

مقایسه میزان جامعیت و مانعیت در ابزارهای کاوش اطلاعات فازی و غیرفازی

فاطمه احمدی*

دانشجوی کارشناسی ارشد علم اطلاعات و دانش شناسی دانشگاه آزاد اسلامی واحد همدان

دکتر مظفر چشمه سهرابی

استادیار گروه علم اطلاعات و دانش شناسی دانشگاه اصفهان

تاریخ پذیرش: ۱۳۹۱/۶/۱۲

تاریخ دریافت: ۱۳۹۱/۳/۱۸

چکیده

هدف: پژوهش حاضر به مقایسه میزان جامعیت و مانعیت در ابزارهای جستجوی فازی و غیرفازی می‌پردازد. علاوه بر این، پژوهش قصد دارد قابلیت‌های موجود در سیستم‌های بازیابی اطلاعات فازی را شناسایی کند.

روش: روش تحقیق از نوع نیمه تجربی است. در این پژوهش سه ابزار جستجوی فازی (گوگل، آلتاویستا، و آسک) و سه ابزار جستجوی غیرفازی (یاهو، ماما، و متاکراولر) بررسی شد. برای تجزیه و تحلیل داده‌ها از روش‌های آمار توصیفی استفاده شد. برای توصیف کمی پارامترها از جداول فراوانی و آمارهای حداقل، حداکثر و میانگین استفاده شد و رسم نمودارها با استفاده از نرم افزار مایکروسافت اکسل ۲۰۰۳ صورت گرفت.

یافته‌ها: نتایج حاکی از آن است که برای ده عبارت جستجو در سه ابزار جستجوی فازی و سه ابزار جستجوی غیرفازی، بیشترین مدارک مرتبط بازیابی شده مربوط به سیستم‌های بازیابی فازی است. در میان سیستم‌های جستجوی فازی، گوگل بیشترین مدارک مرتبط را بازیابی کرده است و در میان ابزارهای جستجوی غیرفازی یاهو در صدر قرار دارد، همچنین ماما کمترین مدارک مرتبط را در میان ۶ ابزار جستجو بازیابی کرده است.

نتیجه گیری: سیستم‌های فازی در مقایسه با سیستم‌های غیرفازی از مانعیت بیشتری برخوردار هستند و در میان سه ابزار جستجوی فازی، آسک بیشترین مانعیت را دارد. در میان ابزارهای جستجوی غیرفازی یاهو در صدر قرار دارد. همچنین ماما کمترین میزان مانعیت را در میان ۶ ابزار جستجو دارد. و بیشترین میزان جامعیت مربوط به سیستم‌های جستجوی فازی است. در میان سیستم‌های جستجوی فازی گوگل بیشترین جامعیت را دارد. در میان ابزارهای جستجوی غیرفازی یاهو در صدر قرار دارد. همچنین ماما کمترین میزان جامعیت را در میان ۶ ابزار جستجو دارد. نتایج این تحقیق می‌تواند برای طراحان پایگاه‌های اطلاعاتی و کاربران مثبت و ارزشمند باشد.

کلیدواژه‌ها: ابزارهای جستجو، بازیابی اطلاعات، جامعیت، مانعیت، منطق فازی، منطق غیرفازی.

مقدمه و بیان مسأله

منطق فازی، یک جهان بینی جدید است که به رغم ریشه داشتن در فرهنگ مشرق زمین، با نیازهای دنیای پیچیده امروز بسیار سازگارتر از منطق ارسطویی است. منطق فازی جهان را آن طور که هست به تصویر می کشد. بدیهی است چون ذهن ما با منطق ارسطویی پرورش یافته، برای درک مفاهیم فازی در ابتدا باید کمی تأمل کنیم، ولی وقتی آن را شناختیم، دیگر نمی توانیم به سادگی آن را فراموش کنیم. دنیایی که ما در آن زندگی می کنیم، دنیای مبهمات و عدم قطعیت است. مغز انسان عادت کرده است که در چنین محیطی فکر کند و تصمیم بگیرد و این قابلیت مغز که می تواند با استفاده از داده های نادقیق و کیفی به یادگیری و نتیجه گیری بپردازد، در مقابل منطق ارسطویی که لازمه آن داده های دقیق و کمی است، قابل تأمل است. منطق فازی بر آن است که بر این مفاهیم غیردقیق و کیفی دلالت کند (کاسکو، ۱۳۸۰).

منطق یا تئوری فازی «یک نوع منطق است که روش های نتیجه گیری در مغز بشر را جایگزین می کند» (ویکی پدیا، ۲۰۰۶). منطق فازی اولین بار در پی تنظیم نظریه مجموعه های فازی به وسیله پروفیسور لطفی زاده (۱۹۶۵) در صحنه محاسبات نو ظاهر شد. کلمه فازی به معنای غیر دقیق، ناواضح و مبهم است. منطق فازی گونه ای بسیار مهم از منطق است و به طور جدی در مقابل منطق دودویی ارسطویی قرار گرفت. منطق فازی که در فرهنگ لغت «شرایط عدم قطعیت و نامعلوم» تعریف شده است، معتقد به ابهام در ماهیت علم است. لطفی زاده اینطور استدلال کرد که بشر به ورودیهای اطلاعاتی دقیق نیازی ندارد بلکه قادر است تا کنترل تطبیقی میان اطلاعات موجود انجام دهد. بنابراین، این منطق در ابتدا به عنوان روشی برای پردازش اطلاعات معرفی شد و برخلاف منطق کلاسیک یا ارسطویی است که تنها دو حالت برای موقعیتهای مختلف قایل است: سیاه و سفید؛ آری و نه؛ روشن و تاریک؛ صفر و یک؛ درست و غلط، حال آنکه قایلان به تفکر فازی معتقد به ابهام در ماهیت علم هستند؛ یعنی همان طور که این ابهام ها در ذهن بشر وجود داشته و بشر با درک و توجه به آنها در ذهن خود پدیده ها را تغییر و مدل سازی می کند، منطق فازی نیز سعی دارد مدلهایی ارائه دهد که ابهام را به عنوان بخشی از سیستم ارائه کند. قوانین علمی گذشته، مثل ریاضیات، فیزیک، و مکانیک نیوتونی، همه بر اساس همین منطق دو ارزشی استوار شده اند، اما بدیهی است که ذهن ما کارهایش را با منطق دیگری انجام می دهد و تصمیم هایش را می گیرد. با کمک منطق فازی می توان شیوه تفکر انسان را به فناوری منتقل کرد (فرخیان، ۲۰۰۶).

مجموعه های فازی برای بازیابی اطلاعات مفیدند، زیرا این مجموعه ها می توانند «موضوع» مدرک را توصیف کنند. به علاوه، از آنجا که در منطق فازی «زبان طبیعی» به جای متغیرهای عددی برای تشریح رفتار و عملکرد سیستم به کار می رود، می توان برای بازیابی اطلاعات در بانکهای اطلاعاتی، به نحو مؤثری از آن بهره جست. واژه «محاسبات تقریبی» بوسیله «لطفی زاده» ارائه شد و در همگرایی روش هایی برای حل مسائلی مفید است که نیاز به نوعی هوش (که از محاسبات کلاسیک ناشی می شود) دارند. «محاسبات تقریبی» یک مجموعه از فنون مناسب برای رفع ابهام، ذهن گرایی، و کلیت موجود در برخی مسائل است. هدف بازیابی اطلاعات مدل سازی، طراحی، و اجرای سیستم هایی است که قادر باشند دسترسی سریع و کارآمد بر پایه محتوا را به مقادیر عظیم اطلاعات، تأمین کنند. هدف یک سیستم بازیابی اطلاعات، برآورد ربط اقلام اطلاعاتی با نیازهای اطلاعاتی یک کاربر (که در قالب یک سؤال بیان شده) است. این کار مشکل و پیچیده ای است؛ زیرا با ذهنیت، ابهام و عدم دقت آکنده است. پردازش و سازماندهی اطلاعات محاسبات تقریبی روش های متفاوتی از قبیل منطق فازی، الگوریتم های ژنتیکی، شبکه های عصبی، مجموعه های نادقیق، و شبکه های بیزی را شامل می شود. مسئله بازیابی اطلاعات، یک حوزه کاربردی معمول برای محاسبات تقریبی است. بعضی از رویکردهای اصلی محاسبات تقریبی در بازیابی اطلاعات از این قرارند: منطق و مجموعه های فازی، ترکیب اطلاعات، استخراج متن، مدل های زبان پرس و جو، و خوشه بندی مدرک؛ شبکه های عصبی: رده بندی و خوشه بندی اسناد و اصطلاحات، و بازیابی چند رسانه ای ها؛ الگوریتم های ژنتیکی: رده بندی مدارک، بازیابی تصویر، بازخورد ربط، و یادگیری پرس و جو؛ تکنیک های احتمالاتی: رتبه بندی، وب کاوی مجموعه های نادقیق و منطق های چند ارزشی: خوشه بندی مدارک؛

شبکه‌های بیزی: مدل‌های بازیابی، رتبه‌بندی، ساخت اصطلاحنامه، و بازخورد ربط (Baez Yates, 2003).

چنانکه رکوردهای به ثبت رسیده در لیزا^۱ نیز نشان می‌دهد، کاربرد عمده مجموعه‌های فازی در کتابداری و اطلاع‌رسانی، در حوزه بازیابی اطلاعات است. اساساً مفهوم ربط به دلیل نامشخص و نسبی بودنش (فتاحی، ۱۳۸۳)، یک مفهوم فازی است. در جریان جستجو و یافتن اطلاعات به رکوردهایی می‌رسیم که نمی‌توان به طور قطع آنها را مرتبط یا کاملاً بی‌ربط با موضوع مورد جستجو تلقی کرد. بنابراین دو مفهوم ربط و فازی را می‌توان قرین و همزاد یکدیگر دانست (Wood & Wilson, 2002).

مکتبی فرد در مقاله خود با عنوان "مدیریت اطلاعات با رویکرد فازی" به این نتیجه رسید که با وجود ابزارها و محملهای جدید اطلاعاتی مانند اینترنت، روز به روز با عدم قطعیت و ابهام بیشتری مواجه می‌شویم. در حقیقت، با فراگیر شدن رسانه‌های اطلاعاتی، با مخاطبان عامی مواجه هستیم که کنترل آنها به راحتی ممکن نیست و وجود و ماهیت آنها برای خدمت دهندگان در حاله‌ای از ابهام قرار دارد. هرچند این کاربران ممکن است خواسته‌های مشخص و دقیقی داشته باشند، اما از آنجا که دانش آنها در سطوح متفاوت قرار دارد و به هیچ وجه قابل اندازه‌گیری نیست، ذخیره، سازماندهی و شیوه‌های بازیابی اطلاعات باید به گونه‌ای باشد که بتواند با در نظر گرفتن ابهامی که در سیستم وجود دارد و با کمک زبان طبیعی، پاسخگویی این خواسته‌ها باشد. زبان طبیعی از آنجا که به ساختار ذهنی انسان نزدیک‌تر بوده و توسط همه کاربران با هر سطح از توانایی و دانش قابل استفاده است، به‌ویژه در شرایطی که خدمات کتابخانه غیرحضوری است و امکان آموزش به استفاده‌کنندگان نیز وجود ندارد، در ایجاد پایگاه‌های اطلاعاتی و طراحی ابزارهای جستجو و ذخیره و بازیابی اطلاعات، بسیار مورد توجه قرار گرفته است و منطبق فازی به تحقق چنین نظامی بسیار کمک می‌کند. به نظر می‌رسد در حوزه فراهم‌آوری و رده‌بندی، منطبق فازی نوپا بوده و عرصه برای پژوهشهای بیشتر، خالی مانده است. اگرچه این دو حوزه به لحاظ ماهیت کار و با توجه به پژوهشهای انجام شده، تا حدودی از منطبق فازی دور هستند، اما به دلیل انعطاف زیادی که در نظریه مجموعه‌های فازی وجود دارد، این علم به باز شدن گره‌های موجود در این دو حوزه نیز کمک خواهد کرد (مکتبی فرد، ۱۳۸۷).

مجموعه‌های کتابخانه‌ها که امروزه کم کم از حالت سخت خارج شده و به سمت دنیای مجازی حرکت می‌کنند، بی‌شک بر بنیادهای فازی بهتر بنا شده و خدمات‌رسانی آنها نیز رضایت بخش‌تر خواهد بود. به‌ویژه در پاسخگویی به نیازهای کاربران، از آنجا که این نیازها ماهیتاً فازی بوده و از فردی به فرد دیگر ممکن است تغییر یابند، استفاده از منطبق فازی برای سامانه‌های اطلاعاتی این امکان را فراهم می‌سازد که با در نظر داشتن نقاط ابهام در درخواست‌های کاربران، دقیق‌ترین اطلاعات موجود در پایگاه را در اختیار آنان قرار دهند (مکتبی فرد، ۱۳۸۷). با توجه به نقش ارزنده‌ای که منطبق فازی می‌تواند در مدیریت اطلاعات داشته باشد، لزوم توجه به این علم در کتابداری و اطلاع‌رسانی، ضروری‌تر به نظر می‌رسد. با توجه به نو بودن موضوع و اینکه تا کنون تحقیق در این خصوص صورت نگرفته است، نتایج تحقیق می‌تواند برای طراحان پایگاه‌های اطلاعاتی و ابزارهای بازیابی اطلاعات و همچنین کاربران بسیار مثبت و ارزشمند باشد. لذا مسأله اساسی این تحقیق بررسی قابلیت‌های ابزارهای جستجوی فازی در بازیابی اطلاعات و در صدد پاسخگویی به سوالات زیر است:

۱. میزان مانعیت و جامعیت در ابزارهای جستجوی غیر فازی چقدر است؟
۲. میزان مانعیت و جامعیت در ابزارهای جستجوی فازی چقدر است؟
۳. آیا در بازیابی اطلاعات مرتبط ابزارهای جستجوی فازی موفق‌تر از ابزارهای جستجوی غیر فازی عمل می‌کنند؟
۴. آیا میزان مانعیت در ابزارهای جستجوی فازی بیشتر از ابزارهای جستجوی غیر فازی است؟
۵. آیا میزان جامعیت در ابزارهای جستجوی فازی بیشتر از ابزارهای جستجوی غیر فازی است؟

^۱ LISA

پیشینه پژوهش

منطق فازی در زمینه‌های متنوع و متفاوتی کاربرد دارد و تقریباً حدود ۱۰ سال پس از ابداع آن به متون کتابداری راه یافت. به گواهی بانک چکیده‌های مقاله‌های کتابداری و اطلاع‌رسانی (لیزا، ۲۰۰۶-۱۹۶۹). مفهوم فازی اولین بار با عبارت «تئوری مجموعه‌های فازی» و در مقاله‌ای تحت عنوان «جستجوی یک فایل در شرایط فازی» در دنیای کتابداری و اطلاع‌رسانی مطرح شد. در این مقاله که در سال ۱۹۷۵ و به زبان فرانسه منتشر شده بود، به کاربرد نظریه و مجموعه‌های فازی در بازیابی اطلاعات پرداخته شده بود. به جز مقاله‌ای که در سال ۱۹۷۶ در خصوص کاربردهای عملی مجموعه‌های فازی در حل مشکلات مربوط به سازماندهی اطلاعات با عنوان «مجموعه‌های پراکنده در نظریه رده‌بندی» به زبان روسی منتشر شد، چند سالی تقریباً تمامی پژوهشها و مقاله‌های انتشار یافته در خصوص منطق فازی در عرصه کتابداری و اطلاع‌رسانی، بر مباحث مرتبط بر بازیابی اطلاعات متمرکز بود، به نحوی که از مجموع ۳۲ مقاله‌ای که در فاصله سالهای ۱۹۷۵ تا ۱۹۹۰ با موضوع فازی منتشر شده، ۲۵ مورد آن به نقش و کاربردهای منطق فازی در ذخیره و بازیابی اطلاعات و راهبردهای جستجو و... اختصاص داشته است. داورپناه در مقاله خود با عنوان "ضرورت‌های نوین بازنگری در ذخیره و بازیابی اطلاعات" بیان می‌کند انتقال از منطق دو ارزشی به چند ارزشی، شیوه تفکر ما را تغییر داده و افق‌های جدیدی را پیش روی ما گشوده است. زمانی که ارزیابی پارادایم‌ها (یک شیوه دیدن جهان) در دستگاه منطقی چند ارزشی صورت گیرد، درست یا غلط بودن هر پارادایم، امری نسبی می‌شود. در نظام سنتی چون همه اعضا باید تمام ویژگی‌های اصلی را که در تعریف آن دسته بیان می‌شود داشته باشند، بنابراین تمام اعضا به طور یکسان نمایانگر دسته یا گروه یا طبقه هستند و هیچ عضوی بر دیگری برتری ندارد (داورپناه، ۱۳۸۴).

در سیستم‌های فازی، مجموعه قوانین^۱ به سیستم امکان انتخاب با عدم قطعیت می‌دهد. در چنین سیستمی قوانین در ابتدا استخراج می‌شود و سپس در طول کار سیستم با توجه به بازخوردی که از کاربر گرفته می‌شود، تغییر خواهد کرد. سیستم بازیابی اطلاعات یک ابزار محاسباتی است که اطلاعات را به شکلی پیاده سازی می‌کند که بعداً بتواند بطور خودکار بازیابی شوند. سیستم‌های بازیابی اطلاعات غالباً، فقط اطلاعات متنی را ذخیره و بازیابی می‌کنند. مطالعات اخیر در زمینه هوشمند سازی بازیابی اطلاعات، به این نتیجه رسیده است که برای بهبود کارایی سیستم بازیابی اطلاعات، احتیاج به تکنیک‌هایی است که محتوای درخواستها و مدارک را مشخص کنند. اخیراً محققان تئوری اطلاعات سعی بر این داشتند که رابطه میان مدرک و درخواستها را مشخص کنند. هدف این است که درخواست کاربر را در مجموعه محلی مدارک پیاده سازی کن (دستغیب، ۱۳۸۴، ص. ۲۵).

دستغیب در مقاله خود با عنوان «مروری بر نمایه سازی متنی، محتوایی و فازی تصاویر رقمی» چنین استنباط می‌کند که توجه به افزایش روزافزون اطلاعات چند رسانه‌ای در وب و اهمیت بازیابی تصاویر در علوم مختلف، کارایی و دقت نظامهای بازیابی تصاویر مهم است. در میان روش‌های مختلف نمایه سازی و بازیابی اطلاعات، نظامهای فازی به دلیل تطبیق با عدم قطعیت ذاتی بازیابی اطلاعات، مقایسه، محاسبه شباهت و ربط را بسیار بهتر از نظامهای قطعی انجام می‌دهند (دستغیب، ۱۳۸۶).

کوان در مقاله خود با عنوان "هستی شناسی بازیابی فازی برای کتابخانه‌های دیجیتالی" آورده‌اند که در سالهای اخیر کتابخانه‌های دیجیتالی دارای پیشرفتهای زیادی در زمینه بازیابی اطلاعات بوده بعضی از کتابخانه‌ها همچنین خود را با تکنیک‌های بازیابی فازی وفق داده‌اند ایشان برای گسترش کتابخانه‌های دیجیتالی دو پیشنهاد داده‌اند اول، استفاده از هستی شناسی به منظور نشان دادن اطلاعات پژوهشگرانه بطوری کتابخانه‌های دیجیتالی قادر به اشتراک گذاری در محیط وب معنایی باشند. دوم، تئوری فازی به منظور اشاعه اطلاعات پژوهشی مبهم و غیر مشخص به اشکال مختلف فازی و سوال فازی (Quan et al., 2007).

ژینگ و احمد در مقاله خود آورده‌اند که: شکل گیری یک سیستم رده بندی به میزان زیادی وابسته به تکنیک خوشه بندی است. اگر خوشه بندی تحت تأثیر، شامل تصمیم گیری سخت (منطقی) باشد، یک تصویر می‌تواند تنها جزئی از یک خوشه باشد.

^۱. Rule base

در این موقعیت‌ها، اگر تصویر خواسته شده متعلق به خوشه ویژه‌ای باشد که در برگزیده تصویر نباشد، ممکن است هرگز کسی قادر به یافتن آن نشود حتی اگر در پایگاه اطلاعاتی وجود داشته باشد. اما اگر از خوشه بندی فازی استفاده شود، تصویر یکی یا خوشه‌های بیشتری را مدنظر قرار می‌دهد. این امر ممکن است احتمال یافتن تصاویر مربوط را افزایش دهد (Xing & Ahmad, 2008).

در منطق فازی روابط مبتنی بر اصول ازپیش تعیین شده نیست، بلکه بر پایه دانش کلی یا بافت است. در نظام فازی چون طیف وجود دارد، می‌توان اعضای طبقه را رتبه بندی کرد. بر همین مبنا ساراسویک (در: پائو، ۱۳۷۹، ص. ۱۱۲) معتقد است در نظام ذخیره و بازیابی اطلاعات، ربط مفهوم و هویتی دو ارزشی نیست. ربط بر مقیاسی مستمر و پیوسته مبتنی است و نباید مفهومی را مربوط یا نامربوط شمرد. در پاسخ به هر پرسش، هر مدرک موجود در بایگانی (فایل) ممکن است احتمالاً مطالبی مربوط را به دست دهد. مدارک بایگانی را می‌توان بر حسب احتمال مربوط بودن، به ترتیب نزولی مرتب کرد. ویژگیهای مهم برای ابزارهای جستجوی مبتنی بر منطق فازی که از متون تخصصی در این زمینه استخراج شده به اختصار در جدول زیر آمده است.

جدول ۱. خصوصیات فازی در ابزارهای جستجوی مبتنی بر منطق فازی

ربط	خصوصیات فازی
انعطاف پذیری	
زبان طبیعی	
درجه عضویت نسبی	
توانایی رویارویی با داده‌های که ممکن است در آنها عدم قطعیت و ناکامل بودن دیده شود	
اصلاح نتایج کاوش	
رتبه بندی رکوردهای بازیابی شده بر اساس احتمال مرتبط بودن با نیاز اطلاعاتی کاربر	
الگوریتم ریشه یابی	
توانایی مدل سازی سیستم‌های خبره چندگانه با تضمین امنیت حفظ اطلاعات	
ذخیره اطلاعات با استفاده از فهرست راهنمای تعریف شده	
الگوریتم کشف لغات	
الگوریتم شناسایی علاقه مندی کاربران در کلمات جستجو شده الگوریتم غلط یابی	
عدم نمایه مجدد صفحاتی که اطلاعات آنها تغییری نداشته، به دلیل داشتن ماژول هوشمند	
الگوریتم رتبه بندی	

روش شناسی تحقیق

این تحقیق با استفاده از روش نیمه تجربی انجام شد. جامعه آماری این پژوهش شامل کلیه ابزارهای جستجوی تحت وب (ابزارها، ابرابزارها، و راهنماهای موضوعی) است که از میان آنها با اخذ نظر متخصصان سه ابزار جستجوی مبتنی بر فازی و سه ابزار جستجوی غیر فازی معروف انتخاب شدند. در این پژوهش، روش جستجوی پیشرفته مبنا قرار گرفت. هر عبارت جستجو در ابزارهای جستجوی منتخب وارد و نتایج نسبتاً زیادی بازیابی شد اما تنها سی رکورد اول به منظور محدود کردن مطالعه مورد بررسی قرار گرفت. عبارت‌ها شناسایی و به منظور تعیین میزان جامعیت و مانعیت در ابزارهای جستجو که از منطق فازی استفاده کرده‌اند و آنهایی که استفاده نکرده‌اند مورد جستجو و بررسی قرار می‌گیرد که دارای سه مرحله است:

در مرحله اول اطلاعات مرتبط موجود به شکل چاپی و الکترونیکی برای مطالعه جمع آوری شد. در مرحله دوم پس از بررسی ابزارهای مختلف جستجو، با متخصصان حوزه کامپیوتر و علم اطلاعات و دانش شناسی مشورت شد و در نهایت برای جستجوی فازی ابزارهای گوگل، آلتاویستا، و آسک و برای جستجوی غیرفازی ابزارهای یاهو، ماما، و متاکراولر انتخاب شدند.

در مرحله سوم، بر اساس ساختار موضوعات در سرعنوان های موضوعی فارسی و به منظور جامع بودن موضوعات انتخابی و رعایت جوانب مختلف کار، سعی شد شصت عبارت جستجو از میان عناوین پایان نامه های دفاع شده فارسی در رشته های مختلف انتخاب شود. سپس ده عبارت که حوزه های مختلف دانش را پوشش می دادند و از نظر شکل موضوع و سیستم ارجاعی برای جستجو مناسب تشخیص داده شدند، انتخاب گردیدند. اضافه بر این، پس از انجام هر جستجو، به منظور تعیین میزان ربط موارد بازیابی شده با عبارت های جستجو، با کمک متخصصان مربوطه میزان ارتباط آنها مشخص شد.

سنجش میزان مانعیت و جامعیت

مانعیت کسر خروجی یک جستجو است که محاسبه آن نیاز به دانستن نتایج مرتبط و نامرتب در یک سری مدارک ارزیابی شده دارد (Clarke & Willet, 1997). بنابراین امکان محاسبه دقیق مانعیت ابزارهای جستجو به آسانی امکان پذیر است. در این تحقیق منظور از مانعیت توانایی نظام در کنار گذاشتن مدارک نامرتب است. به عبارتی دیگر، مانعیت نسبت اسناد بازیابی شده ای است که واقعاً مربوط هستند. بنابراین مانعیت به این صورت تعریف می شود:

$$\text{نسبت مانعیت} = \frac{\text{تعداد مدارک مرتبط بازیابی شده}}{\text{کل مدارک بازیابی شده ابزار کاوش}} \times 100$$

از سوی دیگر جامعیت، توانایی بازیابی یک سیستم برای کسب تمام یا بیشتر مدارک مرتبط در مجموعه است. بنابراین (اندازه گیری آن) به دانستن نه فقط مدارک مرتبط و بازیابی شده بلکه مدارک بازیابی نشده، نیز نیاز دارد (Clarke & Willet, 1997). روش صحیحی برای اندازه گیری دقیق جامعیت ابزارهای جستجو وجود ندارد. زیرا امکان دانستن تعداد کل مدارک مرتبط در پایگاه های اطلاعاتی عظیم وجود ندارد اما کلارک و ولیت سنجش جامعیت سنتی را برای استفاده در محیط وب از طریق ارائه یک ویژگی نسبی به آن، تنظیم کرده اند. در این پژوهش روش به کار گرفته شده توسط کلارک و ولیت را از طریق ادغام کردن نتایج مرتبط برای ابزارهای جستجوی فازی و غیر فازی جهت تشکیل مخرج کسر محاسبه شد. بنابراین مقدار جامعیت نسبی به این صورت تعریف می شود:

$$\text{جامعیت نسبی} = \frac{\text{تعداد مدارک مرتبط بازیابی شده توسط یک ابزار کاوش}}{\text{تعداد کل مدارک مرتبط بازیابی شده توسط شش ابزار کاوش}} \times 100$$

یافته های پژوهش

تجزیه و تحلیل داده های پژوهشی براساس سوالات پژوهش به شرح زیر است:

سوال اول پژوهش: میزان مانعیت و جامعیت در ابزارهای جستجوی غیر فازی چقدر است؟

جدول ۲. میزان مانعیت جامعیت در راهنمای موضوعی یاهو

ردیف	بارت	تعداد مدارک بازنمایی شده با حذف تکراریها	تعداد کل مدارک بازنمایی شده مرتبط	مانعیت (درصد)	جامعیت (درصد)
۱	بررسی وضعیت صادرات زعفران در ایران	۴۸۳	۲۶۸	۵۵	۱۷
۲	نقش فناوری نانو در بسته بندی فرآوردهای غذایی	۶۸۵	۴۵۵	۶۶	۲۴
۳	اخلاق زیست محیطی	۴۷۶	۳۷۱	۷۷	۱۹
۴	فعالیت آنتی اکسیدانی انار	۹۳	۴۹	۵۲	۱۸
۵	مصارف درمانی آلونه‌ورا	۸۹	۶۴	۷۲	۱۷
۶	گروه‌های خونی و نقش آنتی ژن و آنتی بادی در انتقال خون	۲۰۴	۱۳۴	۶۶	۱۹
۷	بیماریهای التهابی شبکیه	۳۲۹	۲۵۹	۷۹	۱۷
۸	سلیمان نبی در آینه شعر فارسی	۱۶۹	۴۵	۲۷	۱۵
۹	بررسی حقوقی و جرم شناسی پول شویی	۱۱۶	۸۵	۷۴	۱۲
۱۰	بررسی وضعیت طب سوزنی در جهان	۴۷۲	۳۶۲	۷۷	۱۴
	جمع کل	۳۱۱۶	۲۰۹۲	۶۷	۱۸

همانطور که از جدول ۲ قابل مشاهده است، در ستون اول ده عبارت انتخاب شده که در راهنمای موضوعی یاهو مورد جستجو قرار گرفته، آمده است که بطور مجزا برای هر عبارت تعداد کل مدارک بازنمایی شده را با حذف تکراری، تعداد کل مدارک بازنمایی شده مرتبط در دو ستون بعدی قرار گرفته و در دو ستون آخر میزان مانعیت و جامعیت برای هر عبارت جستجو شده آورده شده و در ردیف آخر جمع کل مانعیت و جامعیت را برای هر ده عبارت راهنمای موضوعی یاهو محاسبه شده است. یافته‌های جدول حاکی از این است که در میان ده عبارت مورد جستجو عبارت ردیف ۷ با ۷۹٪ از بیشترین میزان مانعیت و عبارت ردیف ۸ با ۲۷٪ از کمترین میزان مانعیت برخوردار است. و برای جامعیت عبارت ردیف ۲ با ۲۴٪ از بیشترین میزان جامعیت و عبارت ردیف ۹ با ۱۲٪ از کمترین میزان جامعیت برخوردارند.

جدول ۳. میزان مانعیت و جامعیت ابر ابزار جستجوی ماما^۱

ردیف	عبارت	تعداد مدارک بازنمایی شده با حذف تکراریها	تعداد کل مدارک بازنمایی شده مرتبط	مانعیت (درصد)	جامعیت (درصد)
۱	بررسی وضعیت صادرات زعفران در ایران	۱۳۵	۵۷	۴۲	۴
۲	نقش فناوری نانو در بسته بندی فرآوردهای غذایی	۱۱۸	۸۰	۶۷	۴
۳	اخلاق زیست محیطی	۱۱۴	۹۵	۸۳	۴
۴	فعالیت آنتی اکسیدانی انار	۰	۰	۰	۰
۵	مصارف درمانی آلونه‌ورا	۲۴	۴	۱۶	۴
۶	گروه‌های خونی و نقش آنتی ژن و آنتی بادی در انتقال خون	۵۲	۲۰	۳۸	۵
۷	بیماریهای التهابی شبکیه	۷۲	۲۷	۳۷	۳
۸	سلیمان نبی در آینه شعر فارسی	۶۰	۹	۱۵	۵
۹	بررسی حقوقی و جرم شناسی پول شویی	۶۰	۲۷	۴۵	۵
۱۰	بررسی وضعیت طب سوزنی در جهان	۳۹۹	۱۹۹	۴۹	۱۱
	جمع کل	۱۰۳۴	۵۱۸	۵۰	۵

همانطور که از جدول ۳ قابل مشاهده است در ستون اول ده عبارت انتخاب شده که در ابر ابزار جستجوی ماما مورد جستجو قرار گرفته شده آمده است که بطور مجزا برای هر عبارت تعداد کل مدارک بازنمایی شده را با حذف تکراریها، تعداد کل مدارک بازنمایی شده مرتبط در دو ستون بعدی قرار گرفته و در دو ستون آخر میزان مانعیت و جامعیت برای هر عبارت جستجو شده در ماما

^۱ . Mamma

محاسبه شده است. و یافته‌های جدول نشان می‌دهد که عبارت ردیف ۴ با میزان ۰٪ کمترین میزان مانعیت و عبارت ردیف ۳ با میزان ۸۳٪ بیشترین میزان مانعیت را در میان ده عبارت مورد جستجو دارند. کمترین میزان جامعیت مربوط به عبارت ردیف ۴ با مقدار ۰٪ و بیشترین میزان جامعیت مربوط به عبارت ردیف ۱۰ با میزان ۱۱٪ است

جدول ۴. مانعیت و جامعیت ابزار جستجوی متاکراولر^۱

ردیف	عبارت	تعداد مدارک بازنمایی شده با حذف تکراری‌ها	تعداد کل مدارک بازنمایی شده مرتبط	مانعیت (درصد)	جامعیت (درصد)
۱	بررسی وضعیت صادرات زعفران در ایران	۴۲۶	۲۵۶	۶۰	۱۵
۲	نقش فناوری نانو در بسته بندی فرآورده‌های غذایی	۳۴۸	۲۸۹	۸۳	۱۲
۳	اخلاق زیست محیطی	۳۴۴	۲۵۹	۷۵	۱۴
۴	فعالیت آنتی اکسیدانی انار	۸۶	۶۱	۷۰	۱۶
۵	مصارف درمانی آلونه‌ورا	۸۵	۷۲	۸۵	۱۶
۶	گروه‌های خونی و نقش آنتی ژن و آنتی بادی در انتقال خون	۱۵۷	۹۷	۶۲	۱۵
۷	بیماریهای التهابی شبکه	۳۳۴	۱۴۹	۴۵	۱۸
۸	سلیمان نبی در آینه شعر فارسی	۱۸۰	۴۲	۲۳	۱۶
۹	بررسی حقوقی و جرم شناسی پول شویی	۱۷۶	۸۶	۴۹	۱۸
۱۰	بررسی وضعیت طب سوزنی در جهان	۵۰۸	۳۵۰	۶۹	۱۵
	جمع کل	۲۶۳۹	۱۶۶۱	۶۲	۱۵

همان‌طور که از جدول ۴ قابل مشاهده است در ستون اول ده عبارت انتخاب شده که در ابر ابزار جستجوی متاکراولر مورد جستجو قرار گرفته، آمده است که بطور مجزا برای هر عبارت تعداد کل مدارک بازنمایی شده را با حذف تکراریها، تعداد کل مدارک بازنمایی شده مرتبط در دو ستون بعدی قرار گرفته و در دو ستون آخر میزان مانعیت و جامعیت برای هر عبارت جستجو شده آورده شده و در ردیف آخر جمع کل مانعیت و جامعیت را برای هر ده عبارت در ابر ابزار متاکراولر محاسبه شده است. یافته‌های جدول نشان می‌دهد که عبارت ردیف ۵ با ۸۵٪ بیشترین میزان مانعیت و عبارت ردیف ۸ با ۲۳٪ کمترین میزان مانعیت را در میان ده عبارت مورد جستجو به خود اختصاص داده‌اند، عبارتهای ردیف ۷ و ۹ با میزان ۱۸٪ بیشترین میزان جامعیت و عبارت ردیف ۲ با ۱۲٪ کمترین میزان جامعیت را در میان ده عبارت مورد جستجو دارند.

جدول ۵. متوسط میزان مانعیت در ابزارهای جستجوی غیر فازی

ردیف	موتور جستجو	مانعیت	مانعیت کل
۱	متاکراولر	۶۲	۶۷
۲	ماما	۵۰	
۳	یاهو	۶۷	

همچنان که از یافته‌های جدول ۵ برمی آید برای سه ابزار جستجوی غیرفازی به طور مجزا درصد مانعیت و در ستون آخر مانعیت کل برای هر سه ابزار جستجوی غیرفازی محاسبه شده است. یافته‌ها حاکی از این است که در میان ابزار جستجوی غیرفازی یاهو با ۶۷٪ از بیشترین میزان مانعیت و ماما با ۵۰٪ از کمترین میزان مانعیت برخوردار است.

جدول ۶. میزان جامعیت در ابزارهای جستجوی غیر فازی

ردیف	موتور جستجو	جامعیت (درصد)	جامعیت کل (درصد)
۱	متاکراولر	۱۵	۱۲
۲	ماما	۵	
۳	یاهو	۱۸	

^۱. Metacrawler

همچنان که از یافته‌های جدول ۶ بر می‌آید برای سه ابزار جستجوی غیرفازی به طور مجزا درصد جامعیت و در ستون آخر جامعیت کل برای هر سه ابزار جستجوی غیرفازی محاسبه شده است. یافته‌ها حاکی از این است که در میان سه ابزار جستجوی غیر فازی یا هو با ۱۸٪ از بیشترین میزان جامعیت و ماما ۵٪ از کمترین میزان جامعیت برخوردار است.

سوال دوم پژوهش: میزان مانعیت و جامعیت در ابزارهای جستجوی فازی چقدر است؟

جدول ۷. میزان مانعیت و جامعیت در ابزار جستجوی گوگل

ردیف	عبارت	تعداد مدارک بازیابی شده با حذف تکراری‌ها	تعداد کل مدارک بازیابی شده مرتبط	مانعیت (درصد)	جامعیت (درصد)
۱	بررسی وضعیت صادرات زعفران در ایران	۷۹۳	۵۲۲	۶۶	۲۸
۲	نقش فناوری نانو در بسته بندی فراورده‌های غذایی	۶۹۲	۵۵۹	۸۰	۲۴
۳	اخلاق زیست محیطی	۷۲۰	۴۷۷	۶۶	۲۹
۴	فعالیت آنتی اکسیدانی انار	۱۵۵	۱۲۹	۸۳	۳۰
۵	مصارف درمانی آلوه‌ورا	۱۲۷	۱۱۰	۸۶	۲۵
۶	گروه‌های خونی و نقش آنتی ژن و آنتی بادی در انتقال خون	۳۰۸	۲۲۴	۷۳	۳۰
۷	بیماریهای التهابی شبکیه	۵۰۶	۳۷۲	۷۴	۲۶
۸	سلیمان نبی در آینه شعر فارسی	۳۵۵	۲۱۰	۶۰	۳۲
۹	بررسی حقوقی و جرم شناسی پول شویی	۳۲۲	۳۰۲	۹۴	۳۳
۱۰	بررسی وضعیت طب سوزنی در جهان	۸۸۹	۵۲۵	۶۰	۲۶
	جمع کل	۴۸۱۳	۳۴۳۰	۷۱	۲۸

همانطور که از جدول ۷ قابل مشاهده است در ستون اول ده عبارت انتخاب شده که در موتور جستجوی گوگل مورد جستجو قرار گرفته، آمده است که بطور مجزا برای هر عبارت تعداد کل مدارک بازیابی شده را با حذف تکراریها، تعداد کل مدارک بازیابی شده مرتبط در دو ستون بعدی قرار گرفته و در دو ستون آخر میزان مانعیت و جامعیت برای هر عبارت جستجو شده محاسبه شده است. با توجه به یافته‌های جدول بیشترین میزان مانعیت را عبارت ردیف ۹ با میزان ۹۴٪ و عبارت‌های ردیف ۸ و ۱۰ با میزان ۶۰٪ کمترین درصد مانعیت را به خود اختصاص داده‌اند، عبارت ردیف ۹ با ۳۳٪ بیشترین میزان جامعیت برخوردار است و کلید عبارت ردیف ۲ با ۲۴٪ کمترین میزان مانعیت را در میان ده عبارت مورد جستجو دارند.

همانطور که از جدول ۸ قابل مشاهده است در ستون اول ده عبارت انتخاب شده که در موتور جستجوی آلتا ویستا مورد جستجو قرار گرفته، آمده است که بطور مجزا برای هر عبارت تعداد کل مدارک بازیابی شده را با حذف تکراریها، تعداد کل مدارک بازیابی شده مرتبط در دو ستون بعدی قرار گرفته و در دو ستون آخر میزان مانعیت و جامعیت برای هر عبارت جستجو شده آورده شده و در ردیف آخر جمع کل مانعیت و جامعیت را برای هر ده عبارت در موتور جستجوی آلتا ویستا محاسبه شده است. در این موتور جستجو عبارت ردیف ۱۰ با میزان (۸۲٪) از بیشترین میزان مانعیت برخوردار است و عبارت ردیف ۸ با میزان (۴۰٪) از کمترین میزان مانعیت را در میان ده عبارت مورد جستجو دارند، بیشترین میزان جامعیت متعلق به بازیابی شده ردیف ۲ با ۲۳٪ و عبارت ردیف ۹ با ۱۲٪ کمترین میزان جامعیت در میان ده عبارت مورد جستجو دارد.

جدول ۸. میزان مانعیت و جامعیت در موتور جستجوی آلتاویستا^۱

ردیف	کلید واژه	تعداد مدارک بازایی شده با حذف تکراریها	تعداد کل مدارک بازایی شده مرتباً	مانعیت (درصد)	جامعیت (درصد)
۱	بررسی وضعیت صادرات زعفران در ایران	۴۸۸	۲۷۲	۵۶	۱۷
۲	نقش فناوری نانو در بسته بندی فرآوردهای غذایی	۶۸۸	۴۸۶	۷۱	۲۳
۳	اخلاق زیست محیطی	۴۷۶	۳۷۱	۷۸	۱۹
۴	فعالیت آنتی اکسیدانی انار	۹۳	۴۹	۵۳	۱۸
۵	مصارف درمانی آلونه‌ورا	۹۰	۶۳	۷۰	۱۷
۶	گروه‌های خونی و نقش آنتی ژن و آنتی بادی در انتقال خون	۲۰۸	۱۳۴	۶۴	۲۰
۷	بیماریهای التهابی شبکیه	۳۶۳	۲۵۶	۷۶	۱۹
۸	سلیمان نبی در آینه شعر فارسی	۱۶۹	۶۸	۴۰	۱۵
۹	بررسی حقوقی و جرم شناسی پول شویی	۱۲۲	۸۹	۷۳	۱۲
۱۰	بررسی وضعیت طب سوزنی در جهان	۴۴۲	۳۶۲	۸۲	۱۳
	جمع کل	۳۱۳۹	۲۱۵۰	۶۸	۱۸

جدول ۹. میزان مانعیت و جامعیت برای موتور جستجوی آسک^۲

ردیف	عبارت	تعداد مدارک بازایی شده با حذف تکراریها	تعداد کل مدارک بازایی شده مرتباً	مانعیت (درصد)	جامعیت (درصد)
۱	بررسی وضعیت صادرات زعفران در ایران	۴۵۰	۳۱۸	۷۰	
۲	نقش فناوری نانو در بسته بندی فرآوردهای غذایی	۳۵۳	۲۹۲	۸۲	
۳	اخلاق زیست محیطی	۳۵۲	۳۰۵	۸۷	
۴	فعالیت آنتی اکسیدانی انار	۸۶	۷۴	۸۷	
۵	مصارف درمانی آلونه‌ورا	۸۹	۸۸	۹۹	
۶	گروه‌های خونی و نقش آنتی ژن و آنتی بادی در انتقال خون	۹۵	۷۵	۷۹	
۷	بیماریهای التهابی شبکیه	۲۸۱	۲۵۵	۹۱	
۸	سلیمان نبی در آینه شعر فارسی	۱۵۵	۴۵	۲۹	
۹	بررسی حقوقی و جرم شناسی پول شویی	۱۷۸	۱۳۹	۷۸	
۱۰	بررسی وضعیت طب سوزنی در جهان	۶۲۹	۴۶۵	۷۴	
	جمع کل	۲۵۵۸	۲۱۶۶	۸۵	

همان‌طور که از جدول ۹ قابل مشاهده است در ستون اول ده عبارت انتخاب شده که در موتور جستجوی آسک مورد جستجو قرار گرفته، آمده است که به طور مجزا برای هر عبارت تعداد کل مدارک بازایی شده را با حذف تکراریها، تعداد کل مدارک بازایی شده مرتبط در دو ستون بعدی قرار گرفته و در دو ستون آخر میزان مانعیت و جامعیت برای هر عبارت جستجو شده محاسبه شده است. یافته‌ها حاکی از این است که عبارت ردیف ۵ با میزان ۹۹٪ بیشترین میزان مانعیت و عبارت ردیف ۸ با میزان ۲۹٪ کمترین میزان مانعیت را در میان ده عبارت مورد جستجو به خود اختصاص داده‌اند و عبارت ردیف ۱۰ با ۱۹٪ بیشترین میزان جامعیت و عبارت ردیف ۶ با ۹٪ کمترین میزان جامعیت را در میان ده عبارت مورد جستجو دارند.

^۱. AltaVista

^۲. Ask

جدول ۱۰. میزان مانعیت در ابزارهای جستجوی مبتنی بر منطق فازی

ردیف	موتور جستجو	مانعیت (درصد)	مانعیت کل (درصد)
۱	گوگل	۷۱	۷۵
۲	آلتاویستا	۶۸	
۳	آسک	۸۵	

همان‌طور که در جدول ۱۰ قابل مشاهده است برای سه ابزار جستجوی مبتنی بر منطق فازی به طور مجزا درصد مانعیت و در ستون آخر مانعیت کل برای هر سه ابزار جستجوی فازی محاسبه شده است. یافته‌های جدول نشان می‌دهد که آسک با (۸۵٪) در صدر قرار دارد و آلتاویستا با (۶۸٪) کمترین میزان مانعیت را در میان سه ابزار جستجوی فازی به خود اختصاص داده است

جدول ۱۱. میزان جامعیت در ابزارهای جستجوی مبتنی بر منطق فازی

ردیف	موتور جستجو	جامعیت (درصد)	جامعیت کل (درصد)
۱	گوگل	۲۸	۲۰٪
۲	آلتاویستا	۱۸	
۳	آسک	۱۵	

همان‌طور در جدول ۱۱ قابل مشاهده است برای سه ابزار جستجوی مبتنی بر منطق فازی به طور مجزا درصد مانعیت محاسبه شده است و در ستون آخر مانعیت کل برای هر سه ابزار جستجوی فازی محاسبه شده است. که یافته‌های جدول نشان می‌دهد که گوگل با (۲۸٪) در صدر قرار دارد و آسک با (۱۵٪) کمترین میزان مانعیت را در میان سه ابزار جستجوی فازی به خود اختصاص داده است.

سوال سوم پژوهش: آیا در بازیابی اطلاعات مرتبط ابزارهای جستجوی فازی موفق‌تر از ابزارهای جستجوی غیرفازی عمل

می‌کنند؟

جدول ۱۲. مدارک بازیابی شده مرتبط توسط سیستم‌های جستجوی فازی و غیرفازی

سیستم‌های بازیابی فازی	ابزار جستجو	جمع کل مدارک بازیابی شده با حذف تکراری‌ها	جمع کل مدارک بازیابی شده مرتبط
سیستم‌های بازیابی فازی	گوگل	۴۸۱۳	۳۴۳۰
	آلتاویستا	۳۱۳۹	۲۱۵۰
	آسک	۲۵۵۸	۲۱۶۶
سیستم‌های بازیابی غیرفازی	ياهو	۳۱۱۶	۲۰۹۲
	ماما	۱۰۳۴	۵۱۸
	متاکراولر	۲۶۳۹	۱۶۶۱

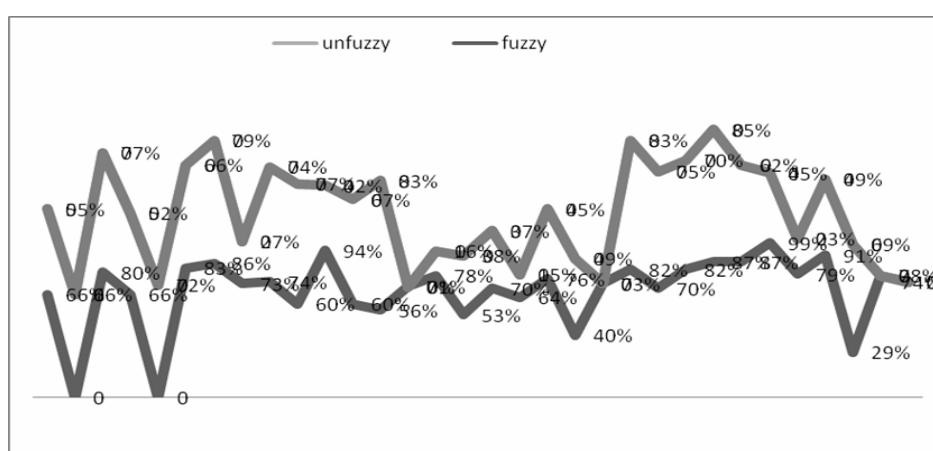
همچنان که از داده‌های جدول ۱۲ برمی‌آید، بیشترین مدارک بازیابی شده مربوط به سیستم‌های جستجوی فازی است. در میان سیستم‌های جستجوی فازی گوگل بیشترین مدارک مرتبط را بازیابی کرده است. در میان ابزارهای جستجوی غیرفازی یاهو در صدر قرار دارد. هم‌چنین ماما کمترین مدارک مرتبط را در میان ۶ ابزار جستجو بازیابی کرده است.

سوال چهارم پژوهش: آیا میزان مانعیت در ابزارهای جستجوی فازی بیشتر از ابزارهای جستجوی غیر فازی است؟

جدول ۱۳. میزان مانعیت در سیستم‌های جستجوی فازی و غیرفازی

مانعیت (درصد)	جمع کل مدارک بازایی شده مرتبط	ابزار جستجو	سیستم‌های بازایی فازی
۷۱	۳۴۳۰	گوگل	سیستم‌های بازایی فازی
۶۸	۲۱۵۰	آلتاویستا	
۸۵	۲۱۶۶	آسک	
۶۷	۲۰۹۲	ياهو	سیستم‌های بازایی غیرفازی
۵۰	۵۱۸	ماما	
۶۲	۱۶۶۱	متاکراولر	

همچنان که از داده‌های جدول ۱۳ برمی‌آید، بیشترین میزان مانعیت مربوط به سیستم‌های جستجوی فازی است. در میان سیستم‌های جستجوی فازی، آسک بیشترین مانعیت را دارد. در میان ابزارهای جستجوی غیرفازی، یاهو در صدر قرار دارد. همچنین ماما کمترین میزان مانعیت را در میان ۶ ابزار جستجو دارد.



نمودار ۱. مقایسه مانعیت سیستم‌های فازی با غیرفازی برای هر عبارت

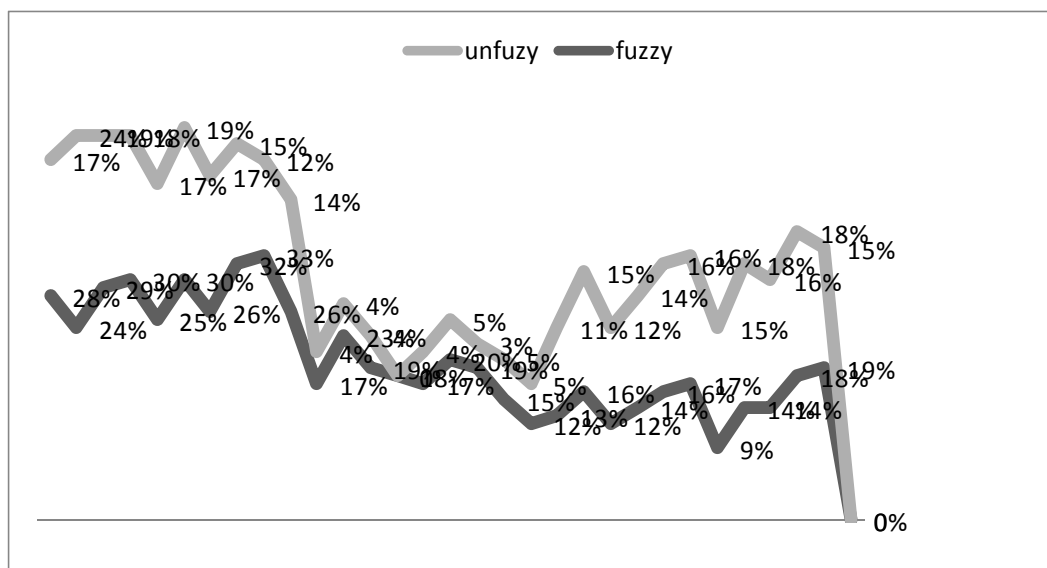
همان‌طور که در نمودار ۱ قابل مشاهده است میزان مانعیت برای هر ده عبارت مورد جستجو در سه ابزار جستجوی فازی در مقایسه با سه ابزار جستجوی غیرفازی دارای مانعیت بیشتری بوده است. این خود حاکی از آن است که مدل فازی می‌تواند باعث افزایش مانعیت در سیستم‌های بازایی اطلاعات باشد.

سوال پنجم پژوهش: آیا میزان جامعیت در ابزارهای جستجوی فازی بیشتر از ابزارهای جستجوی غیرفازی است؟

جدول ۱۴. میزان جامعیت در سیستم‌های جستجوی فازی و غیرفازی

جامعیت (درصد)	جمع کل مدارک بازایی شده مرتبط	ابزار جستجو	سیستم‌های بازایی فازی
۲۸	۳۴۳۰	گوگل	سیستم‌های بازایی فازی
۱۸	۲۱۵۰	آلتاویستا	
۱۵	۲۱۶۶	آسک	
۱۸	۲۰۹۲	ياهو	سیستم‌های بازایی غیرفازی
۵	۵۱۸	ماما	
۱۲	۱۶۶۱	متاکراولر	

همچنان که از داده‌های جدول ۱۴ برمی‌آید، بیشترین میزان جامعیت مربوط به سیستم‌های جستجوی فازی است. در میان ابزارهای جستجوی فازی گوگل بیشترین جامعیت را دارد. در میان ابزارهای جستجوی غیرفازی یاهو در صدر قرار دارد. همچنین ماما کمترین میزان جامعیت را در میان ۶ ابزار جستجو دارد.



نمودار ۲. مقایسه جامعیت سیستم‌های فازی با غیر فازی برای هر عبارت

همانطور که در نمودار ۲ قابل مشاهده است میزان جامعیت در هر ده عبارت مورد جستجو برای سه ابزار جستجوی فازی در مقایسه با سه ابزار جستجوی غیر فازی دارای جامعیت بیشتری بوده است. این خود حاکی از آن است که مدل فازی می‌تواند باعث افزایش جامعیت در ابزارهای بازیابی اطلاعات باشد.

نتیجه‌گیری

این پژوهش حاوی پنج سوال در مورد ابزارهای جستجوی فازی و غیر فازی است.

تحلیل یافته‌های مربوط به سوال اول پژوهش مبنی بر اینکه "میزان مانعیت و جامعیت در ابزارهای جستجوی غیر فازی چقدر است؟" یافته‌ها حاکی از این است که در میان سه ابزار جستجوی غیر فازی، یا هو با ۶۷٪ از بیشترین میزان مانعیت و اما ۵۰٪ از کمترین میزان مانعیت برخوردار است. برای جامعیت سه ابزار جستجوی غیر فازی نتایج چنین بود که در میان سه ابزار جستجوی غیر فازی یا هو با ۱۸٪ از بیشترین میزان جامعیت و اما ۵٪ از کمترین میزان جامعیت برخوردار بود.

تحلیل یافته‌های مربوط به سوال دوم پژوهش مبنی بر اینکه "میزان مانعیت و جامعیت در ابزارهای جستجوی فازی چقدر است؟" یافته‌ها حاکی از این است که آسک با ۸۵٪ در صدر قرار دارد و آلتاویستا با ۶۸٪ کمترین میزان مانعیت را در میان سه ابزار جستجوی فازی به خود اختصاص داده است. و برای جامعیت گوگل با ۲۸٪ در صدر قرار دارد و آسک با ۱۵٪ کمترین میزان مانعیت را در میان سه ابزار جستجوی فازی به خود اختصاص داده است. تحلیل‌های مربوط به سوال سوم پژوهش مبنی بر اینکه "آیا در بازیابی اطلاعات مرتبط ابزارهای جستجوی فازی موفق تر از ابزارهای جستجوی غیر فازی عمل می‌کنند؟" همان‌طور که در جدول ۱۱ نشان دادیم بیشترین مدارک مرتبط بازیابی شده مربوط به سیستم‌های بازیابی فازی است. در میان ابزارهای جستجوی فازی گوگل بیشترین مدارک مرتبط را بازیابی کرده است. و در میان ابزارهای جستجوی غیر فازی یا هو در صدر قرار دارد. همچنین اما کمترین مدارک مرتبط را در میان ۶ ابزار جستجوی بازیابی کرده است.

تحلیل یافته‌های مربوط به سوال چهارم پژوهش مبنی بر اینکه "آیا میزان مانعیت در ابزارهای جستجوی فازی بیشتر از ابزارهای جستجوی غیر فازی است؟" نتایج حاصل شده در جدول ۱۲ حاکی از این بود که بیشترین میزان مانعیت مربوط به ابزارهای جستجوی فازی است. در میان ابزارهای جستجوی فازی آسک با ۸۵٪ بیشترین مانعیت را دارد. در میان ابزارهای جستجوی غیر فازی یا هو با ۶۷٪ در صدر قرار دارد. هم‌چنین اما با ۵۰٪ کمترین میزان مانعیت را در میان ۶ ابزار جستجو دارد.

تحلیل یافته‌های مربوط به سوال پنجم پژوهش مبنی بر اینکه "آیا میزان جامعیت در ابزارهای جستجوی فازی بیشتر از ابزارهای جستجوی غیرفازی است؟" همان‌طور که در جدول ۱۳ نشان دادیم بیشترین میزان جامعیت مربوط به ابزارهای جستجوی فازی است. در میان سیستم‌های جستجوی فازی، گوگل با ۲۸٪ بیشترین جامعیت را دارد. در میان ابزارهای جستجوی غیرفازی یا هو با ۱۸٪ در صدر قرار دارد. هم‌چنین ماما با ۵٪ کمترین میزان جامعیت را در میان ۶ ابزار جستجو دارد.

پیشنهاد‌های کاربردی

حوزه بازیابی اطلاعات باید بیشتر بر روی نیازهای اطلاعاتی کاربران کار کرده و چالش عمده بین متخصصان اطلاع‌رسانی، این نکته است که آیا باید سیستم‌های اطلاعاتی را برای بازیابی بهینه اطلاعات و مدارک، منعطف کرد و یا برای رسیدن به این منظور یعنی حداکثر منابع مرتبط، بر روی کاربران سرمایه گذاری کرد. البته که نظام‌های مبتنی بر منطق فازی و نظام‌های با فناوری‌های بهتر برای انجام جستجو تاثیرات شگرفی بر بهینه‌شدن نتایج بازیابی خواهند داشت.

با توجه به یافته‌های تحقیق در خصوص «میزان بازیابی اطلاعات مرتبط در ابزارهای جستجوی فازی و غیرفازی» پیشنهاد می‌شود:

الف. طراحان و تولیدکنندگان ابزارهای کاوش، از قابلیت‌های جستجوی فازی در ابزارهای کاوش استفاده کنند. زیرا استفاده از قابلیت‌های جستجوی فازی می‌تواند به کاربران کمک کند که به اطلاعات مرتبط‌تری دست یابند.

ب. کاربران بهتر است به منظور دستیابی به اطلاعات مرتبط تر از ابزارهای کاوش فازی استفاده کنند.

منابع

- پائو، میراندالی (۱۳۷۹). مفاهیم بازیابی اطلاعات. (اسدالله آزاد و رحمت‌الله فتاحی، مترجم). مشهد: انتشارات دانشگاه فردوسی مشهد.
- داورپناه، محمدرضا (۱۳۸۴). ضرورت‌های نوین در ذخیره و بازیابی اطلاعات. کتابداری و اطلاع‌رسانی، ۳(۳)، ۶۷-۸۸.
- دستغیب، محمدرضا (۱۳۸۶). مروری برنمایه‌سازی متنی، محتوایی و فازی تصاویر رقمی: نظام بازیابی تصاویر رقمی. بازیابی شده از: www.olumensani.com
- دستغیب، محمدرضا (۱۳۸۲). بهینه‌سازی درخواست کاربر مبتنی بر هوشمندسازی بازیابی اطلاعات به‌وسیله شبکه عصبی. اطلاع‌رسانی، ۱۹(۲)، ۲۴-۲۹.
- فرخیان، سمیه (۲۰۰۶). مقدمه‌ای بر منطق فازی. بازیابی شده از: <http://www.vojoudi.com/uncertainty/fuzzy-logic/farokhian.htm>
- کاسکو، بارت (۱۳۸۰). تفکر فازی. (علی غفاری، عادل مقصودپور، علیرضا پورممتاز، جمشید قسیمی، مترجم). تهران: انتشارات دانشگاه صنعتی خواجه نصیرالدین طوسی.
- منطق فازی. ویکی‌پدیا: دانشنامه آزاد(آنلاین) قابل دسترس در: <http://www.fa.wikipedia.org/wiki>
- مکتبی‌فرد، لیل (۱۳۸۷). مدیریت اطلاعات با رویکرد فازی. کتابداری و اطلاع‌رسانی، ۱۱(۲)، ۱۸۳-۲۰۴.
- Baeza-Yates, Ricardo (2003). Information retrieval in the web: Beyond current search engines *International Journal of Approximate Reasoning*, 34, 97-104.
- Clarke,s & Willett (1997). Estimatiing the recall performance of search service performance. *Aslib Proceedings*, 49 (7), 184 -189.
- Hood, William W., Wilson, Concepcion (2002). Solving Problems in Library and Information Science Using Fuzzy Set Theory. *Library Trends*, 50 (3), 393-405.
- Quan , T. T., Hui, S. C, Fong, A. C. M., Cao, T. H. (2007). Ontology-based fuzzy retrieval for digital library. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 4822 LNCS , 95-98
- Xing, Nan & Ahmad, I. S. (2008). Fuzzy Clustering Paradigm and the Shape-Based Image Retrieval. In: *Proceedings of the Twenty-First International FLAIRS Coference* , 1-2.

