

# مقایسه دقت موتورهای کاوش عمومی و تخصصی پزشکی در بازیابی تصاویر پزشکی\*

نجلا حریری<sup>۱</sup>، حسین وکیلی مفرد<sup>۲</sup>

## مقاله پژوهشی

### چکیده

**مقدمه:** حجم اطلاعات و تصاویر موجود در اینترنت و رشد روز افزون آن، کار جستجوی بهینه‌ی اطلاعات و به‌ویژه تصاویر را برای کاربران مشکل کرده است. در بین ابزارهای جستجو، موتورهای کاوش، بیش‌تر مورد استفاده‌ی کاربران قرار می‌گیرند. هدف از تحقیق حاضر، تعیین میزان دقت موتورهای کاوش عمومی و تخصصی پزشکی در بازیابی تصاویر پزشکی و مقایسه‌ی آن‌ها بوده است.

**روش بررسی:** پژوهش حاضر، کاربردی و از نوع پیمایشی مقایسه‌ای بود که در سال ۱۳۹۰ انجام شد. جامعه‌ی مورد بررسی در این پژوهش، دو موتور کاوش عمومی Yahoo و Google و دو موتور کاوش تخصصی پزشکی OmniMedicalSearch و Trip Database بودند. به‌منظور تعیین و مقایسه‌ی میزان دقت بازیابی تصاویر پزشکی در چهار موتور کاوش مورد بررسی، ده کلیدواژه با مشورت متخصصان پزشکی انتخاب و در این موتورهای کاوش مورد جستجو قرار گرفتند. سپس ده نتیجه‌ی اول هر موتور کاوش استخراج شد و با اخذ نظر متخصصان پزشکی، ربط یا عدم ربط آن‌ها با کلیدواژه‌ی مورد نظر، با استفاده از نرم‌افزار SPSS، روش‌های آمار توصیفی و آزمون مستقل مورد تجزیه و تحلیل قرار گرفتند.

**یافته‌ها:** موتور کاوش Google با دقت ۹۰، دارای بیش‌ترین دقت در بازیابی تصاویر مرتبط و کاملاً مرتبط و سپس موتور کاوش پزشکی OmniMedicalSearch دارای دقتی معادل ۸۶/۸ در بازیابی تصاویر پزشکی بود. تفاوت معنی‌داری بین میزان دقت موتورهای کاوش عمومی و موتورهای کاوش تخصصی پزشکی در بازیابی تصاویر وجود نداشت.

**نتیجه‌گیری:** علی‌رغم همپوشانی برخی از نتایج موتورهای کاوش، هر کدام از موتورهای کاوش مورد بررسی نتایج منحصر به فردی را نیز بازیابی کرده‌اند. این موضوع نشان‌دهنده‌ی تفاوت در الگوریتم نمایه‌سازی منابع موجود در وب، توسط موتورهای کاوش مختلف است و پیشنهاد می‌شود که برای جستجوی تصاویر پزشکی، از چند موتور کاوش عمومی و تخصصی پزشکی به‌طور همزمان استفاده شود.

**واژه‌های کلیدی:** اینترنت؛ ذخیره و بازیابی اطلاعات پزشکی؛ تصاویر پزشکی؛ دقت؛ موتور کاوش

مشکل اصلی فرایند اطلاع‌یابی در اینترنت، اندک بودن حجم مطالب بازیابی شده نیست، بلکه مشکل جدیدی تحت عنوان «بازیابی بیش از حد اطلاعات» پدید آمده است. موتورهای کاوش، مهم‌ترین ابزارهای رایج برای جستجو و بازیابی اطلاعات مورد نیاز از بین حجم عظیم اطلاعات موجود در اینترنت هستند.

دریافت مقاله: ۹۱/۸/۲۷  
اصلاح نهایی: ۹۲/۲/۲۸  
پذیرش مقاله: ۹۲/۳/۲

**ارجاع:** حریری نجلا، وکیلی مفرد حسین. مقایسه دقت موتورهای کاوش عمومی و تخصصی پزشکی در بازیابی تصاویر پزشکی. مدیریت اطلاعات سلامت ۱۳۹۲؛ ۱۰(۶): ۸۳۹-۸۳۰.

### مقدمه

اینترنت به عنوان شاخص‌ترین نماد عصر اطلاعات، طی دهه اخیر با سرعتی حیرت‌انگیز رشد کرده و امروزه به عنوان مناسب‌ترین و پرکاربردترین بستر انتقال اطلاعات الکترونیکی مورد استفاده قرار می‌گیرد. اینترنت، مجموعه‌ای گسترده و ناهمگون از اطلاعات متنی، صوتی و تصویری است. امروزه

<sup>۱</sup> این مقاله حاصل تحقیق مستقل بدون حمایت مالی می‌باشد.  
۱- دانشیار، گروه علم اطلاعات و دانش‌شناسی، دانشگاه آزاد اسلامی، واحد علوم و تحقیقات، تهران، ایران  
۲- مربی، علوم کتابداری و اطلاع‌رسانی پزشکی، دانشگاه علوم پزشکی همدان، دانشکده‌ی پیراپزشکی، همدان، ایران (نویسنده‌ی مسؤول)  
Email: vakili.hn@gmail.com

تصاویر مرتبط که پاسخگوی نیازهای کاربر باشد حایز اهمیت بسیاری است. برای به دست آوردن تصاویر مفید، کاربران نیاز دارند یک ابزار جستجوی کارآمد و با کفایت به کار برند که آن‌ها را برای بازیابی آن‌چه که آن‌ها از مجموعه‌ی بزرگ تصویر نیاز دارند کمک کند. یکی از ابزارهای مهم در این زمینه، موتور کاوش تصویر است (۵). استفاده از ابزارهای کارآمد برای بازیابی دقیق و بهینه‌ی تصاویر برای کاربران، یک ضرورت اجتناب‌ناپذیر است (۶). تصاویر برای استفاده‌کنندگان در برخی موضوعات و به ویژه حوزه‌ی علوم پزشکی دارای اهمیت فوق‌العاده‌ای هستند. در رشته‌های علوم پزشکی از تصاویر در امور آموزشی، پژوهشی و درمانی بسیار استفاده می‌شود تا آن‌جا که برخی از امور آموزشی یا درمانی بدون دسترسی و استفاده از تصاویر مورد نیاز، بسیار مشکل یا غیرممکن می‌گردد.

به منظور جستجو و بازیابی بهینه‌ی اطلاعات و تصاویر، عموماً کاربران در جستجوی موتور کاوشی هستند که تنها اطلاعات کاملاً مرتبط را ارائه دهد. برای نیل به این هدف علاوه بر استفاده از روش‌های جستجوی پیشرفته و عمل‌گرها در موتورهای کاوش عمومی از موتورهای کاوش تخصصی نیز استفاده می‌شود. یک موتور کاوش تخصصی بر روی موضوعاتی خاص، منطقه‌ی جغرافیایی خاص و ... تأکید می‌کند و در نتیجه موتورهای کاوش تخصصی صفحات کم‌تری را نمایه می‌کنند و سعی دارند صفحات نامرتبط را کنار بگذارند (۷). اما باید با انجام بررسی‌هایی مشخص شود که موتورهای کاوش تخصصی نیز تا چه اندازه در رسیدن به این هدف موفق بوده‌اند.

نظر به برخی مشکلات اشاره شده، پژوهش‌هایی در خصوص موتورهای کاوش به منظور جستجو و بازیابی بهینه‌ی اطلاعات انجام شده است. نکته‌ای که کاربران موتورهای کاوش باید مدنظر داشته باشند، این است که این ابزارها چه توانایی‌ها و چه محدودیت‌هایی دارند. سپس با توجه به این نکته جستجوهای خود را انجام دهند. مطالعات مقایسه‌ای مختلفی درباره‌ی موتورهای کاوش اینترنت صورت گرفته و تقریباً هر یک به نتایج متفاوتی دست یافته‌اند (۸). علت این امر نیز تفاوت در

هر کدام از موتورهای کاوش، ویژگی‌های کاوش و بازیابی اطلاعات مختص به خود را دارند و از نظر اندازه، سیاست‌های نمایه‌سازی، دقت و ... با یکدیگر متفاوت‌اند. برخی از موتورهای کاوش نیز در زمینه‌های خاص موضوعی مانند علوم پزشکی به نمایه‌سازی و بازیابی اطلاعات می‌پردازند. بیش‌تر موتورهای کاوش، اطلاعات میلیون‌ها مدرک اینترنتی متنی، تصویری، صوتی و ... 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 کلیه‌ی منابع اینترنتی را ندارند (۱).

از سوی دیگر به دلیل حجم زیاد اطلاعات، تعداد نتایج بازیابی شده برای پرسش‌های متعددی که مطرح می‌شود، به هزاران و گاه میلیون‌ها می‌رسد که البته در میان این موارد بازیابی شده، اقلامی دیده می‌شوند که از نظر محتوایی با کلیدواژه‌ی موردنظر مرتبط نیستند و یا این‌که صفحاتی هستند که دیگر در وب وجود ندارند. موتورهای کاوش عمده، پایگاه‌هایی دارند که شامل میلیون‌ها صفحه است. با این تعداد صفحه، کاربران می‌توانند بخش عمده‌ای از اطلاعات مورد نیاز را بیابند.

طبق بررسی‌های انجام شده استفاده از اینترنت باعث افزایش بهره‌وری پژوهشگران و افزایش تولید اطلاعات شده است (۲)، اما شبکه‌ی وب، منبعی است که اطلاعات بسیار با ارزش و نیز اطلاعات بی‌پهلو و زاید را با هم نگهداری می‌کند. با توجه به تراکم اطلاعات موجود در محیط اینترنت، بازیابی اطلاعات از این ابزار مهم با مشکلات فراوانی روبروست. یک مشکل عمده در تحقیقات وب، تعیین کیفیت اطلاعات بازیابی شده است. بسیاری از سایت‌های اینترنتی و موتورهای کاوش که منابع را انتخاب و مرور می‌کنند، بیش از آن که متکی بر محتوای اطلاعاتی باشند، بر ارزش‌های موضوعی تمرکز دارند (۳). بنابراین، موضوع مهمی که در بحث جستجوی اطلاعات وجود دارد، دقت نتایجی است که موتور کاوش به کاربر عرضه می‌کند. وقتی جستجو برای یافتن تصویر باشد، ممکن است کاربر خصوصیات تصویر مانند رنگ، بافت، نوع تصویر و ابعاد تصویر را مشخص کند. بنابراین نمایه‌سازی، جستجو و بازیابی تصاویر، با اطلاعات متنی متفاوت است (۴) و جستجو و بازیابی

جانشین برای بازیابی مدارک علمی وب بنیاد است. همچنین نتایج نشان داد که مانعیت و جامعیت رابطه‌ی معکوسی با یکدیگر دارند به این معنی که اگر مانعیت افزایش یابد، جامعیت کاهش پیدا می‌کند و برعکس (۱۰).

Veronis در پژوهشی به مقایسه‌ی شش موتور کاوش پرداخت و به این نتیجه رسید که رضایت کاربران از نتایج جستجوها، در سطح پایینی قرار دارد. درجه‌ی ربط مدارک بازیابی شده که توسط شرکت‌کنندگان تعیین شده است، نیز پایین می‌باشد. همچنین نتایج نشان داد که در بین موتورهای کاوش مورد بررسی، موتورهای کاوش Google و Yahoo به ترتیب با درجه‌ی ربط ۲/۹ و ۲/۸، دارای بالاترین درجه‌ی ربط بودند (۱۱).

Vanhecke و همکاران در پژوهشی دو موتور کاوش تخصصی پزشکی Pubmed و Highwire press را مورد بررسی قرار دادند. آن‌ها دقت بازیابی، سرعت بازیابی، تعداد نتایج به دست آمده و امکانات و توانایی‌های این موتورهای کاوش را با هم مقایسه نمودند و به این نتیجه رسیدند که موتور کاوش Highwire press تعداد نتایج بیش‌تر و مرتبط‌تری را بازیابی می‌کند و از نظر کیفیت گرافیکی بهتر است و دسترسی آسان‌تری را به مقالات تمام متن فراهم می‌آورد، در صورتی که موتور کاوش Pubmed نتایج را سریع‌تر ارائه می‌دهد (۱۲).

آزادی در تحقیقی به بررسی میزان دقت موتورهای کاوش وب در بازیابی اطلاعات کتابداری و اطلاع‌رسانی پرداخت. نتایج این مطالعه نشان داد که اختلاف معناداری در میزان دقت موتورهای کاوش مورد بررسی در بازیابی اطلاعات کتابداری و اطلاع‌رسانی وجود ندارد. موتورهای کاوش مورد بررسی از نظر دقت به این ترتیب رتبه‌بندی شدند: HotBot, Infoseek, WebCrawler و Lycos, Excite, Altavista, Google (۱۳).

قاضی میر سعید، حقانی و اکبری در پژوهشی به مقایسه‌ی موتورها و ابر موتورهای کاوش منتخب در بازیابی اطلاعات فیزیوتراپی از شبکه‌ی جهانی وب پرداختند. در این بررسی هفت موتور کاوش و هفت ابر موتور کاوش تعیین و مورد بررسی قرار گرفتند. نتایج این مطالعه نشان داد که موتورهای کاوش All

روش‌ها، الگوریتم‌های نمایه‌سازی، بازیابی و رتبه‌بندی موتورهای کاوش است. به هر حال مطالعه بر روی جنبه‌های کمی و کیفی موتورهای کاوش و میزان دقت آن‌ها در بازیابی اطلاعات، آثار سودمندی در به‌کارگیری از موتورهای کاوش توسط کاربران خواهد داشت.

براساس بررسی صورت گرفته، مطالعات متعددی وجود دارند که به بررسی جامعیت و دقت موتورهای کاوش در بازیابی اطلاعات پرداخته‌اند اما مطالعه‌ای که به دقت در بازیابی تصاویر پزشکی پرداخته باشد، یافت نشد.

Ilic و همکاران در تحقیقی کارایی موتورهای کاوش عمومی و تخصصی پزشکی و کیفیت اطلاعات بازیابی شده از هر یک را مورد بررسی قرار دادند. برای این تحقیق تعداد ۱۸ کلید واژه و عبارت در رابطه با بیماری Androgen Deficiency of Aging Male (ADAM) در چهار موتور کاوش عمومی Excite, Altavista, Google و Yahoo و پنج موتور کاوش تخصصی پزشکی DrKoop, HealthInsite, NHS, MedlinePlus و از مرور ۴۹۲۷ سایت به این نتیجه رسیدند که از مجموع ۳۲۶۷ وب سایت بازیابی شده از موتورهای کاوش عمومی، ۴۷ مورد (۱/۴۴ درصد) و از مجموع ۱۶۶۰ وب سایت بازیابی شده از موتورهای کاوش تخصصی، ۱۰ وب سایت (۰/۶۶ درصد) مرتبط هستند. همچنین کیفیت اطلاعات بازیابی شده از موتورهای کاوش تخصصی اختلاف معناداری با اطلاعات بازیابی شده از طریق موتورهای کاوش عمومی ندارد (۹).

Shafi و Rather در پژوهشی جامعیت و مانعیت پنج موتور کاوش را از نظر بازیابی اطلاعات علمی در زمینه‌ی بیوتکنولوژی بررسی کردند. موتورهای کاوش انتخابی این پژوهش، موتورهای کاوش عمومی HotBot, Altavista, Google و دو موتور کاوش تخصصی بیوتکنولوژی، Scirus و BioWeb هستند. آن‌ها تعداد ۲۰ عبارت جستجو را در سه گروه عبارات ساده، ترکیبی و اصطلاحات پیشرفته‌ی هر موتور کاوش وارد کرده و در هر مورد ۱۰ نتیجه‌ی اول را مورد بررسی قرار دادند. نتایج پژوهش نشان داد که Scirus در بازیابی مدارک علمی بهترین عملکرد را دارد. Google بهترین

مورد جستجو قرار گرفتند. براساس نتایج حاصل از این پژوهش، در بین موتورهای کاوش عمومی مورد بررسی، موتور کاوش Yahoo با دقت ۹۴/۷ و پس از آن Google با دقت ۹۲ درصد، دارای بیشترین دقت و موتور کاوش Ask با دقت ۸۷/۳ درصد، کمترین دقت را دارا بودند. در بین موتورهای کاوش تخصصی پزشکی مورد بررسی نیز، موتور کاوش Trip Database با دقت ۷۵/۷ درصد، بیشترین دقت و موتور کاوش Medology با دقت صفر، کمترین دقت را دارد. از نظر دقت، موتورهای کاوش عمومی و موتورهای کاوش پزشکی با یکدیگر اختلاف معناداری داشته و دقت موتورهای کاوش عمومی بیش از موتورهای کاوش پزشکی بود (۱۷).

همانگونه که اشاره شد مطالعات چندی در زمینه‌ی بازیابی اطلاعات در موتورهای کاوش انجام شده که این مطالعات، بیش‌تر در حوزه‌ی بازیابی اطلاعات متنی بوده است. جستجو و بازیابی تصاویر در حوزه‌هایی نظیر علوم پزشکی کاربرد بسیار دارد (۱۸). تصاویر در همه جای پزشکی نوین حضور دارند. بنابراین شناخت و استفاده از موتورهای کاوش مناسب، برای جستجو و بازیابی تصاویر پزشکی مورد نیاز کاربران در حوزه‌ی علوم پزشکی، کاملاً ضروری است. اما مطالعات اندکی در این زمینه صورت گرفته است. با توجه به اهمیت این موضوع، ضرورت انجام مطالعات بیش‌تر در این زمینه کاملاً احساس می‌شود؛ که مطالعه‌ی حاضر نیز در این راستا صورت گرفته است.

هدف اصلی در این پژوهش، تعیین و مقایسه‌ی میزان دقت موتورهای کاوش عمومی و تخصصی پزشکی در بازیابی تصاویر پزشکی بود.

### روش بررسی

تحقیق حاضر، کاربردی و از نوع پیمایشی مقایسه‌ای بود که در سال ۱۳۹۰ خورشیدی انجام شد. جامعه‌ی مورد بررسی در این پژوهش، دو موتور کاوش عمومی Yahoo و Google و دو موتور کاوش تخصصی پزشکی OmniMedicalSearch و Trip Database بودند که در پایگاه‌هایی مانند Searchenginewatch.com و سایر سایت‌ها به عنوان

the web, AltaVista و Google به ترتیب بیش‌ترین نتایج بازیابی را داشتند و در بین ابر موتورهای کاوش نیز Ixquick دارای بیش‌ترین نتایج بازیابی بود (۱۴).

علیچانی، نیک کار و دهقانی در مطالعه‌ی موتورهای فراموتورهای کاوش را در پاسخ‌گویی به سوالات مرجع آنی، از لحاظ دقت، ریزش کاذب و میزان پیوندهای مرده مقایسه نمودند. نتایج نشان داد موتور کاوش MSN با ۴۵ درصد بازیابی مدارک مربوط و در رتبه دوم Google با ۴۱ درصد بازیابی مدارک مرتبط، بهترین عملکرد را در بین موتورهای کاوش به خود اختصاص داده‌اند. در بین فراموتورهای مورد بررسی نیز به ترتیب Vivisimo، Dogpile و Metacrawler با ۳۶، ۳۲ و ۳۱ درصد بازیابی مدارک مرتبط، در رتبه‌های اول تا سوم قرار می‌گیرند (۱۵).

محمداسماعیل، لفظی قاضی و گیلوری طی پژوهشی به مقایسه‌ی موتورهای و ابرموتورهای کاوش در بازیابی اطلاعات داروشناسی پرداختند. در این پژوهش ۸ کلیدواژه انتخاب و در موتورهای و ابرموتورهای کاوش مورد جستجو قرار گرفتند. از میان نتایج جستجو، ۱۰ نتیجه‌ی اول هر موتور کاوش مورد بررسی قرار گرفتند. نتایج این پژوهش نشان داد که موتور کاوش Yahoo بیش‌ترین مدارک داروشناسی (۳۴ درصد) را بازیابی و رتبه‌ی نخست را به خود اختصاص داد. موتور کاوش AOL با میزان ۶۲ درصد مانعیت و ۲۱ درصد جامعیت، بیش‌ترین مدارک مرتبط را در حوزه‌ی داروشناسی بازیابی نمود. در بین ابر موتورهای کاوش، Dogpile، بیش‌ترین مدارک داروشناسی (۲۲ درصد) را بازیابی و رتبه‌ی نخست را به خود اختصاص داد (۱۶).

صراطی شیرازی طی پژوهشی به بررسی میزان دقت موتورهای کاوش عمومی و تخصصی پزشکی در بازیابی مدارک مربوط به بیماری‌های کودکان پرداخت. در این پژوهش ۱۰ کلیدواژه‌ی مربوط به بیماری‌های کودکان با مشورت پزشکان متخصص کودکان انتخاب شده و در پنج موتور کاوش عمومی Ask، Altavista، Google، MSN و Yahoo و هم‌چنین در پنج موتور کاوش تخصصی پزشکی Healthinsight، Trip Database و Medology، MedlinePlus، Medic8

کاملاً مرتبط با حرف "A"، تصاویر نسبتاً (تا اندازه‌ای) مرتبط با حرف "B"، تصاویر مرتبط با "A+B" و تصاویر غیر مرتبط با "-" نشان داده شدند.

دقت (مانعیت) نیز که نسبت تعداد منابع مرتبط بازیابی شده توسط موتور کاوش به کل منابع بازیابی شده توسط آن است؛ با استفاده از فرمول زیر محاسبه گردید.

$$\text{دقت} = \frac{\text{تعداد منابع (تصاویر) مرتبط بازیابی شده}}{\text{تعداد کل منابع (تصاویر) بازیابی شده}} \times 100$$

جهت تجزیه و تحلیل داده‌های پژوهش در سطح توصیفی از مشخصه‌های آماری نظیر فراوانی و درصد استفاده شد و در سطح استنباطی به منظور مقایسه‌ی بین میزان دقت موتورهای کاوش عمومی و موتورهای کاوش تخصصی پزشکی در بازیابی تصاویر کاملاً مرتبط (A) و تصاویر مرتبط (A+B) پزشکی و تعیین معنی‌داری تفاوت‌ها، از آزمون t مستقل استفاده شد. در این پژوهش از نرم‌افزار SPSS برای تحلیل داده‌ها و قضاوت در مورد آن‌ها مورد استفاده قرار گرفت.

### یافته‌ها

همان‌گونه که اشاره شد تعداد ۳۲۴ تصویر پزشکی با استفاده از ده کلید واژه‌ی ذکر شده توسط چهار موتور کاوش مورد بررسی بازیابی شد. طبق اطلاعات مندرج در جدول ۱، موتور کاوش Google با ۱۰۰ تصویر (۳۰/۹ درصد)، بیش‌ترین و موتور کاوش TripDatabase با ۳۷ تصویر (۱۱/۴ درصد)، کم‌ترین تصاویر پزشکی را بازیابی نمودند.

نمودار ۱، تعداد فراوانی تصاویر پزشکی بازیابی شده براساس هر یک از ۱۰ کلیدواژه‌ی مورد بررسی در چهار موتور کاوش مورد جستجو را نشان می‌دهد. همان‌طور که ملاحظه می‌شود در موتور کاوش Google با جستجوی هر یک از کلیدواژه‌های مورد اشاره، تعداد ۱۰ تصویر مورد انتظار بازیابی شده است. اما در موتور کاوش TripDatabase، با جستجوی کلیدواژه‌های "Mesenchymal stem cell" و "Ascites+abdomen" هیچ تصویری بازیابی نشد.

موتورهای کاوش برتر علوم پزشکی معرفی شده‌اند. به منظور تعیین و مقایسه‌ی میزان دقت بازیابی تصاویر پزشکی در چهار موتور کاوش مورد بررسی، ده کلیدواژه ذیل با مشورت متخصصان پزشکی انتخاب و به منظور جستجو و بازیابی تصاویر در این موتورهای کاوش مورد استفاده قرار گرفتند. لازم به ذکر است که کلیدواژه‌های ذیل توسط متخصصان پزشکی با توجه به نیازهای آنان به تصاویر انتخاب شد و آنان در استفاده از اصطلاح‌نامه‌ها در تعیین کلیدواژه‌ها مختار بودند.

"Hematopoietic stem cell"، "Mesenchymal stem cell"، "Acute lymphocytic leukemia"

"Myelodysplastic syndrome"، "RBC morphology"

+ report، "Structure of nucleic acids"

"Megaloblastic anemia" + atlas Pneumonia + chest X-ray، "Greenstick + fracture + X-ray"

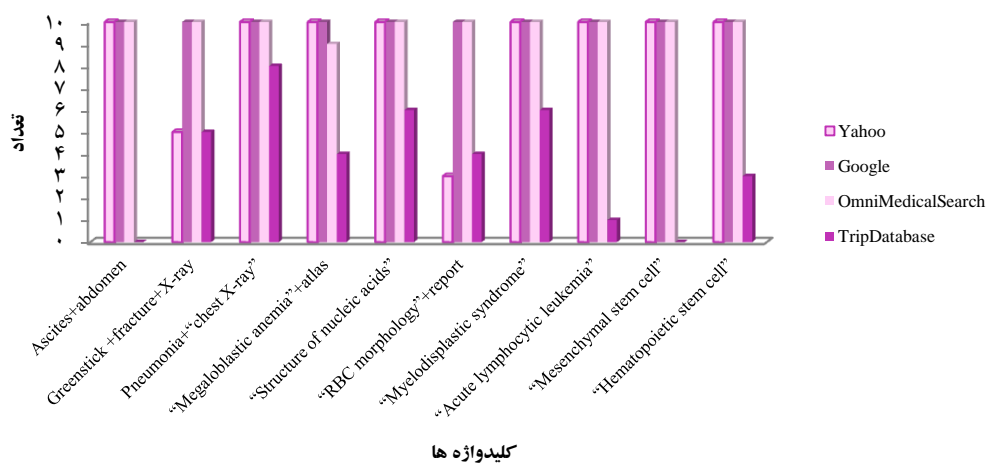
Ascites + abdomen

سپس ۱۰ نتیجه‌ی اول جستجو با کلیدواژه‌ها در هر موتور کاوش جهت بررسی استخراج شد. انتظار می‌رفت تعداد این تصاویر پزشکی ۴۰۰ مورد باشد، اما به دلیل این که در برخی از موتورهای کاوش مورد بررسی، تعداد کل تصاویر بازیابی شده، کم‌تر از ۱۰ تصویر بود، تعداد تصاویر مورد انتظار به ۳۲۴ تصویر کاهش یافت که این مورد از محدودیت‌های این پژوهش بود. در نهایت تعداد ۳۲۴ تصویر پزشکی در این پژوهش مورد بررسی قرار گرفتند. با توجه به این که انتخاب کلیدواژه‌ها توسط متخصصان پزشکی و با توجه به نیاز اطلاعاتی واقعی آن‌ها انجام شده است، در مواردی که انجام کاوش به بازیابی کم‌تر از ۱۰ تصویر منجر شده، کلیدواژه‌های مذکور از پژوهش کنار گذاشته نشده‌اند و محاسبه‌ی دقت با در نظر گرفتن تعداد تصاویر بازیابی شده انجام گرفته است.

در مرحله‌ی بعد تصاویر پزشکی بازیابی شده جهت بررسی و تعیین ارتباط یا عدم ارتباط آن با کلید واژه مورد جستجو در اختیار متخصصان پزشکی قرار گرفت. در این پژوهش برای بررسی میزان ربط هر یک از تصاویر پزشکی بازیابی شده، براساس تحقیق حریری از یک مقیاس سه نقطه‌ای استفاده گردید (۲۰، ۱۹). بدین ترتیب که طبق نظر متخصصان پزشکی، تصاویر

جدول ۱: توزیع فراوانی تصاویر بازیابی شده توسط هریک از موتورهای کاوش

| تعداد | درصد | تصاویر بازیابی شده موتور کاوش |
|-------|------|-------------------------------|
| ۸۸    | ۲۷/۲ | Yahoo                         |
| ۱۰۰   | ۳۰/۹ | Google                        |
| ۹۹    | ۳۰/۶ | OmniMedicalSearch             |
| ۳۷    | ۱۱/۴ | TripDatabase                  |
| ۳۲۴   | ۱۰۰  | جمع                           |



نمودار ۱: تعداد فراوانی تصاویر بازیابی شده براساس کلیدواژه‌های مورد بررسی در چهار موتور کاوش مورد نظر

هم‌چنین دقت موتور کاوش Google در بازیابی تصاویر پزشکی کاملاً مرتبط، ۷۳ درصد و دقت این موتور کاوش در بازیابی تصاویر مرتبط، ۹۰ درصد است.

طبق اطلاعات جدول ۲، در بین دو موتور کاوش عمومی مورد بررسی، موتور کاوش Google با ۷۳ درصد تصاویر کاملاً مرتبط، دارای بیش‌ترین تصاویر پزشکی مرتبط و کاملاً مرتبط می‌باشد.

جدول ۲: توزیع فراوانی میزان ربط و دقت بازیابی تصاویر در موتورهای کاوش عمومی

| موتور کاوش | ربط   |      | تصاویر کاملاً مرتبط (A) |      | تصاویر مرتبط (A+B) |      | تصاویر غیر مرتبط |      | دقت (A+B) |
|------------|-------|------|-------------------------|------|--------------------|------|------------------|------|-----------|
|            | تعداد | درصد | تعداد                   | درصد | تعداد              | درصد | تعداد            | درصد |           |
| Yahoo      | ۴۷    | ۵۳/۴ | ۷۴                      | ۸۴/۱ | ۱۴                 | ۱۵/۹ | ۱۴               | ۱۵/۹ | ۸۴        |
| Google     | ۷۳    | ۷۳   | ۹۰                      | ۹۰   | ۱۰                 | ۱۰   | ۱۰               | ۷۳   | ۹۰        |
| جمع        | ۱۲۰   | ۶۳/۸ | ۱۶۴                     | ۸۷/۲ | ۲۴                 | ۱۲/۸ | ۲۴               | ۶۳/۸ | ۸۷/۲      |

جدول ۳: توزیع فراوانی میزان ربط و دقت بازیابی تصاویر در موتورهای کاوش تخصصی پزشکی

| موتور کاوش        | ربط   |      | تصاویر کاملاً مرتبط (A) |      | تصاویر مرتبط (A+B) |      | تصاویر غیر مرتبط |      | دقت  | دقت   |
|-------------------|-------|------|-------------------------|------|--------------------|------|------------------|------|------|-------|
|                   | تعداد | درصد | تعداد                   | درصد | تعداد              | درصد | تعداد            | درصد | (A)  | (A+B) |
| OmniMedicalSearch | ۵۶    | ۵۶/۶ | ۸۶                      | ۸۶/۹ | ۱۳                 | ۱۳/۱ | ۱۳               | ۱۳/۱ | ۵۶/۵ | ۸۶/۸  |
| TripDatabase      | ۱۹    | ۵۱/۴ | ۳۱                      | ۸۳/۸ | ۶                  | ۱۶/۲ | ۶                | ۱۶/۲ | ۵۱/۳ | ۸۳/۷  |
| جمع               | ۷۵    | ۵۵/۱ | ۱۱۷                     | ۸۶   | ۱۹                 | ۱۴   | ۱۹               | ۱۴   | ۵۵/۱ | ۸۶    |

در بررسی دقت بازیابی تصاویر پزشکی نیز موتور کاوش Google، با دقت ۷۳ در بازیابی تصاویر کاملاً مرتبط و دقت ۹۰، در بازیابی تصاویر مرتبط، بیشترین دقت را داشت. پس از این موتور به ترتیب موتورهای کاوش OmniMedicalSearch، Yahoo و TripDatabase با دقت ۸۶/۸، ۸۴ و ۸۳/۷ قرار دارند. همان‌گونه که قبلاً اشاره شد، مطالعه‌ای بر روی دقت بازیابی تصاویر در موتورهای کاوش یافت نشد، اما در خصوص بازیابی اطلاعات متنی توسط موتورهای کاوش، Shafi و Rather، Veronis و صراطی شیرازی در مطالعات خود به نتایج تقریباً مشابهی دست یافته‌اند (۱۷، ۱۱، ۱۰). نکته‌ی قابل ذکر دیگر این است که موتور کاوش TripDatabase که در مطالعه‌ی حاضر کمترین دقت را در بازیابی تصاویر پزشکی دارا بود، در مطالعه‌ی صراطی شیرازی، در بین موتورهای کاوش تخصصی پزشکی دارای بیشترین دقت در بازیابی اطلاعات متنی مربوط به بیماری‌های کودکان بوده است (۱۷). این مورد می‌تواند نشان‌دهنده‌ی تفاوت بین نتایج بازیابی اطلاعات متنی و تصویری در موتورهای کاوش باشد.

در پژوهش Lopes و Ribeiro که به ارزیابی تطبیقی موتورهای کاوش وب در بازیابی اطلاعات سلامت پرداختند، میزان ربط اطلاعات بازیابی شده از چهار موتور کاوش Bing، Google، Yahoo و Sapo و سه موتور کاوش تخصصی حوزه‌ی سلامت WebMD، Medline Plus، SapoSaude براساس نظرات کاربران و با استفاده از یک مقیاس بررسی ربط دو حالتی (مرتبط و غیرمرتبط) مورد بررسی قرار گرفت. نتایج نشان داد که دقت موتورهای کاوش عمومی از موتورهای کاوش تخصصی حوزه‌ی سلامت بیش‌تر است. Google به ویژه در ۱۰ نتیجه نخست

براساس اطلاعات جدول ۳، در بین دو موتور کاوش پزشکی مورد بررسی، موتور کاوش OmniMedicalSearch با بازیابی ۵۶/۶ درصد تصاویر کاملاً مرتبط، دارای بیشترین تصاویر مرتبط و کاملاً مرتبط پزشکی می‌باشد. هم‌چنین دقت موتور OmniMedicalSearch در بازیابی تصاویر پزشکی کاملاً مرتبط و مرتبط، به ترتیب ۵۶/۵ و ۸۶/۸ درصد است.

نتایج تحقیق نشان داد بین میزان دقت موتورهای کاوش عمومی و موتورهای کاوش تخصصی پزشکی در بازیابی تصاویر کاملاً مرتبط (A) پزشکی تفاوت معنی‌داری وجود ندارد، بدین منظور از آزمون t مستقل استفاده شد. با توجه به نتایج آزمون t مستقل،  $t=1/577$  و  $df=322$  با درجه‌ی اطمینان ۹۵٪، سطح معنی‌داری برابر با ۰/۱۱۶ است، پس فرض صفر تأیید گردید. هم‌چنین نتایج آزمون t مستقل در خصوص تفاوت بین میزان دقت موتورهای کاوش عمومی و موتورهای کاوش تخصصی پزشکی در بازیابی تصاویر مرتبط (A+B) پزشکی نیز  $t=0/314$  و  $df=322$  با درجه‌ی اطمینان ۹۵٪، سطح معنی‌داری برابر با ۰/۷۵۳ است، پس فرض صفر تأیید می‌گردد. بنابراین تفاوت معنی‌داری بین میزان دقت موتورهای کاوش عمومی و پزشکی مورد بررسی در بازیابی تصاویر مرتبط (A+B) پزشکی وجود ندارد.

## بحث

یافته‌های پژوهش نشان داد که در بین چهار موتور کاوش مورد بررسی، تنها موتور کاوش Google، تمامی تصاویر مورد انتظار یعنی ۱۰۰ تصویر (۱۰۰ درصد) را بازیابی نمود. موتور کاوش پزشکی TripDatabase، کمترین تعداد تصاویر مورد انتظار یعنی ۳۷ تصویر (۳۷ درصد) را بازیابی نمود.

هستند و توسط چند موتور کاوش بازیابی شده‌اند، اما هر کدام از موتورهای کاوش نتایج منحصر به فردی را نیز بازیابی کرده‌اند. این موضوع نشان می‌دهد که الگوریتم نمایه‌سازی منابع موجود در وب، توسط موتورهای کاوش مختلف، متفاوت است و از این رو برخی اطلاعات و تصاویر تنها در بانک اطلاعاتی آن موتور کاوش خاص وجود دارد.

### پیشنهادات

با توجه به تفاوت در الگوریتم نمایه‌سازی منابع و تصاویر در موتورهای کاوش مختلف، پیشنهاد می‌شود که برای انجام یک جستجوی کامل و بازیابی بیش‌ترین نتایج و تصاویر، از چند موتور کاوش عمومی و تخصصی پزشکی به‌طور هم‌زمان استفاده شود. همچنین برخی از موتورهای کاوش دارای امکانات ویژه جهت جستجوی تصاویر می‌باشند که استفاده از این موتورها و امکانات جستجوی تصویر آن‌ها به پژوهشگران، متخصصان علوم پزشکی و سایر کاربران توصیه می‌شود. از آن‌جا که ابزارهای کاوش جدید، هر روزه در محیط وب ظهور می‌کنند، پیشنهاد می‌شود سایر موتورهای کاوش تخصصی پزشکی و به ویژه موتورهای کاوش جدید نیز مورد ارزیابی و مقایسه قرار گیرند.

### تشکر و قدردانی

بدین‌وسیله نویسندگان از متخصصان محترمی که در انجام این پژوهش همکاری نمودند و همچنین از داوران محترم که دیدگاه‌های ارزنده‌ای ارائه نمودند، سپاس‌گزاری می‌نمایند.

بهترین نتایج را ارایه داد. همچنین نتایج نشان داد که نیازهای اطلاعاتی در شرایط حساس و خاص دارای دقت بیش‌تری هستند. نتایج به‌دست آمده از این تحقیق در خصوص بازیابی اطلاعات توسط موتورهای کاوش عمومی با نتایج پژوهش حاضر مشابهت دارد (۲۱).

نتایج آزمون آماری  $t$  مستقل نشان داد که اختلاف معنی‌داری بین میزان دقت موتورهای کاوش عمومی و پزشکی مورد بررسی در بازیابی تصاویر پزشکی مرتبط و کاملاً مرتبط وجود ندارد. Ilic و همکاران نیز در تحقیق خود به نتیجه‌ی مشابهی در خصوص بازیابی اطلاعات متنی توسط موتورهای کاوش دست یافتند (۹). البته نتایج پژوهش‌های Lopes و Ribeiro و صراطی شیرازی نشان داد که بین میزان دقت موتورهای کاوش عمومی و پزشکی در بازیابی اطلاعات متنی تفاوت معنی‌دار وجود دارد و میزان دقت موتورهای کاوش عمومی بیش‌تر است (۲۱، ۱۷). در مطالعات اشاره شده برخی موتورهای کاوش تخصصی پزشکی انتخاب شده با موتورهای کاوش پزشکی مورد بررسی در تحقیق حاضر متفاوتند.

### نتیجه‌گیری

نتایج حاصل از این تحقیق مشخص نمود که در بین موتورهای کاوش عمومی، Google و در بین موتورهای کاوش پزشکی OmniMedicalSearch دارای بیش‌ترین و مرتبط‌ترین تصاویر پزشکی در خصوص کلیدواژه‌های مورد بررسی هستند. بررسی تصاویر پزشکی بازیابی شده توسط موتورهای کاوش عمومی و تخصصی پزشکی نشان‌گر آن است که علی‌رغم این‌که برخی از نتایج در بعضی از موتورهای کاوش مشترک

### References

1. Flanagan D. Search Engines [Online]. 1999 [Cited 2011 Dec 20]; Available from: URL: <http://www.home.sprintmail.com/~debflanagan/engines.htm>. 1999.
2. Rezaei Sharifabadi S. Effects of the Internet on Research Activities. Tehran: Ketabdar; 1999. [Book in Persian]
3. Garency M. What Distance Learners Should Know about Information Retrieval on the World Wide Web? Trans. Azadi Gh. Fasname-ketab 2004; 14(4): 169-78. [In Persian]
4. Ze Wang J, Wiederhold G, Firschein O, Wei S.X. Content based Image Indexing and Searching Using Daubechies Wavelets. Int J Digit Lib 1997; 1(4): 311-28.
5. Hassan I, Zhang J. Image Search Engine Feature Analysis. Online Information Review 2001; 25(2): 103-14.
6. Kherfi ML, Ziou D, Bernardi A. Image Retrieval from the World Wide Web: Issues, Techniques and Systems.



- ACM Computing Survey 2004; 36(1): 35-67.
7. King D. Specialized Search Engine: Alternative to the Big Guys [Online]. 2000 [Cited 2011 Nov 5]; Available from: URL: <http://www.infotoday.com/online/OL2000>.
  8. Kousha K. Evaluation Criteria of Internet Search Engines. Informology 2003; 1(1): 79-104.
  9. Ilic D, Bessell TL, Silagy CA, Green S. Specialized Medical Search-Engines Are Not Better Than General Search-Engines in Sourcing Consumer Information about Androgen Deficiency. Hum Reprod 2003; 18(3): 557-61.
  10. Shafi SM, Rather RA. Precision and Recall of Five Search Engines for Retrieval of Scholarly Information in the Field of Biotechnology [Online]. Webology 2005; 2(2). Available from: URL: <http://www.webology.org/2005/v2n2/a12.html>
  11. Veronis J. A Comparative Study of Six Search Engines [Online]. 2006 [Cited 2011 Dec 12]; Available from: URL: <http://www.up.univ-mrs.fr/veronis/pdf/2006-comparative-study.pdf>.
  12. Vanhecke TE, Barnes MA, Zimmerman J, Shoichet S. Pubmed vs. High Wire Press: A Head-to Head Comparison of Two Medical Literature Search Engines. Comput Biol Med 2007; 37(9): 1252-8.
  13. Azadi G. The Scale of Web Search Engines Precision in Information Retrieval of Library and Information Science Discipline. Faslname-ketab 2005; 16(3): 111-22. [In Persian]
  14. Ghazi-Mirsaeed J, Haghani H, Akbari A. A Comparative Study of Search Engines and Metasearch Engines in Retrieving Physiotherapy Information from WWW and Determining Their Overlapping. Health Information Management 2007; 4(1): 11-21. [In Persian]
  15. Alijani R, Nikkar M, Dehghani L. Comparison of Search Engines and Meta-Search Engines in Answering Ready-Reference Questions. Faslname-ketab 2008; 19(2): 171-88. [In Persian]
  16. Mohammadesmaeil S, Lafzighazi E, Gilvari A. Comparing Search Engines and Meta-search Engines in Pharmaceutics Information Retrieval. Health Information Management 2008; 5(2): 121-9. [In Persian]
  17. Serati Shirazi M. A Comparative Study of the Accuracy of General and Specialized Medical Search Engines in Retrieving Documents Related to Children Diseases. Faslname-ketab 2009; 20(1): 77-94. [In Persian].
  18. Dastghaib M. A Survey on Textual, Content-Based, and Fuzzy Indexing Images: Digital Image Retrieval. Faslname-ketab 2006; 16(4): 111-6. [In Persian]
  19. Hariri N. An Investigation on Effectiveness of the "Similar Pages" Feature of Google. Online Information Review 2008; 32(3): 370-8.
  20. Hariri N. Relevance Ranking on Google: Are Top Ranked Results Really Considered More Relevant by the Users? Online Information Review 2011; 35(4): 598-610.
  21. Lopes CT, Ribeiro C. Comparative Evaluation of Web Search Engines in Health Information Retrieval. Online Information Review 2011; 35(6): 869-92.

## A Comparison of the Precision of General and Specialized Medical Search Engines in Medical Images Retrieval\*

Nadjla Hariri<sup>1</sup>; Hosein Vakili Mofrad<sup>2</sup>

### Original Article

#### Abstract

**Introduction:** The huge volume of information and images on the Internet and its increasing growth has made the optimized search performance difficult for users. Among search tools, search engines are mostly used by users. This research was aimed to determine and compare the precision of general and specialized search engines in medical images retrieval.

**Methods:** This applied, comparative approach survey was conducted in 2011. The research sample included 2 general search engines, i.e. Yahoo and Google, and 2 specialized medical search engines, i.e. OmniMedicalSearch and Trip Database. In order to determine and compare the precision of image retrieval in the four search engines, 10 medical keywords were chosen after consulting medical specialists and were searched in the mentioned search engines. Then the first 10 results retrieved by each search engine were extracted and the opinions of medical specialists about relevance of results were received. Descriptive statistics, Independent t-test were used to analyze the data with SPSS.

**Results:** It was found that Google search engine has the most precision (90) in retrieving of relevant medical images and then OmniMedicalSearch search engine has precision to be equal 86/8. Among general search engines and specialized medical search engines there aren't meaningful differences in retrieval of medical images.

**Conclusion:** Despite of some overlap between search engines results, each search engine also was retrieved unique results. This matter indicate that the indexing algorithm of resources on the web different in the various search engines; and it is suggested to use several general and specialized search engines simultaneously to search medical images.

**Keywords:** Internet; Medical Information Storage and Retrieval; Medical Illustration; Precision; Search Engine

Received: 17 Nov, 2012

Accepted: 23 May, 2013

**Citation:** Hariri N, Vakili Mofrad H. A Comparison of the Precision of General and Specialized Medical Search Engines in Medical Images Retrieval. Health Inf Manage 2014; 10(6): 839.

\* This article was an independent research with no financial aid.

1- Associate Professor, knowledge and Information Sciences, Islamic Azad University, Science and Research Branch, Tehran, Iran

2- Lecturer, Medical Library and Information Sciences, Faculty of Allied Medical Sciences, Hamadan University of Medical Sciences, Hamadan, Iran (Corresponding Author) Email: vakili.hn@gmail.com