

بازیابی مؤثر و سودمند اطلاعات در محیط اینترنت

فریبرز درودی^۱

چکیده

بازیابی سودمند اطلاعات در اینترنت کاری دشوار و مشقت‌آور است. در بهره‌گیری از اینترنت و بازیابی بهینه اطلاعات موجود در آن، شناخت راهبردهای جست‌وجو عامل مهمی است. مهم‌ترین رکن در انجام هر جست‌وجوی موفق تدوین راهبرد پرس‌وجوی مناسب با نیاز اطلاعاتی است. هرچه طراحی پرس‌وجو با نیاز اطلاعاتی کاربر تطابق بیشتری داشته باشد، به همان اندازه بازیابی به سوی کامیابی هدایت خواهد شد. فنون کاوش یا تکنیک‌های جست‌وجو، در هدایت درخواست ما و دستیابی به اطلاعات مرتبط تأثیر بسزایی دارد. علاوه بر آن، در بهره‌گیری از اینترنت در صورت به‌دست آوردن اطلاعات خواسته شده، ارزیابی و تحلیل اطلاعات از اهمیت بیشتری برخوردار است. صرف دستیابی به اطلاعات دلیلی بر موفقیت نیست. اطلاعات مستند و مستدل می‌تواند در انجام فعالیت‌های پژوهشی و آموزشی نقشی مؤثر و سودمند داشته باشد.

کلیدواژه‌ها

اینترنت، وب، جست‌جوی اطلاعات، طراحی پرس‌وجو، فنون کاوش، تحلیل و ارزیابی اطلاعات

مقدمه

اینترنت، بازیابی اطلاعات از این ابزار مهم با مشکلات فراوانی روبه‌روست. حجم اطلاعات موجود بر روی اینترنت را بیش از یک صد ترابایت^۲ برآورد کرده‌اند که روند افزایشی دارد. هر ترابایت معادل یک تریلیون (۱,۰۰۰,۰۰۰,۰۰۰,۰۰۰) کاراکتر

امروزه، اینترنت به عنوان مخزن عظیم اطلاعات در جهان شناخته می‌شود. منبعی که اطلاعات بسیار با ارزش و نیز اطلاعات بیهوده و زائد را با هم نگاهداری می‌کند. با توجه به تراکم اطلاعات موجود در محیط

۱. دانشجوی دکترای کتابداری و اطلاع‌رسانی دانشگاه آزاد اسلامی fardoroudi@gmail.com

2. Terabytes

است. سازمان‌ها، مؤسسات، دانشگاه‌ها، مراکز تجاری و بازرگانی، کتابخانه‌ها، ادارات دولتی، فروشگاه‌ها، و اشخاص حقیقی ترجیح می‌دهند که با ایجاد پایگاه در این محیط مجازی به ارائه خدمات و تولید اطلاعات بپردازند. در حال حاضر پس از مراجعه به یک مرکز خدماتی، تجاری و یا علمی بلافاصله با نشانی اینترنتی آنها مواجه می‌شویم که معمولاً بر روی کارت ویزیت آنها ثبت شده است. اکثر این مراکز دارای سایت اختصاصی هستند و اساساً در دنیای معاصر یکی از مؤلفه‌های اعتبار مؤسسات، فعالیت آنها در محیط شبکه و اینترنت است. با نفوذ این ابزار کاربردی در زندگی روزمره، که هر روز برای انجام امور خود به آن رجوع می‌کنیم، بازیابی اطلاعات با دشواری‌های بسیار روبه‌روست؛ و به جرئت می‌توان ادعا کرد که کمتر کسی را می‌توان یافت که با استفاده مستمر از اینترنت، در بازیابی اطلاعات درخواستی، با مشکل روبه‌رو نشده باشد. به‌دست آوردن اطلاعات مفید و مرتبط با نیاز اطلاعاتی کاربر، فعالیتی دشوار است. کاربرانی که از وب استفادهٔ تفننی و سرگرم‌کننده می‌کنند، معمولاً نیازهای اطلاعاتی خود را به نوعی با بازیابی اطلاعات مشابه یا گردش آزاد برآورده می‌سازند. این شیوهٔ جست‌وجو چندان نظام‌مند و ساختاریافته نیست، ولی در برآوردن نیازهای بسیاری از کاربران کارگشاست؛ زیرا در جست‌وجوهای عمومی و فراگیر، کاربران به پاره‌های اطلاعاتی وسیعی دست می‌یابند که به نوعی آنان را

اقناع می‌کند. هنگامی که اطلاعات از ویژگی اخص بودن خارج می‌شود، انطباق نیاز کاربر با اطلاعات فراهم‌آمده ساده‌تر است؛ ولی در آن سوی استفاده از وب، یعنی بهره‌گیری متخصصان، فعالیت دشوار می‌شود. تطبیق درخواست این گروه با اطلاعات به‌دست آمده کار ساده‌ای نیست. عدم رعایت استانداردهای سازماندهی اطلاعات، حجم عظیم اطلاعات موجود، شیوه‌های متفاوت جست‌وجو، الگوریتم‌های مختلف موتورهای جست‌وجو برای بازیابی اطلاعات، و عواملی دیگر سبب شده تا بازیابی اطلاعات در وب به‌خوبی انجام نپذیرد. در این میان مطالعهٔ کاربرد ابزارهای بازیابی اطلاعات و بهره‌گیری مؤثر از آن، استفادهٔ بهینه از موتورهای جست‌وجو، روش‌ها و راهبردهای جست‌وجو، و فنون کاوش و ارزیابی اطلاعات، مؤثر و سودمند است. حال با این پرسش اساسی روبه‌رو می‌شویم که استفادهٔ بهینه و سودمند از اینترنت چگونه میسر است؟

جست‌وجو در محیط اینترنت

بیشترین استفاده از اینترنت در وب صورت می‌پذیرد. جست‌وجوی اطلاعات در وب، همان فرایند ارائهٔ درخواست به مرورگر و دریافت اطلاعات مورد نیاز است. ما همواره در جست‌وجو با معیارهایی مواجه می‌شویم که به ما کمک می‌کنند تا بتوانیم براساس آنها اطلاعات مناسبی را پیدا کنیم. استیسی^۳ چهار نوع جست‌وجوی اطلاعات در محیط وب را معرفی می‌کند که شامل موضوع عمومی^۴،

3. Stacey

4. General topic

اطلاعات مستقیم^۵، اطلاعات غیرمستقیم^۶، و جست‌وجوی منابع^۷ است. در ادامه به توضیح هر یک از انواع جست‌وجوهای نامبرده می‌پردازیم.

جست‌وجوی موضوع عمومی: این شیوه جست‌وجوی اطلاعات از ابتدا برای بازیابی اطلاعات مربوط به یک سؤال ویژه و تخصصی طراحی نشده است. استفاده از این نوع جست‌وجو برای بازیابی اطلاعات تخصصی کارآمد نیست. در برخی موارد که کاربر نیازمند دستیابی به چارچوبی کلی درباره موضوع مورد نظر خود است و دریافت مفهوم اجمالی و فراگیر از موضوع برای او اهمیت دارد، استفاده از این شیوه می‌تواند راهگشا باشد. جست‌وجوی مقدماتی در محیط اینترنت با جست‌وجوی عمومی آغاز می‌شود. این شیوه در عین حال می‌تواند برای دستیابی به اطلاعات تخصصی به عنوان یک پایه آغازین در نظر گرفته شود. ما برای بازیابی اطلاعات دقیق و مستند به شکل‌دهی و تبدیل جست‌وجوی عمومی به دیگر انواع جست‌وجو نیاز داریم و باید آن را در روش‌های دیگری، که در بازیابی اطلاعات از دقت و توجه ویژه برخوردارند، ادغام کرده و در واقع تبدیل کنیم. در این مرحله کاربر با پاره‌ای از مشکلات جست‌وجوی اطلاعات در محیط اینترنت به‌خوبی آشنا می‌شود. راهنماها از ابزارهای مفید در این نوع جست‌وجو هستند. باید توجه داشت که راهنماها جامع، فراگیر، و ذهنی نیستند. نتایج به‌دست آمده از طریق شیوه جست‌وجوی موضوع عمومی، معمولاً نقطه شروع فرایند بازیابی اطلاعات است و

این شیوه به ندرت کاربر را به مرحله نهایی جست‌وجو، یعنی به‌دست آوردن اطلاعات مستند و مرتبط و تخصصی هدایت می‌کند.

جست‌وجوی اطلاعات مستقیم: جست‌وجوی اطلاعات مستقیم موضوعی است با سه قید. اول، پرسش باید به‌خوبی تعریف شود. دوم، برآورده ساختن انتظار منطقی، که در این حالت پاسخ می‌تواند مستقیماً یک سند معتبر در موضوعی خاص را معین کند. و سوم، انگیزه‌ای برای افراد یا سازمان‌ها جهت ارائه پاسخ به سؤال به شیوه پیوسته وجود داشته باشد. به‌طور خاص، یک پرسش در شیوه جست‌وجوی اطلاعات عمومی با شکست روبه‌رو می‌شود، اگر هم پرسش و هم پاسخ به شیوه‌ای موجز و مختصر توضیح داده شود، و نیز اگر اسناد معتبر در موضوع مورد نظر با تمایل برای پاسخ دادن به سؤال برای استفاده‌کنندگان عمومی همراه باشد.

جست‌وجوی اطلاعات غیرمستقیم: پرسشی که در جست‌وجوی اطلاعات غیرمستقیم مطرح می‌شود برای خود کاربر به وضوح مشخص است وی به روشنی می‌داند که به دنبال چه اطلاعاتی است. از سوی دیگر پاسخ پرسش او به شیوه رسمی و مدون در صفحات وب و یا حتی در وب‌سایت نیز یافت نمی‌شود. در واقع این‌گونه پرسش‌ها به‌رغم آنکه برای کاربر واضح است، ولی در عمل جست‌وجوی آنها با مشکل روبه‌روست. در این میان دلایلی وجود دارد که سبب می‌شود تا بازیابی اطلاعات در این شیوه

5. Direct information

6. Derived information

7. Resource search

از جست‌وجوها با تنزل روبه‌رو شود. برخی از این دلایل عبارتند از: عدم انگیزه برای اشتراک اطلاعات، دشواری اشتراک اطلاعات، ناتوانی در اشتراک اطلاعات، و مشکل روزآمدی سریع و فوری.

جست‌وجوی منابع: شیوه بعدی در جست‌وجوی اطلاعات در محیط اینترنت

جست‌وجوی منابع خاص غیرمتنی اطلاعات است. این منابع شامل اطلاعاتی چون فایل‌های دیداری - شنیداری یا برنامه‌های خاص موجود در اینترنت است (۱۴: ۷-۱۴).

همچنین مولر^۸ برای بهره‌گیری مفید و مؤثر از اینترنت نتایج چندین ساله استفاده خود از آن را به صورت هشت مرحله راهبردی برای کاربران مطرح می‌کند که شایسته است هر

جدول ۱. چهار نوع درخواست اطلاعات، ترکیب شناسایی آنها و راهبرد مرتبط و مناسب با هر یک (۱۲: ۸)

نوع اطلاعات	ترکیب شناسایی	راهبرد ربط
موضوع عمومی	سؤال غیردقیق کیفی هدف به دست آوردن مفهوم و چارچوب است	راهنماها مفهوم را فراهم می‌کنند موتورهای جست‌وجو و راهنماها با هم ترکیب می‌شوند تبدیل جست‌وجوی موضوع عمومی به جست‌وجوی اطلاعات مستقیم یا غیرمستقیم
اطلاعات مستقیم	وجود سؤال دقیق با پاسخ فشرده اطلاعات احتمالاً در یک صفحه تهیه شده است هیچ موقعیتی برای اعمال اطلاعات غیرمستقیم وجود ندارد	حدس بخردانه درباره مکان‌یاب جهانی منابع (یا موتور جست‌وجو URL) عبارت‌سازی دقیق درخواست ارزیابی سوگیری در مقابل صحت و درستی
اطلاعات غیرمستقیم	وجود سؤال دقیق با پاسخ فشرده، اما: روزآمد سازی سریع اطلاعات مورد نیاز است. یا: هیچ انگیزه‌ای برای دیگران جهت انتشار اطلاعات نیست یا: اطلاعات ترکیبی از تعداد بیشتری از پاره‌های اطلاعاتی است	بازیابی خودکار صفحات وب شناسایی فراهم‌آوردندگان اطلاعات تحلیل خودکار صفحات وب هنجارسازی و سری‌های زمانی یا تحلیل رگرسیون
جست‌وجوی منبع	هیچ منبع متنی مورد نیاز نیست	شناسایی داده‌های متنی پیوند داده شده به منابع غیرمتنی موتورهای جست‌وجوی استفاده شده برای مکان‌یابی سایت‌های درگاهی (ارتباطی) جهت انواع منابع

کسی آن را رعایت کند. این مراحل عبارتند از: (۱) چه چیزی را می‌خواهم بدانم؛ (۲) آیا موضوع خود را می‌شناسم؛ (۳) اگر موضوع خود را نمی‌دانید، ابتدا جست‌وجو برای منابع سودمند را شروع کنید؛ (۴) با دیگران گفت‌وگو کنید؛ (۵) از دیگران بپرسید؛ (۶) ببینید که دیگران چگونه جست‌وجو می‌کنند؛ (۷) مستقیماً جست‌وجوی اطلاعات را انجام دهید؛ و (۸) [اطلاعات خود را] معین و مشخص سازید (۸: ۱۳-۱۴).

در این رویکرد، ما مسیری معین را برای رسیدن به هدف خود طی می‌کنیم. در مرحله اول آنچه که بیشتر مدنظر است باید به صورت تدوین شده درآید، یا در فکر خود و یا بر روی کاغذ باید عبارت جست‌وجوی خود را مشخص کنیم. اندیشیدن دربارهٔ واژه‌های مترادف و مشابه که احتمالاً بیشتر ما را کمک می‌کند در اینجا مدنظر است. در مرحله دوم آنچه که می‌خواهیم بدانیم این است که اطلاعات مورد نیاز خود را شناسایی کنیم و بدانیم آن را در کجا باید پیدا کنیم، و در واقع جست‌وجوی خود را چگونه شروع کنیم. حیطهٔ موضوعی و گسترهٔ آن را در نظر بگیریم. مرحله سوم به این مطلب می‌پردازد که ما ابتدا نباید به جست‌وجوی اطلاعات خود پردازیم، بلکه باید به جست‌وجوی منابع مناسبی پردازیم که ممکن است اطلاعات مورد نیاز ما را در بر داشته باشند. به رغم آنکه این شیوه بسیار طبیعی و روشن به نظر می‌رسد، اغلب به دست فراموشی سپرده می‌شود. این رویکرد همچنین اهمیت

زیادی در ارزیابی وب‌سایت دارد. این مقوله بر پایهٔ دو پرسش اساسی بنا می‌شود: یکی آنکه چه کسی می‌تواند اطلاعاتی را که من به دنبال آن هستم ارائه دهد؟ و دیگر آنکه چگونه می‌توانم این اطلاعات را به دست آورم؟ مرحله چهارم مبادلهٔ فنون و تجارب با همکاران است؛ در واقع همان بهره‌گیری از توانمندی‌های ارزشمند و راهگشای آنان. مرحله پنجم درخواست از دیگران برای انجام کاری است که ما دربارهٔ آن فعالیت می‌کنیم. بسیاری از متخصصان باتجربه در گروه‌های خبری می‌توانند در این امر به ما کمک کنند. مرحله ششم توجه به این شیوه است که دیگران موضوع مورد علاقهٔ خود را چگونه جست‌وجو می‌کنند. ممکن است که این امر مورد پذیرش قرار گرفته و حتی بسیار سودمند و آموزنده باشد. زایتگیتس گوگل^۹ به ما اجازه می‌دهد تا بتوانیم جست‌وجوهای دیگران را مشاهده کنیم. در مرحله هفتم موتورهای جست‌وجو به صورت مستقیم مورد استفاده قرار می‌گیرند تا نشانی‌های اینترنتی وب‌سایت‌هایی را به دست آورند که شامل اطلاعات مورد نیاز ما هستند. این روش ساده، آسان، و در بسیاری موارد سودمند است. این شیوه برای کاربران نآزموده روشی نامناسب است، زیرا ممکن است که زمان بسیار زیادی را از دست بدهند و نتیجهٔ مناسبی نیز به دست نیاورند. در مرحله آخر نیز تدوین و صورت‌بندی دقیق جست‌وجو مطرح می‌شود. ما باید آنچه را که می‌دانیم به روشی روشن و گویا و با استفاده

از واژگان استاندارد نظام بیان کنیم، زیرا بهترین نرم‌افزارها و موتورهای جست‌وجو نیز نمی‌توانند آنچه را که ما در ذهن داریم تشخیص دهند.

اسپینک و سو^{۱۰} در ارتباط با گزیده‌ای از یافته‌های پژوهشی که دربارهٔ بازیابی اطلاعات به شیوهٔ جست‌وجوی عمومی از طریق موتور جست‌وجوی اکسایت^{۱۱} انجام داده‌اند، به اختصار عوامل مؤثر در ارتقای رهیافت به‌دست آوردن اطلاعات مفید را ارائه می‌دهند. برخی از عوامل مهم مطرح شده عبارتند از: (۱) طول پرسش و بسامد آن، (۲) پرسش بر مبنای عملگرهای بولی، (۳) فرمول‌بندی تازهٔ پرسش، (۴) جست‌وجوی عبارت، (۵) واژگان جست‌وجو: توزیع، (۶) بازخورد ربط، (۷) مرور نتایج، و (۸) جست‌وجوهای متوالی (۱۱). پاره‌ای از نتایج به‌دست آمده از این پژوهش نشان می‌دهد که طول پرسش کاربر با گذشت زمان افزایش یافته است. در سال ۱۹۹۶ طول واژگان به کار رفته در پرسش، در ایالات متحده، انگلستان، و کشورهای اروپایی ۱/۵ واژه بوده است، در حالی که در سال ۱۹۹۹ در کشورهای ایالات متحده و انگلستان ۲/۶ واژه و در کشورهای اروپایی ۱/۹ واژه بوده است. همچنین میزان استفاده از عملگرهای بولی بیشتر شده و عبارات جست‌وجو (واژه‌های محدود شده در گیومه) به ندرت استفاده شده است، تا آنجا که از هر ۱۶ پرس‌وجوی ارائه شده به

نظام، تنها ۱ پرس‌وجو به شیوهٔ عبارتی طرح شده، ولی به درستی به‌کار گرفته شده است. علاوه بر آن بهره‌گیری از بازخورد ربط بسیار اندک بوده است.

ناوارو-پریئو^{۱۲} و دیگران در پژوهش دیگری در ارائهٔ راهبردهای مهم جست‌وجو در محیط وب به سه راهبرد عمده اشاره می‌کنند که عبارتند از:

۱. **راهبرد بالا به پایین**^{۱۳}. این راهبرد هنگامی استفاده می‌شود که کاربران در حوزه‌ای عمومی جست‌وجو می‌کنند و سپس جست‌وجوی خود را از پیوندهای فراهم آمده محدود می‌سازند تا وقتی که آنچه را که به دنبالش بودند پیدا کنند. به‌طور خاص، کسانی که از این راهبرد بهره می‌برند سایت‌های عمومی زیادی را ملاحظه می‌کنند، که شامل سیاههٔ عوامل ساختاریافته در طبقه‌بندی‌های هدفمند است.

۲. **راهبرد پایین به بالا**^{۱۴}. در این راهبرد بر خلاف نوع قبل، کاربران به دنبال کلیدواژهٔ خاصی هستند که با راهنمایی و هدایت به دست آورده‌اند. در استفاده از این راهبرد کاربران مستقیماً کلیدواژه‌های خاص را در موتور جست‌وجو وارد کرده، و سپس نتایج به‌دست آمده را مرور می‌کنند. این کار با بازکردن هر پیوند و سپس بازگشت به سیاههٔ جست‌وجو ادامه پیدا می‌کند تا اطلاعات خواسته شده فراهم شود. این راهبرد اغلب توسط کاربران باتجربه مورد استفاده قرار گرفته و برای جست‌وجوهای اطلاعات

10. Spink & Xu

11. Excite

12. Navarro-Prieto

13. Top-down strategy

14. Bottom-up strategy

صحیح^{۱۵} کاربرد دارد.

۳. راهبرد ترکیبی^{۱۶}. برخی از کاربران از هر دو راهبرد گفته شده به صورت موازی بهره می‌گیرند. جست‌وجو برای اطلاعات خواسته شده را آغاز می‌کنند و در همان زمان در چندین پنجره به دنبال آنها می‌گردند. برخی از آنها راهبردهای متناوبی است که کاربر در طول جست‌وجوی خودش در ذهن دارد. این راهبرد فقط توسط کاربران کارآزموده استفاده می‌شود (۹).

علاوه بر آن باید گفت که در جست‌وجوهای انجام گرفته در محیط وب، حفظ جامعیت^{۱۷} و مانعیت^{۱۸} از اهمیت بیشتری برخوردار است. این مقوله به شیوه جست‌وجو نیز بستگی دارد. در ارتباط با حفظ جامعیت، راهبرد جست‌وجوی عمومی، و سپس محدود ساختن نتایج آن و دسترسی به اطلاعات مطلوب، از اهمیت برخوردار است. این روش به افزایش جامعیت منجر می‌شود. در شیوه دیگر استفاده از کلیدواژه‌های خاص که با دقت انتخاب شده‌اند، و اخص ساختن جست‌وجو از ابتدا مورد نظر قرار می‌گیرد. این روش به افزایش مانعیت می‌انجامد. در باب مقوله راهبرد ترکیبی ضمن توجه به این نکته که روش مذکور خاص متخصصان با تجربه و با مهارت است، باید توجه داشت که استفاده از این شیوه به عنوان یک راهبرد، همراه با شناخت کافی نسبت به تمامی جنبه‌های بازیابی اطلاعات میسر می‌باشد. استفاده از راهبرد افزایش مانعیت در بازیابی اطلاعات

منجر به نتایج مطلوب‌تری می‌شود؛ به‌ویژه آنکه در میان حجم مطالب انبوه و انفجار آمیز اینترنت رویکرد افزایش جامعیت روش مناسبی برای دسترسی به اطلاعات تخصصی به شمار نمی‌آید. البته حفظ جامعیت در شرایطی که حیطه موضوع مورد نظر محدود بوده و تعداد صفحات بازیابی شده اندک باشد، ممکن است به عنوان راهبردی راهگشا مورد استفاده قرار گیرد.

راهبرد جست‌وجو از جنبه دیگری نیز مورد مذاقه قرار گرفته است. بیتس^{۱۹} میان راهبرد جست‌وجو به عنوان حوزه‌ای برای پژوهش، و راهبرد جست‌وجو به عنوان برنامه‌ای برای کاوش تفاوت قائل شد. راهبرد جست‌وجو به منزله زمینه‌ای برای پژوهش عبارت است از مطالعه نظریه‌ها، اصول، و عملیات مربوط به تدوین و کاربرد راهبردها و شیوه‌های جست‌وجو. از این رو، راهبرد جست‌وجو شامل مفهوم خاص‌تری است که به طرح برنامه جست‌وجو مربوط می‌شود. به عنوان بخشی از فرایند بازیابی، راهبرد جست‌وجو عبارت است از فرایندی که از طریق آن فایلی مورد جست‌وجو قرار می‌گیرد تا مدارک متناسب با نیاز کاربر شناسایی شود. این مدارک براساس مجموعه‌ای از معیارهایی که شخص متقاضی مطرح می‌کند بازیابی می‌شود (۲: ۳۱۴).

در زمینه کاربرد شیوه‌های متعدد جست‌وجوی اطلاعات در وب می‌توان به این نکته اشاره کرد که ارائه راهبردهای گوناگون

15. Fact-finding searches

16. Mixed strategy

17. Recall

18. Precision

19. Bates

برای جست‌وجو در وب، زمینه مناسبی برای دستیابی به اطلاعات مرتبط با پرسش کاربر فراهم می‌آورد. بهره‌گیری از تجارب و نتیجه تحقیقات پژوهشگران این حوزه سبب می‌شود تا استفاده‌کنندگان از محیط اینترنت بتوانند اطلاعات ارزشمندی را از دل داده‌های زائد استخراج کنند. طرح رویکرد^{۲۰} بیان شده توسط هر یک از متخصصان، به منظور ایجاد زمینه مناسب برای بهره‌گیری از اطلاعاتی است که در شرایط عادی به‌خوبی بازیابی نمی‌شوند. این شیوه‌ها اگرچه در بسیاری از موارد کاربر را به نتیجه‌ای کاملاً مطلوب نمی‌رساند، ولی شرایط مناسبی را برای بازیابی اطلاعات مرتبط فراهم می‌سازد.

در بسیاری از موارد، مشکلات ناشی از مخدوش بودن شیوه‌های مؤثر جست‌وجو مزید بر علت شده و بر دشواری بازیابی اطلاعات مرتبط می‌افزاید. بسیاری از کاربران اینترنت در انتخاب راهبردهای جست‌وجو دارای نظمی پیوسته نیستند، به‌ویژه آنکه نداشتن خط‌مشی صحیح در فرایند کاوش اطلاعات، خود به معضلی دیگر بدل شده که همواره جست‌وجوهای پیوسته را دچار ریزش کاذب^{۲۱} می‌سازد. در جست‌وجوهای تخصصی، راهبرد جست‌وجوی عمومی کاربرد قابل توجهی ندارد. گرچه ممکن است در شرایطی خاص، استفاده از این راهبرد برای حفظ جامعیت در جست‌وجو مورد توجه قرار گیرد، ولی در مجموع شیوه مناسبی برای جست‌وجوی تخصصی به شمار نمی‌آید. در جست‌وجوهای انجام شده

دستیابی به کلیدواژه‌های معینی، که از بسامد بیشتری در سایت‌های گوناگون برخوردارند، مهم است. در موضوع‌های فنی و مطالب علمی، کلیدواژه‌های استاندارد معمولاً از طریق اصطلاحنامه‌های تخصصی یا تزاروس بهتر قابل بازیابی است؛ ولی ممکن است که در برخی موضوع‌ها، واژگان به‌کار رفته لزوماً با آنچه که اصطلاحنامه‌ها ارائه می‌کنند همخوانی نداشته باشد. در این حالت یکی از آزمون‌های کاربردی برای یافتن این‌گونه کلیدواژه‌های جست‌وجو - که منجر به بازیابی اطلاعات بهتری نیز می‌شود - بهره‌گیری از شمارش صفحات بازیابی شده است. اگر کلیدواژه‌ای با بسامد ۵۰۰۰ صفحه بازیابی شده ظاهر شود، و در مقابل کلیدواژه مترادف با ۱,۲۰۰,۰۰۰ صفحه بازیابی شده همراه باشد، به عنوان یک اصل کلی، به شیوه اکثری - و نه کاملاً و همیشه صحیح - می‌توان بیان کرد که کلیدواژه پربسامد از کاربرد افزون‌تر و پذیرش بیشتری برخوردار بوده است. در این‌گونه موارد با اخص ساختن این واژه‌ها در ترکیب با واژگان مرتبط و مورد درخواست، می‌توان به نتایج مطلوب‌تری دست یافت. در استفاده از راهبردهای جست‌وجو به این نکته نیز باید توجه داشت که هر کاربر متخصصی می‌تواند از بین الگوهای طرح شده با عنایت به نوع اطلاعات و نیاز خود، راهبردی ترکیبی و فراتر از طرح‌های پیشنهاد شده را برای خود برگزیند. پیشنهاد می‌شود که در هر جست‌وجوی که منجر به بازیابی مطلوب نمی‌شود، استفاده از راهبردهای

دیگر نیز آزمون شود. علاوه بر آن همیشه باید به یاد داشت که اطلاعات علمی، فنی، و کاربردی همیشه به رایگان در محیط وب قرار نمی‌گیرد. پس در مواردی نیز پایگاه‌های اطلاعاتی پیوسته که هزینه آنها پرداخت شده، می‌توانند نیاز اطلاعاتی ما را برآورده سازند. همچنین نباید فراموش کرد که بهره‌گیری از این‌گونه پایگاه‌ها نیازمند توانمندی و مهارت در جست‌وجوی اطلاعات است.

طراحی و تدوین پرس‌وجو

استفاده از راهبردهای گوناگون در فرایند جست‌وجو در وب، مستلزم بهره‌گیری صحیح و مناسب از پرس‌وجوهای متناسب با نیاز اطلاعاتی است. ارتباط میان شیوه‌های گوناگون جست‌وجو با طراحی مناسب پرسش در بازیابی اطلاعات سودمند، بیش از سایر عوامل است. طراحی صحیح پرسش همواره از ظرافت خاصی برخوردار است. معمولاً پدیدار شدن هر پرسشی در ذهن با یک نیاز اطلاعاتی همراه است. تبدیل این سؤال به یک عبارت جست‌وجو که قابلیت بازیابی اطلاعات مرتبط در موتور جست‌وجو را داشته باشد، کار آسانی نیست. اگر در طراحی سؤال از واژگان عام بهره‌گیریم، در آن صورت تعداد مدارک بازیابی شده بسیار زیاد خواهد بود، و در مقابل اگر پرسش ما از اصطلاحات بسیار خاص، تخصصی و با ویژگی کاربرد اندک در مدارک علمی و حرفه‌ای مرتبط با یک رشته علمی بهره‌گرفته باشد، اسناد بسیار کمی به دست خواهد آمد.

ایجاد تعادل در طراحی عبارت پرسش که از هر دو سو، یعنی عام و خاص بودن واژگان دچار مشکل نشود، از مهارت‌هایی است که شایسته است در طراحی عبارت جست‌وجو در نظر گرفته شود.

تحلیل پرسش عبارت است از مطالعه فرایند رفتار جست‌وجوی اطلاعات توسط انسان. این مطالعه شامل درک مؤثر نیازهای متقاضیان اطلاعات، روان‌شناسی برقراری رابطه مؤثر میان افراد، فن پرسش، و فرمول‌بندی منطقی درخواست‌های جست‌وجوپذیر برای اطلاعات است. چنین حوزه پیچیده‌ای از مطالعه در برگیرنده اصول و تجربه گفت‌وگو درباره پرسش است که به‌ویژه به محتوای برخی حوزه‌ها از جمله ارتباطات، مهارت‌ها، و روان‌شناسی حل مشکلات توجه دارد. تحلیل پرسش به دو مرحله دیگر تقسیم می‌شود. نخستین مرحله عبارت است از درک فرایند درونی کردن^{۲۲} یک مشکل اطلاعاتی توسط پرسشگر، که نهایتاً به شکل درخواست اطلاعات مطرح می‌شود. دومین مرحله را می‌توان نوعی شفاف ساختن دانست که شامل مجموعه‌ای از پرسش و پاسخ میان مراجعه‌کننده و جست‌وجوگر درباره درخواست مطرح شده است (۱: ۳۰۲-۳۰۳).

علاوه بر آن برای پدید آوردن یک پرسش مناسب در جست‌وجو باید به دو نکته اساسی نیز توجه کرد: اخص ساختن^{۲۳} و درستی. در محیط اینترنت، موتورهای جست‌وجو تمامی محتوای صفحات را نمایه‌سازی می‌کنند.

به علاوه، استفاده از حوزه‌های موضوعی استاندارد که معمولاً در کتابخانه‌ها کاربرد دارد در اینجا مورد استفاده قرار نمی‌گیرد. در این حالت استفاده از واژگان عمومی با بازیابی حجم عظیمی از اطلاعات غیرمرتبط و اضافی همراه خواهد بود. هر قدر از واژگان خاص و اصطلاحات تخصصی استفاده کنیم، در بازیابی اطلاعات به منابع بهتری دسترسی پیدا خواهیم کرد. در ارتباط با موضوعات علمی و خاص، استفاده از واژگان تخصصی و اصطلاحات کاربردی رشته‌های علمی سبب می‌شود تا جست‌وجو در تعداد سایت‌های کمتری انجام شود، و این فرایند باعث دستیابی به اطلاعاتی می‌شود که بیشتر با نیاز ما مطابقت دارد. واضح است که در این روش مانعیت افزایش پیدا می‌کند. در صورتی که بخواهیم از منابع بیشتری استفاده کنیم، بهتر است که از پرسش‌های تخصصی مشابه استفاده کنیم. ولی همچنان که در ابتدا توضیح داده شد، نباید از اصطلاحات بسیار خاص و یا واژگانی که به ندرت به کار می‌روند استفاده کرد.

از سوی دیگر به کارگیری تعداد واژگان در جست‌وجو نقشی بسزا دارد. استفاده از کلیدواژه‌های واحد در اینترنت برای جست‌وجوی اطلاعات امری متداول است، به‌ویژه در کاربرد اختصارات، زیرا امروزه مشابهات فراوانی در رشته‌های مختلف علمی و حرفه‌ای وجود دارد. به کار بردن کلیدواژه واحد، بدون استفاده از توصیفگرهای مناسب در محیط اینترنت سبب بازیابی اطلاعات غیرمرتبط فراوانی خواهد شد. در پژوهشی

که اسپینک و سو در مطالعه تبادلی در محیط سایت اکسایت انجام دادند، مشخص شد که از هر سه پرس‌وجوی ارائه شده، یکی از آنها شامل یک کلیدواژه بوده است. مشکلی که در این مرحله به وجود می‌آید، ضعف معنایی کلیدواژه و عدم توانایی بیان نیاز اطلاعاتی واقعی کاربر است (۱۱).

مشکل دیگری که پرس‌وجوهای تک‌واژه‌ای پدید می‌آورند، عدم توانایی استفاده از امکانات موتورهای جست‌وجو است. موتورهای جست‌وجو میلیون‌ها واژه از وب‌سایت‌ها را نمایه می‌کنند. یک پرس‌وجوی تک‌واژه‌ای نمی‌تواند از این سودمندی بهره‌گیرد، اما در صورتی که پرس‌وجو از سه واژه تشکیل شود، به کاربر این توانایی را می‌دهد تا نمونه مناسبی از اطلاعات را به دست آورد (۱۲: ۲۱۹).

همچنان که پیش‌تر بیان شد، عبارت پرسش علاوه بر اخص بودن، باید صحیح باشد. این بیان در ساده‌ترین سطح، معنای صحت نگارش و املا را می‌رساند؛ ولی از منظر تخصصی عبارت است از آنچه که مورد نیاز است تا نگارش و عبارت‌پردازی دقیقاً مربوط به چیزی باشد که پدیدآورنده اطلاعات فراهم آورده است؛ به تعبیر دقیق‌تر محتوای مربوط به ارائه‌دهنده اطلاعات در محیط وب، و به شیوه پیوسته، اعم از آنکه شخص یا سازمان چنین اطلاعاتی را عرضه کرده باشند. در ارتباط با فعالیت‌های معتبر علمی، فنی، و حرفه‌ای در محیط وب می‌توان به نگارش صحیح آن سایت‌ها اعتماد کرد؛ ولی در برخی موارد ما با نگارش ناصحیح و اشتباه در سایت‌هایی روبه‌رو می‌شویم که انتظار

بهره‌گیری سودمند از آنها را داریم. ممکن است که شما نیز با اشتباهی که در نگارش یک واژه داشته‌اید با حجم قابل توجهی از پیوندهای بازیابی مواجه شده باشید، که این امر دستیابی به اطلاعات را با دشواری روبه‌رو می‌سازد؛ به‌ویژه آنکه در پاره‌ای موارد به دلیل نداشتن نظارت بر صحت و اعتبار منابع اینترنتی به این شیوه میزان قابل توجهی از اطلاعات مرتبط و سودمند در بازیابی ظاهر نمی‌شوند.

از دیگر راهکارهای مفید برای طراحی مناسب پرس‌وجو باید به خوشه‌بندی^{۲۴} پرس‌وجو اشاره کرد. خوشه‌بندی رهیافتی است که به ما اجازه می‌دهد تا بتوانیم با دسته‌بندی سؤالات مورد نظر، به نتایج مفید و مناسبی دست یابیم. اساساً خوشه‌بندی پرسش‌ها، به معنی مرتب کردن آنها در درون رده‌های مشخص با عناوین کلی است و به ما این امکان را می‌دهد تا بتوانیم سرعت و زمان جست‌وجو را بهینه‌سازی کنیم. همچنین سبب می‌شود تا محل قرار گرفتن اطلاعات مرتبط را بهتر پیدا کنیم. خوشه‌بندی، علاوه بر این، با ایجاد ارتباط بین رده‌های کلی، به ایجاد یکپارچگی بین آنها منجر می‌شود. این مهم در ارتباط با طراحی پرس‌وجو به نوعی رده‌بندی موضوعی پرسش‌ها منجر شده و سبب استفاده بهینه از قابلیت‌های سؤالات جست‌وجو در بازیابی اطلاعات مرتبط می‌شود.

ون^{۲۵} و همکاران در پژوهشی که درباره خوشه‌بندی پرسش‌های استفاده‌کننده از یک

موتور جست‌وجو انجام دادند، به نتایج زیر دست یافتند:

- شناسایی عبارات ممکن است که به‌وسیله شناسگر خودکار بر پایه کاربرد نحوی توسعه داده شود،

- شمارش پرسش‌های مشابه می‌تواند وابستگی بیشتر بین واژه‌ها را افزایش دهد، و

- نظام می‌تواند درون یک ابزار تفسیر معنایی واژه، با در نظر گرفتن هر سند انتخاب شده^{۲۶} به‌عنوان یک معنای ممکن از واژه پرسش، توسعه داده شود (۱۳).

همچنین در خوشه‌بندی پرسش ما با دو معیار مهم روبه‌رو هستیم: معیار اول استفاده از محتوای پرسش است. به این معنا که اگر دو پرسش شامل اصطلاحات یکسان یا مشابه باشند، باید براساس نیازهای اطلاعاتی مشابه یا همانند، معین و مشخص شوند. طبیعی است که طول عبارات جست‌وجو، سبب اطمینان اصل اول است. به هر حال کاربران اغلب پرسش‌های کوتاه را به موتورهای جست‌وجو ارائه می‌دهند. یک پرسش نوعی در محیط وب، معمولاً شامل یک یا دو واژه است. در برخی موارد، اطلاعات کافی برای درک نیازهای اطلاعاتی کاربران به درستی وجود ندارد. بنابراین دومین معیار، مشابهت با درک متضمن اسناد خوشه‌بندی شده در بازیابی اطلاعات است. اصلاً اعتقاد بر این است که ارتباط بین اسناد مشترک، گرایش به وابستگی به همان پرسش دارد. در شیوه دیگر ما از مسیر عکس استفاده می‌کنیم. اگر دو

24. Clustering

25. Wen

26. Clicked document

پرسش منجر به انتخاب همان سند شوند (که ما آن را انتخاب سند^{۲۷} می‌نامیم)، پس آنها مشابه هستند. این دو معیار سودمندی خود را به همراه دارند. در استفاده از معیار اول ما می‌توانیم پرسش‌های دارای ساختار مشابه را کنار هم گروه‌بندی کنیم؛ و در استفاده از معیار دوم، از قضاوت کاربر بهره می‌گیریم (۱۳).

در ارائه پرسش به نظام، ما همواره به جست‌وجوی جامع اطلاعات نیاز داریم تا بتوانیم مجموعه کامل و مناسبی از اطلاعات را به دست آوریم. فراهم کردن اطلاعات جامع در انجام یک فرایند پژوهشی و یا آموزشی به کاربر این امکان را می‌دهد که با ابعاد مختلف موضوع آشنا شده و در مراحل پژوهش با دیدگاهی گسترده و فراگیر به این کار مشغول شود. در پاره‌ای موارد ما به فرمول‌بندی دوباره پرس‌وجو نیاز داریم. این شیوه به ما کمک می‌کند تا بتوانیم با تغییر در ترکیب عبارت جست‌وجو به نتایج جدیدی رهنمون شویم.

دلایل ارائه جست‌وجوی جامع با اتکاء بر فرمول‌بندی دوباره به این شرح است: ارزیابی نظرات ذهنی گوناگون در ارتباط با موضوع؛ جبران جهت‌گیری در پاسخگویی به پرسش مرتبط با موضوعی ویژه، ارتقای دیدگاه از محدوده سازمان‌های مختلف یا دیگر فراهم‌آوردگان اطلاعات؛ تصفیه کردن، خواه پاره‌های اطلاعاتی احتمالاً صحیح باشند یا نباشند. فرمول‌بندی دوباره به کاربر این توانایی را می‌دهد تا حیطه گسترده‌ای از سایت‌های مربوط به موضوعی خاص را

به دست آورد. باید دقت کرد که دستیابی‌های موفقیت‌آمیز عوض نمی‌شود، به این معنی که اگر یک پرسش نامناسب انتخاب شده باشد، دستیابی‌های مناسب توسط موتور جست‌وجو تغییر نخواهد کرد. در پاره‌ای از موارد انتخاب یک پرسش مناسب برای شروع جست‌وجو ممکن نیست، بنابراین استفاده از پرسش انتخابی نامناسب، به فرمول‌بندی دوباره منجر می‌شود، تا مسیر طبیعی خود را پیدا کند (۱۲): (۹۸).

در ارتباط با فنون فرمول‌بندی دوباره به دو نکته اساسی باید توجه داشت: اول، انتخاب و کاربرد عبارت جست‌وجو به گونه‌ای باشد که تعداد کمتری از وب‌سایت‌ها در بازیابی ظاهر شود. دیگر آنکه فرمول‌بندی از اصل تکرارپذیری برخوردار باشد. به تعبیر دیگر باید بتوان با فرمول‌بندی‌های متفاوت به جنبه‌های خاص پرسش‌ها و ویژگی آنها وقوف بیشتری یافت. همچنین باید توجه داشت که در این میان نباید از پرسش‌های گوناگون و متفاوت استفاده کرد، بلکه بهتر است که تلاش کاربر بیشتر برای فرمول‌بندی دوباره عبارت جست‌وجوی مناسب باشد، تا آنکه از واژگان و اصطلاحات متعدد و زیاد بهره گرفته شود.

به علاوه استفاده از واژگان پیشنهادی برخی موتورهای جست‌وجو در سنجش صحت املاي کلمات مؤثر است. در برخی موارد ممکن است که کاربر واژه‌ای را به اشتباه در قسمت مربوط به جست‌وجوی مرورگر وارد کرده باشد، در این صورت می‌توان با اصلاح

آن واژه یا عبارت به جست‌وجوی مجدد پرداخت. هنگامی که ما واژه مورد درخواست را به شیوه صحیح به کار می‌بریم، پیشنهاد برنامه در جست‌وجوی برخی موضوع‌های مشابه را نیز باید در نظر داشته باشیم. در آن صورت علاوه بر ملاحظه ارتباط موضوعی با توجه به اطلاعاتی که درباره موضوع داریم، تعداد سایت‌های بازیابی شده نیز معیار مناسبی برای انتخاب ما به شمار می‌رود. یک اصل نسبی درباره انتخاب سایت‌های پیشنهاد شده این است که اگر تعداد سایت‌های بازیابی شده که با انتخاب واژه پیشنهادی به آن دست می‌یابیم، بسیار زیادتر از واژه قبلی یا انتظار ما باشد، در آن صورت می‌توان حدس زد که فرمول‌بندی موضوع مورد نظر، به میزان قابل توجهی به سوی بازیابی سایت‌های عمومی گرایش پیدا کرده است.

از دیگر مقولاتی که در طراحی پرس‌وجو تأثیر دارد، تحلیل گفتار^{۲۸} است. تحلیل گفتار، مطالعه استفاده از زبان در گفتار است، همراه با استفاده از روشی که در آن محتوای گفتار هر موضوعی سازمان یافته، و عبارت زبانی تعیین شود. هدف، مشخص ساختن فرایند پردازش آن بخش‌هایی از زبان است که به عنوان گفتار تفسیر شده است. ما سعی می‌کنیم قواعد زبانی مؤثر در گفتار، معیارهای کاربرد، واج شناختی، دستوری، و معنایی را کشف کنیم. آنچه که در معنای اصطلاح تجزیه و تحلیل عناصر سازنده^{۲۹} در بعد معنایی توضیح داده شده، با مثال‌های ساده‌ای که

اغلب از طریق زبانی برای این منظور مورد استفاده قرار می‌گیرند، عبارتند از: زن-مرد-کودک؛ گاو-نر-گاو ماده-گوساله؛ اسب-نر-مادیان-کره اسب (۱۰: ۱۶). به این ترتیب درک و شناخت واژه‌هایی که از نظر قرابت و ارتباط با هم نزدیک هستند، در جست‌وجو و بازیابی اطلاعات مرتبط مهم می‌باشد. این امر در طراحی پرس‌وجو، به‌ویژه تعیین واژگان هم معنا، از اهمیت زیادی برخوردار است. مثال ساده‌ای که در بالا آمد، در واقع به نوعی طبقه‌بندی واژه‌های خویشاوند، و درک عناصر اصلی واژه در ایجاد ربط^{۳۰} به‌شمار می‌آید. هر قدر تحلیل هم‌معنایی بین واژگان افزایش یابد، در آن صورت دستیابی به عبارت مناسب جست‌وجو ساده‌تر خواهد بود. شناخت ارتباطات مفهومی میان واژه‌ها، در بسیاری از موارد در انتخاب پرس‌وجوی مؤثر برای بازیابی اطلاعات، کمک بسزایی به کاربران می‌کند. این لغات هم‌معنا در برخی موارد می‌تواند با پیوندی که نسبت به واژه‌های خویشاوند خود دارد، پاره‌های اطلاعاتی مرتبط را نیز ارائه دهد و به بازیابی اطلاعات مفید منجر شود.

ارتباط بین واژه‌های جست‌وجو معمولاً از دو نوع معنایی و یا نحوی تشکیل می‌شود. در ارتباط معنایی رابطه خویشاوندی واژگان مدنظر است و ما سعی می‌کنیم تا از واژگان برابر یا جانشین استفاده کنیم. ارتباط معنایی شامل سه نوع است: هم‌ارزی^{۳۱}، سلسله مراتبی^{۳۲}، و دارای شباهت^{۳۳}. رابطه هم‌ارزی خود حاوی

28. Discourse analysis

29. Componential analysis

30. Relevance

31. Equivalence

32. Hierarchical

33. Affinitive

انواع متضادها و مترادف‌ها، شبه‌مترادف‌ها^{۳۴}، پیوستار همانند^{۳۵}، همپوشانی^{۳۶}، املائی برتر^{۳۷}، نام فشرده‌ها و کوتاه‌نوشته‌ها^{۳۸}، اصطلاحات ساخته‌شده و جاری^{۳۹}، ترجمه، و اصطلاحات فنی و عامیانه^{۴۰} است (۱۰: ۳۶). هر یک از موارد بالا علاوه بر ایجاد رابطه معنایی، انواع فنون نگارشی را نیز شامل می‌شود. رعایت این عوامل در طراحی عبارت جست‌وجو اهمیت دارد. به خصوص در برخی سایت‌ها، ممکن است که از کوتاه نوشته‌ها بیشتر استفاده شده باشد. مانند کوتاه‌نوشت «فیفا»^{۴۱} که از عنوان کلی «فدراسیون بین‌المللی فوتبال» کاربرد گسترده‌تری دارد. همچنین برخی واژگان به‌کار رفته در میان شاغلان در یک حرفه که ممکن است با اصطلاح خاص خود آنان شناخته شده باشد؛ و یا ترجمه‌هایی که دقیقاً از زبان مبدأ به زبان مقصد منتقل می‌شود، نظیر رادیو که حتی امکان دارد کاربرد آن از معادل تعیین شده در زبان جاری بیشتر باشد.

رعایت نکات نگارشی و معنایی بدون تردید در طراحی مؤثر عبارت جست‌وجو نقشی تعیین‌کننده دارد. به‌ویژه در زبان فارسی که عدم شیوه مناسب و یکسان نگارش، سبب دشواری در تنظیم پرسش‌های مؤثر برای بازیابی اطلاعات می‌شود. در بسیاری موارد کاربر مجبور است تا چند شیوه نگارشی را همزمان مورد آزمون قرار دهد تا به نتیجه

مورد نظر برسد. نگارش پیوسته و ناپیوسته واژگان ترکیبی یکی از مصادیق آن است. به طور نمونه، نگارش واژه روان‌شناسی، روان‌شناسی، و روانشناسی، که هر یک از این شیوه‌ها در جست‌وجو نتایج مختلفی به همراه خواهد داشت. اگر چه کوشش شده تا در بانک‌های اطلاعاتی برخی از این مشکلات نگارشی مرتفع شود، ولی همچنان در بسیاری از سایت‌های اینترنتی و حتی پایگاه‌های اطلاعاتی این موارد وجود دارد. رعایت جامعیت و مانعیت در جست‌وجو، از عوامل تأثیرگذار در طراحی پرس‌وجو به شمار می‌آید. هر کاربری با توجه به نوع اطلاعات و شیوه جست‌وجوی خود، می‌تواند به کاربرد آنها توجه داشته باشد. راهبرد طراحی پرس‌وجو ایجاب می‌کند که در شرایط متفاوت از هر یک از آنها استفاده مطلوب صورت پذیرد. پیشنهاد می‌شود که در طراحی پرس‌وجو از شیوه بررسی متقابل استفاده شود. در این شیوه می‌توان هنگام فرمول‌بندی دوباره، روش‌های گوناگونی در طراحی و تدوین پرس‌وجو را تجربه کرده و نتایج مختلف به‌دست آمده را مورد سنجش قرار داد. با استفاده از این رویکرد، در هر مرحله، نتایج ذخیره شده و میزان ارتباط پیوندهای بازیابی شده، مثلاً ۱۰ پیوند در صفحه اول با شیوه دیگر مورد سنجش قرار می‌گیرد. می‌توان این مقایسه را در پیوندهای

34. Quasi synonyms

35. Same continuum

36. Overlapping

37. Preferred spelling

38. Acronyms & abbreviations

39. Current and established terms

40. Technical terms & layman terms

41. FIFA=Federation Internationale de Football Association

بازیابی شده دیگر صفحات نیز مورد آزمون قرار داد. حسن این روش آن است که شیوه مناسب در ترکیب واژگان برای بازیابی نوع خاصی از اطلاعات مشخص می‌شود.

فنون کاوش

هنگامی که فرایند کاوش در محیط وب را آغاز می‌کنیم، ممکن است به‌رغم وارد کردن کلیدواژه‌های مختلف با بازیابی اطلاعات نامرتبط روبه‌رو شویم. اساساً ساختار اینترنت به‌گونه‌ای است که سازماندهی مدون و استاندارد اطلاعات در بیشتر سایت‌های آن رعایت نشده است. مشکلات بنیادی اینترنت خود مبحثی جداگانه است که مورد توجه پژوهشگران این رشته نیز قرار گرفته است. به‌رغم وجود تمامی مشکلات، می‌توان با کاربرد فنون مؤثری به بازیابی بهتر اطلاعات پرداخت.

در پژوهشی که کاری^{۴۲} در بهره‌گیری از فنون استفاده از صفحات اینترنتی، در ارتباط با کاربرد صفحات مشاهده شده بر اساس شیوه خود توسعه‌دهنده^{۴۳} برای بازیابی اطلاعات انجام داده، برخی نتایج زیر به‌دست آمده است:

«بیشتر حوزه‌های موضوعی نمونه‌ای که مورد بررسی قرار گرفتند، صفحات وب مربوط به فناوری و اقتصاد بودند؛ وب‌فرم^{۴۴} به طبقه‌بندی جدید نوع‌شناسی

محتوای [اینترنت] افزوده شده است، اما اکثر صفحات فرا-اطلاعاتی هستند؛ حیطة گسترده‌ای از تدابیر مختلف برای گردش در محیط وب به‌دست آمده است، ولی فقط استفاده از شیوه ضربه بر پیوندها^{۴۵} و دکمه به عقب بیشتر از سایر روش‌ها، کاربرد دارد؛ تعقیب^{۴۶} (انتقال خودکار به صفحه دیگر) به عنوان یک راهبرد جدید شناخته شده، اما نشانه‌گذاری^{۴۷} (انتخاب یک گزینه) خیلی کم استفاده شده است؛ هدایت وب با کاربرد تمامی لایه‌های رابط کاربر نمایان شده و شماره‌گذاری صفحه بسیار عمومیت یافته است؛ به‌طور میانگین، هر صفحه‌ای قبلاً بارها بازدید شده، در ویرایش‌های تغییر یافته آن صفحات تشابهاتی نیز وجود داشته است؛ رابطه معنی‌داری از لحاظ آماری بین صفحه و متغیرهای تغییر مکان وجود دارد، اما به نحوی ویژه قوی نیست؛ به‌طور عمده معیار سنجش، کشف رابطه بین موضوع و زبان صفحات وب است» (V).

در استفاده از فنون کاوش باید مد نظر داشت که کاربرد دکمه‌های خاص، امکانات تعبیه شده در ابزارهای بازیابی، استفاده از علامت‌های ویژه، ترکیب‌بندی عبارات و اقداماتی از این دست همگی در بازیابی مؤثر اطلاعات نقش دارند. ابتدا باید گفت که برخی مواقع عدم ارتباط با یک صفحه اینترنتی، ممکن است ناشی از دسترس‌پذیر

شکلی از اطلاعات (گزارش‌های محیطی) را به دریافت کننده ارائه دهد.

42. Kari

43. Self-developers

۴۴. Web form : یک رابط کاربر استاندارد که می‌تواند اطلاعات را از طریق اینترنت بارگذاری کند. یک وب فرم شامل محلی برای وارد کردن پرسش است که استفاده کننده می‌تواند

45. Clicking on links

46. Following

47. Pointing

نبودن آن صفحه باشد. شایع‌ترین مشکلی که در دستیابی به یک وب‌سایت یا صفحه‌ای خاص بروز می‌کند، ظهور پیغام خطا در بازیابی صفحه مورد نظر است. در این‌گونه موارد باید از صحت نشانی وارد شده مطمئن شد. بهترین شیوه برای آزمایش درستی نشانی این است که بخش اصلی تا قسمت دامنه نشانی مکان منابع یکسان^{۴۸} را برای صحت نشانی آزمایش کنیم. در این صورت می‌توانیم با حذف قسمتی از نشانی، که بعد از دامنه اصلی سایت قرار گرفته، به آزمایش دسترس‌پذیری سایت پردازیم. مثلاً هنگامی که نشانی: <http://www.bookworm.com.au/bookworm/compsale.htm> را در نوار نشانی^{۴۹} یک مرورگر اینترنت وارد می‌کنیم و به نتیجه نمی‌رسیم، می‌توانیم قسمت اصلی نشانی سایت یعنی <http://www.bookworm.com.au/> را وارد کرده و مورد آزمایش قرار دهیم. در صورتی که با این نشانی نیز به نتیجه مطلوب نرسیم و پیام خطای نشانی دریافت کنیم، بدان معنی است به دنبال صفحه‌ای هستیم که موتور جست‌وجو قادر به بازیابی آن نیست. ولی اگر بتوانیم به صفحه اصلی وب‌سایت وارد شویم، باید به آزمایش قسمت مورد نظر که با توجه به عنوان صفحه و واژه‌های به کار رفته در نشانی اینترنتی قابل شناسایی است، به آزمایش صحت دسترس‌پذیری صفحه مورد نظر پردازیم. امکانات جست‌وجوی پیشرفته در یک

موتور جست‌وجو به کاربر این امکان را می‌دهد که بتواند با توجه به نیاز اطلاعاتی خود عبارت جست‌وجو را به شیوه‌ای مطلوب سازماندهی کرده و به صورت یک پرس‌وجوی مناسب به نظام بازیابی اطلاعات و موتور جست‌وجو تحویل دهد. در بسیاری از موتورهای جست‌وجو امکاناتی برای عبارت‌سازی پرس‌وجوی ما وجود دارد. مثلاً در سایت گوگل می‌توانیم در بخش جست‌وجوی پیشرفته با محل مخصوص جست‌وجوی عبارتی^{۵۰}، به کاوش دست‌زینیم. وارد کردن واژه‌های متعدد در این قسمت مستلزم قرار دادن آنها در گیومه است. به این معنی که ما دقیقاً به عین عبارت طرح شده نیاز داریم، و این کلمات باید در کنار هم و با ساختار معین جمله‌ای که تشکیل داده‌ایم انجام پذیرد. در صورتی که بخواهیم تمامی واژه‌های ما در جست‌وجو لحاظ شود و یا قصد جست‌وجوی برخی از این کلمات را داشته باشیم، در آن صورت باید از محل‌های تعیین شده که خاص هر یک از موارد نامبرده است استفاده کنیم. این فیلدها شامل «با همه واژگان»^{۵۱} و یا «حداقل یکی از این واژگان»^{۵۲} می‌شود. همچنین در انجام جست‌وجوهای متعدد ما نیاز خواهیم داشت که احیاناً پاره‌ای از کلیدواژه‌های غیرمرتبط را کنار بگذاریم. در این حالت از قسمت «بدون این واژگان»^{۵۳} بهره می‌بریم. در استفاده از ابزارهای بازیابی اطلاعات

48. Uniform Resource Locator (URL)

49. Address bar

50. With the exact phrase

51. With all of the words

52. With at least one of the words

53. Without the words

می‌توانیم بنابر نیاز خود محدودیت‌هایی را در جست‌وجو قائل شویم. یکی از این محدودیت‌ها، به مقولهٔ زبان مدارک باز می‌گردد. چنانچه در جست‌وجوی خود اسنادی به زبان خاص را مدنظر داریم، بهتر است که در فیلد مربوط به زبان، این محدودیت را قائل شویم. در صورتی که محدودیتی را در نظر نگیریم، اسناد بازبایی شده به زبان‌های مختلف خواهد بود. واضح است که در محیط اینترنت بیشتر اسناد و مدارک به زبان انگلیسی است، و معمولاً اکثر مدارک بازبایی شده به این زبان می‌باشد. در برخی رایانه‌ها نصب برنامهٔ نرم‌افزاری حروف ویژهٔ یک زبان به استفادهٔ بهتر از مشخصات نگارشی آن زبان کمک می‌کند. نظیر استفاده از زبان‌های آلمانی و فرانسه که برخی از حروف آنها با نگارش حروف انگلیسی متفاوت است.

در برخی موتورهای جست‌وجو امکان بازبایی مدارک بر اساس قالب ذخیره‌سازی اطلاعات وجود دارد. این امکان نیز از سودمندی خاصی برخوردار است؛ زیرا ما در پاره‌ای از جست‌وجوهای خود به اطلاعاتی نیاز داریم که نوع قالب آن تعیین‌کننده است. به عنوان مثال چنانچه اطلاعات مورد نیاز خود را به شیوهٔ آموزشی و براساس چارچوب ارائهٔ مطلب در نظر داشته باشیم، قالب پاورپوینت^{۵۴} بسیار مناسب است. زیرا اطلاعات ارائه شده از طریق این برنامه، معمولاً به صورت گرافیکی و با طراحی مناسب اسلایدها، همراه با توضیحات مهم و

کوتاه است. بنابراین می‌توان در جست‌وجو قالب خاصی را معین ساخت. بر همین اساس امکان بازبایی منابع دیداری یا شنیداری نظیر ویدئو، عکس یا قطعات موسیقی نیز وجود دارد.

در جست‌وجوهای صورت گرفته در اینترنت، برخی اوقات به اطلاعاتی دست می‌یابیم که از نظر روزآمدی نیاز اطلاعاتی ما را برآورده نمی‌سازند. در بعضی مواقع، جدید بودن اطلاعات به تولید و ارائهٔ آن در مدت زمانی بسیار کوتاه‌تر از زمان دستیابی ما به اطلاعات بستگی دارد^{۵۵}. استفاده از امکان محدودسازی زمان در ابزارهای کاوش اینترنت، این امکان را به ما می‌دهد که در بازبایی اطلاعات مورد نیاز، به اسناد بسیار جدید و روزآمد دسترسی داشته باشیم. در آن صورت اغلب با انتخاب گزینه‌های نزدیک زمانی به جست‌وجوی اطلاعات می‌پردازیم. آشکار است که در این شیوه بسیاری از اطلاعات مرتبط را از دست می‌دهیم، ولی در عوض به منابعی دست می‌یابیم که از نظر معیار زمانی کاملاً جدید و تازه هستند.

از دیگر امکانات دستیابی به اطلاعات، تعیین قرار گرفتن کلیدواژه‌های ما در موقعیت‌های گوناگون در وب‌سایت‌هاست. قرار گرفتن آن در نشانی مکان منابع یکسان، عنوان، پیوندها، متن، و یا صرف نظر از معین ساختن مکان دقیق کلیدواژه، و صرف وجود آن در وب‌سایت، توانایی بازیافت اطلاعات بهتری به ما می‌دهد. راهبرد آن نیز بر اساس عام و خاص بودن، نوع کاربرد، بسامد استفاده،



اصطلاح یا واژه علمی و حرفه‌ای، پیوندهای موازی و برخی معیارهای دیگر از بازیابی کم به سوی بیشتر در نوسان است. همچنین در برخی موارد با تعیین دامنه فعالیت سایت‌ها (نظیر: آموزشی: edu، سازمانی: org، تجاری: com، و نظایر آن) می‌توان به جست‌وجوی اطلاعات پرداخت. این شیوه در پاره‌ای از موارد کاربرد مناسبی پیدا می‌کند. در مواقعی که ما قبلاً یک نشانی اینترنتی را دیده‌ایم و بخشی از آن را به خاطر داریم، یا در هنگام جست‌وجو برای به دست آوردن اطلاعات ممکن است که بتوانیم بخشی از نشانی مکان منابع یکسان را حدس بزنیم، که در آن صورت استفاده از فیلد مربوط یعنی نشانی مکان منابع یکسان مفید است.

امکان مناسب دیگری که در برخی از موتورهای جست‌وجو وجود دارد، محدوده جغرافیایی منابع است. معمولاً این ابزار به ما اجازه می‌دهد تا در حیطه جغرافیایی خاصی به جست‌وجو پردازیم. معین کردن منابع اطلاعاتی در محدوده مکانی مشخص، امکان به دست آوردن اطلاعات منطقه‌ای را افزایش می‌دهد. به عنوان مثال در موتور جست‌وجوی هات بات^{۵۶} در بخش جست‌وجوی پیشرفته می‌توانیم در سایت‌های با دامنه بازرگانی، آموزشی، دولتی، نظامی، شبکه‌ای، و سازمانی در منطقه آمریکای شمالی به جست‌وجو پردازیم.

استفاده از امکان بازیابی سایت‌های مشابه، در مواردی که سایت مناسبی را بازیابی کرده‌ایم، می‌تواند مؤثر باشد. در

برخی موارد یافتن سایت‌هایی که به این سایت پیوند داده‌اند برای استفاده از مطالبی که احتمالاً قرابت موضوعی یا مفهومی دارند، می‌تواند کمک خوبی به حساب آید. در پاره‌ای از جست‌وجوها برای به دست آوردن اطلاعات علمی، ممکن است با برخی نتایج مغایر با معیارهای اخلاقی مواجه شویم. در این صورت می‌توان با بهره‌گیری از امکان صافی^{۵۷} این مشکل را تا حدودی برطرف ساخت. مانند مطالعه درباره رفتار جنسی انسان که می‌تواند به طور ناخواسته به بازیابی برخی سایت‌های نامناسب منجر شود.

بعضی از موتورهای جست‌وجو این امکان را در اختیار کاربر قرار می‌دهند که با جست‌وجو در وب‌سایت‌های مورد نظر، به میزان محدودی از نتایج مربوط به یک سایت دسترسی داشته باشد. به عنوان مثال حداکثر تا دو صفحه از یک وب‌سایت در نتیجه بازیابی ظاهر شود، که در صورت نیاز از طریق همان صفحات می‌توان به جست‌وجوی بیشتر پرداخت. در مواردی نیز ممکن است کاربر تمایل داشته باشد که تمامی صفحات مرتبط در نتیجه جست‌وجو ظاهر شود، که در آن صورت می‌تواند اطلاعات بیشتری از یک سایت را مرور کند.

از دیگر شیوه‌های مناسب برای بهره‌مندی سودمند و مناسب از امکانات جست‌وجو، استفاده از راهنماهای تنظیم شده موجود در وب است که بهره‌گیری از آن، به انجام جست‌وجوهای بهتر منجر می‌شود. این‌گونه راهنماها جایگزین مهارت‌های جست‌وجو

تحلیل نتایج و ارزیابی

بازیابی اطلاعات با اتکا بر روش‌شناسی مبتنی بر ارزشیابی علمی کارآمد، در سال‌های اخیر مورد توجه قرار گرفته است. این روش‌شناسی کمک‌های بسیاری به ما کرده است، و اکنون نیز برای بهینه ساختن ساختارهای خود نیازمند توسعه و پژوهش است، زیرا ماهیت جست‌وجوی اطلاعات متحول شده است. به علاوه، این روش‌شناسی در سایر رشته‌ها نیز، بدون درک عمیق از شالوده و محدودیت‌های آن، با پذیرش روبه‌رو شده است. بیشتر پژوهشگران بازیابی اطلاعات از سنجش‌های کلاسیک در ارتباط با جامعیت و مانعیت برای ارزشیابی نظام‌ها و روش‌های بازیابی اطلاعات بهره می‌گیرند، ولی در تحلیل نتایج ضرورت دارد مقوله آموزش مهارت‌ها در نظر گرفته شود تا با ارزیابی آن بتوان به نتایج بهتری دست یافت.

در مهارت‌های آموزش اینترنت دو بخش عمده وجود دارد: فراهم‌آوری اطلاعات و ارزیابی. در فراهم‌آوری اطلاعات، ما کوشش می‌کنیم تا کسانی را که توانایی آموختن مهارت‌های ضروری برای استفاده از فناوری را دارند، تقویت کنیم، نظیر دانستن ماهیت اینترنت (از چه چیزی تشکیل شده است)، آگاهی از چگونگی کارکرد مرورگر، و درک پیام‌های اینترنتی. ما همچنین تلاش می‌کنیم تا افرادی را که به شیوه‌ای مؤثر مطالبی را در اینترنت پیدا می‌کنند و نیز اطلاعاتی را که

نیستند، ولی تا حدودی می‌توان از راهنمایی‌های آنان بهره جست. به‌طور نمونه برخی از این راهنماها با طرح جدولی، نوع نیاز اطلاعاتی را با طرح سؤالات خاص مشخص کرده و راهبرد جست‌وجوی مناسب آن را از طریق موتورهای جست‌وجو و یا وب‌سایت‌هایی ارائه می‌دهند که این‌گونه اطلاعات را بهتر بازیابی می‌کنند. این‌گونه ارشادات استفاده از وب، می‌تواند زمینه‌ای مطلوب برای بهره‌گیری از مهارت‌های ضروری بازیابی اطلاعات دیجیتال فراهم آورد و به نوعی در شناخت بهتر موتورهای جست‌وجو مؤثر واقع شود. برای مثال می‌توان در سایت نودل تولز^{۵۸} چنین اطلاعات راهنمای به‌دست آورد (۴).

به‌صورت کلی می‌توان گفت که در بازیابی اطلاعات مناسب، استفاده از نشانگرها و علامت‌های خاص تأثیر زیادی دارد. قرار دادن عبارت مورد نظر داخل گیومه، استفاده از عملگرهای بولی، نشانگر همجواری^{۵۹}، نزدیک‌یابی^{۶۰}، استفاده از علامت جانشین^{۶۱}، استفاده از علامت‌های پرانتز، کروشه، و برخی امکانات دیگر در بازیابی اطلاعات مرتبط، سودمند است. همچنین باید توجه داشت که هر یک از علامت‌ها و فیلدهای ارائه شده، در جای خود، به‌درستی و بر مبنای توصیه‌های پیشنهادی مورد بهره‌برداری قرار گیرد. معمولاً می‌توان با مطالعه بخش کمکی هر موتور جست‌وجو از امکانات تعبیه شده برای ساختن یک پرس‌وجوی بهینه استفاده کرد.

58. <http://www.noodletools.com>

59. Adjacency

60. Proximity

61. Wild card

آنها به دست آورده‌اند و مورد ارزیابی قرار می‌دهند، حمایت کنیم (۱۲: ۲۱۶).

برای سنجش معیارهای ارزیابی پنج معیار بنیادی وجود دارد: صحت، اعتبار، عینیت^{۶۲}، رواج، و شمول^{۶۳}. معمولاً به این پنج معیار ارزشیابی دو فن دیگر افزود می‌شود که عبارت است از: دستیابی، و هدایت و طراحی^{۶۴} (۸: ۹۰). در ادامه به توضیح مختصر درباره هر یک از عوامل نامبرده می‌پردازیم.

«صحت» در معنای اولیه یعنی بدون اشتباه بودن. از آنجاکه در محیط شبکه هر کسی می‌تواند به راحتی اطلاعاتی را ارائه دهد، صحت از اهمیت بالایی برخوردار است. هنگامی که با آثار یک نویسنده ناشناس و یا یک پدیدآورنده سازمانی ناشناس در شبکه روبه‌رو می‌شویم، توانایی تشخیص درستی مطالب ارائه شده در ارزیابی بسیار مهم است. یک بخش از معیارهای ارزیابی در این قسمت به غلط‌های آشکار و اشتباهات صریح برمی‌گردد، ولی بخش دیگر که دشوارتر است به مشکلات ساختاری، مفهومی یا معنایی مطالب مربوط می‌شود.

ارزیابی «اعتبار» بسیار مشکل‌تر است، به خصوص اگر ما با موضوع بازیابی شده آشنایی زیادی نداشته باشیم. شاید بتوان گفت که معیار اصلی ارزیابی همین اعتبار است.

در «عینیت»، یکی از مباحث اصلی به تعصبات و سوگیری مربوط می‌شود و در

«رواج» ارتباط بیشتر با روزآمدی است. همچنین معیار «شمول» در ارتباط با درصدی از مردم است که به رسانه‌های ارتباطی دسترسی دارند. علاوه بر آن توانایی علمی نویسنده در مبحث ارائه شده و ارتباط رشته تحصیلی و تخصصی او با موضوع مورد توجه قرار می‌گیرد. افزون بر آن، پرداختن به این مطلب که پوشش موضوعی و ژرف نگری در مفهوم تا چه میزان بیان شده، حائز اهمیت است.

«دستیابی» نیز به سهولت دسترسی و فراهم آمدن اطلاعات سودمند در جست‌وجو ارتباط دارد. ولی «هدایت» به عنوان عاملی مهم و بنیادی در بحث ارزیابی مطرح می‌شود.

هدایت در یک صفحه وب برای راهنمایی و کمک به ما جهت اطلاع از محتوای آن وجود دارد. این امکان باید بتواند به چنین پرسش‌هایی پاسخ دهد: من کجا هستم؟ چگونه می‌توانم اینجا را پیدا کنم؟ چگونه می‌توانم از این صفحه خارج شوم؟ چه کسی می‌تواند به من کمک کند؟ هدایت باید همیشه برای عرضه به بازار، جهت آنچه که طراحی شده است، مناسب باشد. همچنین باید سازگار و قابل درک باشد. عامل هدایت نیز باید ترجیحاً سلسله مراتبی و به راحتی قابل تشخیص باشد. علاوه بر آن لازم است که عملکرد روشنی داشته و از ساختار سریع و قابل دسترس در هر صفحه برخوردار باشد.

62. Objectivity

63. Coverage

64. Navigation & design

برخی پیوندها باید به صورت دائمی و قابل دسترس در هر صفحه وجود داشته باشند مانند جست‌وجو، تماس گرفتن، و کمک کردن (۸: ۱۰۱) همچنین مدت زمان اجرای برنامه^{۶۵}، پرسش، و اصطلاحات طرح شده از عوامل مؤثر در سنجش و تحلیل نتایج به‌دست آمده به شمار می‌آیند.

از دیگر عوامل مهم در ارزیابی و تحلیل نتایج، سوگیری است. این عامل در محیط اینترنت و در حیطه‌های موضوعی متعدد وجود دارد. برخی اوقات دلیل آن را می‌توان در انگیزه‌ها و یا روش‌شناسی فراهم‌آوردگان اطلاعات جست‌وجو کرد. این پدیده در سایت‌های اینترنتی یکی از عوامل اصلی خدشه‌دار شدن اطلاعات به شمار می‌آید. بررسی مختصات سایت‌ها از جنبه‌های گوناگون می‌تواند ما را در تعیین صحت اطلاعات یاری رساند. اساساً سایت‌های آموزشی و پژوهشی در محیط وب، که معمولاً به مؤسسات و سازمان‌های علمی و حرفه‌ای وابستگی دارند، نسبت به سایت‌های دیگر از صحت و درستی بیشتری برخوردارند. سوگیری در اطلاعات به‌دست آمده در شبکه، یکی از معضلات قابل توجه برای استناد به چنین اطلاعاتی به شمار می‌آید. برای تعیین تعصب‌های علمی یا حرفه‌ای در میان متون ارزیابی شده، بهتر است که از یک راهبرد بدون بهره‌جست. با طرح پرسش‌هایی می‌توان بهتر به مقصود رسید. سؤالاتی که در پی می‌آید برای

رسیدن به پاسخ سودمند است: ایده یا عقیده مطرح شده به چه کسی تعلق دارد؟ این اندیشه شامل چه اجزاء و ویژگی‌هایی است؟ شمول اطلاعات ارائه شده به چه میزان است؟ اطلاعات ارائه شده تا چه حد با مبانی علمی پیش از خود همبستگی و قرابت دارد؟ همچنین باید توجه داشت که در بررسی این پدیده، سوگیری از سوی جست‌وجوکننده با سوگیری اطلاعات بازیابی شده کاملاً متفاوت است و در بررسی نیز دارای ریشه‌ها و زمینه‌های جداگانه است. تعیین میزان سوگیری و صحت اطلاعات را می‌توان به شیوه تخمینی براساس بررسی سایت و فراهم‌آورنده اطلاعات مورد نظر مشخص ساخت.

در ارزیابی بازیابی اطلاعات در محیط وب، هرگز نباید فراموش کنیم که بخشی از این اطلاعات در قسمتی از وب قرار دارد که دسترس‌پذیری آن بسیار اندک است. بر این اساس کاوش اطلاعاتی به رغم رعایت تمامی جنبه‌های فنی و حرفه‌ای، در برخی موارد با شکست روبه‌رو می‌شود؛ و یا اینکه بازیابی اطلاعات با ضعف‌های چشمگیری روبه‌روست. حتماً باید توجه داشته باشیم که ما علاوه بر وب مرئی^{۶۶} با بخشی به نام وب نامرئی^{۶۷} نیز روبه‌رو هستیم.

وب مرئی آن بخش از وب است که می‌تواند مورد جست‌وجو قرار گیرد، و بنابراین با استفاده از موتورهای جست‌وجو و راهنماها قابل مشاهده است. به هر حال

محتوای پایگاه داده‌ها به ندرت در نتایج به دست آمده در موتورهای جست‌وجو ظاهر می‌شود. جست‌وجوگرهای خودکار^{۶۸} موتورهای جست‌وجو نمی‌توانند به جداول اطلاعاتی پایگاه‌های داده وارد شده و اطلاعات آن را استخراج کنند. بنابراین محتوای این پایگاه‌های داده برای آنها وب نامرئی محسوب می‌شوند. این صفحات که نمی‌توانند در نمایش نهایی ظاهر شوند، بخشی از وب نامرئی به‌شمار می‌آیند. به دلیل این واقعیت که بسیاری از پاسخ‌ها ممکن است که با انجام یک جست‌وجوی ساده از طریق یکی از موتورهای جست‌وجوی اصلی بازیابی شود، ما گرایش داریم تا فراموش کنیم که برخی چیزهای مهم به اشکال گوناگون وجود دارد که وب نامرئی، وب پنهان^{۶۹}، یا وب عمیق^{۷۰} نامیده می‌شود (۸: ۷۹). علاوه بر آن شیوه‌های ورود اطلاعات در محیط وب، طراحی وب‌سایت‌ها، استفاده از قلم‌های نامناسب، رنگ‌ها، برچسب‌ها، و سایر عواملی که به هر علت، باعث می‌شوند تا موتورهای جست‌وجو نتوانند در شناسایی اطلاعات آن موفق شوند، در حیطه وب نامرئی قرار می‌گیرند.

همچنین باید توجه داشت که برخی از انواع اطلاعات در محیط اینترنت کمتر وجود دارد. بسیاری از متون که در کتابخانه‌ها به راحتی قابل بازیابی است، در وب به

سختی یافت می‌شود. انواع منابع اطلاعاتی بر اساس قالب تعیین شده در مکان‌هایی غیر از محیط شبکه با سهولت بیشتری قابل دستیابی است. افراد در مواردی تمایل ندارند که پاره‌ای از اطلاعات و نوشته‌های خود را در عرصه شبکه قرار دهند. افراد باتجربه در ارتباط با دریافت برخی از منابع اطلاعاتی معمولاً به وب مراجعه نمی‌کنند. اطلاعات شخصی افراد، دست‌نوشته‌ها، یادداشت‌های خصوصی، و اسنادی نظیر آن کمتر در محیط شبکه وجود دارد. تشخیص اینکه چه نوع منابعی را نمی‌توان از طریق وب به دست آورد، بیشتر حاصل تجربه و کارکردن با ابزارهای جست‌وجوی اطلاعات پیوسته است. در همین رابطه بیستروم و یارولین^{۷۱} انواع منابع اطلاعاتی را به سه دسته تقسیم می‌کنند:

- واقع‌گرا^{۷۲}: که شامل ثبت‌کننده‌ها به صورت فهرست‌ها و فایل‌های دستی و رایانه‌ای و نیز پایگاه‌های داده تجاری می‌شود.
- مسئله‌گرا^{۷۳}: شامل دست‌اندرکاران، نظیر کسانی که پیشنهادی ارائه می‌دهند، یا با فعالیت اجرایی تأثیر می‌گذارند؛ و اسناد رسمی مانند برنامه زمانی، صورت جلسه‌ها، نامه‌ها، درخواست‌ها، یادداشت‌ها، نقشه‌ها، و اسناد برنامه‌ریزی منتشر نشده می‌شود.
- همگانی^{۷۴}: شامل متخصصان مانند

68. Spiders

69. Hidden Web

70. Deep Web

71. Byström and Järvelin

72. Fact-oriented

73. Problem-oriented

74. General-purpose

همکاران مطلع؛ متون و منابع مانند کتاب‌ها، گزارش‌ها، مجلات، روزنامه‌ها؛ مجموعه‌های شخصی چون یادداشت‌های فردی و محاسبات؛ و نظایر آن می‌شود (۳).

یکی از مهارت‌های سودمند اطلاع‌یابی، تشخیص این نکته است که چه مطالبی در وب کمتر به دست می‌آید. شاید همان‌طور که مهارت کتاب نخواندن، یا کنار گذاشتن منابع غیرمفید، اهمیت زیادی در توانایی ما برای مطالعه و خواندن دارد؛ مهارت جست‌وجو نکردن منابع یافت نشدنی در وب، به اندازه مهارت‌های جست‌وجوی اطلاعات در شبکه از ارزش و اهمیت برخوردار باشد. به تعبیر دقیق‌تر باید بدانیم چه نوع منابع و اطلاعاتی را می‌توان از وب به دست آورد و چه نوع منابع یا اطلاعاتی در وب وجود ندارد، یا احتمال بازیابی آن ضعیف است.

از سوی دیگر می‌توان گفت که بازیابی اطلاعات در وب تا حد زیادی به شیوه فعالیت خود ما به عنوان تولیدکنندگان و ارائه‌دهندگان اطلاعات بستگی دارد. به هر میزان که ما در وب‌سایت‌های خود از اصول طراحی صحیح و سازماندهی استاندارد بهره‌گیریم، به همان نسبت به بازیابی اطلاعات سودمند کمک کرده‌ایم. توصیه‌های زیادی برای رعایت اصول همکاری در وب ارائه شده، که رعایت آنها برای همه کسانی که با این محیط سر و کار دارند، مفید و مؤثر است. همچنان که مولر (۸: ۲۳-۲۴) بیان

می‌کند رعایت برخی قواعد سبب می‌شود تا دیدگاه واقعی‌تری نسبت به شبکه و کاربران آن داشته باشیم. وی سپس بیان می‌کند که قوانین مهم رفتار اطلاع‌یابی در محیط اینترنت که در سایت گوگل کاربرد دارد و ما می‌توانیم از آن استفاده کنیم به این شرح است:

- هرگز فراموش نکنیم کسانی که آن سوی شبکه قرار دارند انسان هستند،
- نپنداریم که اشخاص به جای سازمان خود صحبت می‌کنند،
- نسبت به آنچه که درباره دیگران می‌گویید دقت کنید،
- گزیده گویند،
- نیک نگارش کنید،
- از توصیف موضوعی مناسب استفاده کنید،
- درباره مخاطبان خود به خوبی اندیشه کنید،
- نسبت به رضایت و ناخرسندی (دیگران) دقت کنید،
- فقط یک پیام ارسال کنید،
- به جای ارسال پیگیری از پست نامه استفاده کنید،
- پیگیری خود را به صورت خلاصه بیان کنید،
- از ارجاعات مناسب بهره بگیرید،
- ارسال پیام برای گروه خبری^{۷۵} یا پاسخ را برجسته سازید،
- طول خط را محدود سازید و از

به‌کارگیری کاراکتر کنترل پرهیزید،
- از یوزنت^{۷۶} به عنوان یک رسانه تبلیغی
استفاده نکنید، و
از ارسال مطلب به گروه‌های خبری
چندگانه^{۷۷} اجتناب کنید.

همچنین در ارتباط با استفاده از واژگان
در بازیابی مؤثر از وب، برخی از نتایجی که
بیتس در مقاله خود پیشنهاد می‌کند به این
قرار است:

- از رده‌بندی خردشونده به‌جای
سلسله‌مراتبی استفاده کنید،

- درک آنچه که رده‌بندی اطلاعات را
مشخص می‌کند و واژگانی از معادل‌های
جهان فیزیکی را توسعه دهید و از واژه
گمراه‌کننده هستی‌شناسی^{۷۸} استفاده نکنید، و
- از واژگان خاص طراحی شده در [نظام]
بازیابی اطلاعات، به جای واژگان زبان عمومی
استفاده کنید (۲).

از عوامل دیگر در بحث ارزیابی - که
اهمیت زیادی نیز دارد - باید به استفاده
از مدل‌های مفهومی^{۷۹} در جست‌وجوهای
نظام‌مند اشاره کرد. در مطالعه برخی پدیده‌ها،
باید آنها را در شرایط و ابعاد گوناگون
مورد بررسی قرار داد. بنابراین نمی‌توانیم
جست‌وجوی اطلاعات را فقط از جنبه علمی
مورد مذاقه قرار دهیم، بلکه باید فعالیت
سایر متخصصان و افراد تازه‌کار نیز مد نظر
قرار گیرد. در این‌گونه مدل‌ها مراحل شروع

کار جست‌وجو از ابتدای فعالیت با انتخاب
کلیدواژه‌های آن شروع شده، و با طی
فرایند تفحص اطلاعات، با دستیابی به نتایج
رضایت‌بخش به انتها می‌رسد.

از جنبه نظری ما به استفاده از مدل‌های
مفهومی نیاز داریم تا بتوانیم در جست‌وجو و
بازیابی اطلاعات، و نیز انجام برخی اهداف
پژوهشی از آنها بهره‌گیریم. در توصیف و
شرح روابط بین عوامل جست‌وجو، مدل‌های
مفهومی این فرایند را به نیکوترین شکلی
نمایش می‌دهند. پاره‌ای از متخصصان
اطلاع‌رسانی در موضوع طراحی و کاربرد
این مدل‌ها مطالعاتی انجام داده‌اند. به عنوان
نمونه در مدل بیستروم و یارولین^(۳)، عواملی
چون مجاری استفاده شده و نیز دلایل استفاده
مطرح می‌شود. آنان همچنین عوامل موقعیتی
را در ارتباط با نقش آموزش و تجربه، و نیز
ارزیابی منابع استفاده شده مورد توجه قرار
داده‌اند.

در چارچوب مدل رفتاری الیس^{۸۰} که
توسط ویلسون^{۸۱} توضیح داده شده، با شروع
فعالیت سه عملکرد مرورکردن، رابطه پیوسته،
و هدایت‌کردن آغاز می‌شود. در این مدل،
سپس، مراحل رفتاری تعیین افتراق، استخراج
اطلاعات، ارزیابی، و به‌دست آوردن اطلاعات
مناسب قرار می‌گیرد (۱۴).

در مدل اینگورسن^{۸۲} در تبادل بین
پرس‌وجو، درخواست، و مدل‌ها رابطه‌ای

76. Usenet

77. Multiple newagroup

78. Ontology

79. Conceptual models

80. Ellis

81. Wilson

82. Ingwersen

متقابل وجود دارد. رابط کاربر، فضای ادراکی استفاده‌کننده و محیط سازمانی/ اجتماعی نیز از عوامل تأثیرگذار به شمار می‌آیند (۵).

چارچوب یک مدل مفهومی بر اساس معیارهای مؤثر بر آن تدوین می‌شود. در استفاده از مدل‌ها بهتر است که به انواع مختلفی از آنان نظر داشت و جنبه‌های خاص هر یک را مورد بررسی و مطالعه قرار داد. در این میان سنجش مدل‌ها از اهمیت زیادی برخوردار است.

هنگامی که دو مدل مفهومی رقیب با هم مقایسه می‌شوند، معیارهایی که به شیوه عملی برای سنجش امتیاز آنها استفاده می‌شود، عبارتند از:

• سادگی^{۸۳}: در شرایط مساوی، سادگی بهتر است؛

• صحت^{۸۴}: درستی و صراحت در مفاهیم رضایت‌بخش است؛

• دامنه^{۸۵}: دامنه گسترده‌تر بهتر است، زیرا در شرایط مساوی [موضوع] خاص‌تر را پوشش می‌دهد؛

• توان نظام‌مندی^{۸۶}: توانایی سازماندهی مفاهیم، روابط و داده در شیوه نظام‌مند معنادار رضایت‌بخش است؛

• توان تبیینی (توضیحی)^{۸۷}: توانایی توضیح و پیش‌بینی پدیده‌ها رضایت‌بخش است؛

• پایایی (قابلیت اعتماد)^{۸۸}: توانایی، درون‌حیطه‌ای از مدل‌ها، برای فراهم‌آوردن ارائه درست در امتداد حیطه کاملی از موقعیت‌های ممکن؛ و

• سودمندی^{۸۹}: توانایی پیشنهاد مسائل برای حل، و فرضیه‌هایی برای آزمایش رضایت‌بخش است (۶).

نتیجه‌گیری

ارائه راهبردهای متفاوت برای جست‌وجو در محیط وب، بستر مناسبی برای فراهم‌آوردن اطلاعات مرتبط با پرسش کاربر ایجاد می‌کند. تجارب و روش‌های به‌دست آمده از تحقیقات پژوهشگران، به بهبود بهره‌گیری استفاده‌کنندگان از محیط اینترنت می‌انجامد. به‌کارگیری این روش‌ها به کاربران این توانایی را می‌دهد که بتوانند اطلاعات باارزشی را از درون داده‌های بسیار کسب کنند. کاربرد هریک از رویکردهای ارائه شده توسط متخصصان، حاصل مطالعات و تجربیات گسترده‌ای است که می‌تواند به دستیابی‌های موفقیت‌آمیز منجر شود. این راهبردها توان ایجاد زمینه مناسب برای جست‌وجو و نیز بازیابی اطلاعات مورد نیاز استفاده‌کنندگان را به همراه دارد. می‌توان گفت که شناخت و انتخاب هریک از روش‌های جست‌وجو به نوع و قالب اطلاعاتی جست‌وجوگر بستگی دارد.

83. Simplicity

84. Accuracy

85. Scope

86. Systematic power

87. Explanatory power

88. Reliability

89. Fruitfulness

کوشش تمامی این روش‌ها، در نهایت به اخص ساختن جست‌وجو و دستیابی به اطلاعات مرتبط منجر می‌شود. روش‌های سودمند جست‌وجو با طراحی پرس‌وجوی مناسب ارتباط نزدیکی دارند. تلاش اصلی برای پدیدآوردن یک پرس‌وجو، به نحوه شناسایی صحیح نیاز اطلاعاتی کاربر، و بازگرداندن آن در قالب یک عبارت جست‌وجوی مناسب برای فرایند کاوش باز می‌گردد. در این میان عوامل بسیاری دخالت دارند. تحلیل پرسش، اخص‌سازی، صحت، رعایت توازن میان جامعیت و مانعیت، تعداد واژگان به‌کار رفته، خوشه‌بندی پرسش‌ها، تحلیل گفتار، و فرمول‌بندی دوباره پرس‌وجو از جمله عواملی است که در ایجاد یک عبارت پرسش مناسب تأثیری بسزا دارد. برای عملی ساختن فرایند یک عبارت جست‌وجو در کاوش وب، باید از فنون مؤثر کاوش بهره گرفت. فنون کاوش رعایت نکات کاربردی و کلیدی در استفاده از مرورگرهای اینترنت است که در موتورهای جست‌وجو از آنها استفاده بهینه‌ای می‌شود. استفاده از نشانگرها، علامت‌های خاص، فیلدها، و پیوندهای تعبیه شده در موتورهای جست‌وجو، می‌تواند به بازیابی اطلاعاتی منتهی شود که ارتباط بیشتری با نیاز کاربر دارد. به بیان دیگر فنون کاوش، کاربردپذیری عبارت جست‌وجو را افزایش می‌دهد. تمامی مهارت‌های لازم برای کسب اطلاعات در محیط وب که منجر به بازیابی اطلاعات مرتبط می‌شود، نیازمند مهارت تکمیلی دیگری است که «ارزیابی

اطلاعات فراهم‌آمده» نام دارد. استفاده از اطلاعات به‌دست آمده در محیط وب در صورتی می‌تواند در فرایند آموزش و پژوهش به‌کار گرفته شود که از اعتبار لازم برخوردار باشد. شناخت عوامل و نشانه‌های سوگیری و تعصب در اطلاعات موجود در وب، خدشه و نادرستی در اطلاعات، سست بودن پایه‌های بنیادی اطلاعات فراهم‌آمده، و نیز توان ارزشیابی اطلاعات با معیارهای علمی شناخته شده، به کاربران کمک می‌کند تا در انتخاب اطلاعات به‌دست آمده از وب، با دقت و ظرافت بیشتری عمل کنند. در ارزیابی اطلاعات موجود در محیط وب توان تشخیص اطلاعات مناسب برای استفاده، به اندازه قدرت تعیین انواع اطلاعات نامناسب و یا اطلاعاتی که در محیط وب نمی‌توان به‌دست آورد، از اهمیت برخوردار است. استفاده مؤثر و کارآمد از اینترنت - و به‌صورت خاص محیط وب - با ترکیب توانایی‌های چندگانه میسر است. شناخت و به‌کارگیری راهبردهای جست‌وجو، طراحی پرس‌وجوی مناسب، بهره‌گیری از فنون کاوش مؤثر، به‌علاوه توان ارزشیابی اطلاعات به‌صورت مستمر، همگی سبب استفاده بهینه از وب خواهد شد.

منابع

۱. پائو، میراندا لی. مفاهیم بازیابی اطلاعات. ترجمه اسدالله آزاد و رحمت‌الله فتاحی. مشهد: دانشگاه فردوسی، مؤسسه چاپ و انتشارات، ۱۳۷۹.
2. Bates, Marcia J. "After the Dot-

8. Müller, Jeanne Froidevaux. *A librarian's guide to the internet: searching and evaluating information*. Oxford: Chandos Publishing, 2003.

9. Navarro-Prieto, Raquel; Scaife, Mike; Rogers, Yvonne. "Cognitive strategies in WEB searching". In *5th Conference on Human Factors & the Web (Gaithersburg, Maryland, USA)*, 1995. [on-line]. Available: <http://zing.ncsl.nist.gov/hfweb/proceedings/navarro-prieto/> [19 Feb. 2006].

10. Pandey, R.C. *Information retrieval systems: a linguistic study*. Delhi, Abhijeet Publications, 2003.

11. Spink, Amanda; Xu, Jack L. "Selected results from a large study of Web searching: the excite study". *Information Research*, Vol.6, No.1 (2000). [on-line]. Available: <http://InformationR.net/ir/6-1/paper90.html>. [18 Feb. 2006].

12. Stacey, Alison; Stacey, Adrian. *Effective information retrieval from the internet: an advanced user's guide*. Oxford: Chandos publishing, 2004.

13. Wen, Ji-Rong; Nie, Jian-Yun; Zhang, Hong-Jiang. "Clustering user queries of a search engine". 2001. [on-line]. Available: <http://www10.org/drom/papers/368>. [18 Feb. 2006].

14. Wilson, T.D. "Models in information behaviour research". *Journal of*

Bomb: getting web information retrieval right this time". *First Monday*, Vol.7, No.7 (July 2002). [on-line]. Available at: http://www.firstmonday.org/issues/issue7_7/bates/. [25 Feb. 2006].

3. Byström, K.; Järvelin, K. "Task complexity affects information seeking and use". *Information Processing & Management*, Vol.31, No.2 (1995): 191-213.

4. "Information literacy: search strategies, Choose the Best Search for Your Information Need". 2006 [on-line]. Available: <http://www.noodletools.com/debbie/literacies/information/5locate/adviceengine.html>. [27Feb. 2006].

5. Ingwersen, P. "Cognitive perspectives of information retrieval interaction". *Journal of Documentation*, Vol.52, No.1 (1996): 3-50.

6. Järvelin, K.; Wilson, T.D. "On conceptual models for information seeking and retrieval research". *Information Research*, Vol.9, No.1 (2003): 163. [on-line]. Available: <http://InformationR.net/ir/9-1/paper163.html>. [18 Feb. 2006].

7. Kari, J. "Web information seeking by pages: an observational study of moving and stopping". *Information Research*, Vol,9 No.4 (2004): 183. [on-line]. Available: <http://InformationR.net/ir/9-4/paper183.html>. [25 Feb. 2006].

Documentation, Vol.55, No.3 (1999):
249-270.

تاریخ دریافت: ۱۳۸۴/۱۲/۱۷

