کرده است مطابقت دارد؛ یا صفحه‌ای است که با تمامی معیارهای مدنظر کاربر در توصیف نیاز اطلاعاتی او مطابقت دارد. برای پرسشهای راهبری، مقیاسهای مختلفی را که در پژوهش «هاوکینگ و کراسول» (2005) تشریح شده است، برای داده‌ها محاسبه شده است. به منظور محاسبه عملکرد بازیابی برای پرسشهای راهبری، از سنجه‌های بازیابی متعددی نظیر سنجه موفقیت N^1 و رتبه معکوس میانگین^۲ که در پژوهش‌های مک‌فارلان (2007)، کراسول و هاوکینگ^۳ (2005) به کار گرفته شده، استفاده شد.

یافته‌ها

نمودار ۱، تعداد و درصد سؤالهای راهبری پاسخ داده شده توسط پنج موتور جستجوی گوگل، یاهو، بینگ، اسک و موتور جستجوی آوال را نشان می‌دهد. از کل پرسشهای ارائه شده به موتورهای جستجو، بیشترین درصد بازیابی با $95/8\%$ (۹۶ سوال) متعلق به موتور جستجوی گوگل می‌باشد. ضعیف‌ترین عملکرد با $59/2\%$ (۶۸ سوال) و $64/9\%$ (۷۴ سوال) به ترتیب به دو موتور جستجوی بینگ و یاهو تعلق دارند.

1. Success N.
2. Mean reciprocal rank.
3. Craswell and Hawking.



نمودار ۱. عملکرد هر یک از موتورهای جستجو در پاسخ به سؤالاتها

از مجموع ۱۰۰ سؤال راهبری ارائه شده به موتورهای جستجو، تنها پاسخ ۶۰٪ آنها توسط همه موتورهای جستجو بازیابی شده است (جدول ۱). از سویی دیگر، هیچ یک از موتورهای جستجوی مذکور قادر به یافتن پاسخی برای ۲٪ از پرسشها نبوده‌اند. جزئیات مربوط به عملکرد کلی موتورهای جستجو، در جدول ۱ قابل مشاهده است.

جدول ۱. عملکرد کلی موتورهای جستجو

بدون پاسخ		پاسخ داده شده		پرسشها موتورهای جستجو
درصد	فراوانی	درصد	فراوانی	
۲٪	۲	-	-	هیچ کدام
-	-	۲٪	۲	یک موتور جستجو
-	-	۲٪	۲	دو موتور جستجو
-	-	۲۴٪	۲۴	سه موتور جستجو
-	-	۱۰٪	۱۰	چهار موتور جستجو
-	-	۶۰٪	۶۰	پنج موتور جستجو

مقایسه عملکرد موتورهای کاوش در هدایت کاربران ... / ۷۱

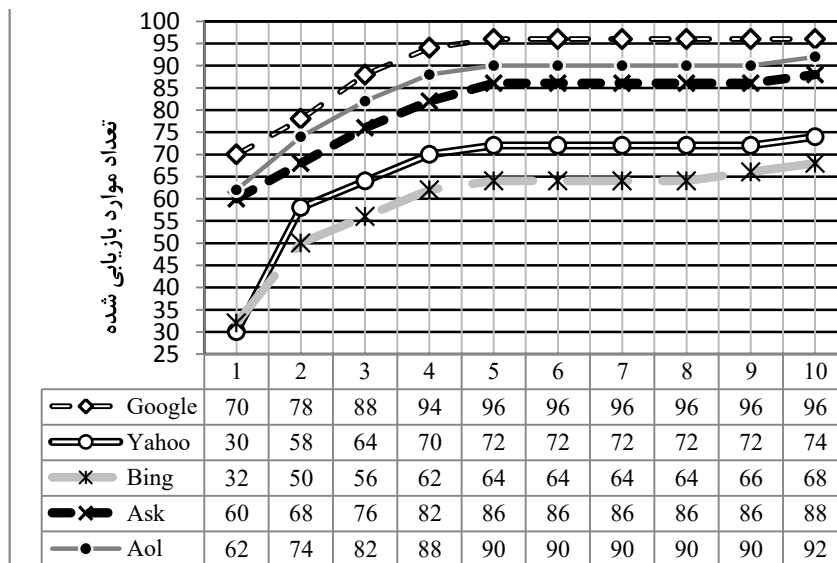
داده‌های جدول ۲ نشان می‌دهد در مجموع موتور جستجوی گوگل با ۹۶٪ بهترین عملکرد را در بازیابی به خود اختصاص می‌دهد و موتور جستجوی آوال و اسک به ترتیب با ۹۲٪ و ۸۸٪ در رتبه‌های بعدی قرار دارند. در موتور جستجوی گوگل از مجموع ۹۶ مورد پاسخ بازیابی شده ۷۲/۹۲٪ آن در رتبه اول فهرست نتایج بازیابی شده قرار دارد و از لحاظ قرار گرفتن نتایج در رتبه اول فهرست نتایج بازیابی شده، بهترین عملکرد را به خود نسبت داده است. دو موتور جستجوی اسک و آوال به ترتیب با ۶۸/۱۸٪ و ۶۷/۳۹٪ مورد پاسخ بازیابی شده در رتبه اول فهرست نتایج در مقام‌های بعدی قرار دارند. همچنین، موتور جستجوی بینگ و یاهو به ترتیب با ۴۷/۰۶٪ و ۴۰/۵۴٪ در قرار دادن صفحات درخواستی در رتبه اول موفق عمل کرده‌اند و در مقام‌های بعدی قرار دارند.

جدول ۲. تعداد کلی نتایج بازیابی شده در رتبه‌های اول تا دهم

Aol			Ask			Bing			Yahoo			Google			موتور جستجو
رتبه	تعداد	درصد	رتبه	تعداد	درصد	رتبه	تعداد	درصد	رتبه	تعداد	درصد	رتبه	تعداد	درصد	
رتبه ۱	۷۰	-	۶۲	۶۸/۱۸	-	۶۰	۴۷/۰۶	-	۳۲	۴۰/۵۴	-	۳۰	۷۲/۹۲	-	۷۰
رتبه ۲	۷۸	۸	۷۴	۷۷/۲۷	۸	۶۸	۷۳/۵۳	۱۸	۵۰	۷۸/۳۸	۲۸	۵۸	۸۱/۲۵	۸	۷۸
رتبه ۳	۸۸	۱۰	۸۲	۸۶/۳۶	۸	۷۶	۸۲/۳۵	۶	۵۶	۸۶/۴۹	۶	۶۴	۹۱/۶۷	۱۰	۸۸
رتبه ۴	۹۴	۶	۸۸	۹۳/۱۸	۶	۸۲	۹۱/۱۸	۶	۶۲	۹۴/۵۹	۶	۷۰	۹۷/۹۲	۶	۹۴
رتبه ۵	۹۶	۲	۹۰	۹۷/۷۳	۴	۸۶	۹۴/۱۲	۲	۶۴	۹۷/۳۰	۲	۷۲	۱۰۰	۲	۹۶
رتبه ۶	۹۶	۰	۹۰	۹۷/۷۳	۰	۸۶	۹۴/۱۲	۰	۶۴	۹۷/۳۰	۰	۷۲	۱۰۰	۰	۹۶
رتبه ۷	۹۶	۰	۹۰	۹۷/۷۳	۰	۸۶	۹۴/۱۲	۰	۶۴	۹۷/۳۰	۰	۷۲	۱۰۰	۰	۹۶
رتبه ۸	۹۶	۰	۹۰	۹۷/۷۳	۰	۸۶	۹۴/۱۲	۰	۶۴	۹۷/۳۰	۰	۷۲	۱۰۰	۰	۹۶
رتبه ۹	۹۶	۰	۹۰	۹۷/۷۳	۰	۸۶	۹۷/۰۶	۲	۶۶	۹۷/۳۰	۰	۷۲	۱۰۰	۰	۹۶
رتبه ۱۰	۹۶	۰	۹۲	۱۰۰	۲	۸۸	۱۰۰	۲	۶۸	۱۰۰	۲	۷۴	۱۰۰	۰	۹۶

درصد بازیابی	%۹۶	%۷۴	%۶۸	%۸۸	%۹۲
-----------------	-----	-----	-----	-----	-----

نمودار ۲ تعداد دفعات قرار گرفتن پاسخ پرسشهای راهبری در رتبه اول تا دهم، برای هر یک از موتورهای جستجو را نشان می‌دهد. همان‌طور که در نمودار مذکور قابل مشاهده است، در اکثر این موتورهای جستجو از رتبه‌های چهارم به بعد شیب نمودار بسیار کم و گاه صفر می‌گردد. در برخی از این موتورهای جستجو نظیر گوگل و آوال و اسک، تعداد نتایج ظاهر شده در رتبه‌های اول بالاست و از رتبه‌های چهارم به بعد در تعداد نتایج مورد نظر بهبود چندانی حاصل نمی‌شود. در دو موتور جستجوی یاهو و بینگ، میزان نمایان شدن صفحات درخواستی در رتبه‌های دوم قابل توجه است.



نمودار ۲. عملکرد موتورهای جستجو در بازیابی صفحات درخواستی در رتبه‌های اول تا دهم

چنان که در جدول ۳ ملاحظه می‌شود، با در نظر گرفتن ۵ و ۱۰ نتیجه اول فهرست نتایج بازیابی شده، کمترین درصدها به ترتیب با $\frac{27}{14}\%$ و $\frac{37}{14}\%$ به موتور جستجوی گوگل و بیشترین درصدها به ترتیب با $\frac{14}{10}\%$ و $\frac{146}{67}\%$ به موتور جستجوی یاهو تعلق دارند. این بدان معناست که با در نظر گرفتن تعداد نتایج بیشتر، بهبود قابل توجهی در نتایج بازیابی شده توسط موتور جستجوی گوگل حاصل نشده

_____ مقایسه عملکرد موتورهای کاوش در هدایت کاربران ... / ۷۳

است بر عکس، در موتور جستجوی یاهو در تعداد نتایج بازیابی شده، بهبود نسبتاً خوبی حاصل شده است. نسبت یافته‌ها در موتورهای که درصد پایین‌تری از نتایج دلخواه را در رتبه اول قرار می‌دهند، بالاتر است.

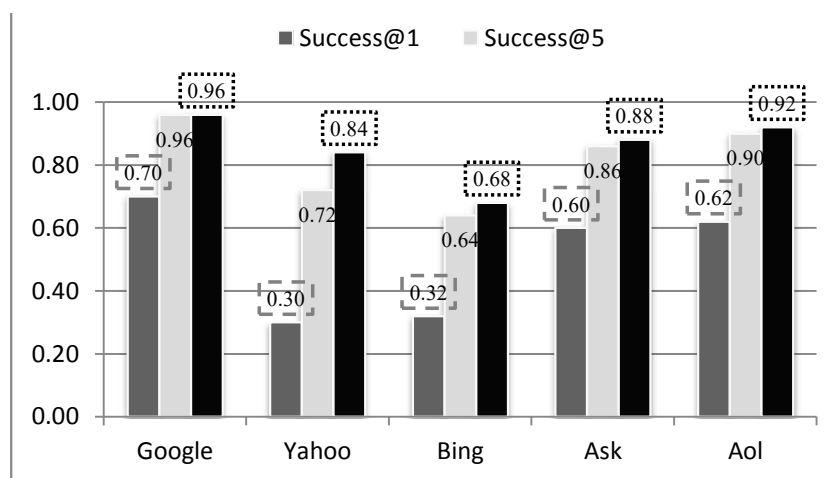
جدول ۳. نسبت بازیابی موتورهای کاوش با در نظر گرفتن تعداد نتایج مختلف

موتور جستجو	Google	Yahoo	Bing	Ask	Aol
نسبت یافته‌ها ^۱	درصد	درصد	درصد	درصد	درصد
اولین به پنجمین	۳۷/۱۴٪	۱۴۰٪	۱۰۰٪	۴۳/۳۳٪	۴۵/۱۶٪
اولین به دهمین	۲۷/۱۴٪	۱۴۶/۶۷٪	۱۱۲/۵٪	۴۶/۶۷٪	۴۸/۳۹٪
پنجمین به دهمین	۰	۲/۷۸٪	۶/۲۵٪	۲/۳۳٪	۲/۲۲٪

برای محاسبه عملکرد بازیابی برای پرسشهای راهبری، از شاخصهای بازیابی متعددی نظیر سنجه موفقیت N و رتبه معکوس میانگین که در پژوهشهای مک‌فارلان (2007)، کراسول و هاوکینگ (2005) به کار گرفته شده، استفاده می‌کنیم. نمودار ۳ نشان‌دهنده شاخص موفقیت N برای بیان میزان اثربخشی موتورهای جستجوست. همان‌طور که در نمودار مذکور مشاهده می‌شود، با در نظر گرفتن نتیجه اول، موتور جستجوی گوگل دارای بهترین و موتور جستجوی یاهو دارای بدترین شاخص موفقیت می‌باشد (۰/۷۰ و ۰/۳۰). با در نظر گرفتن ۵ نتیجه اول، موتورهای جستجوی گوگل و بینگ به ترتیب با ۰/۹۶ و ۰/۶۴ دارای بهترین و بدترین میزان بازیابی هستند و با در نظر گرفتن ۱۰ نتیجه اول، موتور جستجوی گوگل و بینگ با ۰/۹۶ و ۰/۶۸ دارای بهترین و بدترین میزان اثربخشی می‌باشند. مقادیر به دست آمده برای موتورهای جستجو که در جدول ۴ نتایج بیان شده برای سنجه موفقیت N را تأیید می‌نمایند. علاوه

1. Gain ratio.

بر آن، رتبه معکوس میانگین پنج موتور جستجوی مورد استفاده در این پژوهش، در جدول ۴ نشان داده شده است. بیشترین مقدار متعلق به موتور جستجوی گوگل (۰/۷۹) و کمترین مقدار متعلق به موتور جستجوی بینگ (۰/۴۵) می‌باشد.



نمودار ۳. میزان موفقیت در ارائه نتایج توسط موتورهای جستجو

جدول ۴. وضعیت بازیابی موتورهای جستجو در رتبه‌های مختلف

موتورهای کاوش	میانگین رتبه	رتبه اول		رتبه اول تا پنجم		رتبه اول تا دهم	
		اول	دوم	اول	پنجم	اول	دهم
Google	۰/۷۹	۰/۷۰	۰/۳۰	۰/۹۶	۰/۰۴	۰/۹۶	۰/۰۴
Yahoo	۰/۴۸	۰/۳۰	۰/۷۰	۰/۷۲	۰/۲۸	۰/۸۴	۰/۳۶
Bing	۰/۴۵	۰/۳۲	۰/۶۸	۰/۶۴	۰/۳۶	۰/۶۸	۰/۳۲
Ask	۰/۶۹	۰/۶۰	۰/۴۰	۰/۸۶	۰/۳۴	۰/۸۸	۰/۱۲
Aol	۰/۷۳	۰/۶۲	۰/۳۸	۰/۹۰	۰/۱۰	۰/۹۲	۰/۰۸

بنابراین، با توجه به مقادیر شاخص موفقیت N و رتبه معکوس میانگین به دست آمده برای موتور جستجوی گوگل، فرضیه پژوهش مبنی بر اینکه «در بازیابی پاسخ پرسشهای راهبری، موتور جستجوی گوگل نسبت به سایر موتورهای جستجوی مورد استفاده در این پژوهش بیشترین اثربخشی را دارد» تأیید می‌گردد.

بحث و نتیجه‌گیری

یافته‌های پژوهش نشان می‌دهند سه موتور جستجوی گوگل، آوال و اسک تقریباً با ۹۶، ۹۱ و ۸۷٪ موفقیت در بازیابی، عملکرد نسبتاً خوبی در بازیابی پاسخ پرسشهای راهبری به زبان فارسی دارند، اما دو موتور جستجوی یاهو و بینگ تقریباً با ۳۵ و ۴۷٪ عدم موفقیت، عملکرد ضعیفی در بازیابی پاسخ پرسشهای راهبری به زبان فارسی دارند. یافته‌های ما در مورد عملکرد موتور جستجوی گوگل در بازیابی سؤالهای راهبری، با یافته‌های لواندوسکی (2008) همسوست، اما با نتایج ارائه شده برای موتور جستجوی یاهو همسویی ندارد. این نکته که موتور جستجوی گوگل در بازیابی پاسخ پرسشهای راهبری به زبان دیگری غیر از زبان انگلیسی عملکرد خوبی دارد، حایز اهمیت است. از بین ۱۰۰ سوال ارائه شده به موتورهای جستجو، تنها ۲ سؤال وجود داشت که هیچ یک از موتورهای جستجو مورد استفاده در این پژوهش قادر نبودند پاسخی برای آنها بیابند. از سویی دیگر، تنها ۶۰٪ کل پرسشها توسط همه موتورهای جستجو پاسخ داده شدند. هر چند درصد سؤالهای که پاسخ آنها توسط همه پنج موتور جستجو بازیابی می‌شوند نسبتاً بالاست، باید به این نکته توجه داشت که حدود ۴۰٪ سؤالهای مورد استفاده در این پژوهش هستند که تنها بعضی از موتورهای جستجو توانستند به آن پاسخ گویند. شاید یکی از دلایل آن نمایه نشدن صفحات وب فارسی توسط آن موتورهای جستجو باشد. یکی از نکات کاربری قابل استنباط از یافته‌های این بخش آن است که کاربران هنگام مواجه شدن با ناکامی در استفاده از موتور جستجویی خاص، برای نیل به پاسخ مناسب می‌توانند از دیگر موتورهای جستجو نیز استفاده کنند.

نکته دیگری که از نتایج به دست آمده قابل استنباط است، آن است که نتایجی که توسط گوگل و اسک و آوال به دست آمده‌اند، اغلب در رتبه اول صفحه نتایج قرار دارند. علاوه بر آن، همه موتورهای جستجوی مورد نظر در این پژوهش از نتیجه اول به

نتیجه پنجم پیشرفت دارند. به عبارت دیگر، آنها تقریباً در بیش از ۹۰٪ موارد قادرند صفحات درخواستی خود را بدون مرور کردن نتایج ارائه شده در پایین فهرست نتایج ارائه شده به کاربر در اولین صفحه نتایج نشان دهند. این مسئله در بازیابی سؤالهای راهبری با بهره‌گیری از موتورهای جستجو دارای اهمیت بسزایی است، زیرا کاربران به بررسی نتایجی که پایین‌تر از وسط اولین صفحه نمایش قرار دارند (گرانکا و همکاران، ۲۰۰۴)، یا حتی نتایجی که بعد از ۱۰ نتیجه اول ارائه می‌گردند، تمایل چندانی ندارند. با در نظر گرفتن ۱۰ نتیجه اول، عملکرد آوال و اسک در بازیابی پاسخ سؤالهای راهبری به عملکرد موتور جستجوی گوگل بسیار نزدیک می‌باشند. اما اگر رتبه‌های اول مد نظر قرار گیرد، عملکرد این موتورهای جستجو ضعیف‌تر است و موتور جستجوی آوال درصد کمتری از صفحات درخواستی را بازیابی می‌کند. اگر چه موتورهای جستجوی یاهو و بینگ در قرار دادن صفحات درخواستی در رتبه اول فهرست رکوردهای بازیابی شده موفق عمل نکرده‌اند (۴۰/۵۴٪ و ۴۷/۰۶٪) با این حال، با در نظر گرفتن رتبه‌های دوم، عملکرد این موتورها بهبود قابل توجهی می‌یابد و در صورتی که کل ۱۰ نتیجه بازیابی شده در نظر گرفته شود، این موتورها نیز عملکرد نسبتاً بالایی دارند. با توجه به اهمیت قرار دادن صفحه درخواستی در مکان اول، این نتیجه قابل استنباط است که مشکل بسیاری از موتورهای جستجو، پیدا کردن صفحه درخواستی نیست، بلکه رتبه‌بندی آن بر اساس نیت کاربر است.

هر چند دو موتور جستجوی یاهو و بینگ در مقایسه با سه موتور جستجوی دیگر مورد استفاده در این پژوهش عملکرد ضعیف‌تری دارند، با نگاهی به نتایج ارائه شده در جدول ۲ می‌توان عملکرد آنها را خوب قلمداد کرد. درحالی که این موتورهای جستجو در مورد ارائه نتایج در رتبه اول نسبت به گوگل، اسک و آوال عملکرد بدتری داشته‌اند، اما با در نظر گرفتن چند نتیجه بعد آنها بهبود می‌یابد تا جایی که نسبت یافته‌های این موتورهای جستجو نشان‌دهنده افزایش بازیابی صفحات درخواستی در رتبه‌های بعدی در فهرست ۵ و ۱۰ نتیجه ارائه شده در صفحه اول می‌باشد (جدول ۳).

طبق نتایج به دست آمده در این پژوهش، از آنجا که در بازیابی پرسشهای راهبری به زبان فارسی توسط موتورهای جستجو اکثر نتایج در رتبه اول تا پنجم (جدول ۲)

صفحه نتایج قرار می‌گیرند، جستجوگری که دنبال یافتن پاسخ برای یک پرسش راهبری می‌باشد، چنانچه جواب مورد نظر خود را میان چند نتیجه اول پیدا نکند، نگاه کردن به صفحات بیشتر در فهرست نتایج بی‌حاصل است. بنابراین، چنین به نظر می‌رسد که جستجوگر باید دوباره پرسش خود را صورت‌بندی کند و از هدر دادن وقت برای مرور سایر نتایج بپرهیزد.

همان‌طور که انتظار می‌رفت، نسبت یافته‌ها در موتورهایی که درصد بالاتری از نتایج دلخواه را در رتبه اول فهرست نتایج قرار می‌دهند، پایین‌تر است. اکثر نتایجی که توسط گوگل و آوال و اسک به دست آمده‌اند، در رتبه اول صفحه نتایج قرار داشته‌اند و به همین علت نسبت یافته‌ها از رتبه اول نسبت به کل رتبه‌های بررسی شده بسیار پایین است. هر سه موتور جستجو از نتیجه اول به نتیجه دوم پیشرفت داشته‌اند، در حالی که موتورهای جستجوی یاهو و بینگ در مورد ارائه نتایج در رتبه اول نسبت به گوگل، اسک و آوال عملکرد بدتری داشته‌اند، اما با در نظر گرفتن چند نتیجه بعد آنها بهبود می‌یابد تا جایی که نسبت یافته‌های این موتورهای جستجو (از اول به دهم) به ترتیب ۱۴۶/۶۷٪ و ۱۱۲/۵٪ است که نشان‌دهنده افزایش بازیابی صفحات درخواستی در رتبه‌های بعدی در فهرست ۱۰ نتیجه ارائه شده در صفحه اول است. به عبارت دیگر، در این دو موتور جستجو نتایج اغلب در انتهای فهرست نتایج بازیابی شده ظاهر می‌شوند. نکته دیگر آنکه در هر پنج موتور جستجوی مورد استفاده در این پژوهش، با آنکه نسبت یافته‌ها از اولین نتیجه به پنجمین نتیجه افزایش چشمگیری دارد، نسبت یافته‌ها از پنجمین نتیجه به دهمین نتیجه افزایش بسیار اندکی دارد. این امر حکایت از آن دارد که اغلب نتایج بازیابی شده در این موتورهای جستجو، در رتبه‌های اول تا پنجم ظاهر می‌شوند.

در کل، با توجه به نتایج به دست آمده در این پژوهش برای موتورهای جستجو، موتور جستجوی گوگل در بازیابی پاسخ پرسش‌های راهبری نیز همانند پرسش‌های اطلاعاتی دارای بهترین عملکرد است (جدول ۴ و نمودار ۳). این امر می‌تواند دلیلی برای پرسش باشد که چرا کاربران موتور جستجوی گوگل را به سایر موتورهای جستجو ترجیح می‌دهند.

منابع

- آزادی، قاسم (۱۳۸۴). میزان دقت موتورهای کاوش وب در بازیابی اطلاعات کتابداری و اطلاع‌رسانی. فصلنامه کتاب، ۶۳: ۱۱۱-۱۲۲.
- کمیجانی، احمد (۱۳۸۱). مقایسه کارایی موتورهای کاوش عمومی و تخصصی وب در بازیابی اطلاعات کشاورزی. پایان‌نامه کارشناسی ارشد کتابداری و اطلاع‌رسانی، دانشکده روانشناسی و علوم تربیتی، دانشگاه تهران.
- Brooks, T.A. (2003), Web Search: how the Web has changed information retrieval. *Information Research*, 8(3), paper no. 154. Available at: <http://informationr.net/ir/8-3/paper154.html>. (accessed 9 July 2012).
 - Craswell, N. and Hawking, D. (2005), Overview of the TREC-2004 Web Track. Available at: trec.nist.gov/pubs/trec13/papers/WEB.OVERVIEW.pdf (accessed 9 July 2012).
 - Granka, L.A., Joachims, T. and Gay, G. (2004), Eye-tracking analysis of user behavior in WWW search, *Proceedings of Sheffield SIGIR, Twenty-Seventh Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, 56: 478-479. Available at: http://www.summarization.com/~radev/767w10/papers/Week09/search_log_mining/Granka04.pdf (accessed 9 July 2012).
 - Hariri, N. (2011), Relevance ranking on Google: Are top ranked results really considered more relevant by the users?, *Online Information Review*, 35(4): 598 – 610. Available at: <http://www.emeraldinsight.com/journals.htm?articleid=1944421&show=abstract> (accessed 9 July 2012).
 - Hawking, D. and Craswell, N. (2004), The Very Large Collection and Web Tracks. Available at: http://es.csiro.au/pubs/trecbook_for_website.pdf (accessed 9 July 2012).
 - Jansen, B. J, Booth, D.L and Spink A. (2007), Determining the user intent of web search engine queries. In: *Proceedings of the 16th international conference on World Wide Web*, ACM, New York, USA, 1149–1150. Available at: <http://dl.acm.org/citation.cfm?id=1242739>. (accessed 9 July 2012).
 - Jansen, B.J., Booth, D.L., and Spink, A. (2008), Determining the informational, navigational, and transactional intent of Web queries, *Information Processing & Management*, 44(3): 1251-1266.
 - Kang, I.H., and Kim, G. (2003), Query type classification for web document retrieval. In: *SIGIR 2003: Proceedings of the 26th annual international ACM SIGIR conference on Research and development in information retrieval*, 64–71. ACM Press, New York. Available at: <http://wing2.ddns.comp.nus.edu.sg/downloads/keyphraseCorpus/161/161.pdf>. (accessed 9 July 2012).
 - Keane, M.T., O'Brien, M. and Smyth, B. (2008), Are people biased in their use of search engines?, *Communications of the ACM*, 51(2):49-52. Available at: <http://irserver.ucd.ie/dspace/bitstream/10197/1643/3/MOB.ACM.v3>. (accessed 9 July 2012).
 - Kilgour, F.G. (1999), Retrieval effectiveness of surname-title-word searches for known items by academic library users, *Journal of the American Society for Information Science*, 50(3): 265-270. Available at:

- [http://onlinelibrary.wiley.com/store/10.1002/\(SICI\)1097-4571\(1999\)50:3%3C265::AID-ASI9%3E3.0.CO;2-R/asset/9ftp.pdf?v=1&t=h9e1mks2&s=02ed9b8c651d6268b600ccb033cf62408e6bd4](http://onlinelibrary.wiley.com/store/10.1002/(SICI)1097-4571(1999)50:3%3C265::AID-ASI9%3E3.0.CO;2-R/asset/9ftp.pdf?v=1&t=h9e1mks2&s=02ed9b8c651d6268b600ccb033cf62408e6bd4). (accessed 9 July 2012).
- Lewandowski, D. (2006), Query types and search topics of German Web search engine users, *Information Services & Use*, 26(4): 261-269. Available at: <http://eprints.rclis.org/bitstream/10760/10126/1/isu2006.pdf>. (accessed 9 July 2012).
 - Lewandowski, D. (2008), The Retrieval Effectiveness of Web Search Engines: Considering Results Descriptions, *Journal of Documentation*, 64(6): 915 - 937. Available at: <http://www.emeraldinsight.com/journals.htm?issn=0022-0418&volume=64&issue=6&articleid=1752241&show=pdf>. (accessed 9 July 2012).
 - Lewandowski, D. and Höchstötter, N. (2008), Web Searching: A Quality Measurement Perspective. Available at: http://eprints.rclis.org/bitstream/10760/9595/1/LewHoech_Preprint.pdf. (accessed 9 July 2012).
 - Lewandowski, D., Drechsler, J. and Mach S.V. (2012). Deriving Query Intents from WebSearchEngineQueries. Available at: http://fileadmin/user_upload/lewandowski/doc/JASIST_Query_Intents_Preprint.pdf. (accessed 9 July 2012).
 - MacFarlane, A. (2007), Evaluation of Web search for the information practitioner, *Aslib Proceedings: New Information Perspectives*, 59(4-5): 352-366. Available at: www.emeraldinsight.com/journals.htm?articleid=1626450. (accessed 9 July 2012).
 - Westerveld, T., Kraaij, W. and Hiemstra, D. (2002), Retrieving web pages using content, links, URLs, and anchors. In: *The Tenth Text REtrieval Conf (TREC-10)*, NIST SP 500-250: 663-672. Available at: