

بررسی فرآیند اطلاعاتیابی کاربران در...

نظامهای بازیابی اطلاعات خودکار

علی گزنی¹

چکیده

در سیستمهای جهانی بازیابی اطلاعات، بهره‌مندی از مهارت جستجو و استفاده از اطلاعات یکی از اجزای مهم زندگی می‌باشد. اما علی‌رغم اهمیت جستجوی اطلاعات، بسیاری از جستجوگران به شیوه آزمایش و خطا عمل کرده و از روش خاص و ساخت یافته‌ای جهت جستجوی اطلاعات استفاده نمی‌کنند. جستجو با این سبک یا نتیجه‌ای دربر نداشته یا تعداد مدارک زیاد و نامرتب را مورد بازیابی قرار می‌دهد. بر پایه مطالعات انجام شده مقاله حاضر سعی در ارائه یک دید نظام‌یافته از فرآیند بازیابی اطلاعات دارد که آن به جستجوگران کمک خواهد کرد تا بتوانند به برنامه‌ریزی و انجام جستجوهای موفق در راستای برطرف کردن نیازهای اطلاعاتی خود و استفاده از این سیستمها بپردازند. استفاده از این دید در آموزش جستجوگران و طراحی سیستمهای بازیابی اطلاعات بسیار حائز اهمیت می‌باشد.

واژه‌های کلیدی: فرآیند بازیابی اطلاعات - نظامهای بازیابی اطلاعات خودکار - اطلاع‌یابی کاربران - جستجوی اطلاعات

مقدمه

گسترش سریع شبکه جهانی وب، فنآوری اطلاعات، خدمات و نظامهای اطلاعاتی پیوسته بسط افق فکری بشر را به همراه داشته است. مطابق آمارهای اعلام شده از جانب

¹ عضو هیئت علمی کتابخانه منطقه‌ای علوم و تکنولوژی شیراز.

شرکت سیوی لنس^۱ تا تاریخ ۱۰ جولای سال ۲۰۰۰ معادل ۲/۱ میلیارد صفحه اطلاعاتی در وب شمارش شده است و هر روزه حدود ۷ میلیون صفحه اطلاعاتی جدید به آن افزوده می‌شود. همچنین بنابر پیش‌بینی این شرکت تا اوایل سال ۲۰۰۱ تعداد این صفحات به ۴ میلیارد صفحه اطلاعاتی خواهد رسید. طبق بررسی‌های انجام شده حدود ۸۳٪ این اطلاعات تجاری، ۶٪ علمی و آموزشی و ۱۱٪ را سایر زمینه‌ها تشکیل می‌دهند.

سرویس دهندگان اطلاعاتی بر روی وب با استفاده از نرم‌افزارهای خود موسوم به خزندگان وب^۲ سراسر این شبکه تار عنکبوتی را مورد کاوش قرار داده، به جمع‌آوری اطلاعات می‌پردازند. این اطلاعات پس از جمع‌آوری، سازماندهی شده و به وسیله موتورهای کاوش و فهرستهای راهنما در دسترس عموم قرار می‌گیرد (گزنی، ۱۳۸۰).

باید توجه داشت که در حال حاضر رایج‌ترین الگوهای مورد استفاده در نظامهای بازیابی اطلاعاتی الگوهای تطبیقی^۳ هستند، امروزه غالب پژوهشگران در مورد محدودیت این الگوها اتفاق نظر دارند (سالتون، ۱۹۹۲^۴)، مسیر حرکت در این الگوها یک‌سویه و بر روی خود نظام و نه کاربران متمرکز می‌گردد. بنابراین بازیابی اطلاعات با استفاده از ابزارهای موجود نیازمند دانش و مهارت است. در عصر حاضر برخورداری از دانش استفاده از این اطلاعات یکی از اجزای مهم زندگی می‌باشد. اما علی‌رغم اهمیت جستجوی

1. Cyveillance
2. Web Crawler

— بلکین و کرافت (۱۳۸۷) الگوهای تطبیقی (Match Paradiagm) را بازیابی و نمره‌گذاری مدارک بازیابی شده با استفاده از فنون مختلفی همانند جستجوی بولی، محاسبه احتمالات، مدل‌های برداری، کوه‌سازی و ژله‌ها، شبکه‌های عصبی و ... معرفی می‌کنند.

❖ در مورد محدودیت الگوهای تطبیقی سالتون (۱۹۹۲) می‌گوید:

الف) فرمول جستجو ممکن است واقعاً بازگوکننده نیاز اطلاعاتی کاربر نباشد.
ب) روشهایی که جهت متمایز کردن اطلاعات مرتبط از نامرتبط به کار می‌روند کامل نیستند.

اطلاعات در عصر حاضر بسیاری از جستجوگران به شیوه آزمایش و خطا عمل کرده و از روش خاص و ساخت یافته‌ای جهت جستجو استفاده نمی‌کنند. جستجو با این سبک یا نتیجه‌ای دربر نداشته یا تعداد مدارک زیاد و نامرتب را مورد ارزیابی قرار می‌دهد. این حالت می‌تواند برای کلیه کاربران و هر نوع اطلاعاتی در سطح دنیا مصداق داشته باشد. همچنین عدم شناخت فرآیند جستجوی اطلاعات کاربران، طراحی غیربهینه نظام‌های بازیابی اطلاعات را نیز به همراه خواهد داشت.

هدف مقاله

جستجو از مجموعه‌ای اجزای مرتبط با یکدیگر تشکیل شده است. نتیجه نهایی جستجو می‌تواند متأثر از نحوه انجام هر یک از این اجزا باشد. هدف مقاله حاضر بررسی اجزای جستجو به منظور ارائه یک دید نظام یافته از جستجو برای کاربران و طراحان نظام‌های بازیابی اطلاعات می‌باشد. ارائه این اجزا به صورت یک مدل طبقه‌بندی شده علاوه بر اینکه کاربران را در برنامه‌ریزی و انجام جستجوهای موفق یاری می‌کند، برای طراحان نظام‌های بازیابی اطلاعات نیز این فرصت را فراهم می‌آورد تا بتوانند نظام‌های تعاملی بهتر و بهینه‌تری را با درک این فرآیند طراحی کنند.

پیشینه

تحقیقات اخیر نشان داد جستجوی کاربران در یک دوره زمانی برای حل یک مسئله دارای یک چرخه یا توالی می‌باشد. مدل ویلسون¹ (1999) نشان داد که رفتار اطلاع‌یابی یک رفتار جهت‌دار است که هدف نهایی آن حل مسئله می‌باشد. اسپینک و ساراسیویس² (1997) عناصر مهم در جستجوی یک مرحله‌ای شامل انواع بازخوردها³ و راهبرد انتخاب مؤثر کلمات را مشخص کردند.

مطالعات هوانگ⁴ (1992)، بیولیو و روبرتسون⁵ (1992) نیز نشان داد جستجوی کاربران دارای یک توالی می‌باشد. مطالعات اسپینک (1996) نشان داد، توالی در بازیابی

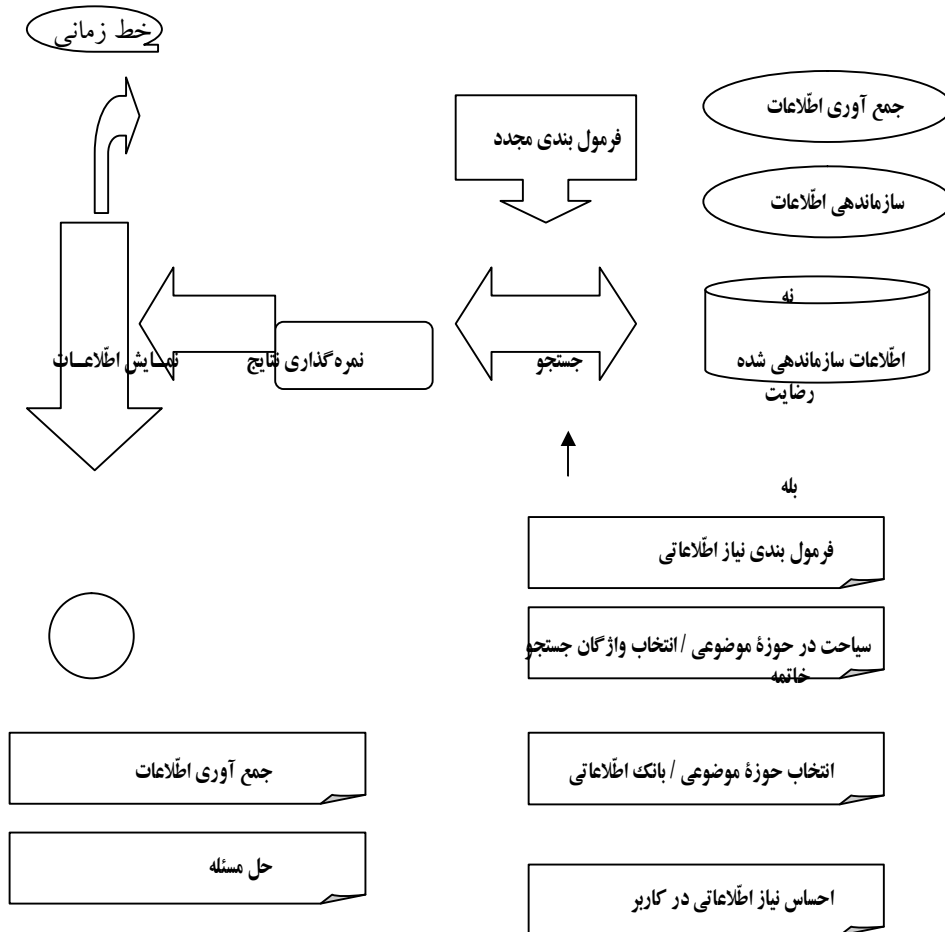
-
1. Willson
 2. Spink & Saracevic
 3. Feedbacks
 4. Huang
 5. Beaulieu & Robertson

اطلاعات یکی از جنبه‌های اساسی در رفتار کاربرانی می‌باشد که به دنبال اطلاعات برای حل مسئله می‌باشند. تحقیقات اخیر اسپینک، بیتمن و گریس دورف^۱ (1998) بر بازیابی اطلاعات در وب نیز مشخص کرد که جستجوی کاربران برای حل مسئله در یک دوره زمانی دارای یک توالی یا چرخه می‌باشد.

ارائه مدل ساخت یافته جستجو

یافته‌های حاصل از مدل بوک استین^۲ (1983)، مدل بویس و کرافت^۳ (1985)، مدل کوهلثا^۴ (1993)، بلکین، کول، استین و ثیل^۵ (1995)، ساراسیوس (1996)، الیس^۶ (1997) و اینگ ورسن^۷ (1992، 1996) و مطالعات انجام شده توسط پنی من^۸ (1975) و چاپمن^۹ (1981) در زمینه جستجوی اطلاعات نشان داد که، جستجوگران بانکهای متنی و کتابشناختی از یک چرخه دستورات به صورت مشترک استفاده می‌کنند، یک مدل که نشان دهنده فرآیند بازیابی اطلاعات باشد ارائه شد. این مدل در شکل شماره ۱ قابل مشاهده می‌باشد. در ادامه به بررسی اجزای مختلف این مدل می‌پردازیم:

-
1. Bateman & Greisdorf
 2. Bookstein
 3. Boyce & Kraft
 4. Kuhlthau
 5. Belkin, Cool, Stein & Theil
 6. Ellis
 7. Ingwersen
 8. Penniman
 9. Chapman



شکل شماره ۱: توالی یا چرخه جستجوی کاربران

۱- بازیابی اطلاعات

لنکستر^۱ (۱۹۷۹) بازیابی اطلاعات را عبارت از فرآیند جستجو در میان مجموعه‌ای از مدارک با هدف تعیین آن دسته از مدارک که در حیطه موضوعی درخواست شده باشند می‌داند.

1. Lancaster

چهار عنصر اصلی در این تعریف قابل مشاهده می‌باشد:

1. نیاز اطلاعاتی که نقش محرک در بازیابی را ایفا می‌کند.
 2. هدف از بازیابی اطلاعات که همان یافتن اطلاعات مورد نیاز می‌باشد.
 3. جستجو در میان مجموعه‌ای از اطلاعات ذخیره شده صورت می‌پذیرد، رکوردهای اطلاعاتی با استفاده از اطلاعات تولید شده توسط پدیدآورندگان ایجاد می‌گردد و سپس در بانک‌های اطلاعاتی سازماندهی می‌شوند. بدون آنکه بدانند در آینده چه کسانی و در چه شرایطی آنها را مورد مطالعه قرار می‌دهند.
 4. فرآیند جستجو، فرآیندی تعاملی می‌باشد، تعامل به این معنا که کاربر بر فرآیند جستجو کنترل داشته و در این مسیر دست به انتخاب می‌زند. پرسش کاربر ممکن است در طی این فرآیند با توجه به بازخوردهای حاصل از نتایج بازیابی یک سیر تکاملی را طی کند و انجام انتخاب مستلزم صرف هزینه، وقت و تلاش می‌باشد.
- کنترل فرآیند بازیابی و انتخاب اطلاعات مرتبط، دو جزء لازم در بازیابی اطلاعات می‌باشند. کپی کردن یک فایل بر روی دیسکت، بازیابی اطلاعات نامیده نمی‌شود و یا نگاه کردن به یک برنامه اخبار تلویزیون، بازیابی اطلاعات به حساب نمی‌آید، زیرا در طول آن بیننده هیچ کنترلی بر آنچه که پخش می‌شود، ندارد. یک کتابخانه، بهترین مثال برای نشان دادن بازیابی اطلاعات می‌باشد. جستجوگر کل مجموعه را برای بازیابی اطلاعات مورد مطالعه قرار نمی‌دهد، او اغلب مدارکی را انتخاب می‌کند که بالاترین درجه ارتباط را با نیازهای اطلاعاتی او دارند.

2- نظامهای بازیابی اطلاعات

نظامهایی که به منظور بازیابی و پردازش بر روی داده‌های ساختارنیافته طراحی شده و به لحاظ نوع سازماندهی، ساختار پایگاهها، استراتژی بازیابی، گروههای سرویس‌گیرنده و فرآیندی که در طی آن به درخواستهای اطلاعاتی کاربران پاسخ می‌دهند منحصر به فرد می‌باشند (گزنی، 1380).

3- جمع‌آوری اطلاعات

در این مرحله تصمیم‌گیری در مورد ایجاد بانک اطلاعاتی، دامنه، حوزه موضوعی

و محتوای اطلاعاتی بانک صورت می‌گیرد (هارتر¹ 1986).

مراحل ایجاد یک بانک اطلاعاتی عبارتست از:

1. تصمیم‌گیری برای ایجاد بانک اطلاعاتی براساس نوع نیازهای اطلاعاتی موجود.
2. تعیین حدود و دامنه موضوعی بر اساس اینکه چه اطلاعاتی، در چه سطحی و در کدام حوزه موضوعی باید جمع‌آوری گردد.
3. تعیین ساختاربانک مبنی بر اینکه چه مشخصه‌هایی (قلم‌های اطلاعاتی) باید برای ایجاد یک بانک اطلاعاتی تعریف شود.
4. انتخاب و گردآوری اطلاعات.
5. درون‌داد اطلاعات به‌صورت درون‌داد صحیح، اصلاح، روزآمد کردن و نگهداری اطلاعات، که خود مستلزم صرف وقت و هزینه می‌باشد. در کنار بانک اطلاعاتی فایل‌های کمکی (همانند فایل مقلوب) ساخته شده و به صورت مرتب روزآمد می‌شوند. این فایلها برای کنترل عملیات جستجو به کار می‌روند.
6. کنترل اطلاعات، اعم از جمع‌آوری اطلاعات و برطرف کردن خطاهای موجود در اطلاعات درون‌داد شده است. کنترل کیفیت و صحت اطلاعات در این مرحله صورت می‌گیرد.

4- سازماندهی اطلاعات

هر نظام اطلاعاتی دارای یک مبنای خاص برای تجزیه و تحلیل اطلاعات می‌باشد که نظام بر اساس آن به تفسیر اطلاعات و مطابقت بین اقلام و درخواستهای اطلاعاتی پرداخته و بدین ترتیب بازیابی صورت می‌گیرد. این تجزیه و تحلیل، سازماندهی اطلاعات نامیده می‌شود (گزنی، 1380).

آنچه از نقطه‌نظر کاربر باید درک شود این است که نظامهای بازیابی اطلاعات هر مدرک را به صورت مجموعه‌ای از واژگان در نظر می‌گیرند، محل قرار گرفتن این واژگان در یک مدرک، تعداد رخداد و مجاورت آنها از عوامل اصلی هستند که مرتبط بودن یک مدرک با واژه خاصی را مشخص می‌کنند. مدارکی که نزدیک تر به پرسش تشخیص داده

شوند، نمره بالاتری را به خود اختصاص داده و در مجموعه نتایج بازیابی شده در اولویت بالاتری جهت مشاهده قرار می‌گیرند.

5- احساس نیاز اطلاعاتی در فرد

فرآیند بازیابی اطلاعات هنگامی آغاز می‌گردد که فرد خلأیی در دانش خود حس کند. در این حالت اطلاعات و فهم آنها می‌تواند برای حل مشکل مورد توجه قرار گیرد. حس ابهام و سردرگمی فکری از مسائلی است که در اولین مراحل بازیابی اطلاعات مشاهده می‌گردد. در اولین مراحل جستجو تردید معمولاً به صورت یک نقطه منفی در نظر گرفته می‌شود که می‌خواهیم هر چه سریع‌تر برطرف شود.

6- انتخاب حوزه موضوعی

به منظور حل مشکل یا مسئله یک یا چند حوزه موضوعی و متناسب با بانکهای اطلاعاتی مختلف انتخاب می‌شوند. بعد از انجام انتخاب حس ابهام قدری کمتر شده و یک حس خوش‌بینی به وجود می‌آید که می‌تواند سرآغاز جستجو باشد.

باید بعد از تعیین نیاز اطلاعاتی، محل جستجو مشخص گردد. تعیین صحیح محل جستجو از اهمیت ویژه‌ای برخوردار است. به عنوان مثال اگر به دنبال یک مقاله علمی در زمینه پزشکی هستید، بانکهای اطلاعاتی مدلاین خیلی بهتر از موتور کاوش آلتاویستا جوابگوی جستجوی شما هستند. جستجوگران حرفه‌ای می‌دانند که برای هر جستجو چه منبعی را باید مورد جستجو قرار دهند. بنابراین هر جستجوگر باید در حیطه تخصصی خود با منابع مختلف جستجو آشنا باشد.

7- سیاحت

این مرحله مشکل‌ترین مرحله برای کاربران بوده و مرحله‌ای است که توسط تهیه‌کنندگان سرویس‌های اطلاعاتی و طراحان نظامهای بازیابی اطلاعات به خوبی درک نشده است. هدف در این مرحله بررسی اطلاعات کلی درباره مسئله به منظور گسترش دامنه فهم کاربر از مشکل می‌باشد. در ابتدای این مرحله احساس ابهام و سردرگمی افزایش می‌یابد. ناتوانی کاربر در بیان نیازهای اطلاعاتی ارتباط بین نظام و کاربر را مشکل می‌کند.

اما در طی این مرحله کم‌کم ذهن از موضوع مورد نظر آگاهی پیدا کرده، بر روی مسئله متمرکز می‌گردد. درحقیقت کاربر در این مرحله مسیر حرکت خود را مشخص می‌کند. به منظور مشخص شدن مسئله و اینکه سیاحت یک نوع از انواع جستجو می‌باشد، به بررسی انواع جستجوگرها می‌پردازیم.

انواع جستجو

مورس^۱ (1970) مشکل جستجو در یک بانک اطلاعاتی را با مشکل جستجو برای یک کشتی بر روی دریا، یک زیردریایی دشمن، یک کشتی گمشده و بازماندگان یک کشتی شکسته مقایسه کرده و با توجه به ماهیت جستجو این دسته‌بندی را ارائه می‌دهد:

الف. جستجو برای افراد یا اشیای خاص و مشخصی صورت می‌گیرد و محل جستجو نیز معین می‌گردد و ما انتظار داریم از روی علائم مشخصی محل مورد نظر را پیدا کنیم.

ب. جستجو برای هدف مشخصی صورت می‌گیرد. بنابراین با مشاهده هدف می‌توانیم آن را تشخیص دهیم (انسان - چادر - کشتی - قایق)، اما از محل هدف اطلاع نداریم و در یک پهنه وسیع به دنبال هدف می‌باشیم.

ج. جستجوگر به دنبال هدف خاص و مشخصی نمی‌باشد. جستجو برای یک نوع صورت می‌گیرد که ممکن است یکی یا بیشتر از یکی باشد مثل یک اردو یا یک باند هلی کوپتر.

د. جستجوگر هدف خاصی ندارد فقط می‌خواهد بداند در یک منطقه چه چیزهایی وجود دارد، شاید برای نقشه‌برداری، پیدا کردن آب، گیاهان غیر معمولی، جانوران خاص یا معادن.

با بررسی این موارد می‌توان چهار حالت مختلف را برای جستجو در بانکهای اطلاعاتی در نظر گرفت:

الف. جستجو برای اطلاعاتی معین در حوزه‌های معین.

ب. جستجو برای اطلاعاتی معین در یک حوزه نامعین.

پ. جستجو برای اطلاعاتی نامعین در حوزه‌های نامعین.

ت. سیاحت در یک بانک اطلاعاتی.

الف. جستجو برای اطلاعاتی معین در حوزه‌های معین

از یک طرف اطلاعات مورد جستجو ساخت یافته^۱ و دارای حوزه‌های معین و مشخص می‌باشد و از طرف دیگر کاربر معمولاً از نیاز اطلاعاتی خود آگاهی داشته و می‌داند که چه فیلدهایی را برای انجام جستجو انتخاب و چه مقادیری را برای جستجو به این فیلدها اختصاص دهد. درون‌داد و برون‌داد اطلاعات ساخت یافته با قالب‌های معینی صورت می‌پذیرد و حوزه‌های معینی در این قبیل بانکها مشخص و اطلاعات طبقه‌بندی شده‌اند. در این حالت کاربر دقیقاً می‌داند که چه رکوردی را باید مورد جستجو قرار دهد. جستجوگر در صورت مشاهده رکوردهای اطلاعاتی می‌تواند آنها را به راحتی تشخیص دهد. جستجو برای وضعیت صورتحساب محصولی که از شماره آن محصول آگاهی داریم و می‌توانیم با درون‌داد شماره آن به وضعیت صورتحسابهای موردنظر دست یابیم از این نمونه می‌باشد.

ب. جستجو برای اطلاعاتی معین در یک حوزه نامعین

در این حالت حوزه اطلاعاتی مورد جستجو در زمره حوزه‌های ساخت یافته قرار می‌گیرد. در این حالت کاربر از نیاز اطلاعاتی خود آگاهی داشته و می‌داند به دنبال چه اطلاعاتی می‌باشد. اطلاعات ساخت نیافته از دامنه نامعین و حوزه بسیار وسیعی برخوردار می‌باشد. کاربرد واژه نامعین به این دلیل است که حداقل استاندارد یا کنترل بر پدیدآورندگان این اطلاعات وجود دارد. اطلاعات موردنظر می‌تواند در رکوردهای مختلف و پراکنده‌ای وجود داشته باشد. در این حالت کاربر برخلاف جستجو در حوزه‌های معین به دنبال اطلاعات خاصی می‌باشد تا اینکه به دنبال یک سری رکوردهای خاص باشد. و دقیقاً نمی‌داند برای جستجوی اطلاعات خود باید از چه مقادیری استفاده کند.

پ. جستجو برای اطلاعاتی نامعین در حوزه‌های نامعین

جستجوگر به دنبال یافتن موضوعی عمومی و نامعین در یک حوزه نامعین می‌باشد. جستجوی وضعیت یک کمپانی برای سفارش کالا یا یافتن روشهایی برای حل یک معادله

دیفرانسیل از این نمونه می‌باشد. راهی برای تشریح و بیان موضوع مطابق شیوه‌های مرسوم در فرمول‌بندی نیاز اطلاعاتی وجود ندارد. دلیلی مبنی بر اینکه فکر کنیم این اطلاعات در یک رکورد پیدا خواهد شد، وجود ندارد. ممکن است اطلاعات مهمی مورد بازیابی قرار گیرد ولی با بررسی سطحی نتایج جستجو نمی‌توانیم به وجود آنها پی ببریم. در این حالت کاربر باید به شناسایی جنبه‌های مختلف موضوع پردازد و بداند که در حیطه مورد نظر چه اطلاعاتی و چه جنبه‌هایی موجود می‌باشد. در این حالت کاربر باید با صرف وقت و هزینه به تکرار بازیابی اطلاعات پردازد تا به نتایج دلخواه دست یابد.

ت. سیاحت در یک بانک اطلاعاتی

هدف، آگاهی از نوع اطلاعات موجود در بانک اطلاعاتی می‌باشد. در اینجا کاربر به دنبال رفع نیاز اطلاعاتی خود نیست یا به عبارت دیگر به دنبال یافتن پاسخی برای پرسش اطلاعاتی خود نمی‌باشد بلکه می‌خواهد به بررسی سطحی قسمتهایی از بانک اطلاعاتی پردازد. به عنوان مثال وقتی در یک کتابخانه به دنبال کتاب خاصی هستیم نگاهی گذرا به دیگر کتابها می‌تواند موجب آگاهی ما از سایر اطلاعات موجود در کتابخانه گردد این مصداق در مورد بانکهای اطلاعاتی نیز صادق می‌باشد.

8 - تبیین نیاز اطلاعاتی

در این مرحله، فرمول‌بندی نیاز اطلاعاتی، نقطه عطفی در فرآیند بازیابی می‌باشد. به صورتی که:

1. افزایش فهم و درک در حیطه مسئله موجب کاهش حس ابهام می‌گردد.
 2. شکل دادن به اطلاعاتی که در مرحله سیاحت بر روی آن متمرکز شده‌ایم، وظیفه آن است.
 3. ذهن روشن‌تر شده و یک دیدگاه نسبت به مسئله پیدا می‌کنیم.
- به منظور بیان و انتقال نیاز اطلاعاتی به نظام بازیابی اطلاعات از مجموعه‌ای واژگان استفاده می‌گردد. ارتباط منطقی بین مجموعه واژه‌ها با استفاده از عملگرهای موجود در نظام صورت می‌گیرد. مجموعه واژگان انتخاب شده تأثیر به‌سزایی در میزان مرتبط‌بودن نتایج بازیابی شده ایفا می‌کند. انتخاب واژگان جستجو نیازمند دانش پایه در موضوع مورد جستجو می‌باشد. در مواقعی که این دانش پایه وجود نداشته باشد سیاحت در بانک

اطلاعاتی می‌تواند بسیار مفید باشد. سیاحت در بانک اطلاعاتی می‌تواند با مشاهده قسمتهای مختلف بانک اطلاعاتی با اجرای جستجوهای مختلف و بررسی نتایج جستجو به منظور انتخاب واژگان مناسب صورت گیرد. بتس¹ (1989) این سیاحت را حتی خارج از محدوده جستجوی جاری مهم و ضروری می‌داند.

در تعیین مجموعه واژگان می‌توان از واژه‌نامه‌ها و اصطلاح‌نامه‌ها نیز بهره جست. تعیین واژگان اعم، اخص، هم‌رتبه، املائی مختلف واژگان، مترادفها، متضادها و بررسی ریخت‌شناسی (Morphology) واژه می‌تواند در تکمیل مجموعه واژگان تأثیر به‌سزایی داشته باشد.

در انتخاب مجموعه واژگان باید همیشه به دنبال واژگانی بود که نویسندگان بدنه صفحات خود را به وسیله آن به تحریر درآورده‌اند یا به بیان دیگر هنگامی که یک مدرک مورد بازیابی قرار می‌گیرد برآنیم که چه واژگانی در بدنه مدرک وجود داشته باشد. انتخاب مجموعه واژگان می‌تواند یک سیر تکاملی داشته باشد. بعد از انتخاب واژگان مناسب، باید آنها را با استفاده از عملگرهای AND برای واژه‌های لازم، NOT برای منع واژگان، OR برای حالت اختیاری بودن یا نبودن واژگان، عملگرهای مجاورت برای تعیین مجاورت و نزدیکی دو یا چند واژه به یکدیگر و کوه‌سازی به منظور جستجوی املاهای مختلف کلمه با یکدیگر ترکیب کرد. همان‌طوری که قبلاً نیز به آن اشاره شد فرآیند انتخاب واژگان یک فرآیند تکاملی می‌باشد که در طی فرآیند جستجو به مرور زمان با شکل‌گیری بیشتر دانش موضوعی کامل‌تر می‌گردد.

9- جستجو

در این مرحله کاربر به اجرای جستجو و بازیابی نتایج جستجو می‌پردازد. با افزایش فهم موضوعی کاربر و به منظور بازیابی اطلاعات مرتبط‌تر فرمول‌بندی مجدد صورت گرفته و این چرخه تا احساس رضایت کاربر از اطلاعات موجود ادامه می‌یابد. در ادامه به بررسی دو مفهوم نتایج نمره‌گذاری شده و اندازه‌گیری مشابهت که در این مرحله از اهمیت برخوردارند می‌پردازیم:

1- نتایج نمره‌گذاری شده

در این روش نیز براساس تعداد تکرار اعضای فرمول جستجو در رکوردهای بازیابی شده، به هر کدام از رکوردها یک نمره اختصاص می‌یابد و در نهایت رکوردها برحسب همین نمره در خروجی مرتب می‌شوند. این نمره نشان‌دهنده میزان ارتباط رکورد بازیابی شده حاضر با درخواست اطلاعاتی کاربر می‌باشد.

2- اندازه‌گیری مشابهت

دسترسی آسان کاربر به حداقل یک رکورد که بالاترین میزان ارتباط، با درخواست اطلاعاتی او را به همراه داشته باشد بسیار حائز اهمیت می‌باشد. اولین رکوردی که به این صورت توسط نظام مورد بازیابی قرار گیرد می‌تواند اساسی برای جستجوهای بعدی کاربر باشد. اگر کاربر چنین رکوردی را بازیابی کرد و به جهت مرتبط بودن با درخواست اطلاعاتی آن را تأیید کرد، آن‌گاه نظام به تجزیه و تحلیل رکورد می‌پردازد و با معیارهای ارزشیابی متن و دادن ضریب وزنی به هر یک از واژه‌ها اقدام به انجام یک جستجوی دیگر می‌کند و رکوردهای مشابه را مورد بازیابی قرار می‌دهد. در این روش کاربر نیازی به بازنویسی دوباره فرمول جستجو ندارد.

10- جمع‌آوری اطلاعات

تعامل کارآمد و سودمند میان کاربر و نظام بازیابی اطلاعات، جمع‌آوری اطلاعات را به همراه دارد. کاربر در این زمان به جمع‌آوری اطلاعات مرتبط می‌پردازد.

11- حل مسئله

در این مرحله بازیابی اطلاعات به پایان می‌رسد، مسئله مورد نظر حل می‌گردد و احساس رضایت از جستجو حاصل می‌گردد.

تفکر انتقادی

هرچند بحث تفکر انتقادی به عنوان جزئی از فرآیند اطلاع‌یابی جایگاهی ندارد ولی به لحاظ اینکه مقاله حاضر کاربران و جستجوی آنها در محیط‌های الکترونیکی و به ویژه اینترنت را مورد نظر قرار داده بحث حاضر نیز مطرح می‌گردد. اینترنت برای تبادل اطلاعات بدون کنترل بر چگونگی آن طراحی شده است. استاندارد یا قانونی وجود ندارد که نوع و کیفیت اطلاعات موجود بر روی اینترنت را کنترل کند. سایتهای وبی ممکن است

اطلاعات علمی، عقاید، داده‌ها، ایده‌ها، تبلیغات، اطلاعات تجاری و... را دربر نداشته باشند. هر کسی می‌تواند یک صفحه‌خانگی بسازد و آن را بر روی اینترنت قرار دهد و در این مسیر فرد کنترل کاملی بر نحوه نگارش سایت دارد. علاوه بر این محتوای اطلاعاتی صفحات می‌تواند با توجه به ماهیت تبلیغات، تجاری بودن، صرفه‌جویی در پول و ایجاد پول و مسائل رقابتی سوء، گرفته و یک‌طرفه باشد. بنابراین در برخورد با این صفحات باید به ارزیابی آنها پرداخت و یک دیدگاه انتقادی داشت.

نتیجه‌گیری

بر اساس پژوهش‌های مختلفی که در مقاله حاضر نیز به آنها اشاره شده است و مدل‌های مختلفی که مورد بررسی قرار گرفتند مشخص گردید که جستجوی اطلاعات یک فرآیند نظام‌یافته می‌باشد و باید گفت که این فرآیند دارای مراحل زیر می‌باشد:

- 1- احساس نیاز اطلاعاتی کاربر
- 2- انتخاب حوزه موضوعی / بانک اطلاعاتی
- 3- سیاحت در حوزه موضوعی / انتخاب واژگان جستجو
- 4- فرمول‌بندی نیاز اطلاعاتی
- 5- جمع‌آوری اطلاعات
- 6- حل مسئله

احساس نیاز اطلاعاتی سرآغاز انجام جستجو می‌باشد. به دنبال آن کاربر باید محل جستجوی خود را که از اهمیت ویژه‌ای برخوردار است به درستی انتخاب کند. در این میان نظام‌بازیبی اطلاعات باید اطلاعاتی در زمینه حوزه موضوعی بانک، تعداد رکوردها، جاری بودن و... را در اختیار کاربران قرار دهد و هر یک از کاربران نیز در حیطه تخصصی خود با محلهای جستجوی موردنیازشان آشنا باشند. بعد از تعیین محل، اگر جستجوگر از دانش موضوعی در زمینه جستجو برخوردار نباشد، باید قبل از انجام جستجو به سیاحت در بانک و سایر محلهای موردبحث در مقاله پردازد تا بتواند به انتخاب واژگانی مناسب برای جستجو و فرمول‌بندی نیاز اطلاعاتی خود پردازد. داشتن اطلاع از ترکیبات و دستورات سیستم و نحوه سازماندهی اطلاعات در انجام جستجوهای موفق از اهمیت برخوردار می‌باشد. فرآیند جستجو یک فرآیند تعاملی در یک سیر تعاملی می‌باشد. جمع‌آوری

اطلاعات و انجام جستجوهای مختلف به منظور حل مسئله می‌تواند یک مرحله نبوده و در یک دوره زمانی صورت گیرد.

آگاهی جستجوگران از فرآیند و مراحل مختلف جستجوی اطلاعات، تغییر دیدگاه آنها را به همراه خواهد داشت. جستجوگران باتجربه می‌دانند که انجام جستجو مستلزم صرف وقت و هزینه می‌باشد، بنابراین باید به صورت نظام‌یافته و در حداقل زمان حداکثر نتایج را به دست آورند. از سوی دیگر طراحان نظامهای بازیابی اطلاعات نیز باید با آگاهی از فرآیند جستجو نظامهای بهتری را طراحی کنند.

پیشنهادها

1. جستجوگران باید قبل از انجام جستجو با فرآیند آن آشنا بوده و نحوه انجام آن را به صورت نظام‌یافته بدانند و زمینه را برای اطلاع‌رسانی و آموزش بهتر فراهم نمایند.
2. همان‌گونه که غالب پژوهشگران حتی پیشگامان الگوهای تطبیقی نیز معتقد هستند، طراحی نظامهای بازیابی اطلاعات بدون در نظر گرفتن کاربر به عنوان یک عضو فعال موجب کاهش کارایی نظام می‌گردد. بنابراین طراحان نظامهای بازیابی اطلاعات می‌توانند با مطالعه فرآیند حاضر طراحی بهینه‌تری را انجام دهند.

منابع

- Bates, M. 1989. The Design of Browsing and Berry picking Techniques for On-line Search Interface. **On-line Review**. (13):407-424.
- Belkin, N. J., Cool, C., Stein, A., & Theil, S. 1995. Cases, Scripts, and Information-Seeking Strategies: On the Design of Interactive Information Retrieval Systems. **Expert System With Application**. 9(3): 379-395.
- Belkin, N. J. and Croft, W. B. 1987. Retrieval Techniques. **Annual Review of Information Science & Technology** (22):109-145.

Bookstein, A. 1983. Outline of a General Probabilistic Retrieval Model. **Journal of Documentation**. 39 (2): 63-72.

Boyce, B. R. and Kraft, D.H. 1985. Principles and Theories in Information Science. In: **Annual Review of Information Science and Technology**. Vol. 20. White Plains, New York: Knowledge Industry Publications.

Chapman, J. L. 1981. A State Transition Analysis of On-line Information-Seeking Behavior. **Journal of the American Society for Information Science**. 35(2):107-116.

Ellis, D. 1997. The Dilemma of Measurement in Information Retrieval. **Journal of the American Society for Information Science**. 47(1), 23-36.

Harter, S. P. 1986. **On-line Information Retrieval, Concepts, Principles, and Techniques**. San Diego: Academic Press.

Huang, M. H. 1992. **Pausing Behavior of End-Users in On-line Searching**. Unpublished Doctoral Dissertation. University of Maryland.

Ingwersen, P. 1992. **Information Retrieval Interaction**. London: Taylor Graham.

Ingwersen, P. 1996. Cognitive Perspectives of Information Retrieval Interaction: Elements of a Cognitive IR theory. **Journal of Documentation**. 52(1): 3-50.

Kuhlthau, C. C. 1993. **Seeking Meaning: A Process Approach to Library and Information Services**. Norwood, Ablex Publishing.

Lancaster, F. W. 1979. **Information Retrieval Systems: Characteristics, Testing, and Evaluation**. 2nd ed. New York: Wiley.

Morse, P.M. 1970. Search Theory and Browsing. **Library Quarterly**. 40(4):391-408.

Penniman, W.D. 1975. **A Stochastic Process Analysis of On-line Behavior**. Proceeding of the American Society for Information Science 38th Annual Meeting. Washington D.C: ASIS.

Robertson, S. E., & Hancock-Beaulieu, M. M 1992. On the Evaluation of IR System. **Information Processing and Management**. 28(4):457-466.

Salton, G. 1992. The State of Retrieval System Evaluation. **Information Processing & Management**. (28): 441-449.

Saracevic, T. 1996. Modeling Interaction in Information Retrieval (IR): A Review and Proposal. **Proceedings the Annual Meeting of the American Society for Information Science**. 33: 3-9.

Saracevic, T. 1997. Extension and Application of the Stratified model of information retrieval interaction. **Proceedings of the Annual Meeting Society for Information Science**.

Spink, A., Greisdorf, H., & Bateman, J. 1998. From Highly Relevant to not Relevant: Examining Different Region of Relevance. **Information Processing and Management**. 34(3): 257-274.

Spink, A. 1996. A Multiple Search Session Model of End-User Behavior: An Exploratory Study. **Journal of the American Society for Information Science**. 46(8): 603-609.

Spink, A. 1997. Interaction in Information Retrieval (IR): Successive Searches by Users Over Time. **National Science Foundation 1997 Power Grant Proposal**.

Wilson, T.D. 1999. Models in Information Behavior Research. **Journal of Documentation**. 55 (3): 249-270.

گزنی، علی. 1380. «موتورهای کاوش و دسترسی به اطلاعات وب». **خبرنامه انفورماتیک**، سال شانزدهم، شهریور و مهر 1380، شماره 79.

گزنی، علی. 1380. «طراحی سیستم‌های بازیابی اطلاعات بهینه در نرم‌افزارهای

بررسی فرآیند اطلاع‌یابی کاربران در ... & ۱۸

کتابخانه‌ای و اطلاع‌رسانی»، **علوم اطلاع‌رسانی**، دوره 16، شماره 1 و 2.

