



میزان دقت موتورهای کاوش وب در بازیابی اطلاعات کتابداری و اطلاع‌رسانی

قاسم آزادی^۱

چکیده

هدف اصلی این مقاله مقایسه میزان دقت موتورهای کاوش وب در بازیابی اطلاعات کتابداری و اطلاع‌رسانی است. برای این منظور کلیدواژه‌های تخصصی حوزه کتابداری و اطلاع‌رسانی در ۱۰ موضوع مختلف از مجله Library Trends انتخاب شد. این کلیدواژه‌ها در ۷ موتور کاوش جستجو شدند. در هر جستجو ۱۰ نتیجه اول ارزیابی شد. معیارهای تعیین میزان دقت عبارت بودند از: حضور کلیدواژه‌ها در عنوان مدرک بازیابی شده، همجواری کلیدواژه‌های مورد جستجو، حضور کلیدواژه‌ها در بخش ابرنشانه‌های کلیدواژه‌ای صفحات وب، و بسامد کلیدواژه‌ها در صفحات وب بازیابی شده.

نتایج نشان داد که اختلاف معناداری در میزان دقت موتورهای کاوش مورد بررسی در بازیابی اطلاعات کتابداری و اطلاع‌رسانی وجود ندارد. موتورهای کاوش از نظر دقت به این ترتیب رتبه‌بندی شدند: اینفوسیک، هات‌بات، گوگل، آلتا ویستا، اکسایت، لایکاس و وب کراولر. در ضمن موتورهای کاوش هات‌بات، اکسایت، اینفوسیک، آلتا ویستا، لایکاس، گوگل و وب کراولر به ترتیب بیشترین کاربرد ابرنشانه‌های کلیدواژه‌ای را داشتند. میزان خطای مشاهده شده در ۱۰ نتیجه اول در هر جستجو، معیار دیگری بود که از آن برای مقایسه موتورهای کاوش وب در بازیابی اطلاعات کتابداری و اطلاع‌رسانی استفاده شد. در این پژوهش موتورهای کاوش آلتا ویستا و اینفوسیک (به‌طور مشترک)، اکسایت، گوگل و لایکاس (به‌طور مشترک)، وب کراولر و هات‌بات به ترتیب کمترین میزان خطا را داشتند.

کلیدواژه‌ها

موتورهای کاوش، اینترنت، دقت، بازیابی اطلاعات، کتابداری و اطلاع‌رسانی

مقدمه

فهرستی از صفحات و منابع ارائه می‌شوند. مدارکی که بیشترین ارتباط را با موضوع داشته باشند در ابتدای فهرست به نمایش درمی‌آید (۱: ۸۷).

امروزه در حوزه بازیابی اطلاعات در وب به اولویت‌سنجی اینکه چه صفحات یا سایت‌هایی ابتدا نمایش داده شوند، توجه زیادی شده و در این زمینه مطالعات متعددی شده است. این موضوع، خود به یکی از مهم‌ترین عرصه‌های رقابت موتورهای کاوش تبدیل شده است. زیرا آنها سعی دارند، در ابتدا نتایج مرتبط‌تر با کلیدواژه‌های جستجو را نمایش دهند تا محبوبیت و اعتبار بیشتری در میان جامعه استفاده‌کنندگان مبتدی و حرفه‌ای پیدا کنند. مهم‌ترین معیارهایی که موتورهای کاوش از آن برای رتبه‌بندی نتایج خود استفاده می‌کنند عبارتند از: کلیدواژه‌ها در عنوان صفحه، ابرنشانه‌ها^۹، بسامد کلیدواژه‌ها، و همجواری کلیدواژه‌ها (۳: ۵۹-۶۰).

بیان مسئله

هر کدام از موتورهای کاوش، ویژگی‌های کاوش و بازیابی اطلاعات مختص خود را دارند و از نظر اندازه، سیاست‌های نمایه‌سازی، دقت، و شکل ظاهری صفحات جستجو با یکدیگر متفاوتند. بیشتر موتورهای کاوش اطلاعات میلیون‌ها مدرک اینترنتی را گردآوری کرده‌اند، با وجود این حتی بهترین و عظیم‌ترین آنها نیز نمی‌توانند بیش از نیمی از مدارک موجود در اینترنت را پوشش دهند و قابلیت جستجوی کلیه منابع اطلاعاتی موجود در شبکه اینترنت را ندارند (۹).

به همین دلیل آشنایی و شناخت ابزارها و موتورهای کاوش در اینترنت برای جستجوی دقیق و بازیابی منابع و مدارک مرتبط امری اجتناب‌ناپذیر است.

در این پژوهش، از طریق مرور منابع مختلف چاپی و الکترونیکی، هفت موتور کاوش عمومی که پرستفاده‌ترین بود انتخاب شد تا با استفاده از آنها، میان دقت موتورهای کاوش وب در بازیابی اطلاعات حوزه کتابداری و اطلاع‌رسانی مقایسه شود.

با درج کلیدواژه‌های تخصصی و رعایت اصول جستجو در هر یک از موتورهای کاوش، مشخص می‌شود کدام یک از آنها مرتبط‌ترین اطلاعات را ارائه می‌دهد. بدین طریق می‌توان موتورهای کاوش را رتبه‌بندی کرد و به کتابداران و اطلاع‌رسانان در انتخاب موتورهای کاوش مناسب برای این حوزه یاری رسانید.

اینترنت ابزاری برای تبادل اطلاعات در همه حوزه‌هاست. کتابخانه‌ها فهرست‌های رایانه‌ای خود را در اینترنت عرضه می‌کنند و فروشندگان پایگاه‌های اطلاعاتی منابع اطلاعاتی خود را از طریق اینترنت در دسترس قرار می‌دهند. سازمان‌های ملی، بین‌المللی و آموزشی، تحقیقاتی، تجاری و تقریباً هر سازمان و مؤسسه‌ای، اطلاعات گوناگونی را در اینترنت و به‌ویژه در وب در دسترس قرار می‌دهند.

اطلاعات قابل دسترسی از طریق اینترنت در قالب صفحات^۲ وب منتشر می‌شوند. برای دسترسی به این اطلاعات باید از محل استقرار آنها آگاه بود. محل فیزیکی این صفحات بر روی رایانه‌هایی است که سرور^۳ یا خدمت‌دهنده نامیده می‌شوند و در نقاط مختلفی از جهان پراکنده‌اند. به همین دلیل برای هر جستجو نمی‌توان به‌طور مستقیم به اصل صفحات مراجعه و اطلاعات مورد نیاز را بازیابی کرد.

برای حل این مشکل ابزارهای جستجو در اینترنت پدید آمده‌اند که مجموعه‌ای از اطلاعات منظم شده صفحات وب را در بردارند و در واقع هنگام جستجو در اینترنت مستقیماً به یک سایت‌ها^۴ و صفحات وب مراجعه نمی‌شود بلکه فقط در نمایه‌ها و فهرست‌های مختلف جستجو می‌شود. ابزارهای جستجو در اینترنت به چند گروه عمده تقسیم می‌شوند: فهرست‌های موضوعی^۵، موتورهای کاوش^۶ و ابرموتورهای کاوش^۷.

بیشتر افراد اصطلاح موتور کاوش را به‌طور عام برای همه ابزارهای جستجو به کار می‌برند. در صورتی که این ابزارها کاملاً یکسان نیستند و تفاوت‌های بارزی با یکدیگر دارند. موتورهای جستجو خدمت‌دهندگان اینترنتی هستند که با گردآوری اطلاعات کتابشناختی مدارک و منابع اینترنتی، امکان جستجو و بازیابی این مدارک و منابع را فراهم می‌آورند. موتورهای جستجو، اطلاعات مورد نیاز خود را با استفاده از نرم‌افزارهای ویژه و با کمترین دخالت مستقیم نیروی انسانی گردآوری و منظم می‌کنند (۹).

جستجو در موتورهای کاوش با درج کلیدواژه‌ها توسط کاربران در کادر جستجو^۸ آغاز می‌شود. موتور جستجو نیز پایگاه اطلاعاتی از پیش ساخته خود را جستجو می‌کند. در هر جستجو با توجه به اجزای سؤال مطرح شده منابع موجود در پایگاه اطلاعاتی موتور کاوش درجه‌بندی می‌شوند. سپس

تعریف عملیاتی اجزای مسئله

موتور کاوش: برنامه‌ای نرم‌افزاری است که با استفاده از کلیدواژه‌ها، منابع اطلاعاتی را از طریق اینترنت جستجو می‌کند و سیاهه‌ای از مدارک واجد آن کلیدواژه را ارائه می‌دهد. اطلاعات کتابداری و اطلاع‌رسانی: منظور مقالات انتخاب شده در حوزه کتابداری و اطلاع‌رسانی است که از طریق کلیدواژه جستجو شده‌اند. دقت: در این تحقیق منظور از دقت، میزان ارتباط نتایج با کلیدواژه‌های مورد جستجو توسط موتورهای کاوش است. معیارهایی که موتورهای کاوش برای ارائه اطلاعات مرتبط به کار می‌برند عبارتند از:

(۱) حضور کلیدواژه‌ها در عنوان مدرک بازاریابی شده؛

(۲) همجواری کلیدواژه‌های مورد جستجو؛

(۳) حضور کلیدواژه‌ها در بخش ابرنشان‌های کلیدواژه‌ای صفحات وب؛

(۴) بسامد کلیدواژه‌ها در صفحات وب بازاریابی شده. هر چقدر این معیارها در مدارک بازاریابی شده بیشتر رعایت شود، منابع مرتبط‌تر خواهند بود.

خطا: منظور از خطا در این پژوهش، منابع بازاریابی شده تکراری، یعنی مدارکی که بیش از یک بار در ده نتیجه اول نمایش داده می‌شود، و نیز پیوندهای کور و غیرفعال با پیغام‌های زیر است:

- This site is no longer available for the we apologize inconvenience.
- Cannot find server
- Server error
- Http 404 not found

پرسش‌های اساسی

۱. آیا اختلاف معناداری از لحاظ دقت در بازاریابی مدارک میان موتورهای کاوش مورد بررسی وجود دارد؟
۲. کدام یک از موتورهای کاوش مورد بررسی در زمینه کتابداری و اطلاع‌رسانی منابع مرتبط‌تری را بازاریابی می‌کنند؟
۳. کدام یک از موتورهای کاوش بیشترین میزان استفاده از ابر نشان‌های کلیدواژه‌ای را برای توصیف صفحات وب دارند؟
۴. کدام یک از موتورهای کاوش کمترین میزان خطا را در بازاریابی اطلاعات دارند؟

هدف و فایده پژوهش

هدف اصلی، مقایسه توانمندی موتورهای کاوش در بازاریابی اطلاعات حوزه کتابداری و اطلاع‌رسانی از شبکه جهانی وب است.

اهداف فرعی این تحقیق نیز از این قرار است:

۱. تعیین بهترین موتور جستجو از نظر دقت در بازاریابی اطلاعات کتابداری و اطلاع‌رسانی.
۲. تشخیص امکانات و ویژگی‌های هر یک از موتورهای جستجوی مورد بررسی به صورت تجربی و عملی. با انجام این پژوهش مشخص خواهد شد که کدام یک از موتورهای کاوش در حوزه کتابداری و اطلاع‌رسانی دقت بیشتری دارند و منابع مرتبط‌تری را بازاریابی می‌کنند. بر این اساس موتورهای کاوش رتبه‌بندی خواهند شد و متخصصان رشته کتابداری و اطلاع‌رسانی به بهترین موتور جستجو مراجعه می‌کنند و دقیق‌ترین نتایج را دریافت خواهند کرد. این کار به صرفه‌جویی در وقت برای جستجو کمک زیادی خواهد کرد.

جدول ۱. مجموع امتیازات موتورهای کاوش از نظر دقت، کاربرد ابرنشان، و کمبود خطا

موتورهای کاوش	امتیاز دقت	کاربرد ابرنشان	تعداد خطا
آلتاویستا	۴۳۲/۱	۱۸	۴
کسایت	۴۲۹/۸	۲۱	۵
ینفوسیک	۵۶۸	۲۰	۴
گوگل	۴۶۶	۱۴	۵
لایکاس	۴۰۴/۹	۱۷	۵
وب کرولر	۴۰۲/۲	۱۳	۶
هات بات	۵۰۳/۸	۲۳	۹

متغیرهای اساسی پژوهش

متغیر مستقل: عملکرد موتورهای کاوش انتخابی وب.
متغیر وابسته: میزان بازیابی اطلاعات مرتبط در حوزه کتابداری و اطلاع‌رسانی.

1. Rural Libraries and Information services (1995, No1) "rural Public Library" + "Information services"

2. Services to Remote Users (1998, No1) "Information Services" + "digital library" + "remote users"

3. Marketing of library and Information services (1995, No3) "Marketing" + "library and Information services" + "Political Processes"

4. The library Bill of Rights (1996, No1) "library Bill of Rights" + "School library" + "Media programs"

5. The Role of professional Associations (1997, No2) "Professional associations" + "library and Information science"

6. Resource sharing in a changing environment (1997, No3) "electronic resources" + "materials selection" + "collection development"

7. How classifications work: Problems and challenges in an Electronic Age (1998, No4) "Classification systems" + "Library" + "Electronic age"

8. Networked scholarly Publishing (1995, No4) "electronic Publishing" + "databases" + "evolution"

9. Buildings, Books and Bytes: Perspectives on the Benton (1997, No1)

پس از ورود کلیدواژه‌ها و جستجو در موتورهای کاوش مورد بررسی از میان نتایج هر موتور، ده نتیجه اول در هر جستجو برای ارزیابی به عنوان جامعه مورد مطالعه در این بخش انتخاب شدند.

روش پژوهش

در این تحقیق از دو روش متفاوت استفاده شده است: برای پیدا کردن موتورهای کاوش عمومی وب که کاربرد بیشتری دارند و نیز معیارهایی که موتورهای کاوش وب از آن استفاده می‌کنند تا دقت خود را در بازیابی مدارک بالا ببرند از روش سندی یا کتابخانه‌ای^{۱۰}؛ و برای سنجش میزان دقت موتورهای کاوش از روش مقایسه‌ای^{۱۱} استفاده شده است.

جامعه مورد مطالعه

موتورهای کاوش زیر برای این تحقیق انتخاب شدند:

- آلتاویستا - Altavista (<http://www.AV.com>)
- اکسایت - Excite (<http://www.Excite.com>)
- گوگل - Google (<http://www.Google.com>)
- هات بات - Hot Bot (<http://www.Hotbot.com>)
- اینفوسیک - Infoseek (<http://www.Infoseek.Go.com>)
- لایکاس - Lycos (<http://www.Lycos.com>)
- وب کراولر - Webcrawler (<http://www.Webcrawler.com>)

این ابزارهای کاوش در دسته موتورهای کاوش عمومی و جزء پر استفاده‌ترین آنها هستند. راهنماهای موضوعی، ابرموتورهای کاوش، و موتورهای کاوش کم‌استفاده در این جامعه وارد نشده‌اند. برای تعیین موتورهای کاوش پر استفاده پیشینه تحقیق بررسی شد و هفت موتور کاوش که بیشترین حضور را داشتند، برای این پژوهش انتخاب شدند.

برای واقعی بودن کلیدواژه‌های به کار رفته در جستجو جامعه مورد مطالعه در این بخش را مقالات تخصصی رشته کتابداری و اطلاع‌رسانی مجله Library Trends تشکیل می‌دهد زیرا این مجله هر شماره خود را به موضوع ویژه‌ای در زمینه کتابداری و اطلاع‌رسانی اختصاص می‌دهد. مجلات ۱۹۹۵ تا ۱۹۹۹ که به صورت چاپی در اختیار بود، مبنای انتخاب کلیدواژه قرار گرفت. در ابتدا عنوان مقاله و سپس شماره و تاریخ و بعد کلید واژه‌های استخراج شده ارائه می‌شود:

شیوه گردآوری اطلاعات

کلیدواژه‌های انتخابی به هر کدام از موتورهای کاوش داده شد. برای دستیابی به بهترین نتیجه، راهبرد جستجو یا به عبارت دیگر شیوه به‌کارگیری کلیدواژه‌ها به نحوی تنظیم شد که از بیشترین قابلیت‌های هر موتور کاوش در ترکیب کلیدواژه‌ها استفاده شود. کلیدواژه‌ها با استفاده از عملگرهای بولی و ریاضی و امکانات جستجوی عبارتی و نیز عملگرهای کوتاه‌سازی، با یکدیگر ترکیب و برای جستجو به هر یک از موتورهای کاوش داده شدند. در چند جستجو از بخش جستجوی ساده موتورهای کاوش، و در چند جستجو از بخش پیشرفته آنها استفاده شد. به این ترتیب که اگر برای جستجوی کلیدواژه‌هایی قرار شد از جستجوی ساده استفاده شود این عمل در همه موتورهای جستجو یکسان باشد و اگر جستجوی پیشرفته مدنظر بود برای همه موتورها از جستجوی پیشرفته استفاده شود. از کل نتایج ارائه شده توسط هر موتور جستجو، ده نتیجه اول ارزیابی شد.

برای تعیین میزان ربط منابع بازیابی شده چهار معیار در نظر گرفته شد:

۱. حضور کلیدواژه‌ها در عنوان مدارک: اگر هر یک از کلیدواژه‌های مورد جستجو در عنوان صفحه وب بازیابی شده ظاهر شود به ازای هر کلیدواژه یک امتیاز تعلق خواهد گرفت.

۲. همجواری کلیدواژه‌های مورد جستجو: زمانی که کلیدواژه‌های انتخابی، ترکیبی یا عبارتی بود، اگر هر یک از کلیدواژه‌ها به‌تنهایی در عنوان صفحه ظاهر می‌شد، یک امتیاز می‌گرفت و هرگاه کلیدواژه‌ها با همان ترکیبی که جستجو شده بود در عنوان ظاهر می‌گردید، علاوه بر امتیاز حضور هر کلیدواژه در عنوان، یک امتیاز همجواری هم کسب می‌کرد.

۳. حضور کلیدواژه‌ها در بخش ابرنشان‌های

توصیفی صفحات وب: به ازای حضور هر یک از کلیدواژه‌های مورد جستجو در بخش ابرنشان‌های کلیدواژه‌ای صفحه، یک امتیاز تعلق می‌گرفت.

۴. بسامد (فراوانی) کلیدواژه‌ها: در این معیار تعداد حضور کلیدواژه‌های مورد جستجو در صفحات بازیابی شده مشخص می‌شد و تعداد کلیدواژه‌ها بر تعداد صفحات منابع بازیابی شده تقسیم و در واقع امتیاز بسامد بر صفحه مشخص می‌شد. امتیازاتی که هر یک از موتورهای جستجو با توجه به معیارهای بالا کسب می‌کردند در مجموع نشان‌دهنده میزان ربط منابع بازیابی شده با توجه به موضوعات مورد جستجو تلقی می‌شد.

برای تعیین میزان استفاده موتورهای کاوش از ابرنشان‌های کلیدواژه‌ای برای توصیف صفحات وب، امتیازهای آنها در بخش ابرنشان‌های کلیدواژه‌ای صفحات وب بازیابی شده، شمارش گردید.

پیوندهای تکراری یعنی مدارکی که بیش از یک بار در ده نتیجه نخست تکرار شده است و پیوندهای کور و غیرفعال، خطا تلقی شدند. پیوندهای وبی که پس از درخواست شامل خطای ۴۰۴ و خطا در خدمت‌دهنده باشد، پیوندهای کور در نظر گرفته شدند. خطای ۴۰۴ زمانی به‌وجود می‌آید که مدرک مورد نظر به محل دیگری منتقل شده، تغییر نام داده و یا حذف شده است، یا محل قرارگیری مدرک بر روی شبکه توسط موتور جستجو به‌درستی ثبت نشده است. نسبت پیوندهای کور و غیرفعال در موتورهای کاوش، نشان‌دهنده مقیاسی برای چگونگی روزآمدسازی اطلاعات در پایگاه‌های اطلاعاتی هر یک از موتورهای کاوش و نشان‌دهنده میزان کنترل آنان بر این پیوندها در فواصل زمانی مطلوب است.

پیشینه پژوهش

پیشینه در ایران

نوی در سال ۱۳۸۰ در پایان‌نامه خود با عنوان «مطالعه

جدول ۲. نتایج حاصل از اجرای آزمون F برای سنجش معناداری تفاوت نمره میانگین عملکرد موتورهای کاوش

گروه‌ها	نمره میانگین	مقدار F	سطح معنادار (Sig.)
میان گروهی	۳۶۴/۴۸	۰/۵۰۹	۰/۷۹
درون گروهی	۷۱۵/۹۰		

مقایسه‌ای ابر موتورهای جستجو در بازیابی اطلاعات کتابداری و اطلاع‌رسانی از شبکه جهانی وب»، ۱۲ ابرموتور جستجوی عمده وب را با استفاده از کلیدواژه‌های تخصصی کتابداری و اطلاع‌رسانی با هم مقایسه کرد. این مقایسه در ۱۰ نتیجه اول انجام شد. در ارزیابی منابع بازیابی شده شش ستون تعیین شده بود:

(۱) کاملاً مرتبط؛

(۲) مرتبط؛

(۳) تا حدودی مرتبط؛

(۴) نامرتب؛

(۵) موارد عدم موفقیت؛

(۶) موارد تکراری.

نتایج به دست آمده نشان داد که حتی با استفاده از بهترین ابرموتور جستجو، فقط حدود ۳۰ درصد از منابع بازیابی شده از اینترنت در زمینه کتابداری و اطلاع‌رسانی، مرتبط هستند. همچنین یافته‌ها نشان داد که ابرموتور جستجوی سی فور^{۱۲} بیشترین تعداد منابع وب را در گروه‌های موضوعی مختلف بازیابی می‌کند اما بیشترین درصد منابع مرتبط بازیابی شده در گروه‌های موضوعی مختلف از ابرموتور سی نت^{۱۳} به دست آمد. در نگاه دقیق‌تر، بیشترین تعداد منابع وب بازیابی شده از ابرموتورهای جستجوی مورد بررسی، در گروه موضوعی «فناوری اطلاعات و ارتباطات» و بیشترین درصد منابع مرتبط بازیابی شده در ابرموتورهای جستجوی مورد بررسی نیز در همین گروه به دست آمد (۶).

کمیجانی در ۱۳۸۱ در پایان‌نامه خود تحت عنوان «مقایسه کارایی موتورهای کاوش عمومی و تخصصی وب در بازیابی اطلاعات کشاورزی» ۷ موتور کاوش عمومی وب شامل آلتاویستا، اکسایت، هات‌بات، اینفوسپیک، گوگل، لایکاس و یاهو و ۲ موتور کاوش تخصصی کشاورزی وب شامل آگری‌سرف و وب‌آگری را در بازیابی اطلاعات تخصصی حوزه کشاورزی با هم مقایسه کرد. این مقایسه نشان داد که در معیار ارتباط و اعتبار سه موتور کاوش گوگل، لایکاس و یاهو به ترتیب رتبه اول تا سوم را داشتند و آخرین رتبه‌ها به موتورهای کاوش تخصصی کشاورزی تعلق داشت.

این پژوهش نشان داد که موتورهای کاوش عمومی در هر دو معیار ارتباط و اعتبار که به نظر می‌رسد به ترتیب مهم‌ترین معیارهای مورد توجه از طرف کاربران متخصص کشاورزی باشند با اختلاف چشمگیری بر موتورهای کاوش تخصصی

کشاورزی برتری دارند. به نحوی که موتورهای کاوش کشاورزی همیشه در انتهای جدول رتبه‌بندی جای داشتند و این بیانگر آن است که این موتورها هنوز به سطح قابل قبولی برای مراجعه کاربران کشاورزی در امور پژوهشی دست نیافته‌اند (۲).

پیشینه در خارج

چو و رزنتال^{۱۴} در سال ۱۹۹۶ از دانشکده علوم کتابداری و اطلاع‌رسانی دانشگاه لایت‌آیلند در شهر نیویورک با ارائه مقاله‌ای با عنوان «موتورهای کاوش برای شبکه جهانی وب: مطالعه مقایسه‌ای و روش ارزیابی» سه موتور کاوش آلتاویستا، اکسایت و لایکاس را در زمینه ارائه امکانات جستجو مانند به‌کارگیری عملگرهای بولی و کوتاه‌سازی در جستجو، نوع فیلدهای جستجو، جستجوی واژه و عبارت و همچنین عملکرد آنها در بازیابی در مواردی نظیر میزان مانعیت و زمان پاسخگویی، ارزیابی و مقایسه کردند. آنها این جستجوها را با استفاده از سوالات واقعی مطرح شده در بخش مرجع کتابخانه انجام دادند. این پژوهشگران به این نتیجه رسیدند که موتور کاوش آلتاویستا در ارائه امکانات جستجو و عملکرد بازیابی برتری قابل ملاحظه‌ای نسبت به دو موتور جستجوی دیگر داشت، در صورتی که میزان پوشش صفحات وب در موتور کاوش لایکاس نسبت به هر موتور دیگر بالاتر بود (۷).

شلیکتینگ و نیلسن^{۱۵} در ۱۹۹۶ چهار موتور جستجوی پر استفاده در وب جهانگستر شامل لایکاس، آلتاویستا، اکسایت و اینفوسپیک را ارزیابی کردند. برای این کار پنج عضو هیئت علمی سیاهه‌ای از چهار تا شش کلیدواژه برای یک موضوع تحقیقاتی فراهم کردند. بعد از اجرای این کلیدواژه‌ها در موتورهای جستجو، ده سایت بازیابی شده اول ارزیابی شدند. برای تحلیل نتایج، دو معیار برای هر موتور جستجو در نظر گرفته شد. اولین معیار، حساسیت موتور جستجو را در پیدا کردن اطلاعات مفید می‌سنجید و معیار دوم، میزان محدودیت یا آزادی موتور جستجو در تعیین اینکه چه سایت‌هایی را در بر می‌گیرد مشخص می‌کرد. 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، اینفوسپیک با ۱۲ مورد و آلتاویستا با ۹ مورد پیوند مربوط در ردیف‌های بعدی قرار داشتند (۱۹).

آشکارترین نتیجه این بود که عملکرد موتورهای جستجو از وضعیت مطلوب بسیار دور است و هیچ یک از موتورهای جستجو به سطح قابل قبولی از عملکرد در مورد این درخواست‌های کاملاً مشخص و علمی نرسیده‌اند.

لبدوف^{۱۶} در سال ۱۹۹۷ با استفاده از هشت کلیدواژه در زمینه شیمی و فیزیک، هشت موتور جستجو را از لحاظ بازیابی اطلاعات علمی سنجید و نتیجه گرفت هر چه حجم پایگاه اطلاعاتی مورد جستجو انبوه‌تر باشد، احتمال به دست آوردن اطلاعات علمی از طریق جستجو در آن نیز بیشتر است. این پژوهش در جستجوی کلیدواژه‌های شیمی و فیزیک، موتور کاوش آلتاویستا و هات‌بات را بهترین موتور در بازیابی اطلاعات معرفی می‌کند (۱۴).

در ۱۹۹۷، لیتون و سریواستاوا^{۱۷} از گروه رایانه دانشگاه ایالتی وینونا و دانشگاه مینوتا میزان مانعیت در بازیابی اطلاعات به وسیله پنج موتور جستجوی آلتاویستا، هات‌بات، لایکاس، اکسایت و اینفوسیک را بررسی کردند. آنان جستجوهای خود را با استفاده از ۱۵ موضوع متفاوت انجام دادند و پس از جستجو، ۲۰ نتیجه اول در هر موتور کاوش را برای تعیین ارتباط آنها با موضوعات مطرح شده بررسی کردند. آنها با تجزیه و تحلیل آماری به این نتیجه رسیدند که موتورهای کاوش آلتاویستا، اکسایت و اینفوسیک به ترتیب در رتبه‌بندی نتایج مرتبط بهترین موتورهای جستجو هستند (۱۵).

هسه-یی^{۱۸} در ۱۹۹۸ هشت موتور کاوش را با به کارگیری بیست و یک سؤال مرجع واقعی و یا سؤال موضوعی ارزیابی کرد. هر سؤال در هر موتور دوبار جستجو شد. به جستجوگران کلیدواژه‌ها و پارامترهای اساسی درباره اینکه چگونه هر موتور کاوش باید جستجو شود، داده شد. جستجوگران قضاوت خود را در ارزیابی ربط ده مدرک اول بازیابی شده به کار بردند. چهار متغیر: دقت، تکراری بودن، مرتبط‌ترین مدرک و رتبه‌بندی براساس ربط، برای ارزیابی موتورهای جستجو به کار رفتند. موتورها به طور جداگانه برای هر نوع سؤال ارزیابی شدند. بهترین عملکرد برای سوالات مرجع این تکست^{۱۹} بود و برای سوالات موضوعی، اینفوسیک بهترین عملکرد بود (۱۲).

تومه یولو^{۲۰} در ۱۹۹۹ در یک بررسی با عنوان «آیا ابر موتورهای جستجو بهتر جستجو می‌کنند؟» ضمن مقایسه ۴ ابر موتور جستجوی متاکراولر، اینترنت اسلوت^{۲۱}، داگ پایل^{۲۲}، سایبر ۱۱۴^{۲۳} با دو موتور جستجوی هات‌بات و

آلتاویستا در توانایی بازیابی اطلاعات چنین بیان کرد: با وجود اینکه برخی موتورهای جستجو یقیناً رابط‌های کاربری بهتری در مقایسه با موتورهای جستجوی انفرادی دارند، این تحقیق نشان می‌دهد که ابرموتورهای جستجو نسبتاً بهتر کار می‌کنند و صرف نظر از اینکه کاربر از کلیدواژه‌های خاص یا عمومی استفاده کند، شاید ابرموتورهای جستجو همان اطلاعاتی را پیدا کنند که موتورهای جستجوی انفرادی بازیابی می‌کنند اما از آنجا که هیچ موتوری دقیقاً پایگاه‌های تحت پوشش موتورهای دیگر را نمایه‌سازی نمی‌کند، کاربر با استفاده از یک ابرموتور جستجو شانس خود را در ارزیابی اطلاعات مرتبط بالا می‌برد (۲۱).

گانگ وو و جی لی^{۲۴} (۱۹۹۹) از دانشگاه ایالتی وین، هفت موتور جستجوی آلتاویستا، اینفوسیک، اکسایت، نورثرن لایت، یاهو، هات‌بات، و مدیکال ورلدسرج^{۲۵} را مقایسه کردند تا کارایی آنها در پاسخگویی به سوالات کاربران علوم بهداشتی بسنجند. آنها سه موتور جستجوی اینفوسیک، آلتاویستا و اکسایت را 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 دارای بالاترین تعداد پیوندهای غیرفعال بودند. هات‌بات پوشش جامع و ویژگی‌های جستجوی پیشرفته 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 داد که حتی با استفاده از بهترین موتورهای کاوش فقط نیمی از موارد بازیابی شده، مرتبط خواهند بود (۱۰).

گوردون و پاتک^{۲۶} (۱۹۹۹) هفت موتور جستجو و یک راهنمای موضوعی (ياهو) را با به کارگیری سوالات واقعی استفاده‌کنندگان و ارزیابی آنها در مورد ربط مدارک بازیابی شده آزمایش کردند. از کاربران خواسته شد که اطلاعات مورد نیاز را شرح دهند، مهم‌ترین واژه‌ها یا عبارات را در توصیف‌هایشان مشخص کنند، هر مترادف و واژه مربوطی که آنها فکر می‌کردند مفید خواهد بود مشخص کنند و جستجوییشان را به شکل درخواست بولی درآورند. بیست نتیجه اول از بهترین جستجو برای هر موتور برای ارزشیابی چاپ و به طور تصادفی مرتب شد و به کاربران داده شد که با

۴ معیار آنها را ارزیابی کنند.

این چهار معیار عبارت بودند از: کاملاً مرتبط، تاحدودی مرتبط، تاحدودی نامرتب و کاملاً نامرتب. این پژوهشگران به این نتیجه رسیدند که اثربخشی بازیابی موتورهای جستجو نسبتاً پایین است و موتورهای جستجو آلتاویستا و آپن تکست بهترین و هات‌بات و یا هو بدترین عملکرد را داشتند (۱۱).
مارتینز و سانچز^{۲۷} در مقاله‌ای با عنوان «مقایسه‌ای بین ابزارهای جستجوی اینترنت» ده موتور کاوش را برای تحقیق انتخاب کردند. در این پژوهش امکانات مختلف جستجو مانند عملگرهای بولی، کوتاه‌سازی، استفاده از پرانتز و امکانات پیشرفته نظیر اصلاح نتیجه جستجو، جستجو در فیلدهای خاص، امکان محدود کردن موضوع و نیز زبان کنترل شده، 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 یا هو و موتور کاوش نورثرن لایت قابل مشاهده است. اگرچه موتور کاهش نورثرن لایت از نظر درجه مقبولیت آخرین رتبه را دارد ولی در فهرست رتبه‌بندی امکانات جستجو اولین رتبه را از آن خود کرد. حال آنکه یا هو که از رایج‌ترین و عامه‌پسندترین ابزارهای کاوش است، از لحاظ بازیابی اطلاعات رتبه بالایی کسب نکرده است (۱۶).

تجزیه و تحلیل یافته‌ها

در این پژوهش برای تجزیه و تحلیل آماری نتایج و تعیین معنادار بودن اختلاف میان میزان دقت موتورهای کاوش وب از برنامه نرم‌افزاری اکسل^{۲۸} استفاده شد.
برای جوابگویی به سوالات اساسی پژوهش و در مجموع رسیدن به هدف تحقیق باید مجموع امتیازاتی که موتورهای کاوش وب در بازیابی اطلاعات کتابداری و اطلاع‌رسانی طی ده جستجو کسب کرده‌اند با هم مقایسه شود. این امتیازات از حضور کلیدواژه‌ها در عنوان مدرک، همجواری کلیدواژه‌های مورد جستجو، حضور کلیدواژه‌ها در بخش ابرنشانه‌های کلیدواژه‌ای صفحات وب، و میزان حضور یا بسامد کلیدواژه‌ها در متن مدرک تشکیل شده است.
شیوه امتیازدهی به این ترتیب بود که ابتدا کلیدواژه‌های انتخابی با به‌کارگیری امکانات جستجو و عملگرهای مختلف در هر یک از موتورهای کاوش جستجو می‌شد. ۱۰ نتیجه

اول از کل نتایج بازگردانده شده توسط هر موتور برای ارزیابی و مقایسه در نظر گرفته شد. ابتدا به عنوان مدرک بازیابی شده توجه می‌شد، اگر کلیدواژه‌ای که برای جستجو انتخاب شده بود در عنوان مدرک موجود بود به تعداد حضور آنها یک امتیاز تعلق می‌گرفت. بعضی از کلیدواژه‌ها به صورت عبارتی و ترکیبی بودند. اگر کلیدواژه‌ها در عنوان مدرک به همان صورت ترکیبی یا عبارتی ظاهر می‌شد گذشته از امتیاز ظهور در عنوان، یک امتیاز همجواری نیز دریافت می‌کرد. پس از این مرحله به مدرک بازیابی شده مراجعه می‌شد پس از کلیک کردن بر روی عنوان، متن مدرک نمایش داده می‌شد. برای اینکه بتوان تشخیص داد که چقدر از کلیدواژه‌ها در بخش ابرنشانه‌های توصیفی صفحات وب وجود دارند به گزینه View و Source مراجعه می‌شد. پس از آنکه صفحه مربوط به زبان اچ.تی.ام.ال باز می‌شد و بخش ابرنشانه‌های کلیدواژه‌ای که با عبارت

<meta name = "Keywords" = content = ...> مشخص می‌شود توجه می‌شد تا میزان حضور کلیدواژه‌ها مشخص شود. به تعداد حضور کلیدواژه‌ها در این قسمت نیز یک امتیاز داده می‌شد. پس از این مرحله باید تعداد حضور کلیدواژه‌ها در متن مدرک بازیابی شده مشخص می‌شد. برای رسیدن به این منظور به گزینه Edit و on this page Find مراجعه شد. با این عمل هر کلیدواژه مورد جستجو در متن تشخیص داده می‌شد. برای اینکه تراکم کلیدواژه‌ها در مدرک مشخص گردد تعداد کلیدواژه‌های شمارش شده بر تعداد صفحات آن مدرک تقسیم می‌شد. امتیازاتی که این موتورها در موارد مختلف به دست می‌آوردند با یکدیگر جمع می‌شد و مجموع امتیازاتی که بدین طریق حاصل شد ملاک مقایسه قرار گرفت.

با مشخص شدن نحوه امتیازدهی میزان دقت، یافتن میزان کاربرد ابرنشانه‌های کلیدواژه‌ای برای توصیف صفحات وب، و نحوه پیش آمدن خطا در بازیابی اطلاعات، جستجوهای ده گانه به این منظور شروع شد.

در بررسی ۱۰ نتیجه اول از هر جستجو در هر موتور کاوش، اگر مدرکی بیش از یک بار تکرار می‌شد یک امتیاز منفی می‌گرفت. اگر پیوندی در میان مدارک بازیابی شده پیغام خطا می‌داد به این علت که آن صفحه وب تغییر نام داده یا نشانی آن تغییر کرده یا به هر علتی حذف شده بود و یا اینکه محل قرارگیری مدرک بر روی شبکه به درستی توسط موتور کاوش ثبت نشده بود و امکان دسترسی به آن صفحه ممکن نبود

27.Martinez, A.M & Sanche, E.F

28.Excel

امتیاز منفی دریافت می‌کرد (این موارد هر کدام پیغام خاص خود را دارد). مجموع امتیازات حاصل شده از این موارد، میزان خطای موتورهای کاوش وب را در بازیابی اطلاعات نشان می‌داد. هر چقدر امتیاز حاصل شده از این موارد بیشتر می‌شد میزان خطای بالا را در موتور مورد نظر نشان می‌داد و هر چقدر تعداد امتیازات کمتر می‌شد نشانگر میزان خطای پایین در موتور کاوش مورد نظر بود. طبق بررسی‌ها و شمارش میزان خطاها در ده نتیجه اول در هر جستجوی انجام شده، مشخص شد که موتورهای کاوش آلتاویستا و اینفوسیک (به‌طور مشترک)، اکسایت، گوگل و لایکاس (به‌طور مشترک)، وب کراولر و هات‌بات به ترتیب کمترین میزان خطا را در بازیابی اطلاعات مربوط به حوزه کتابداری و اطلاع‌رسانی داشتند. به این معنا که دو موتور کاوش آلتاویستا و اینفوسیک به‌طور مشترک و موتور کاوش هات‌بات به ترتیب دارای پایین‌ترین و بالاترین میزان خطا در بررسی‌های انجام شده بودند. نسبت پیوندهای کور و پیغام‌های خطا در موتور کاوش، میزان روزآمدسازی اطلاعات را در پایگاه‌های اطلاعاتی هر یک از موتورهای کاوش و میزان کنترل آنها را بر این پیوندها نشان می‌دهد. بنابراین هر چقدر میزان خطاها بیشتر باشد روزآمدسازی و کنترل پایگاه‌های اطلاعاتی بر مدارک کمتر است و برعکس هر چقدر تعداد خطاها کمتر باشد کنترل و روزآمدسازی در پایگاه‌های اطلاعاتی بهتر و سریع‌تر انجام می‌شود.

امتیازاتی که هر یک از موتورهای کاوش مورد بررسی با توجه به معیارهای از پیش تعیین شده کسب کردند مشخص کرد که موتورهای کاوش اینفوسیک، هات‌بات، گوگل، آلتاویستا، اکسایت، لایکاس و وب‌کراولر به ترتیب رتبه‌های بالاتری کسب کردند. به عبارتی در میان موتورهای کاوش هفت‌گانه، اینفوسیک بالاترین میزان دقت و وب‌کراولر پایین‌ترین میزان دقت را در بازیابی منابع تخصصی مربوط به رشته کتابداری و اطلاع‌رسانی داشتند.

برپایه میزان استفاده موتورهای کاوش وب مورد بررسی از ابرنشان‌های کلیدواژه‌ای در صفحات وب و مقایسه آنها، مشخص شد که هات‌بات، اکسایت، اینفوسیک، آلتاویستا، لایکاس، گوگل و وب‌کراولر به ترتیب بیشترین میزان استفاده از ابر نشان‌های کلیدواژه‌ای را در صفحات وب خود داشتند. از این منظر موتور کاوش هات‌بات و موتور کاوش وب‌کراولر به ترتیب بالاترین و پایین‌ترین امتیاز را داشتند. تشخیص اینکه چه موتورهای کاوشی بیشتر از ابرنشان‌های کلیدواژه‌ای برای توصیف صفحات وب خود استفاده می‌کنند از این جهت

اهمیت دارد که موتورهای کاوش، کلیدواژه‌های مورد جستجو را در بخش ابرنشان‌های کلیدواژه‌ای جستجو می‌کنند و صفحاتی را که این کلیدواژه‌ها را دربرمی‌گیرد از نظر میزان ربط در اولویت قرار می‌دهند و آنها را در ابتدای فهرست نتایج خود نشان می‌دهند.

نتایج اصل از اندازه‌گیری میزان دقت، میزان کاربرد ابرنشان‌ها و تعداد خطاها در هر یک از موتورهای کاوش در جدول یک آمده است.

در این پژوهش، برای آزمون معناداری تفاوت نمره میانگین عملکرد موتورهای کاوش در بازیابی اطلاعات کتابداری و اطلاع‌رسانی، از آزمون پارامتری F استفاده شد و نتیجه این آزمون نشان داد که تفاوت مشاهده شده میان نمره میانگین عملکرد موتورهای کاوش در بازیابی اطلاعات در هیچ‌کدام از دو سطح کمتر از ۰/۰۵ و کمتر از ۰/۰۱ وجود ندارد (جدول ۲).

از طرفی، نتیجه آزمون شفه^{۲۹} نیز که برای مقایسه‌های چندگانه به کار می‌رود، نشان‌دهنده آن است که به دلیل نبود تفاوت معنادار، نمره میانگین عملکرد موتورهای کاوش در بازیابی اطلاعات کتابداری و اطلاع‌رسانی، همه گروه‌های هفت‌گانه موتورهای جستجو در یک طبقه جای می‌گیرند.

پس از اینکه مجموع امتیازات حاصل شده توسط هر موتور کاوش مشخص شد، این امتیازات مبنای قضاوت برای تعیین بهترین موتور کاوش و اولویت‌بندی آنها از نظر دقت در بازیابی منابع اطلاعاتی کتابداری و اطلاع‌رسانی قرار گرفت و طبیعتاً مجموع امتیازات بیشتر در هر موتور نشان‌دهنده دقت بیشتر آن بود و اینکه این موتور کاوش دقت بالاتری از بقیه موتورهای کاوش در بازیابی اطلاعات کتابداری و اطلاع‌رسانی دارد.

نمودار ۱، امتیازات هر یک از موتورهای کاوش را در بازیابی اطلاعات کتابداری و اطلاع‌رسانی از منظر چند معیار نشان می‌دهد.

نتایج این نمودار، بر امتیاز کل و طبیعتاً رتبه هر یک از این موتورها نیز دلالت می‌کند. همان‌طور که نمودار نشان می‌دهد موتورهای کاوش اینفوسیک، هات‌بات، گوگل، آلتاویستا، اکسایت، لایکاس و وب‌کراولر به ترتیب بیشترین رتبه را از حیث عملکردشان در بازیابی اطلاعات کتابداری و اطلاع‌رسانی کسب کرده‌اند. به عبارت دیگر، موتور کاوش اینفوسیک بالاترین رتبه و موتور کاوش وب‌کراولر پایین‌ترین رتبه را از نظر دقت در بازیابی منابع کتابداری و اطلاع‌رسانی به خود اختصاص داده‌اند.



از این مقایسه براساس مجموع امتیازات و نهایتاً رتبه بندی موتورهای کاوش هفت گانه صورت گرفت. بر این اساس هر قدر امتیازات کسب شده از تعداد خطاها کمتر باشد، رتبه موتور کاوش مورد نظر پایین تر خواهد بود و هر قدر تعداد خطاها کمتر باشد رتبه موتور کاوش از این لحاظ بالاتر خواهد رفت. نتایج این مقایسه مشخص کرد که موتورهای کاوش آلتاویستا و اینفوسیک (به طور مشترک)، اکسایت، گوگل و لایکاس (به طور مشترک)، وب کرولر و هات بات به ترتیب کمترین میزان خطا را در بازبازی اطلاعات مربوط به کتابداری و اطلاع رسانی داشته اند. به عبارتی روشن تر، موتورهای کاوش آلتاویستا و اینفوسیک (به طور مشترک) و هات بات به ترتیب کمترین و بیشترین میزان خطا را در بازبازی اطلاعات مرتبط با حوزه کتابداری و اطلاع رسانی داشته اند. نمودار ۳ گویای دقیق تر این موضوع است.

پیشنهاد های پژوهش

مبنای قضاوت برای تعیین اینکه کدام یک از موتورهای کاوش بیشترین استفاده از ابر نشانه های کلیدواژه ای را داشتند، تعداد کلیدواژه هایی بود که صفحه ا.ح. تی.ام.ال مدارک بازبازی شده شمارش شده بود. پس از اینکه تعداد کل ابر نشانه های کلیدواژه ای در مدارک بازبازی شده در جستجوی ده گانه مشخص شد، موتورهای کاوش با هم مقایسه شدند، نتایج این مقایسه که در نمودار ۲ آمده است، بر این امر دلالت دارد که موتورهای کاوش هات بات، اکسایت، اینفوسیک، آلتاویستا، لایکاس، گوگل و وب کرولر به ترتیب بیشترین امتیاز و طبیعتاً بالاترین رتبه را از نظر بیشترین استفاده شان از ابر نشانه های کلیدواژه ای به خود اختصاص داده اند. به عبارتی موتور کاوش هات بات بالاترین رتبه و موتور کاوش وب کرولر پایین ترین رتبه را از لحاظ بیشترین استفاده از ابر نشانه های کلیدواژه ای برای توصیف صفحات وب به خود اختصاص داده اند.

کمترین خطا در ده نتیجه اول در هر جستجو معیار دیگری بود که از آن برای مقایسه موتورهای کاوش وب در بازبازی اطلاعات کتابداری و اطلاع رسانی استفاده شد. نتایج حاصل



براساس یافته‌های تحقیق، پیشنهادهای زیر ارائه می‌شود:

۱. برای جستجوی منابع کتابداری و اطلاع‌رسانی و دریافت نتایج دقیق‌تر استفاده از موتورهای کاوش زیر به ترتیب اولویت پیشنهاد می‌شود: اینفوسیک، هات‌بات، گوگل، آلتاویستا، اکسایت، لایکاس و وب‌کراولر.
۲. برای اینکه هنگام بازیابی اطلاعات با پیغام‌های خطا و پیوندهای کور و غیرفعال کمتری مواجه شویم و در درجه نخست به‌کارگیری موتورهای کاوش آلتاویستا، اینفوسیک، اکسایت، گوگل و لایکاس و سپس موتورهای کاوش وب‌کراولر و هات‌بات پیشنهاد می‌گردد.

پیشنهاد برای پژوهش‌های آتی

۱. پیشنهاد می‌شود میزان دقت موتورهای کاوش وب در بازیابی اطلاعات در حوزه‌ها و رشته‌های دیگر بررسی شود.
۲. پیشنهاد می‌گردد در تحقیقی میزان دقت راهنماهای موضوعی وب در بازیابی اطلاعات حوزه کتابداری و اطلاع‌رسانی یا حوزه‌های دیگر بررسی شود.
۳. پیشنهاد می‌شود در پژوهشی به مقایسه دقت موتورهای کاوش و راهنماهای موضوعی وب در حوزه‌های مختلف علوم پرداخته شود.

منابع

۱. داودیان، فرشته. «ابزارهای جستجو در اینترنت». پیام کتابخانه دوره یازدهم، ۴ (زمستان ۱۳۸۴).
۲. کميجانی، احمد. «مقایسه کارایی موتورهای کاوش عمومی و تخصصی وب در بازیابی اطلاعات کشاورزی»

پایان نامه کارشناسی ارشد کتابداری و اطلاع‌رسانی، دانشکده روانشناسی و علوم تربیتی، دانشگاه تهران، ۱۳۸۱.

۳. کوشا، کیوان. ابزارهای کسائوش اینترنت: اصول، مهارت‌ها و امکانات جستجو در وب. تهران: نشر کتابدار، ۱۳۸۱.
۴. لارج، اندرو؛ تد، لوسی؛ هارتلی، ریچارد. جستجوی اطلاعات در عصر اطلاعات: اصول و مهارت‌ها. ترجمه زاهد بیگدلی. تهران: کتابدار، ۱۳۸۲.
۵. لیو، جی‌یان. «راهنمای ابر موتورهای کاوش». ترجمه کیوان کوشا. پژوهشنامه اطلاع‌رسانی، دوره سوم، ۵ (۱۳۷۸).

۶. نبوی، فاطمه. «مطالعه مقایسه‌ای ابر موتورهای جستجو در بازیابی اطلاعات کتابداری و اطلاع‌رسانی از شبکه جهانی وب». پایان نامه کارشناسی ارشد کتابداری و اطلاع‌رسانی، دانشکده مدیریت و اطلاع‌رسانی، دانشگاه علوم پزشکی ایران، ۱۳۸۰.

7. Chu, Hetting; Rosenthal, Marilyn. "Search engines for the world wide web: A comparative study and evaluation methodology". 1996. [on-line]. Available: <http://www.asis.org/annual-96/electronic-proceedings/chu.html>.

8. Farrelly, Glen. "Search Engines: Evolution and Revolution". 1999. [on-line]. Available: <http://www.webhome.indirect.com/~gleujenn/search/history1.htm>.



نمودار ۲. نمودار اولویت بندی موتورهای کاوش از حیث میزان خطا در بازیابی اطلاعات

16. Martinez, A.M.; Sarchez E.F. "Comparing Internet Search tools". Proceedings of the 23rd. International online information meeting, 1999.

17. "Meta Search Engines". 2000. [on-line]. Available: <http://www.sc.edu/beaufort/library/lesson2.html>.

18. Nicholson, Scott. "A proposal for Categorization and nomenclature for web search tools". *In Internet searching and Indexing: The subject approach*. Edited by Alan R. Thomas and James R. Shearer. England: The Haworth Press, 2000.

19. Schlichting, A.; Nilsen, E. "Signal detection analysis of www Search engines". 1996. [on-line]. Available: <http://www.microsoft.com/usability/webconf/schlichting/schlichting.htm>.

20. Sonnenreich, Wes. "A History of search Engines". 2000. [on-line]. Available: <http://www.cobbusa.com/sehistory.htm>.

21. Tomaiuolo, Nicholas. "Are metasearch engines better searches?" *Searcher: The Magazine for data base Professionals*, No. 1 (Jan. 1999)

تاریخ دریافت: ۱۳۸۴/۲/۲۸

9. Flanagan, Debbie. "Search Engines". 1999. [on-line]. Available: <http://www.home.sprintmail.com/~debflanigan/engines.htm>

10. Gang wu, Jie Li, "comparing web search engine performance in searching consumer health information: evaluation and recommendations". *Bulletin of medical association*, Vol. 84, No. 4 (1999).

11. Gordon, M.; Pathak P. "Finding Information on the world wide web: The retrieval effectiveness of search engines". *Information proceeding and management*. Vol. 35, No. 2 (1999).

12. Hsieh-ye, Ingrid, "Internet: organization and searching the retrieval power of selected search engines: How well do they address General reference questions and subject questions?" *Reference Librarian*, No. 60 (1998).

13. Koster, Martijn. "Robots in the web: Threat or treat". 1997. [on-line]. Available: <http://www.robotstxt.org/wc/threat or treat.html>.

14. Lebedey, Alexander. "Best search engines for finding scientific information in the web". 1997. [on-line]. Available: <http://www.chem.msu.su/eng/comparison.html>.

15. Leighton, Vernon; Srivastava, J. "Precision among world wide web search services (search engines): Altavista, Excite, Hotbot, Indseek, Lycos". 1997. [on-line]. Available: <http://www.Winona.msus.edu/library/webind2.htm>.