

میزان جامعیت و مانعیت بازیابی تصاویر در موتورهای کاوش Bing و Google Image

علی جلالی دیزجی*

دانشیار گروه علم اطلاعات و دانش شناسی دانشگاه علامه طباطبائی

رضا خالقی

کارشناسی ارشد علم اطلاعات و دانش شناسی دانشگاه علامه طباطبائی

سید مهدی طاهری

استادیار گروه علم اطلاعات و دانش شناسی دانشگاه علامه طباطبائی

تاریخ پذیرش: ۱۳۹۵/۹/۹

تاریخ دریافت: ۱۳۹۵/۳/۱۶

چکیده

هدف: هدف پژوهش، تعیین میزان جامعیت و مانعیت دو موتور کاوش بینگ و گوگل در بازیابی محتوای محور تصاویر است. **روش شناسی:** روش پژوهش شبه آزمایشی است، جامعه پژوهش، تصاویر پایگاه‌های دو موتور کاوش بینگ و گوگل، و نمونه شامل جستجوی ۱۵ تصویر منتخب در هر موتور کاوش است. همه منابع بازیابی شده با جستجوی محتوای محور گردآوری شده و جامعیت و مانعیت نتایج هر موتور کاوش با فرمول ربط محاسبه و میانگین درصدها به دست آمده است. فرضیه‌های پژوهش با آزمون یومن ویتنی^۱ بررسی شده‌اند. **یافته‌ها:** یافته‌ها نشان می‌دهد جامعیت گوگل با ۸۸/۷۳٪ دارای رتبه‌ی بالاتری در میزان بازیابی نسبت به موتور کاوش بینگ با جامعیت ۲۰/۸۶٪ است. اما موتور کاوش بینگ با ۹۶/۸۶٪ مانعیت، رتبه‌ی بالاتری در میزان دقت نسبت به موتور کاوش گوگل با ۹۴/۸۰٪ داشته است. بین میزان جامعیت دو موتور کاوش، تفاوت معناداری با اطمینان ۹۵٪ وجود دارد اما بین میزان مانعیت آنها تفاوت معناداری موجود نیست.

کلیدواژه‌ها: بازیابی محتوایی تصاویر، گوگل، بینگ، جامعیت، مانعیت

مقدمه و بیان مسئله

در عصر حاضر شاهد رشد روزافزون تولید علم و فناوری هستیم. بر این اساس انواع اطلاعات در قالب‌های مختلف در حال افزایش است و هر روز حجم‌تر می‌شود. این روند، کاربران اطلاعات را با سردرگمی و مشکل در بازیابی اطلاعات مناسب مواجه می‌سازد. گسترش وب جهانی و افزایش حجم انواع اشیاء چندرسانه‌ای (متنی، تصویری، صوتی، ویدئویی، پویانمایی^۱ و جز آن) سبب شده تا متخصصان نمایه‌سازی برای سازمان‌دهی این اطلاعات حجم، فعالیت‌های خود را بر پایه‌ی شبکه جهانی توسعه دهند (Hidderley & Rafferty, 2005).

بی‌شک، این حجم عظیم اطلاعات، نیاز به سازمان‌دهی دارد تا هنگام جستجو، دسترسی به سرعت و سهولت انجام پذیرد. رسالت اصلی هر کتابخانه یا محیط اطلاعاتی، در دسترس قرار دادن اطلاعات موجود بر اساس نیاز کاربران است. در راستای تحقق این وظیفه، اطلاعاتی که در کتابخانه‌ها یا محیط‌های اطلاعاتی ذخیره می‌شوند؛ باید بر مبنای مجموعه‌ای از اصول، روش‌ها و ابزارها نمایه گردند تا به بهبود بازیابی اطلاعات ذخیره شده کمک کنند. در طول تاریخ تحول و گسترش کتابخانه‌ها و مراکز اطلاع رسانی، به منظور در دسترس قرار دادن اطلاعات، رده بندی و نمایه سازی به عنوان دو فعالیت مهم سازماندهی اطلاعات به کار رفته است. بنابراین، نمایه‌سازی یکی از فعالیت‌های اصلی در زمینه پردازش اطلاعات است و در بازیابی اطلاعات تأثیر چشمگیری دارد (Hidderley & Rafferty, 2005).

اکنون بسیاری از وب‌گاه‌ها، برنامه‌ای برای جستجوی درون وب‌گاه خود طراحی می‌کنند که همچون فایل‌های تمام‌متن می‌توانند در نتایج جستجوی خود دارای مدرک نامربوط و به‌اصطلاح همراه با ریزش کاذب باشند. اگر در چنین وب‌گاه‌هایی نمایه‌ای شبیه آن‌چه در انتهای کتاب‌ها دیده می‌شود، وجود داشته باشد کاربر می‌تواند به سرعت و با استفاده از لیست نمایه به مدخل مورد نظر خود وارد شده و با سرعت بالا و با صرف زمان اندکی به مدارک مورد نیاز و حتی مدارک مرتبط با آن نیز دسترسی یابد. بعضی از ابزارهای جستجوی وب سعی می‌کنند در مرور وب‌گاه‌ها توسط افراد متخصص صورت پذیرد و این متخصصان پس از تحلیل محتوای آن، کلیدواژه‌های مناسب را انتخاب کرده و بر اساس لیست موضوعی ویژه خود آن را در محل موضوعی قرار می‌دهند. این ابزارها در واقع برای استفاده کنندگان یک راهنمای موضوعی را فراهم می‌آورند. انجام این کار در اصل بدین معناست که موتورهای کاوش به‌طور معمول برای هدایت فرد به وب‌گاه، تمام صفحات آن را نمایه کرده است، ولی از سوی دیگر یک راهنمای موضوعی بسیار شبیه به پیوند درونی آن تلقی می‌شود (Tyner, 2001).

همان‌گونه که اشاره شد در محیط وب جهانی انواع اطلاعات چندرسانه‌ای ذخیره و هرکدام به شیوه‌ای نمایه‌سازی و قابل دسترس می‌شوند. یکی از انواع اطلاعات موجود در وب، تصاویر به عنوان یکی از مقبول‌ترین داده‌های موجود در وب هستند. وب منبع مهمی برای یافتن تصاویر است. جستجو و بازیابی تصاویر حوزه‌ای از نمایه‌سازی است که همواره با چالش‌های عمده‌ای مواجه بوده است (کوکبی، ۱۳۹۲).

نظام بازیابی مورد استفاده برای پایگاه‌های اطلاعاتی تصاویر، دارای پیچیدگی‌های خاص خود است و پژوهش‌های دربارۀ ساختار، عملکرد و شیوه‌های ذخیره و بازیابی پایگاه و موتور کاوش تصاویر می‌تواند، افق‌های روشن‌تری پیش پای طراحان این نوع نظام‌های بازیابی قرار دهد. مطالعه در این زمینه نه تنها موجب کاهش ابهامات و چالش‌های موجود در بازیابی تصاویر می‌شود، بلکه سهولت بازیابی و دسترسی کاربران نیازمند به تصاویر را تضمین کرده و کیفیت و ربط تصاویر به‌دست آمده از موتورکاوش را بهبود و افزایش می‌دهد.

بنابراین، به عنوان زیر مجموعه‌ای از نظام ذخیره و بازیابی، جامعیت و مانعیت در بازیابی تعداد زیادی تصویر از یک پایگاه تصویر مسئله مهمی محسوب می‌شود. از زاویه دید دیگر، تصاویر در عین اینکه غنی هستند، می‌توانند در آینده برای پژوهشگران

¹ Animation

مفید باشند. برای مثال، مجموعه عکس‌هایی از چشم‌انداز یک خیابان یک سده پیش ممکن است برای مورخان مفید باشد. اما این تصاویر هم اغلب حاوی اطلاعات آلوده‌ای هستند.

تصاویر به عنوان یکی از انواع اطلاعات موجود در شبکه جهانی وب، همانند سایر انواع اطلاعات مانند متن، صوت، ویدئو و غیره روزبه‌روز در حال افزایش است. روزانه میلیون‌ها تصویر در پایگاه‌های موتورهای کاوش و شبکه‌های اینترنتی قرار می‌گیرد و به خاطر حجیم بودن آن‌ها مشکل دسترسی به‌وجود می‌آید که لازم است به شکلی سازمان‌دهی شوند تا دسترسی برای همه با سهولت و در سریع‌ترین زمان ممکن محقق شود. امروزه نمایه‌سازی به عنوان یکی از روش‌های مطلوب سازمان‌دهی اطلاعات، علاوه بر شکل سنتی آن یعنی نمایه‌های انتهایی کتاب، کارایی خود را برای اطلاعات چندرسانه‌ای موجود در پایگاه‌های اطلاعاتی موتورهای کاوش اینترنتی در شبکه اینترنت نیز نشان داده است.

نمایه‌سازی در پایگاه‌های اطلاعاتی موتورهای کاوش تصاویر به منظور ارائه پرس‌وجوی لازم برای تصاویر انجام می‌شود. هر تصویری مشخصه منحصر به فردی دارد؛ بنابراین نمایه‌سازی تصاویر می‌تواند از طریق مقایسه مشخصات آن‌ها انجام شود که از خود تصاویر استخراج می‌شوند.

به جز موتورهای کاوش تصویری پایگاه‌های اطلاعاتی خاص موجود در وب جهانی، بینگ^۱، گوگل^۲ و ای‌اوال^۳ موتورهای کاوش عمومی تصاویر موجود در آن هستند که عملکرد بازیابی آنها چندان مشخص نیست. به این دلیل، هدف این پژوهش تعیین، ارزیابی یا شناسایی عملکرد موتورهای کاوش مذکور بود اما پژوهشگر در حین پژوهش متوجه شد که موتور کاوش ای‌اوال از قابلیت جستجوی تصاویر به روش محتوایی برخوردار نیست. یعنی نه تنها امکان جستجوی تصویر به تصویر ندارد بلکه در قسمت محدودسازی جستجو نیز امکان مشخص کردن ویژگی‌های محتوایی تصویر شامل رنگ، بافت و شکل وجود ندارد؛ به همین دلیل این موتور کاوش در مرحله اجرا از ارزیابی حذف و دو موتور کاوش گوگل و بینگ مقایسه شد که از این امکان بهره‌مند بودند. حال با توجه به مطالب اشاره شده، پژوهشگر قصد دارد به صورت تطبیقی، میزان جامعیت و مانعیت این دو موتور کاوش تصویر را مورد بررسی قرار دهد.

پرسش‌های پژوهش

۱. میزان جامعیت موتور کاوش بینگ در بازیابی تصاویر چقدر است؟
۲. میزان مانعیت موتور کاوش بینگ در بازیابی تصاویر چقدر است؟
۳. میزان جامعیت موتور کاوش گوگل ایمج در بازیابی تصاویر چقدر است؟
۴. میزان مانعیت موتور کاوش گوگل ایمج در بازیابی تصاویر چقدر است؟
۵. کدام یک از موتورهای کاوش، از جامعیت و مانعیت بیشتری نسبت به موتور کاوش دیگر برخوردار است؟

فرضیه‌های پژوهش

۱. تفاوت معناداری بین جامعیت بازیابی تصاویر در موتورهای کاوش بینگ و گوگل ایمج وجود دارد.
۲. تفاوت معناداری بین مانعیت بازیابی تصاویر در موتورهای کاوش بینگ و گوگل ایمج وجود دارد.

¹ Bing

² Google

³ American Online (AOL)

روش پژوهش

پژوهش حاضر از نظر ماهیت کاربردی است و به لحاظ هدف و گردآوری داده‌ها از روش‌های پژوهش وب سنجی و تطبیقی استفاده کرده است. یافته‌های پژوهش از طریق کاوش در موتورهای جستجوی مورد نظر "بینگ" و "گوگل" و مشاهده مستقیم منابع بازیابی شده گردآوری شده است. لازم به ذکر است که قرار بر مقایسه بازیابی تصویر در سه موتور کاوش به روش محتوایی بود. اما پژوهشگر بنا به دلایل مذکور در مقدمه، فقط به مطالعه و مقایسه دو موتور کاوش گوگل و بینگ پرداخت که از این امکان بهره مند بودند. لازم به یادآوری است که هر چند قبلاً درباره موتور کاوش گوگل مطالعه‌ای صورت گرفته است اما تاکنون از لحاظ بازیابی تصاویر هیچ پژوهش مقایسه‌ای بین آن و موتور کاوش پرمرآجه و پر استفاده بینگ انجام نشده است.

جامعه پژوهش را تصاویر ذخیره شده موجود در پایگاه‌های تصویر دو موتور کاوش بینگ و گوگل ایمج تشکیل می‌دهد. بدین منظور ۱۵ گروه تصویری با موضوع گروه داعش^۱ در رابطه با جریانات سیاسی آنها و صحنه‌های مختلف مربوط به این گروه، مانند هلاکت داعشیان، جهاد نکاح، تجاوز به زنان، چهره‌ی سرکرده‌ی داعش، سلاح‌های داعش، کشتار کودکان، پرچم داعش، مبارزان داعش، آوارگان سوری، اسکناس‌های داعش، خوردن گوشت حیوانات وحشی توسط داعش، مانورهای داعش، انفجارها، زنان داعشی و صحنه‌هایی از جنایت‌های این گروه انتخاب و از هر گروه یک عکس به صورت تصادفی انتخاب و بر روی هر دو موتور کاوش به روش جستجوی تصویر به تصویر (جستجو با نمونه) انجام شد. بنابراین، نمونه پژوهش متشکل از ۱۵ مورد جستجو در هر یک از موتورهای کاوش مورد مطالعه با تصاویر منتخب بود. انتخاب تعداد ۱۵ مورد جستجو به دلیل جلوگیری از حجیم‌تر شدن پروژه و میزان عملی بودن اجرای آن در زمان مقرر پژوهش بوده است. هدف اجرایی پژوهش مقایسه صرف تصاویر نبود بلکه می‌بایست تعداد معتابه و قابل توجهی از نتایج بازیابی از نظر ربط مورد بررسی قرار گیرند و افزایش بیش از حد تعداد تصاویر و در نتیجه نتایج بازیابی می‌توانست افزایش زمان پژوهش و احتمال اشتباه در بررسی دقیق میزان ربط را در پی داشته باشد. از این رو، همه نتایج بازیابی شده بدون استفاده از نمونه‌گیری مطالعه و به صورت سرشماری مورد بررسی قرار گرفت. به هنگام بررسی ربط، نتایج بازیابی شده از هر تصویر در هر یک از موتورهای کاوش مورد شمارش قرار گرفت و نتایجی که از نظر معیارهای شکل، رنگ، بافت، موضوع و پیام تصویر با نمونه هم‌خوانی و ربط بیشتری را نشان می‌داد به عنوان نتایج مرتبط انتخاب شدند. بدین صورت نتایج دارای ربط از نتایج نامربوط توسط خود پژوهشگر تشخیص داده شد و فراوانی هر کدام تعیین گردید.

تجزیه و تحلیل داده‌های پژوهش

با استفاده از فرمول ربط در نتایج هر یک از تصاویر، جامعیت و مانعیت تعیین و محاسبه شد. ربط با دو معیار اصلی دقت (مانعیت) و بازیافت (جامعیت) سنجیده می‌شود و فرمول ربط به صورت زیر است (حریری و دیگران، ۱۳۹۲؛ پائو، ۱۳۷۹):

$$\text{دقت} = \frac{\text{تعداد مدارک بازیابی شده مربوط}}{\text{تعداد کل مدارک بازیابی شده از پایگاه}}$$

$$\text{بازیافت} = \frac{\text{تعداد مدارک بازیابی شده مربوط}}{\text{تعداد کل مدارک موجود در پایگاه}}$$

^۱ در انتخاب این موضوع بیش از آن که مسئله سیاسی مد نظر باشد مسئله انسانی، اهمیت اعمال غیرانسانی این گروه و مطرح بودن آن در اذهان کاربران اینترنتی در طول انجام پژوهش مطرح بوده است. اهمیت سیاسی و امنیتی آن در منطقه، به روز بودن موضوع و همچنین به تصویر کشیدن جنایت‌های این گروه نیز در انتخاب موضوع موثر بوده است.

در نتایج بازیابی شده، تصاویری که از نظر معیارهای مورد نظر با تصویر نمونه هم‌خوانی نداشت ریزش کاذب محسوب و از تصاویر مرتبط با تصویر نمونه جدا شدند و در نهایت میانگین جامعیت و مانعیت در هر موتور کاوش محاسبه شد. برای محاسبه جامعیت در نتایج بازیابی شده‌ی هر تصویر نمونه، تعداد تصاویر بازیابی شده‌ی مرتبط یک موتور کاوش بر مجموع تعداد تصاویر بازیابی شده مرتبط همه موتورهای کاوش مورد مطالعه تقسیم شد و برای محاسبه مانعیت، تعداد مدارک بازیابی شده مرتبط بر تعداد کل مدارک بازیابی شده مرتبط و نامرتب همان موتور کاوش تقسیم شد. برای رسم نمودارها و همچنین محاسبه میانگین جامعیت و مانعیت در هر یک از موتورهای کاوش از نسخه ۲۰۱۰ نرم افزار مایکروسافت اکسل استفاده گردید.

برای آزمون فرضیه‌ها، داده‌های مربوط به جامعیت و مانعیت هر یک از موتورهای کاوش، در نسخه ۲۰ نرم افزار آماری اسپاس وارد شد. با توجه به اینکه فرضیه‌های مطرح شده به مقایسه میانگین دو جامعه مستقل می‌پردازد و استفاده از آزمون پارامتری مقایسه میانگین (آزمون تی دو جامعه مستقل) به دلیل نرمال نبودن توزیع دو جامعه یا نمونه‌های کوچک امکان‌پذیر نیست، بنابراین، در تحلیل بررسی صحت و سقم فرضیه‌های تحقیق، بدلیل تعداد اندک نمونه (۱۵ در هر گروه)، آزمون یو مان ویتنی بر روی داده‌ها انجام گرفت.

پاسخ به پرسش‌های پژوهش

در این بخش به پرسش‌های پژوهش پاسخ داده می‌شود. بیان این نکته ضروری است که برآورد جامعیت مطلق یک موتور کاوش با حجم عظیمی از اطلاعات، خارج از توان هر پژوهشگری است بنابراین، متغیر جامعیت در این پژوهش به عنوان جامعیت نسبی شناخته می‌شود.

پرسش ۱. میزان جامعیت موتور کاوش بینگ در بازیابی تصاویر چقدر است؟

جدول ۱. میزان جامعیت موتور کاوش بینگ در هر تصویر و میانگین جامعیت نسبی

سوال	۱	۲	۳	۴	۵	۶	۷	۸	۹	۱۰	۱۱	۱۲	۱۳	۱۴	۱۵
شاخص آماری															
درصد جامعیت	۱	۱۱	۳۷	۲۰	۷	۱۷	۲۶	۱	۱۱	۷۹	۷	۱۸	۱۰	۳۲	۳۶
میانگین نسبی جامعیت (درصد)	۲۰/۸۶														

داده‌های مربوط به میزان درصد جامعیت موتور کاوش بینگ در تصاویر مورد جستجو و میانگین نسبی جامعیت آن در جدول ۱ تنظیم و ارائه شده است. با محاسبه میانگین درصدهای جامعیت بدست آمده از نتایج مربوط به جامعیت این موتور کاوش در ۱۵ تصویر، میانگین نسبی جامعیت موتور کاوش ۲۰/۸۶ درصد به دست آمد و به عنوان پاسخ این پرسش گزارش شد.

پرسش ۲. میزان مانعیت موتور کاوش بینگ در بازیابی تصاویر چقدر است؟

جدول ۲. میزان مانعیت موتور کاوش بینگ در هر تصویر و میانگین مانعیت کلی

سوال	۱	۲	۳	۴	۵	۶	۷	۸	۹	۱۰	۱۱	۱۲	۱۳	۱۴	۱۵
شاخص آماری															
درصد مانعیت	۱۰۰	۱۰۰	۹۳	۹۴	۱۰۰	۹۵	۱۰۰	۱۰۰	۹۴	۹۲	۸۷	۱۰۰	۱۰۰	۱۰۰	۹۸
میانگین مانعیت	۹۶/۸۶														

داده‌های مربوط به میزان درصد مانعیت موتور کاوش بینگ در تصاویر مورد جستجو و میانگین مانعیت کلی آن در جدول ۲ تنظیم و ارائه شده است. با محاسبه میانگین درصدهای مانعیت بدست آمده از نتایج مربوط به مانعیت این موتور کاوش در ۱۵ تصویر، میانگین مانعیت کلی ۹۶/۸۶ درصد در این موتور کاوش به دست آمد و به عنوان پاسخ این پرسش گزارش شد.

پرسش ۳. میزان جامعیت موتور کاوش گوگل در بازیابی تصاویر چقدر است؟

جدول ۳. میزان جامعیت موتور کاوش گوگل در هر تصویر و میانگین جامعیت نسبی

سوال	۱	۲	۳	۴	۵	۶	۷	۸	۹	۱۰	۱۱	۱۲	۱۳	۱۴	۱۵
شاخص آماری	۹۹	۸۵	۶۴	۸۷	۹۳	۷۷	۹۵	۹۸	۹۵	۹۰	۹۳	۹۶	۹۶	۸۰	۸۳
درصد جامعیت															
میانگین نسبی جامعیت (درصد)	۸۸/۷۳														

داده‌های مربوط به میزان جامعیت موتور کاوش گوگل در تصاویر مورد جستجو و میانگین نسبی جامعیت آن در جدول ۳ تنظیم و ارائه شده است. با محاسبه میانگین درصدهای جامعیت بدست آمده از نتایج مربوط به جامعیت این موتور کاوش در ۱۵ تصویر، میانگین نسبی جامعیت ۸۸/۷۳ درصد در این موتور کاوش به دست آمد و به عنوان پاسخ این پرسش گزارش شد.

پرسش ۴. میزان مانعیت موتور کاوش گوگل در بازیابی تصاویر چقدر است؟

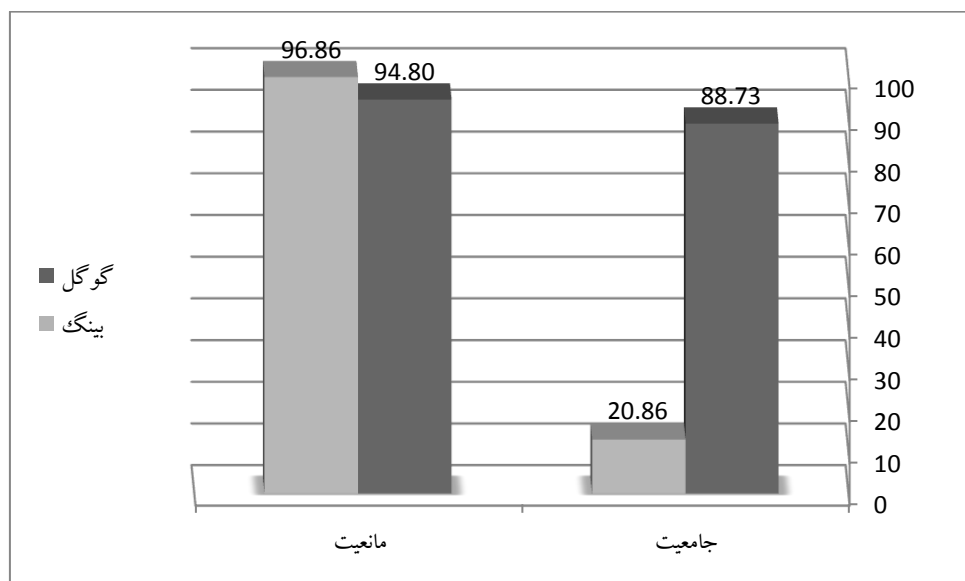
جدول ۴. میزان مانعیت موتور کاوش گوگل در هر تصویر و میانگین مانعیت کلی

سوال	۱	۲	۳	۴	۵	۶	۷	۸	۹	۱۰	۱۱	۱۲	۱۳	۱۴	۱۵
شاخص آماری	۹۲	۹۳	۹۲	۹۷	۹۸	۹۹	۹۵	۸۲	۹۴	۹۷	۹۶	۹۸	۹۶	۹۴	۹۹
درصد مانعیت															
میانگین مانعیت	۹۴/۸۰														

داده‌های مربوط به میزان مانعیت موتور کاوش گوگل در تصاویر مورد جستجو و میانگین مانعیت کلی آن در جدول ۴ تنظیم و ارائه شده است. با محاسبه میانگین درصدهای مانعیت بدست آمده از نتایج مربوط به مانعیت این موتور کاوش در ۱۵ تصویر، میانگین مانعیت کلی ۹۴/۸۰ درصد در این موتور کاوش به دست آمد و به عنوان پاسخ این پرسش گزارش شد.

پرسش ۵. کدام یک از موتورهای کاوش، از جامعیت و مانعیت بیشتری نسبت به موتور کاوش دیگر برخوردار است؟

به منظور پاسخ به این پرسش و امکان مقایسه، داده‌های به دست آمده از محاسبه درصد میانگین جامعیت و مانعیت دو موتور کاوش در نمودار ۱ و جدول ۵ ارائه شد.



نمودار ۱. میزان جامعیت و مانعیت موتورهای کاوش مورد مطالعه

درصد میانگین جامعیت نسبی و درصد میانگین مانعیت کلی موتورهای کاوش مورد مطالعه در نمودار ۱ تنظیم و ارائه شده است. میانگین جامعیت نسبی موتور کاوش گوگل $88/73$ درصد و موتور کاوش بینگ $20/86$ درصد بدست آمده است. میانگین مانعیت کلی موتور کاوش گوگل $94/80$ درصد و موتور کاوش بینگ $96/86$ درصد بدست آمده است.

جدول ۵. مقایسه میزان شاخص های جامعیت و مانعیت موتورهای کاوش مورد مطالعه

شاخص	جامعیت (درصد)	مانعیت (درصد)
موتورهای کاوش		
گوگل	$88/73$	$94/80$
بینگ	$20/86$	$96/86$

یافته های جدول ۵ نشان می دهد که موتور کاوش گوگل از بیشترین جامعیت نسبت به بینگ و موتور کاوش بینگ از بیشترین مانعیت نسبت به گوگل برخوردار است. اما اینکه وضعیت کدام یک در انتخاب موتور کاوش مورد استفاده توسط کاربران تاثیرگذار خواهد بود؛ در نتایج مربوط به فرضیه های پژوهش مشخص و بیان خواهد شد.

آزمون فرضیه های پژوهش

برای آزمودن فرضیه ها از نرم افزار آماری اسپاس اس، و آزمون یو مان ویتنی استفاده شده است. در اینجا به ارائه فرضیه های پژوهش بر اساس نتایج بدست آمده از آزمون آنها می پردازیم. پس از ورود داده ها به نسخه ۲۰ نرم افزار اسپاس اس، خروجی های به دست آمده در زیر سرفصل فرضیه مربوط به شرح زیر گزارش شد:

فرضیه ۱. تفاوت معناداری بین جامعیت بازیابی تصاویر در موتورهای کاوش مورد مطالعه وجود دارد.

به منظور آزمون فرضیه اول پژوهش درباره معنادار بودن تفاوت بین جامعیت بازیابی تصاویر در موتورهای کاوش مورد مطالعه فرض صفر و مخالف به صورت زیر در نظر گرفته شد:

فرض صفر: بین جامعیت بازیابی تصاویر در موتورهای کاوش مورد مطالعه تفاوت معناداری وجود ندارد.
فرض مخالف: بین جامعیت بازیابی تصاویر در موتورهای کاوش مورد مطالعه تفاوت معناداری وجود دارد.

جدول ۶. آمار توصیفی داده‌های مربوط به جامعیت

تعداد	میانگین	انحراف معیار	کمترین	بیشترین	درصدها	
					۲۵	۵۰ (متوسط)
۳۰	۵۴/۸۰۰۰	۳۷/۷۶۶۴۳	۱/۰۰	۹۹/۰۰	۱۵/۵۰۰۰	۷۵
					۷۰/۵۰۰۰	۹۳/۰۰۰۰

جدول ۷. نتایج آزمون یومان ویتنی برای فرضیه اول

متغیرها	تعداد	رتبه میانگین	مجموع رتبه‌ها	مقدار یومان ویتنی	مقدار z	مقدار معناداری
بینگ	۱۵	۸/۱۳	۱۲۲/۰۰			
گوگل	۱۵	۲۲/۸۷	۳۴۳/۰۰	۲/۰۰۰	۴/۵۸۶	۰/۰۰۱
جمع	۳۰					

در نتایج آزمون یومان ویتنی فرضیه پژوهش، سطح معناداری بدست آمده برابر با ۰/۰۰۱ و مقداری کوچک‌تر از مقدار سطح معناداری مفروض (۰/۰۵=آلفا) با ضریب اطمینان ۹۵ درصد بوده است. مقدار یومان ویتنی بدست آمده برابر ۲/۰۰۰ و مقدار Z برابر ۴/۵۸۶ است. با توجه به این که مقدار سطح معناداری بدست آمده ۰/۰۰۱ کوچک‌تر از مقدار مفروض (۰/۰۵=آلفا) است، پس فرض صفر رد و فرض مخالف تایید می‌شود و می‌توان نتیجه گرفت که تفاوت معناداری در میزان جامعیت موتورهای کاوش مورد مطالعه وجود دارد. با توجه به این که نمره جامعیت موتور کاوش گوگل برابر ۸۸/۷۳ درصد و موتور کاوش بینگ ۲۰/۸۶ درصد است، این تفاوت به نفع موتور کاوش گوگل است و در نتیجه رتبه بالایی را نسبت به موتور کاوش رقیب دارد. یعنی تفاوت معناداری بین جامعیت دو موتور کاوش مورد مطالعه وجود دارد و جامعیت موتور کاوش گوگل بسیار بالاتر از موتور کاوش بینگ است.

فرضیه ۲. تفاوت معناداری بین مانعیت بازیابی تصاویر در موتورهای کاوش مورد مطالعه وجود دارد.

به منظور آزمون فرضیه دوم پژوهش درباره معنادار بودن تفاوت بین مانعیت بازیابی تصاویر در موتورهای کاوش مورد مطالعه فرض صفر و مخالف به صورت زیر در نظر گرفته شد:

فرض صفر: بین مانعیت بازیابی تصاویر در موتورهای کاوش مورد مطالعه تفاوت معناداری وجود ندارد.
فرض مخالف: بین مانعیت بازیابی تصاویر در موتورهای کاوش مورد مطالعه تفاوت معناداری وجود دارد.

جدول ۸. آمار توصیفی داده‌های مربوط به مانعیت

تعداد	میانگین	انحراف معیار	کمترین	بیشترین	درصدها	
					۲۵	۵۰ (متوسط)
۳۰	۹۵/۸۰۰۰	۴/۲۲۹۶۲	۸۲/۰۰	۱۰۰/۰۰	۹۳/۷۵۰۰	۷۵
					۹۶/۰۰۰۰	۱۰۰/۰۰۰۰

جدول ۹. نتایج آزمون یومان ویتنی برای فرضیه دوم

متغیرها	تعداد	رتبه میانگین	مجموع رتبه‌ها	مقدار یومان ویتنی	مقدار z	مقدار معناداری
بینگ	۱۵	۱۸/۳۷	۲۷۵/۰۰			
گوگل	۱۵	۱۲/۶۷	۱۹۰/۰۰	۷۰/۰۰۰	۱/۷۸۵	۰/۰۷۴
جمع	۳۰					

در نتایج آزمون یومان ویتنی فرضیه پژوهش، سطح معناداری به دست آمده برابر با ۰/۰۷۴ و مقداری بزرگتر از مقدار سطح معناداری مفروض (۰/۰۵=آلفا) با ضریب اطمینان ۹۵ درصد بوده است. مقدار یومان ویتنی بدست آمده برابر ۷۰/۰۰۰ و مقدار Z برابر ۱/۷۸۵ است. با توجه به این که مقدار سطح معناداری بدست آمده ۰/۰۷۴ بزرگتر از مقدار مفروض (۰/۰۵=آلفا) است، پس فرض مخالف رد و فرض صفر پذیرفته می‌شود و می‌توان نتیجه گرفت که تفاوت معناداری در میزان مانعیت موتورهای کاوش مورد مطالعه وجود ندارد. به عبارت دیگر تفاوت معناداری در بین میزان مانعیت موتور کاوش گوگل (۹۴/۸۰) و مانعیت موتور کاوش بینگ (۹۶/۸۶) وجود ندارد.

نتیجه‌گیری

با ورود و نفوذ فناوری‌های اطلاعاتی به ویژه شبکه‌ی جهانی وب در عرصه‌های مختلف زندگی و افزایش اهمیت تصویر در حوزه‌های مختلف، لزوم پرداختن به تصاویر و نقش آنها بیش از پیش احساس می‌شود. اگر چه به نظر می‌رسد که حوزه‌ی تصاویر، موضوعی است که شاید در گذشته هم بیش از هر نوع اطلاعاتی مورد اهمیت بوده و هر تصویر نقش هزاران کلمه را بر عهده داشته است، اما در عین حال متوجه می‌شویم که موضوع ذخیره و بازیابی تصاویر مورد غفلت واقع شده و در سال‌های اخیر توجه متخصصان این حوزه را بیش از پیش به خود جلب کرده است. اکنون بیشتر معلوم شده است که تصویر در فرهنگ، آموزش، مسائل اجتماعی، پزشکی و حوزه‌های دیگر چه نقش موثر و پر اهمیتی ایفا می‌کند. در حال حاضر افراد در هر عرصه‌ای، چه در محیط‌های سنتی و چه در محیط‌های پیشرفته و نوین، بیش از هر نوع اطلاعات به دنبال تصاویر دلخواه خود هستند. افزایش حجم تصاویر لحظه به لحظه در شبکه جهانی وب مشکلات سازماندهی و دسترسی را پدید می‌آورد.

اکنون پرداختن به موضوع بازیابی اطلاعات هر لحظه احساس می‌شود. تا به حال موضوع بازیابی تصاویر، بیشتر در حد نظریه و مقدمات باقی مانده و به ندرت پژوهش‌هایی صورت گرفته که به صورت کاربردی به این مسئله پردازد. در پژوهش حاضر سعی بر آن بوده که قدمی هر چند کوچک در این حوزه برداشته شود.

بدیهی است که در هر پژوهش به بخش کوچکی از یک حوزه پرداخته می‌شود و مورد بررسی قرار می‌گیرد و یافته‌های یک پژوهش فقط می‌تواند پاسخگوی پرسش‌های همان پژوهش باشد و در حیطه و دامنه همان پژوهش تعمیم داده شود. در تحقیق حاضر نیز یافته‌هایی که در مورد بازیابی محتوایی تصاویر در دو موتور کاوش بینگ و گوگل گزارش شد؛ تنها بر بخش محتوایی تصاویر صادق است در حالی که ممکن است یافته‌های مذکور در بازیابی مفهومی تصاویر صدق نکند. با توجه به یافته‌های پژوهش درباره دو متغیر جامعیت و مانعیت می‌توان چنین نتیجه گرفت که میزان جامعیت دو موتور کاوش بینگ و گوگل به ترتیب ۲۰/۸۶ و ۸۸/۷۳ است و آزمون فرضیه اول مربوط به جامعیت، تفاوت معناداری را بین دو موتور کاوش نشان می‌دهد و این تفاوت به نفع موتور کاوش گوگل است. بنابراین، می‌توان نتیجه گرفت که با توجه به یافته‌ها، موتور کاوش گوگل برای کاربرانی که به دنبال جامعیت بیشتری هستند گزینه مناسبی است و عملکرد بهتری نسبت به موتور کاوش رقیب دارد به عبارتی دیگر، موتور کاوش بینگ برای کاربرانی که به دنبال جامعیت بیشتری هستند، گزینه‌ی مناسبی نیست. آزمون فرضیه دوم مربوط به مانعیت، نشان می‌دهد که میزان مانعیت دو موتور کاوش بینگ و گوگل به ترتیب ۹۶/۸۶ و ۹۴/۸۰ است و تفاوت معناداری بین دو موتور کاوش

وجود ندارد. لذا می‌توان استنتاج کرد که هر دو برای کاربرانی که به دنبال دقت بیشتر در بازیابی تصاویر هستند گزینه خوبی است اما با توجه به میزان بالای درصد مانعیت موتور کاوش بینگ نسبت به گوگل، این موتور کاوش (بینگ) برای کاربرانی که به دنبال مانعیت بیشتری در بازیابی هستند ارجح است. در بازیابی اطلاعات، مطلوب این است که میزان جامعیت و مانعیت هم سطح باشند اما معمولا در بازیابی اطلاعات از موتورهای کاوش و همچنین موتورهای کاوش مورد مطالعه هرچه میزان جامعیت بیشتر می‌شود مانعیت کمتر و هرچه میزان مانعیت بیشتر می‌شود جامعیت کمتر می‌شود. در نتایج پژوهش حاضر نیز چنین است؛ میزان جامعیت موتور کاوش گوگل بیشتر است در نتیجه مانعیت کمتری نسبت به موتور کاوش بینگ دارد و همچنین موتور کاوش بینگ میزان جامعیت کمتری دارد اما از مانعیت نسبتا بیشتری نسبت به موتور کاوش گوگل برخوردار است. امید است انجام پژوهش حاضر گام کوچکی برای شناخت وضع موجود در این حوزه‌ی موضوعی باشد و برای جامعه‌ی بزرگ علمی و پژوهشی کمک کوچکی ارائه کند و راه را برای پژوهش‌های آینده هموارتر نماید.

منابع و مآخذ:

- پائو، میراندا لی (۱۳۷۹). مفاهیم بازیابی اطلاعات. ترجمه اسدالله آزاد و رحمت الله فتاحی. مشهد: دانشگاه فردوسی مشهد.
- حریری، نجلا؛ باب الحوائجی، فهیمه؛ فرزندی پور، مهرداد؛ نادى راوندی، سمیه (۱۳۹۲). معیارهای ارزیابی ربط در نظام های بازیابی اطلاعات: دانسته ها و ندانسته ها. پژوهشنامه پردازش و مدیریت اطلاعات، ۳۰ (۱)، ۱۹۹-۲۲۱.
- کوکبی، مرتضی؛ رحیمی، صالح؛ عصاره، فریده؛ و نوروزی، علیرضا (۱۳۹۲). تأملاتی بر نمایه‌سازی تصاویر: یک تصویر ارزشی برابر هزار واژه. تحقیقات اطلاع‌رسانی و کتابخانه‌های عمومی، ۱۹ (۲)، ۲۵۷-۲۷۶.
- Hidderley, R., & Rafferty, P. (2005). *Indexing multimedia and creative works: The problems of meaning and interpretation*. New York: Ashgate Publishing Company.
- Tyner, Ross (2001). *Sink or swim: internet search tools & techniques*. Retrieved October 8, 2014, from <http://www.ouc.bc.ca/libr/connect96/search.html>.