



معاونت تحقیقات و فناوری

جزوه مهارت های اطلاع یابی و جستجو

برگرفته از:

- ۱- کوشا، کیوان. ابزارهای کاوش در اینترنت. تهران: کتابدار.
- ۲- موغلی، علیرضا و علیچانی، رحیم. جستجوی اطلاعات علمی و پژوهشی در وب. تهران: چاپار.

فهرست مطالب

۴	چکیده
۴	مقدمه
۶	شبکه جهانی وب
۷	چرا باید به اطلاعات در شبکه جهانی وب اهمیت داد؟
۸	شبکه "وحشی" وب یا شبکه جهانی "انتظار"
۱۰	تقسیم بندی جستجو ها و انواع آن
۱۰	الف. جست و جوی منابع معلوم
۱۲	ب. جست و جوی منابع نامعلوم
۱۳	استراتژی جست و جو
۱۳	فرایند جست و جو
۱۵	ساختار ابزارهای باز یابی اطلاعات
۱۵	ویژگی های مشترک جستجو
۱۶	انواع جستجو
۱۶	جستجوی ساده و جستجوی پیشرفته
۱۶	جستجو از طریق عملگرهای بولی
۱۸	جستجو از طریق عملگر نزدیک یابی یا مجاورتی
۱۹	کوتاه سازی کلید واژه ها
۲۰	ریشه سازی
۲۰	جستجوی کلید واژه در عنوان صفحات وب
۲۱	جستجوی پیوندهای فرامتنی در صفحات دیگر
۲۲	جستجوی حوزه سایت ها
۲۲	جستجوی عبارتی
۲۲	جستجوی میدانی
۲۳	جستجوی زبان طبیعی
۲۴	جستجوی استنادی
۲۶	جستجوی فرایببندی
۲۶	محدودیت زبانی

۲۶.....	حساسیت به حروف بزرگ و کوچک
۲۷.....	محدود کردن زمانی
۲۷.....	توانایی های مرور
۲۸.....	رتبه بندی نتایج جست و جو
۲۹.....	چگونگی انجام یک جستجوی کارآمد
۳۰.....	خلاصه مطالب
۳۱.....	نکات کلیدی

چکیده

هدف از این تحقیق ارتقاء سطح مهارت کاربران در زمینه جستجو در وب می باشد. از این رو هدف از این کلاس آشنایی بیشتر پرسنل با مهارت ها و استراتژی های جستجو می باشد. امروزه، بدلیل گستره وسیع کاربرد وب سایت ها و تنوع فراوان آن ها برای جستجو مطالب و دستیابی به نتایج مطلوب مورد نظر و همچنین برای سهولت جستجو و دسترسی به اطلاعات در اینترنت، امکاناتی تحت عنوان "ابزارهای کاوش اینترنت" توسعه یافتن، از این رو بررسی مفاهیم و مهارتهای اطلاع یابی و همچنین چگونگی انتخاب استراتژی های مختلف جهت جستجو می تواند کاربران را در زمینه راهنمایی و هدایت نماید. بنابراین آگاهی و آموزش در مورد این بیماری اهمیت بسیار بالایی دارد.

مقدمه

بشر همواره به دنبال کشف پدیده های اطراف و افزایش آگاهی خویش نسبت به آنها بوده است. انسان ها در خلال زندگی، مجبور به انجام فعالیت های مختلفی هستند که بدون شناخت و درک جهان پیرامون خود، امکان پذیر نیست.

برای دستیابی به سطح قابل قبولی از شناخت و درک پدیده ها، به "اطلاعات" نیاز است و "اطلاع یابی"، همواره به عنوان یکی از بنیادی ترین نیازهای بشر برای توسعه و پیشرفت مطرح بوده است. این نیاز در کلیه تمدن های بشری قابل مشاهده و بررسی است. در تمدن های کهن، دانشمندان و فیلسوفان که قشر خاصی از جامعه بودند، استفاده کنندگان اصلی اطلاعات به شمار می آمدند. اما امروزه، این حد و مرز به طور کامل برداشته شده و اطلاع یابی به عنوان یکی از اساسی ترین نیازهای افراد در جوامع پیشرفته یا به تعبیر دیگر "جوامع اطلاعاتی" مطرح است.

اطلاع یابی، خود مستلزم فراگیری مهارت ها و داشتن ابزارهای خاصی به منظور جستجو و دسترسی به اطلاعات مورد نظر است. به این ترتیب، ابتدا نظام های ذخیره و بازیابی اطلاعات نوشتاری مانند کتابشناسی ها، نمایه نامه ها و چکیده نامه ها و سپس نظام های الکترونیکی نظیر پایگاه های اطلاعاتی کتاب شناختی و تمام متن به منظور دسترسی هر چه بهتر، موثرتر و سریعتر به اطلاعات توسعه یافتند. ظهور شبکه های پیوسته و فناوری دیسک فشرده، تحولی بزرگ در فرایند ذخیره و بازیابی اطلاعات و نیز تولید و انتشار منابع اطلاعاتی علمی و فنی در قالب دیجیتال پدید آورد. از این رو، با توجه به قابلیت ها و امکانات نوین جستجو و بازیابی اطلاعات در منابع و پایگاه های اطلاعاتی

الکترونیکی، داشتن دانش و مهارت های ویژه برای استفاده موثر از آنها تحت عنوان " مهارت های اطلاع یابی" بیش از گذشته ضروری شد.

کتابداران و متخصصان اطلاع رسانی که وظیفه اصلی در دسترس قراردادن اطلاعات مورد نیاز استفاده کنندگان را بر عهده داشتند، در گذشته ای نه چندان دور تقریباً تنها قشری از جامعه به شمار می رفتند که دانش و مهارت لازم را برای جستجو در این گونه منابع داشتند. همچنین، کتابخانه ها و مراکز اطلاع رسانی تنها مکانهایی محسوب می شدند که در آنجا منابع و پایگاه های اطلاعاتی وجود نداشت و آنها اغلب ترجیح می دادند به کتابخانه ها یا مراکز اطلاع رسانی مراجعه کنند و از کتابداران یا متخصصان اطلاع رسانی برای دسترسی به اطلاعات مورد نظرشان کمک بگیرند. بدیهی است در چنین شرایطی، تعداد نسبتاً کمی از استفاده کنندگان نیاز به فراگیری اصول و مهارت های اطلاع یابی را در خود احساس می کردند.

ظهور اینترنت و سپس رشد و گسترش روزافزون " شبکه جهانی وب "، تغییر و تحول شگرفی در فرایند اطلاع یابی به وجود آورد و باعث شد که تعداد بسیار زیادی از استفاده کنندگان برای اولین بار از خانه یا محل کار خود به تنهایی از طریق کامپیوتر شخصی، امکان جستجو و دسترسی به حجم زیادی از اطلاعات را پیدا کنند. از یک طرف، استفاده از محیط گرافیکی شبکه جهانی وب، بسیار جذاب و ساده به نظر می رسید و از سوی دیگر، قیمت کامپیوترهای شخصی و هزینه اتصال به اینترنت دائماً کاهش یافت. همین امر باعث شد تا بسیاری از افراد بتوانند جستجوی اطلاعات در این محیط دیجیتالی را تجربه کنند. به این ترتیب، برخلاف گذشته که استفاده کنندگان، نیاز کمتری به فراگیری مهارت های اطلاع یابی در خود احساس می کردند و اغلب از کتابداران یا متخصصان اطلاع رسانی برای دسترسی به اطلاعات کمک می گرفتند، برای اولین بار مجبور شدند که بدون حضور یک واسط انسانی (کتابداران و متخصصان اطلاع رسانی) و داشتن دانش کافی درباره اصول و مهارت های جستجو و بازیابی اطلاعات، به کاوش اطلاعات بپردازند.

ظهور چنین وضعیتی برای بسیاری از افراد، در حکم رهایی از قید و بندهای فرایند اطلاع یابی به شیوه سنتی محسوب می شد. زیرا آنها دیگر مجبور نبودند به کتابخانه ها مراجعه کنند و نیازهای اطلاعاتی خود را با کتابداران در میان بگذارند.

برای سهولت جستجو و دسترسی به اطلاعات در اینترنت، امکاناتی تحت عنوان " ابزارهای کاوش اینترنت " توسعه یافتند که استفاده بهینه و موثر از آنها مستلزم فراگیری دانش و مهارت های جدیدی بود. بتدریج، فراگیری اصول و مهارت های اطلاع یابی در اینترنت برای بسیاری از

جستجوگران اطلاعات اهمیت خاصی پیدا کرد و سپس با افزایش شتابناک حجم اطلاعات در وب، حیاتی شد.

به نظر می‌رسد که شبکه جهانی وب از یک سو به "کتابخانه دیجیتال" مبدل شده است که از طریق آن می‌توان به طیف وسیعی از منابع، پایگاه‌ها و خدمات اطلاع‌رسانی دست یافت و از سوی دیگر حجم بسیار زیادی از اطلاعات در این محیط وجود دارد که از اعتبار، صحت و کیفیت لازم برخوردار نیست. از این رو "سواد اطلاعاتی" که مجموعه‌ای از توانایی‌ها و مهارت‌ها برای جستجو، ارزیابی و استفاده موثر از اطلاعات است، به یکی از عناصر اطلاع‌یابی در اینترنت مبدل شده است. داشتن شناخت کافی از ابزارها و امکانات جستجو در شبکه جهانی وب از جمله مباحث مهمی است که نه تنها کتابداران و متخصصان اطلاع‌رسانی، بلکه بسیاری از استفاده‌کنندگان نهایی باید برای رویارویی با حجم فزاینده اطلاعات، از آنها آگاهی داشته باشند.

شبکه جهانی وب

شبکه جهانی وب در اوایل دهه ۱۹۹۰ در آزمایشگاه فیزیک ذرات اروپا واقع در سوئیس توسط تیم برنرز لی اختراع شد. این شبکه پس از گذشت چند سال، رشد و گسترش شتابناکی به خود گرفت و به عنوان یکی از مهم‌ترین رسانه‌های قرن بیستم مطرح شد. همراه با افزایش حجم اطلاعات قابل دسترس از طریق این محیط، تعداد استفاده‌کنندگان آن نیز به طور فزاینده‌ای زیاد شد. از یک سو، تعداد استفاده‌کنندگان و از سوی دیگر، حجم اطلاعات قابل دسترس از طریق شبکه جهانی وب به طور شگفت‌آوری در حال افزایش است. به طوری که هیچ‌کس حتی مخترع وب هم ظهور چنین محیط پویا و در حال رشدی را تصور نمی‌کرد. در حال حاضر، بسیاری از سازمان‌ها، نهادها، انجمن‌ها، دانشگاه‌ها، شرکت‌های دولتی و خصوصی و حتی افراد معمولی دارای سایت وب هستند یا در نظر دارند سایت خود را بزودی در دسترس قرار دهند. شاید بتوان این طور استنباط کرد که هویت بسیاری از اشخاص حقیقی و حقوقی با داشتن صفحه یا سایت وب مستقل ارتباط مستقیم پیدا کرده است.

برخی عقیده دارند که رشد و توسعه شبکه جهانی وب، بر جوامع اطلاعاتی، تاثیرات مثبتی دارد و برخی دیگر بر این باورند که چنین پدیده‌ای، جوامع اطلاعاتی را با چالش‌ها و مشکلات جدیدی روبرو ساخته است. مشکل اصلی این نیست که اطلاعات کمی از طریق محیط وب قابل‌بازیابی است؛ بلکه بازیابی بیش از حد اطلاعات موجب شده است که در شبکه جهانی وب، برخلاف پایگاه‌های

اطلاعاتی کتابشناختی و تمام متن سنتی، نتوان به طور مناسب و کارآمد به اطلاعات مورد نظر دست یافت. در واقع، شبکه جهانی وب، فاقد یک ساختار نظام مند و منطقی برای ذخیره و بازیابی اطلاعات است و بر محتوای اطلاعات قابل دسترس از طریق آن هیچ گونه نظارت کیفی اعمال نمی شود. امروزه، شاهد آن هستیم که هر کسی از هر کجا می تواند به اینترنت متصل شود و به جستجوی انواع اطلاعات شامل متن، صوت، تصویر و ویدئو پردازد یا به تولید و اشاعه اطلاعات در هر زمینه که علاقه دارد مبادرت ورزد. اما جوامع اطلاعاتی همچنان با یک چالش مهم یعنی رشد شگفت آور اطلاعات و عدم نظارت کیفی بر آنها در محیطی غیر قابل کنترل، بی ثبات و بی نظم روبرو هستند. آنچه مسلم است، همیشه دسترسی به حجم زیادی از اطلاعات به معنای دسترسی بهتر به اطلاعات مورد نظر نیست. جوامع اطلاعاتی امروزی، نیاز بیشتری به اطلاعات مرتبط و با کیفیت دارند که دسترسی به این دو مقوله از طریق محیط وب به هیچ وجه ساده و آسان نیست.

چرا باید به اطلاعات در شبکه جهانی وب اهمیت داد؟

۱- منحصر بفرد بودن

برخی از منابع اطلاعاتی در شبکه جهانی وب، نمونه چاپی یا الکترونیک (روی دیسک فشرده یا شبکه های پیوسته) ندارند.

۲- روزآمد بودن

یکی از ویژگیهای مهم شبکه جهانی وب نسبت به رسانه های چاپی، انعطاف پذیری بیشتر آن برای روزآمدسازی اطلاعات در مدت زمان کوتاهتر است امروزه اطلاعات زیادی در وب نظیر MEDLINE, ERIC و غیره هر هفته یا هر ماه روزآمد می شوند.

۳- رایگان بودن

برخی منابع اطلاعاتی در وب رایگان هستند و این انگیزه ای قوی برای جستجو در این محیط پدید آورده است.

۴- سهولت جستجو و بازیابی اطلاعات

سهولت جستجو و بازیابی اطلاعات از طریق کلیدواژه ها یکی از مهمترین دلایل مراجعه به شبکه جهانی وب است.

۵- دسترسی به اطلاعات چند رسانه ای

یکی از مهمترین قابلیت های وب ، امکان جستجو و بازیابی حجم بسیار زیادی از اطلاعات چند رسانه ای شامل صوت ، تصویر و ویدئو است.

شبکه " وحشی "وب یا شبکه جهانی " انتظار "

متخصصان کتابداری و اطلاع رسانی، سال های متمادی تلاش کردند برای سازماندهی اطلاعات، یک نظام منطقی را توسعه دهند، تا به این ترتیب بتوان به طور مناسب و کارآمد به اطلاعات مورد نظر دست یافت. از این رو، نظام های رده بندی (مانند دیویی و کنگره) در بسیاری از کتابخانه ها مورد استفاده قرار گرفتند و بشر توانست انواع اطلاعات را تحت کنترل خود درآورد.

با رشد فزاینده اطلاعات سازماندهی نشده در وب، چالشهای جدیدی برای دسترسی مناسب و کارآمد به منابع اطلاعاتی در این محیط پدید آمده است. در محیط وب، هیچ استاندارد مشخصی مانند (فهرست نویسی، رده بندی، کنترل کتابشناختی و کنترل واژگان) برای ذخیره سازی و بازیابی منطقی اطلاعات، نظیر آنچه در دنیای چاپی یا پایگاه های اطلاعاتی سنتی با آنها آشنا هستیم، وجود ندارد. بسیاری از صفحات وب، حتی مولف، ناشر و تاریخ انتشار ندارند. فرض کنید در بزرگترین کتابخانه جهان حاوی چند میلیارد صفحه به دنبال اطلاعات هستید که بدون هیچ نظم منطقی روی قفسه های الکترونیکی مجازی یا به عبارت دیگر " سایبراسپیس " قرار گرفته اند. از طرف دیگر، شبکه جهانی وب برخلاف کتابخانه های چاپی، یک فهرست یا پایگاه مرکزی مشخص، ثابت و پایدار ندارد و هر کسی می تواند هر مطلبی را با هر محتوایی در آن منتشر کند، بدون آنکه هیچ گونه نظارت کیفی و کمی بر اطلاعات منتشر شده وجود داشته باشد.

مشکل اساسی تر این است که صاحبان سایت های وب، هیچ گونه الزام و تعهدی برای در دسترس قرار دادن صفحات یا سایت های خود در مدت زمان طولانی ندارند. بنابراین، سایت یا صفحه وبی که برای مثال، امروز قابل دسترس بوده است، فردا شاید دیگر وجود نداشته باشد.

یکی دیگر از مشکلات محیط وب، تغییر نشانی سایت ها به منابع است که به دلایل مختلف روی می دهد. نشانی دسترسی به صفحات یا سایت های وبی که قبلا به آنها مراجعه کردیم، بدون هیچ گونه اطلاعی تغییر کند و هیچ مکانیسمی هم برای دسترسی به نشانی جدید پیش بینی نشود. این موضوع مانند این است که محل نگهداری یک کتاب در قفسه کتابخانه بدون هیچ اصول منطقی تغییر کند که خود به معنای عدم دسترسی به آن منبع است. با وجود مشکلاتی از این قبیل، دور از انتظار

نیست که برخی از منتقدان شبکه جهانی وب به چنین محیط بی نظم، بی ثبات و غیر قابل کنترلی، عنوان "شبکه وحشی وب" اطلاق کنند.

یکی دیگر از مشکلات دسترسی به اطلاعات از طریق وب بویژه در کشورهای در حال توسعه، سرعت پایین خطوط ارتباطی و مخابراتی برای ارسال و دریافت داده ها از طریق اینترنت است. شاید شما هم تجربه چند دقیقه منتظر ماندن برای مشاهده یک صفحه وب را داشته باشید. سرعت پایین اینترنت می تواند بر رفتار اطلاع یابی استفاده کنندگان، تاثیرات منفی داشته باشد. به طوری که ممکن است اشتیاق جستجو در این محیط را از دست بدهند یا فرایند جستجو را نیمه کاره رها کنند.

"شبکه جهانی انتظار" عنوانی است که در برخی متون به منظور اشاره به سرعت پایین خطوط اینترنت برای ارسال و دریافت داده ها به کار رفته است و هم اینک نیز کشور ما تا حدود زیادی از آن رنج می برد.

فرایند اطلاع یابی در محیط وب، می تواند از طریق "جستجو" یا "مرور" دنبال شود. این دو شیوه اطلاع یابی به ترتیب منجر به ظهور دو نوع ابزار کاوش یعنی "موتورهای کاوش" و "راهنماهای موضوعی وب" شده است. به صورت نظری، بازیابی اطلاعات در موتورهای کاوش از طریق امکانات جستجو در راهنماهای موضوعی از طریق مرور ساختار سلسله مراتب موضوعی امکان پذیر است. با وجود این، هم اکنون شاهد هستیم که بسیاری از راهنماهای موضوعی نظیر Open Directory و Yahoo! بخش جستجوی جداگانه ای دارند و برخی موتورهای کاوش نظیر Google و Alta Vista علاوه بر امکانات جستجو، از پایگاه دیگر راهنماها به ترتیب (Look Smart و Open Directory) برای دسترسی به اطلاعات استفاده می کنند. به عبارت دیگر، برخی ابزارهای کاوش اینترنت برای آسان کردن فرایند اطلاع یابی و دسترسی به اطلاعات مرتبط در محیط وب، از هر دو شیوه بازیابی اطلاعات استفاده می کنند.

امکانات جستجو و بازیابی اطلاعات در ابزارهای کاوش، یکی از مهم ترین و اصلی ترین معیارهای انتخاب و ارزیابی آنها هستند. فرض کنید که یک ابزار کاوش، پایگاه اطلاعاتی بزرگی دارد، اما برای بازیابی دقیق و کارآمد اطلاعات، امکانات جستجوی محدودی دارد. بدیهی است که در بسیاری از موارد، استفاده از چنین ابزار کاوشی به بازیابی "هزاران" صفحه وب منجر می شود. که تقریباً هیچ استفاده کننده ای، توانایی و فرصت بررسی آنها را نخواهد داشت. در حقیقت، تنها راه جلوگیری از

غرق شدن در انبوهی از اطلاعات نامرتب یا برعکس، از دست دادن اطلاعات مرتبط در محیط وب، استفاده بهینه از امکانات جستجو و بازیابی اطلاعات در ابزارهای کاوش مختلف است. افزایش سرعت دسترسی به اطلاعات و جلوگیری از صرف زمان و نیروی انسانی، اهمیت نظام های جستجو و بازیابی اطلاعات در ابزارهای کاوش را دو چندان کرده است. اغلب طراحان و عرضه کنندگان ابزارهای کاوش هم تا حدود زیادی به اهمیت این قضیه برای انجام کاوش های موفق در محیط وب پی برده اند. به طوری که در حال حاضر، برخی ابزارهای کاوش نسبت به چند سال گذشته، امکانات جستجو و بازیابی نسبتاً بهتری در اختیار کاربران اینترنت قرار داده اند. حال این سوال مطرح است، چه امکانات و قابلیت هایی برای جستجوی اطلاعات در شبکه جهانی وب وجود دارد؟ و هر یک از آنها چه کاربرد و نقشی در بازیابی اطلاعات دارند؟

تقسیم بندی جستجو ها و انواع آن

جستجوهای که در نظام های اطلاعاتی الکترونیکی انجام می شود را می توان به اعتبار نوع اطلاعاتی که درخواست می شود و روشی که در هنگام جست و جو به کار گرفته می شود تقسیم بندی کرد. در ذیل هر دو نوع جست و جو معرفی و توضیح داده می شوند. برای جلوگیری از تداخل و سردرگمی به جستجوهای که براساس نوع اطلاعات درخواستی است "انواع جست و جو" و جستجوهای که براساس روش های مختلف است "استراتژی های جست و جو" گفته خواهد شد. در یک تقسیم بندی کلی می توان جست و جوهای که در نظام های اطلاعاتی صورت می گیرد را به دو دسته جست و جو برای "منابع معلوم" و جست و جو برای "منابع نامعلوم" تقسیم کرد که نوع اخیر را همچنین می توان "جست و جوی موضوعی" نام داد.

الف. جست و جوی منابع معلوم

در جست و جوی مواد مشخص، همان طور که از نام آن برمی آید، کاربر سعی دارد اطلاعاتی درباره چیزی که هم اکنون معلوم است فراهم آورد. متداول ترین نمونه این گونه جستجو، جست و جوی است که در یک پایگاه اطلاعاتی کتابشناختی، تمام متن، موتور جست و جو یا اپک کتابخانه به منظور یافتن جزئیات بیشتر از یک اثر و حتی یافتن متن کامل آن انجام می شود. در این نوع جستجو کاربر می داند - یا کمابیش احساس می کند که می داند - به دنبال چه می باشد. برای مثال نام نویسنده ای را در اختیار دارد و قصد یافتن اثر/آثاری خاص از او می باشد. به این ترتیب نام مولف را در فیلد مرتبط وارد و آن را جست و جو می کند تا آن مورد را بازیابی کند. در نوع دیگری از جست

و جوی معلوم کاربر نام مقاله، کتاب، مترجم و یا هر اطلاعات مشخص دیگری داشته باشد و آن را جست و جو کند.

اگر چه این نوع جست و جو در ظاهر آسان به نظر می رسد، اما بسیاری مواقع کاربران خصوصاً کاربران غیر مسلط و ناآشنا با نظام های اطلاعاتی الکترونیکی با مشکلات زیادی مواجه می شوند. برای مثال ممکن است کاربر دیکته صحیح نام مولفی را که به دنبال اثری از او است به همراه داشته باشد و براساس نوشته ای یا حدسیات خود، نام او را وارد نظام می کند. روشن است که بسیاری از نظام ها از جمله خارجی و داخلی مجهز به تصحیح دیکته کلید واژه ها نیستند و در چنین مواردی بسیار دیده می شود که حتی اثر مورد نظر کاربر در نظام اطلاعاتی می باشد، اما کاربر با توجه به اشتباهی که در نگارش و تایپ داشته است به منبع مورد نظر خود دسترسی پیدا نمی کند. این وضعیت در پرسش های لاتین زبان که توسط کاربران مطرح می شود بیشتر خودنمایی می کند. بسیاری از کاربران دیکته غیر صحیح نام لاتین یک مولف را وارد می کنند و انتظار دارند که نظام پاسخ بدهد که نتیجه غیر از انتظار خواهد بود.

از دیگر مشکلاتی که در این زمینه رخ می دهد این است که حتی در جستجوی معلوم، نظام های اطلاعاتی الکترونیکی به نحو قابل توجهی متفاوت از هم عمل می کنند. برای مثال مادامی که یک نظام اطلاعاتی پیوسته همچون ISI از کاربران درخواست می کند که نام مولف به صورت مقلوب (نام خانوادگی + حرف اول از نام کوچک مولف وارد شود)، نظامی دیگر همچون LIBRARY NET از کاربران خود درخواست می کند در هنگام اجرای این نوع جست و جو نام مولف به صورت مغلوب اما کامل (نام خانوادگی + نام کوچک مولف به صورت کامل) نوشته شود. عدم اجرای این دستورات منجر به بازیابی نخواهد شد، حتی اگر اثر مورد نظر در سیستم وجود داشته باشد. بنابراین راهکارهایی که پیشنهاد می شود عبارتند از:

۱- انجام جست و جو و بازیابی توسط میانجی اطلاعاتی

۲- آموزش راهکارهای جست و جو توسط میانجی اطلاعاتی به کاربر

۳- مطالعه راهکارهای جست و جو و از جمله مطالعه منوی HELP هر پایگاه اطلاعاتی توسط

کاربر

در جست و جوی منابع معلوم این واقعیت وجود دارد که هنگامی که مدرک مورد نظر کاربر پیدا شد، جست و جو پایان یافته است.

ب. جست و جوی منابع نامعلوم

جست و جوی منابع نامعلوم یا جستجوی موضوعی، بسیار بیشتر از جست و جوی منابع معلوم انجام می شود. به واقع اکثریت جستجوهای که در کتابخانه ها و مراکز اطلاع رسانی و حتی توسط کاربران نهایی در هر مکانی انجام می شود به دنبال مدارکی خاص نمی باشد، بلکه به دنبال این است در یک زمینه موضوعی خاص اطلاعاتی از نظام اطلاعاتی بازیابی شود. این نوع جست و جو تلاش می کند محل اطلاعات مربوط به یک موضوع را مشخص کند، بنابراین این جست و جو را می توان "جست و جوی باز" نیز نام نهاد، زیرا معمولا جستجوگر در آغاز جست و جو درباره آن چه ممکن است از مخزن اطلاعات بازیابی شود یا چه زمانی همه مواد مرتبط پیدا می شوند، هیچ نظری ندارد، اگر چه برخی از کتابداران مرجع خبره بر پایه تجربه خود نسبت به بعضی از موضوعات و با توجه به نظام های اطلاعاتی که در دسترس دارند حدسیاتی اولیه دارند. برای مثال ممکن است موضوعی برای جست و جو از طرف کاربر مطرح شود و کتابدار با توجه به مواردی از جمله نو بودن موضوع یا اندک کارهایی که روی آن موضوع انجام شده است، اظهار نظر کند که در این رابطه منابع اندک می باشد یا برعکس آن موضوعی که بسیار بر روی آن کار شده است را به راحتی تشخیص دهد.

بعضی مواقع جست و جویی که در ابتدا به صورت یک جست و جوی معلوم خود را نشان داده است، هنگام مصاحبه مرجع یا در فرایند جست و جو به جستجوی نامعلوم تبدیل می شود. برای مثال کاربری به دنبال کلیه آثار از یک مولف خاص می گردد و در این زمینه نیز هیچ ایده خاصی ندارد، یعنی اثری خاص را طلب نمی کند. مطالعات مختلف در این زمینه نشان می دهد که برخی از کاربران کتابخانه ای بر این باور هستند که یک شخص خاص در زمینه ای که به دنبال اطلاعات آن هستند صاحب نظر می باشد یا فکر می کنند فقط آن مولف مورد نظر در آن زمینه فعالیت می کند، در چنین مواقعی کتابدار مرجع باید این آگاهی را به کاربر بدهد که ممکن است انجام یک جست و جوی نامعلوم یا موضوعی به بازیابی منابع بسیار بیشتری برای کاربر منتهی شود. البته برخی مواقع قبولاندن چنین مطلبی به کاربران مشکل می باشد، اما کتابدار با انجام یک جست و جوی موضوعی می تواند موارد بازیابی شده را در مقابل دید کاربر قرار داده و در مقابل عکس العمل او جستجو را ادامه دهد.

استراتژی جست و جو

در حال حاضر استراتژی های مختلفی برای جست و جوی منابع الکترونیکی در دسترس هستند. این استراتژی ها در ادامه معرفی و تشریح می شوند. در این میان یافته ها نشان می دهد که اکثریت نظام های اطلاعاتی خارجی و داخلی در حال حاضر از جست و جوی منطق بولی استفاده می کنند، این استراتژی به نحوی غالب است که استراتژی های دیگر را تحت تاثیر کامل قرار داده است. نظام های بازیابی اطلاعات اعم از پایگاه های اطلاعاتی و موتورهای جستجو میلیون ها رکورد را در خود جای داده اند و این در حالی است که روش های جستجو از ابزاری به ابزار دیگر فرق می کند. بعضی مواقع نظام های بازیابی اطلاعات در پاسخ به یک جستجو تعداد زیادی رکورد را بازیابی می کنند، در حالی که در جستجوی دیگر ممکن است یا به طور کلی رکوردی بازیابی نشود و یا اینکه تعداد رکوردهای بازیابی شده اندک باشد. توضیحات بعدی این نکات مبهم را روشن کند.

فرایند جست و جو

پائو (۱۹۸۹) چهار مرحله را قبل از این که کاربر به تعامل با نظام اطلاعاتی برای بازیابی اطلاعات مورد نیاز پردازد تشخیص داده است. در مرحله نخست، مرحله مشکل اطلاعاتی، کاربر به این بینش می رسد که یک کاستی اطلاعاتی وجود دارد. این کاستی اطلاعاتی می تواند برای حل یک مشکل خاص باشد، می تواند نیاز به اطلاعات زمینه ای باشد و یا می تواند نیاز برای روزآمد شدن در موضوع یا زمینه ای خاص باشد. هنگامی که مشکل مشخص گردید، کاربر مشکل اطلاعاتی را به صورت نیاز اطلاعاتی فرموله بندی می کند، به عبارت دیگر مشخص می کند چه چیزی باید برای حل مشکل اطلاعاتی دانسته شود. در مرحله بعد کاربر نیاز اطلاعاتی خود را به صورت سوالی مطرح می کند که می تواند زمینه جست و جو در نظام اطلاعاتی را باعث شود. این سوال سپس به یک تقاضای اطلاعاتی تبدیل می شود که برای نظام اطلاعاتی قابل فهم می باشد. این مرحله آخر می تواند در شکل یک یا چندین کلید واژه، یک سوال یا جمله ای باشد. که برای نظام قابل فهم باشد. بر پایه آنچه که کاربر در تعامل با نظام اطلاعاتی به دست می آورد ممکن است نیاز داشته باشد به مرحله یا مراحل قبل برگشته و نیاز اطلاعاتی خود را مورد بررسی قرار دهد و جست و جو را مجدداً انجام دهد. مارچونینی (۱۹۹۵) وظایفی که کاربر بعد از مشخص کردن تقاضای اطلاعاتی باید انجام دهد را به صورت زیر بیان می دارد:

✓ انتخاب نظام اطلاعاتی

- ✓ هم خوان کردن تقاضای اطلاعاتی با روش جست و جو در آن نظام
- ✓ انجام جست و جو
- ✓ ارزیابی نتایج
- ✓ استخراج اطلاعات
- ✓ عکس العمل در برابر اطلاعات بازیابی شده
- ✓ تکرار جست و جو (در صورت نیاز)
- ✓ نتیجه گیری

زومر مراحل بیان شده توسط مارچيونینی را به صورت زیر بسط دهد:

- بعد از فرموله کردن پرسش به چند طریق می توان آن را در نظام اطلاعاتی مطرح کرد. برای مثال با تایپ کردن کلید واژه ها، انتخاب کلید واژه ها از نمایه (در صورت وجود) یا بیان پرسش به صورت یک پرسش یا جمله طبیعی
- ارزیابی نتایج نیاز به قالب خروجی روشن، منظم و انعطاف پذیر دارد. برای مثال قالب خروجی می تواند شامل بازیابی حداقل/حداکثر فیلدها و ارائه می تواند در قالب تاریخ(از قدیم به جدید یا از جدید به قدیم) یا براساس عنوان و یا ربط باشد.
- استخراج اطلاعات نیز می تواند به دو شیوه اساسی باشد که عبارتند از ذخیره الکترونیک اطلاعات یا چاپ آنها.

بنا به عقیده زومر مراحل فرموله کردن پرسش (که شامل انجام جست و جو می شود)، ارزیابی نتایج و استخراج اطلاعات بخش های اساسی در مبحث بازیابی اطلاعات می باشند. زومر همچنین به نقش صفحه/صفحات کمک در نظام های اطلاعاتی توجه می کند. از آنجا که کاربران اغلب با مشکلاتی در هنگام استفاده از نظام های اطلاعاتی مواجه هستند، صفحات کمک برای حل این مشکلات نقش اساسی دارند. از نظر زومر صفحات کمک باید در همه جای نظام های اطلاعاتی موجود باشند به گونه ای که کاربران به راحتی به آنها دسترسی داشته، به راحتی آنها را درک کنند و بتوانند آن را به مرحله اجراء بگذارند.

اگر چه مراحل چهار گانه گفته شده توسط پائو برای کمک به کاربران مهم می باشد اما مرحله تبدیل سوال به پرسش همخوان با مدل / مدل های جست و جوی نظام از نظر طراحان نظام های اطلاعاتی بسیار مهم می باشد. در ادامه اصول عمومی جست و جو در نظام های اطلاعاتی فعلی را نشان خواهد

داد. به طور عمده دو روش جست و جو که یکی مبتنی بر کلیدواژه و دیگری مبتنی بر زبان طبیعی می باشد بیان خواهد شد.

ساختار ابزارهای بازیابی اطلاعات

آگاهی از چگونگی نمایه سازی مدارک توسط ابزارهای بازیابی اطلاعات می تواند به درک بهتر بازیابی و یادگیری اصول آن کمک کند. در پایگاه های اطلاعاتی و راهنماها، روش نمایه سازی به این صورت است که بعد از مطالعه مدرک توسط نمایه ساز چندین کلیدواژه به آن اختصاص داده می شود، این کلیدواژه ها خصوصا در پایگاه های اطلاعاتی در فیلدهای خاص نظیر عنوان، مولف، موضوع و از این قبیل قرار می گیرند. به این ترتیب نمایه سازی کلید واژه ای کاربر را قادر می سازد که در صورتی که همان کلید واژه ها را وارد نظام کند همان مدارک بازیابی شوند. در موتورهای جستجو عمل نمایه سازی به صورت خودکار انجام می شود و از موتوری به موتور دیگر فرق می کند. برخی از آنها نظیر موتور جستجوی *Alta vista* از روش نمایه سازی متن کامل پیروی می کند که براساس آن تمامی کلمات موجود در مدرک، نمایه سازی می شود. ولی بیشتر موتورهای جستجو فقط کلمات مهم را در متن و عنوان تشخیص را داده و مدرک را بر آن اساس و اصول دیگری که دارند، نمایه سازی می کنند.

ویژگی های مشترک جستجو

فهم ویژگی های مشترک جستجو به لحاظ کاربرد آن در بیشتر نظام های بازیابی اطلاعات اهمیت دارد. به علاوه هر ابزار جستجو معمولا دارای دو گزینه می باشد که یکی به جستجوی ساده و دیگری به جستجوی پیشرفته اختصاص دارد. همچنین در هر کدام پیش فرض هایی برای جستجو تعیین شده است که در صورتی که کاربر آنها را تغییر ندهد جستجو براساس آنها انجام خواهد شد. نظام های بازیابی اطلاعات گستره ای از ویژگی های جستجو را پشتیبانی می کنند که به وسیله آنها می توان جستجو را محدود نمود. البته تمامی نظام های بازیابی اطلاعات پیوسته در این زمینه یکسان عمل نمی کنند. ممکن است آنها تنها برخی از این ویژگی ها را داشته باشند و برخی دیگر را نداشته باشند. این مطلب خصوصا در مورد موتورهای جستجو مصداق پیدا می کند که نمایه سازی در آنها به صورت خودکار انجام می شود. برخی از امکانات جستجو مانند محدود کردن جستجو به فیلد خاص مانند مولف، ناشر، سال نشر، شماره انتشار و یا منبعی خاص امکان پذیر نمی باشد.

بهترین راه برای آگاهی از پوشش این ویژگی ها در هر نظامی اعم از پایگاه های اطلاعاتی یا موتورهای جستجو مراجعه به گزینه کمک می باشد که در آن ویژگی هایی که نظام بازیابی آنها را پشتیبانی می کند و همچنین روش های جستجو داده می شود. از این رو در ادامه به بحث و بررسی پیرامون انواع مختلف جستجو می پردازیم و سپس سایر امکانات جستجو را معرفی می کنیم.

انواع جستجو

جستجوی ساده و جستجوی پیشرفته

همان طور که گفته شد هر ابزار جستجو معمولا دارای دو گزینه جستجوی ساده و جستجوی پیشرفته است. در جستجوی ساده کلمه یا عبارت جستجو تایپ می شود با فشار دکمه Enter یا Search و یا Go نتایج جستجو ارائه می شود. این شیوه جستجو با بازیافت زیاد و دقت کم همراه می باشد. اما در جستجوی پیشرفته امکانات بیشتری نظیر اپراتورهای بولی و محدودگرهایی مانند محدودیت زبانی، مکانی، زمانی، قالب منابع ارائه می شود که سبب می شود نتایج جستجو با دقت بیشتری همراه باشد. معمولا در جستجوی ساده و پیشرفته امکان استفاده از اپراتورهای بولین وجود دارد.

جستجو از طریق عملگرهای بولی

جورج بول، ریاضی دان انگلیسی، در سال ۱۸۴۷ منطق ریاضی را ارائه کرد که امروزه اساس بازیابی اطلاعات در بسیاری از پایگاههای اطلاعاتی و همچنین ابزارهای کاوش اینترنت است. از طریق منطق بول می توان مفاهیم مختلف (کلید واژه ها) را به منظور گسترش دادن یا محدود کردن دامنه جستجو با یکدیگر ترکیب کرد و به نتایج کاوش دقیق تر و مرتبط تری در یک پایگاه اطلاعاتی دست یافت. برای این منظور، از عملگرهای خاصی استفاده می شود که عبارتند از:

AND ، OR و NOT

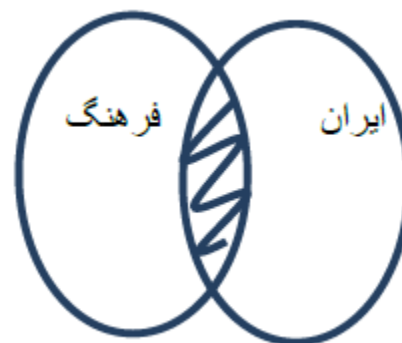
اغلب ابزارهای کاوش وب، امکان جستجو در پایگاه اطلاعاتی خود را از طریق عملگرهای بول در اختیار قرار می دهند. عدم آگاهی از کاربرد و نحوه استفاده از این عملگرها در ابزارهای کاوش مختلف، یکی از مهم ترین دلایل بازیابی بیش از حد اطلاعات است. بنابراین، کاربران اینترنت باید قبل از هر

چیز، کاربرهای عملگرهای بولین را برای گسترش دادن یا محدود کردن دامنه جستجو بخوبی بشناسند.

عملگر AND

عملگر AND یا "و" برای محدود کردن دامنه جستجو از طریق ترکیب کلید واژه های مختلف به کار می رود. برای مثال، اگر به دنبال اطلاعاتی درباره "فرهنگ ایران" هستیم، می توان از این عملگر بخوبی استفاده کرد.

با وارد کردن عبارت جستجوی فوق در ابزارهای کاوشی که عملگر AND را مورد پشتیبانی قرار می دهند، کلیه صفحات وبی که در آنها کلید واژه های [Iran و Culture](#) به کار رفته است، بازیابی خواهند شد. در واقع، عملگر AND قید "باید" را تداعی می کند. یعنی "همه" کلید واژه هایی که از طریق این عملگر با یکدیگر ترکیب می شوند، "باید" در نتایج کاوش وجود داشته باشند (شکل ذیل).



عملگر OR (ضرب منطقی)

OR اپراتور دیگری می باشد که وارد کردن آن بین دو یا تعداد کلیدواژه ها به معنای گسترش حوزه جستجو می باشد. هنگامی که این استفاده قرار گیرد به معنای این است که تمامی مدارکی که هر کدام صرفاً در آنها آمده است بازیابی شود. هنگامی که از اپراتورهای مختلف بولین در یک جستجو استفاده می شود بایستی روشی که تو در توی نامیده می شود مورد استفاده قرار گیرد. بر این اساس کلید واژه هایی که می خواهیم یکی از آنها هر بار بازیابی شوند در پرانتز قرار می گیرند.

Liver AND (cat OR dog OR horse) NOT mouse

عملگر NOT یا BUT (تفریق منطقی)

سومین اپراتور بولین که کمتر از بقیه استفاده می شود اپراتور NOT می باشد که BUT نیز نامیده می شود از این اپراتور هنگامی استفاده می کنیم که قصد حذف واژه ای را در جستجو داشته باشیم.

علاوه بر اپراتورهای گفته شده، اغلب نظام های بازیابی اطلاعات مبتنی بر منطق بولین در حال حاضر اپراتورهای دیگری را نیز ارائه می کنند که در ادامه به شرح آنها پرداخته می شود:

جستجو از طریق عملگر نزدیک یابی یا مجاورتی

گاهی ممکن است استفاده از عملگر AND برای بازیابی کلید واژه های مختلف، صرفاً به این دلیل که در یک صفحه وب وجود دارند، کافی نباشد. به عبارت دیگر، استفاده از عملگر AND ممکن است منجر به بازیابی کلیدواژه ها در " هرکجای متن " شود، بدون آنکه هیچ ارتباط مفهومی با یکدیگر داشته باشند (برای مثال یک کلیدواژه در سطر اول و دیگری در سطر انتهایی). این مشکل بویژه در نظام های بازیابی اطلاعات که فاقد نمایه سازی موضوعی هستند، بیشتر مشاهده می شود. فرض کنید به اطلاعاتی درباره " علم اطلاع رسانی " (information science) نیاز دارید اما پس از ترکیب کلید واژه های مذکور از طریق عملگر متوجه می شوید که در AND بسیاری از موارد، این دو کلید واژه هیچ ارتباط مفهومی با یکدیگر ندارند: "science and technology advancement in the western world is related to explosion of information"

همان طور که در مثال فوق مشاهده می شود، با وجود اینکه کلید واژه های Science و Information در یک صفحه وب بازیابی شده اند، اما با نیاز اطلاعاتی ما (Information science) مرتبط نیستند.

برای جلوگیری از بازیابی اطلاعات غیر مرتبط یا به عبارت دیگر " ریزش کاذب " در چنین مواردی، می توان از امکان کاوش دیگری تحت عنوان جستجوی نزدیک یابی استفاده کرد. استفاده از این امکان کاوش، بویژه در جستجوی آزاد در متن کاربرد فراوانی دارد. زیرا از طریق آن می توان کلید واژه ها و مفاهیم مورد نظر را در " کنار هم " بازیابی کرد. به تعبیر دیگر، از طریق جستجوی نزدیک یابی می توان ارتباط مکانی و جایگاه کلید واژه ها را در کنار هم تعیین کرد.

اغلب پایگاههای اطلاعاتی سنتی از عملگرهای NEAR یا ADJ برای اجرای جستجوی نزدیک یابی استفاده می کنند. ظاهراً ابزارهای کاوش اصلی، آن طور که باید و شاید این قابلیت مهم را مورد توجه قرار نداده اند، به طوری که تنها موتور کاوش Alta Vista و AOL امکان استفاده از عملگر نزدیک

یابی را در اختیار گذاشته اند. فرض کنید در نظر دارید کلید واژه های Nuclear و Power را در صفحات وب در کنار هم بازیابی کنید. به این ترتیب در موتور کاوش Alta Vista، می توان کلید واژه های مورد نظر را تا " ۱۰ کلمه واسط " آنها مورد جستجو و بازیابی قرار داد:

Nuclear NEAR Power

در مثال زیر، نتیجه کاوش کلید واژه های Nuclear و Power از طریق عملگرهای AND و NEAR در موتور کاوش Alta Vista با هم مقایسه شده است:

Nuclear AND Power	۹۷۰۴۲۸	تعداد نتیجه بازیابی شده با فرمول مقابل
Nuclear NEAR Power	۳۲۳۶۸۴	تعداد نتیجه بازیابی شده با فرمول مقابل

کوتاه سازی کلید واژه ها

کوتاه سازی عمل اضافه کردن علامتی به ابتدا یا انتهای ریشه یک کلمه برای بازیابی حالات مختلف آن کلمه می باشد. کوتاه سازی خصوصا برای بازیابی حالات مفرد و جمع یک کلمه مفید می باشد. عمل کوتاه سازی این امکان را فراهم می کند که با وارد کردن بخشی از کلمات، مشتقات مختلف آنها را نیز در فرایند کاوش بازیابی کنیم و به این ترتیب، دامنه جستجو را گسترش دهیم. اکثر ابزارهای کاوش اینترنت، امکان کوتاه سازی کلید واژه ها را از طریق علامت ستاره (*) در اختیار می دهند. برای مثال، جستجوی کلید واژه Iran منجر به بازیابی کلیه مشتقات کلمه Iran، مانند Iranians, Iranian, Iran's و ... خواهد شد. استفاده از عملگر کوتاه سازی، هنگامی اهمیت پیدا می کند که یک کلمه (اغلب ریشه اصلی آن) مشتقات زیادی داشته باشد و کاوشگر بخواهد کلیه آنها را در فرایند جستجو مدنظر قرار دهد. با استفاده از قابلیت کوتاه سازی، دیگر نیازی به وارد کردن همه مشتقات در کادر جستجو نیست و به این ترتیب، در زمان وارد کردن کلید واژه ها نیز صرفه جویی می شود.

نکته مهم در استفاده از عملگر کوتاه سازی، مکانی است که عملگر کوتاه سازی باید به کار رود. برای نمونه، اگر بخواهیم همه کلیه کلیدواژه های Feminists, Feminist, Feminine, Females, Female و Feminism را در فرایند جستجو بازیابی کنیم باید عملگر کوتاه سازی را بعد از حرفی به کار ببریم که تا آنجا با دیگر مشتقات، "مشابهت" دارد که در مثال فوق Fem است. در مثال زیر، نتیجه جستجوی جداگانه کلید واژه های bird و birds و استفاده از عملگر کوتاه سازی bird در موتور کاوش Alta Vista نشان داده شده است:

Bird 2265162
Birds 2192320
Bird* 6247191

همان طور که مشاهده می شود، استفاده از عملگر کوتاه سازی منجر به بازیابی تعداد بیشتری از صفحات وب شده است. برخی ابزارهای کاوش نظیر Excite و yahoo! عملگر کوتاه سازی را به طور " خودکار " اجرا می کنند و برخی دیگر نظیر Northern Light به طور خودکار شکل " جمع " کلمات را در فرایند بازیابی اطلاعات پوشش می دهند (مانند Boy و Boys).

یکی از محدودیتهای استفاده از عملگر کوتاه سازی این است که در برخی موارد ممکن است منجر به " ریزش کاذب " یا به تعبیر دیگر بازیابی اطلاعات ناخواسته یا غیر مرتبط شود. برای مثال، استفاده از عملگر کوتاه سازی بعد از کلمه cit به منظور پوشش کلمات cities و city منجر به بازیابی مفاهیم نامرتبب دیگری نیز خواهد شد. مانند:

Citadel, citrate, citric citadels, citify, citriculture,
Citation, citizen, citrine citations, citizenry, citrone,
Cite, citizens, citronella citizenship, citrus, cited.

ریشه سازی

استفاده از این اپراتور به کاربر اجازه می دهد تا کلماتی را بازیابی کند در یک یا دو حرف با یکدیگر فرق دارند. برای مثال کاربر ممکن است بخواهد جستجویی را در رابطه با کلیدواژه های women و woman انجام دهد. به جای اینکه هر دوی این کلیدواژه ها را وارد کند و بین آنها از اپراتور OR استفاده کند از این اپراتور استفاده می کند به این ترتیب که را وارد می کند و این باعث بازیابی هر دو حالت این کلمه می گردد. از مواردی که این اپراتور در آنها کاربرد دارد برای کلماتی است که انگلیسی بریتانیایی و انگلیسی آمریکایی متفاوت است. استفاده از این اپراتور هر دو حالت را بازیابی می کند برای مثال colo*r هر دو حالت کلمه colour و color را بازیابی می کند. همچنین هنگامی که کاربر از املا صحیحی کلید واژه ها ی که وارد می کند مطمئن نیست نیز این اپراتور کاربرد دارد. نتیجه جستجوی ne?t بازیابی تمام کلمات nest، neat ، next خواهد بود.

جستجوی کلید واژه در عنوان صفحات وب

یکی از مهم ترین و کارآمدترین روش های دسترسی به اطلاعات در محیط وب، جستجوی کلید واژه ها در عنوان صفحات وب است. یکی از پیش فرض هایی که حتی در رسانه های چاپی هم می توان آن را مطرح کرد، این است که عنوان یک صفحه وب، تا حدود بسیار زیادی می تواند نمایانگر محتوای اطلاعات موجود در آن باشد. با چنین پیش فرضی، می توان ادعا کرد که جستجوی کلید واژه های مختلف در عناوین صفحات وب می تواند کاوشگران را به اطلاعات مرتبط در محیط وب هدایت کند.

جستجوی پیوندهای فرامتنی در صفحات دیگر

قابلیت جستجوی پیوندهای فرامتنی در صفحات وب دیگر، به ما امکان می دهد که دریابیم کدام سایت ها، نشانی یک صفحه یا سایت وب مشخص را در صفحه خود گنجانده اند. به تعبیر دیگر، جستجوی پیوندهای فرامتنی (Hypertext Links) در سایر صفحات وب، نوعی تحلیل استنادی به شمار می آید.

موتور کاوش Alta vista و Google از دستور Link: برای اجرای جستجوی پیوندهای فرامتنی در صفحات دیگر استفاده می کنند. برخی دیگر از ابزارهای کاوش نظیر HotBot, Fast Search و Lycos برای انجام این نوع جستجو، فهرست انتخاب یا کادر جستجوی جداگانه ای دارند. برای مثال، می خواهید از این موضوع آگاهی حاصل کنید که نشانی سایت وب " صدا و سیمای جمهوری اسلامی ایران " یا " خبرگزاری جمهوری اسلامی ایران " در قالب یک پیوند فرامتنی در کدام صفحات وب آورده شده است. کافی است که " نشانی دسترسی " سایت وب مورد نظر را بعد از این دستور وارد کنید. در مثال زیر، نتیجه کاوش نشانی سایت وب صدا و سیمای جمهوری اسلامی ایران و نیز خبرگزاری جمهوری اسلامی ایران مبتنی بر قابلیت جستجوی پیوندهای فرامتنی در صفحات دیگر به ترتیب در موتور کاوش Alta Vista نشان داده شده است:

Link:www.irib.com 1466

Link:www.irna.com 4867

همان طور که مشاهده می شود، در زمان انجام این کاوش، به ترتیب ۱۴۶۶ و ۴۸۶۷ صفحه وب دارای پیوند فرامتنی به سایت " صدا و سیمای جمهوری اسلامی ایران " و " خبرگزاری جمهوری اسلامی ایران " بودند. تجزیه و تحلیل استنادی سایت ها، مهمترین کاربرد این امکان جستجو است. از این شیوه می توان دریافت که نشانی یک سایت مشخص در چند صفحه یا سایت وب دیگر آورده

شده است یا به تعبیر دیگر، سایر صفحات وب به آن "استناد" کرده اند. شاید بتوان این پیش فرض را مطرح ساخت که هر چه تعداد استنادها به یک صفحه یا وب سایت بیشتر باشد، آن صفحه از ارزش یا اهمیت بیشتری برخوردار است چنین پیش فرضی یکی از مهمترین معیارهای رتبه بندی نتایج کاوش در موتورهای کاوش Google به شمار می آید که از آن تحت عنوان Link Popularity نام برده می شود.

جستجوی حوزه سایت ها

به صورت قراردادی، هر کشوری حوزه خاصی در محیط وب دارد که نمایانگر وابستگی سایت های آن کشور است.. حوزه سایت هر کشوری بجز آمریکا، از طریق ۲ حرف مشخص می شود. برای مثال ir،uk،fr و de به ترتیب معرف حوزه سایت های کشورهای ایران، انگلستان، فرانسه و آلمان هستند که در انتهای نشانی دسترسی آنها آورده می شوند. در مثال زیر، حوزه سایت ها با حروف سیاه مشخص شده اند:

<http://www.nli.ir>

کتابخانه ملی جمهوری اسلامی ایران

جستجوی عبارتی

جستجوی عبارتی شامل کلید واژه هایی است که بایستی کنار یکدیگر بیایند. از این اپراتور هنگامی استفاده می شود که کلید واژه های مورد نظر کاربر بایستی حتما در کنار یکدیگر و همچنین به نظمی که وارد می کند پشت سرهم بیایند. بیشتر نظام های بازیابی اطلاعات از قبیل پایگاه های اطلاعاتی و موتورهای اپراتور را پشتیبانی می کنند. برای انجام آن به صورت معمول بایستی کلید واژه را در نظمی که می خواهیم در بین دو گیومه قرار دهیم مانند "search engines" جستجوی عبارتی یکی از مفیدترین ویژگی های جستجو می باشد که شانس بازیابی اقلام مربوط تر را افزایش می دهد.

جستجوی میدانی

جستجوی میدانی توانایی محدود کردن جستجو به فیلدی خاص می باشد به طور معمول فیلدهای مختلفی در نظام های بازیابی اطلاعات در نظر گرفته می شود که کاربر از آن طریق می تواند جستجوی خود را محدود کند. مهم ترین این فیلدها شامل موضوع، چکیده و مولف می باشد. همچنین در موتورهای جستجو این فیلدها شامل موارد دیگری از قبیل آدرس صفحه شبکه ای نیز می شود. توانایی محدود کردن جستجو به فیلدی معین میزان ربط مدرک بازیابی شده را افزایش می دهد.

جستجوی واژه kidney محدود به فیلد "عنوان" را در پایگاه از طریق sciencedirect نشان می دهد.

The image shows a search interface with the following elements:

- Navigation tabs: All (selected), Journals, Books, Reference Works, Images.
- Links: Advanced search | Expert search.
- Search tips: ? Search tips.
- Search for: A text input field containing "kidney".
- Search in: A dropdown menu set to "Title".
- AND: A dropdown menu set to "AND".
- Search in: A text input field (empty).
- Search in: A dropdown menu set to "Abstract, Title, Keywords".
- Refine your search: A list of checkboxes including Journals, Books, All, My Favorites, Subscribed publications, and Open Access articles.

جستجوی زبان طبیعی

جستجوی مبتنی بر منطق بولین به عنوان موثرترین راه برای بازیابی اقلام مربوط در نظام های بازیابی اطلاعات پیوسته در میان کتابداران به عنوان واسطه های اطلاعاتی مورد پذیرش قرار گرفته است. اما امروزه جستجو با زبان طبیعی نیز جایگاهی هر چند کوچک را به خود اختصاص داده است. برخی از میزبان ها نظیر Ebsco و برخی از موتورهای جستجو نظیر ASK آن را وارد نظام خود کرده اند.

جستجو به زبان طبیعی بر پایه پردازش به زبان طبیعی بنا شده است که در سال های اخیر در کانون توجه نظام های بازیابی اطلاعات و تعامل انسان و رایانه قرار گرفته است . فرایند جستجو به زبان طبیعی بر این مبنا بنا شده است که کاربران نهایی در درک اصول استفاده از اپراتورهای بولین دچار مشکل هستند و بنابراین اگر آنها به همان گونه که به صورت طبیعی سوالی را مطرح می کنند بتوانند سوال خود را در یک نظام وارد کنند احتمالاً می توانند نتایج بهتری را کسب نمایند . در جستجو به زبان طبیعی، کلید واژه یا مفاهیم کلیدی و مهم از متن سوال استخراج نمی شوند بلکه سوال به همان صورت مطرح شده به سیستم داده می شود .
مثال های زیر نمونه ای از جستجو به زبان طبیعی می باشد .

What is the distance in mile from the earth to the sun?

Who was the first Afro-American to win the Nobel Prize in literature?

What is the definition of opaque?

نظام های جستجو به زبان طبیعی اقلام مربوط را به صورتی متفاوت بازیابی می کنند، هر چند فراهم آورندگان این گونه خدمات کمتر در رابطه با الگوریتم های که در این رابطه استفاده می کنند اطلاعاتی را ارائه می دهند . ولی این نظام ها از دسته ای از روش های پیچیده برای شمارش واژه ها (تعداد واژه ها در یک مدرک، تکرار آن واژه ها) حذف کلماتی که مهم نیستند و از این قبیل استفاده می کنند . هر چه یک مفهوم بیشتر در یک مدرک بازیابی شود وزن بیشتری به آن داده می شود . به این ترتیب این روش های آماری مربوط به این مفهوم برای بازیابی مربوط ترین منابع مورد استفاده قرار می گیرد .

نمونه هایی از موتورهای جستجو به زبان طبیعی عبارتند از:

ASK (www.ask.com)

Homerweb (www.homerweb.com)

Brainboost (www.brainboost.com)

Vrosh ([www.vroosh. Com/adv-search.html](http://www.vroosh.Com/adv-search.html))

جستجوی استنادی

بدون شک بازیابی مدارک موجود در یک نظام بازیابی اطلاعات از طریق جستجوی کلید واژه ای موثر می باشد ولی این تنها راه حل نمی باشد . هنگامی که نویسندگان آثاری از قبیل مقالات و کتابها به منابعی که استفاده کرده اند، چه در متن و چه در ماخذ، استناد می کنند این عمل جستجوی منابع را برای آنها افزایش می دهد . قدرت جستجوی استنادی در این است که منابع

مربوطی که تحقیقات جدید را تحت تاثیر قرار داده اند، را مشخص می کند. موسسه اطلاعات علمی سال ها است که این روش جستجو را به کار گرفته است. در سال ۱۹۹۷ که موسسه اطلاعات علمی خدماتی را با نام شبکه علم راه اندازی نمود، این روش جستجو بهتر و کارا تر در اختیار کاربران نهایی قرار گرفت. اگر چه به نظر می رسد که وب محیطی طبیعی را برای جستجوی استنادی فراهم کرده است، ولی این مفهوم جدید نمی باشد. حدود نیم قرن پیش، در اواسط دهه پنجاه میلادی یوجین گارفیلد ایده تولید اطلاعات علمی با استفاده از این روش را معرفی نمود. به جای جستجوی موضوعی در بیشتر نمایه ها، نمایه های استنادی گارفیلد دسترسی به مقالاتی را فراهم کرد که به مقاله ای مربوط استناد کرده بودند.

به سخنی دیگر جستجوی استنادی به این صورت عمل می کند که معمولا نویسندگان در رشته ها و موضوعات مختلف به آثار مربوط به اثری که در حال تالیف آن هستند استناد می کنند و سپس آن استنادات را در بخش منابع و ماخذ می آورند، حال این منابع می توان برای بازیابی منابع مربوط استفاده نمود. به علاوه با مراجعه به آثار استناد شده می توان از منابع و ماخذ آنها منابع مربوط دیگری را بازیابی نمود. این فرایند در برخی پایگاه های اطلاعاتی مورد استفاده قرار گرفته است. به عنوان مثال، با شخص کردن مقاله ای مربوط به سال ۲۰۱۹، اگر این مقاله توسط افرادی دیگری در سال های بعد مورد استناد قرار گرفته باشد مشخص می شود. همچنین می توان این عمل را به صورت زنجیروار و بدون انتها ادامه داد. به عنوان مثال مقاله ای که توسط بورمن نوشته شده است؛ با مراجعه به پایگاه ISI می توان تعیین نمود بعد از تاریخ فوق چه مقالاتی به این مقاله استناد کرده اند و از آن به عنوان منبع استفاده کرده اند.

جستجوی استنادی در پایگاه ISI

جستجوی استنادی با اینکه می تواند به بازیابی مدارک مربوط منتهی شود ولی به جهت هزینه بسیار بالایی که دارد کمتر توسط نظام های بازیابی اطلاعات مورد استفاده قرار گرفته است. بهترین و بزرگترین نمونه جستجوی استنادی در پایگاه های مختلف اطلاعاتی موسسه اطلاعات علمی انجام می شود که مجموعه ای قریب به ۹۰۰۰ مجله بدون محدودیت موضوعی را پوشش می دهد. پایگاه های اطلاعاتی دیگر از قبیل Science direct, Scopus و Ebsco نیز به صورت محدود جستجوی استنادی را برای مدارک موجود در پایگاه های خود فراهم می کنند.

جستجوی فرایبندی

موفقیت موتور جستجوی گوگل که نتایج جستجو را براساس تعیین فرایبوندهایی که به یک صفحه شبکه ای پیوند خورده اند بازیابی و به نمایه می گذارد، مفهوم جستجوی استنادی را وارد دنیای منابع رایگان شبکه ای نمو نظریه بنیانی فرایبوند مانند جستجوی استنادی می باشد. چه کسی به این صفحات استناد کرده است یا به آن پیوند زده است؟ با این روش جستجو می توان صفحات را بازیابی نمود که در واقع خود را به صفحه مورد نظر پیوند داده اند. برای انجام این نوع جستجو بعد از حصول اطمینان از وجود چنین توانایی موتور جستجوی مورد نظر می توان به صفحه جستجوی پیشرفته مراجعه و در محل مشخص شده برای این نوع جستجو عمل کاوش را انجام داد. تصویر زیر بخشی از صفحه جستجوی پیشرفته موتور گوگل می باشد که در آن محلی که می توان این نوع جستجو را انجام داد نشان داده شده است. به عنوان مثال آدرس صفحه خانگی یوجین گارفیلد در قسمتی که Links نوشته شده است وارد شده است و با عمل جستجو می توان مشخص کرد چه مدارکی در شبکه خود را به این صفحه پیوند داده اند.

محدودیت زبانی

برخی از نظام های بازیابی اطلاعات مدارکی که به زبان های مختلف تالیف شده است را نمایه می کند. این خصوصاً در رابطه با موتورهای جستجو مصداق پیدا می کند که مدارک غیرانگلیسی زبان را نیز نمایه می کنند. محدودیت زبانی در واقع توانایی محدود کردن جستجو به زبان مورد نظر کاربر می باشد البته چنانکه گفته شد این توانایی از قبل بایستی در نظام در نظر گرفته شده باشد. در رابطه با پایگاه های اطلاعاتی و موتورهای جستجوی مهم حتی می توان زبان صفحه رابط را نیز تغییر داد. به این ترتیب کاربر می تواند جستجوی آسان تری را انجام دهد. به عنوان مثال در زمان جستجوی عبارت مورد نظر، در قسمت تعیین زبان مدارکی که قرار است بازیابی شوند، زبان انگلیسی را انتخاب کنیم؛ در این صورت فقط مدارکی که به زبان انگلیسی ذخیره شده اند، بازیابی می شوند.

حساسیت به حروف بزرگ و کوچک

این ویژگی عمدتاً به زبان های لاتین و جستجوی آنها ربط پیدا می کند. برخی از رابط ها، در زمان جستجو با حروف کوچک و یا بزرگ (به ویژه در مورد عملگرهای بولی) نتایج متفاوتی را ارائه می دهند. در برخی از رابط ها، باید عملگرهای بولی را با حروف بزرگ استفاده نماید و اصطلاحاً رابط

نسبت به حروف کوچک و بزرگ حساسیت دارد. در برخی دیگر استفاده از حروف کوچک و بزرگ هیچ تفاوتی ندارند و رابط هیچ حساسیتی نسبت به حروف کوچک و بزرگ ندارند. حساسیت به حروف بزرگ و کوچک ویژگی مهمی می باشد. خصوصا هنگامی که کاربر به دنبال بازیابی اسامی خاص باشد. برخی از ابزارهای جستجو این ویژگی را پشتیبانی می کنند. به عنوان مثال در زمان استفاده از اپراتورهای ولی در پاپ مد، باید این اپراتورها را با حروف بزرگ بنویسیم تا نظام پاپ مد آنها را اعمال نماید.

محدود کردن زمانی

یکی دیگر از ویژگی های مهم ابزارهای بازیابی اطلاعات، توانایی آنها در محدود کردن جستجو به زمانی خاص می باشد. به این ترتیب کاربر هم می تواند جستجوی خود را دقیقا به زمانی خاص مثلا یک سال یا حتی یک ماه محدود کند و یا اینکه دوره ای خاص شامل چندین ماه یا سال را محدود کند. محدود کردن زمانی در موتورهای جستجو توانایی پایگاه های اطلاعاتی را ندارد ولی می توان جستجو را به مدارکی که در زمانی معین از قبیل ماه یا سالی خاص به نمایه موتور جستجو افزوده شده اند، بازیابی نمود. همچنین هنگام نمایش اقلام بازیابی شده، ابزارهای بازیابی اطلاعات امکاناتی را در اختیار کاربر قرار می دهد که از آن جمله است تعداد اقلامی که در هر صفحه نمایش داده می شود که کاربر می تواند تعداد این اقلام را تغییر دهد. امکان دیگر، دسته بندی اقلام بازیابی شده براساس زمان ورود آنها به نظام یا میزان ربط آنها می باشد. همچنین پایگاه های اطلاعاتی توانایی ذخیره موقت اقلام مورد نظر کاربر را تا اتمام یک جلسه جستجو فراهم می کند. به عنوان مثال در زمان جستجو می توان در قسمت اعمال محدودیت زمانی، سال های ۲۰۰۰ تا ۲۰۰۵ را انتخاب نماییم. در این صورت تنها مدارکی بازیابی می شوند که تاریخ انتشار این فاصله زمانی را دارا می باشند.

توانایی های مرور

هنگامی که جست و جو به اتمام می رسد، توانایی های مرور این فرصت را در اختیار کاربر قرار می دهد که از بین اقلامی که بازیابی شده اند آنچه را که مطرح نیازهای او است انتخاب کند. دو روش برای نمایش خلاصه اقلامی که هر یک پرسش بوده اند وجود دارد. نمایش خطی و نمایش گرافیکی از خلاصه بازیابی شده، کاربر می تواند اقلام و محدوده هایی که نمایش پیدا می کنند را انتخاب کند. نظام های بازیابی اطلاعات امروزی همچنین حرکت راحت بین موارد بازیابی شده را فراهم می آورند..

این واقعیت که در صفحه نتایج معمولاً مدرک های بسیاری بازیابی می شوند که با نیازهای اطلاعاتی کاربر منطبق نیستند، توانایی های مرور می تواند به کاربر این یاری را برساند که بهترین ها را انتخاب کند.

رتبه بندی نتایج جست و جو

در بسیاری از نظام های اطلاعاتی اعم از رایگان و غیر رایگان موجود، کلید واژه که توسط کاربر به نظام وارد شده است پررنگ تر نشان داده می شوند. اقلام بازیابی به نظم های مختلف ارائه می شوند، رایج ترین نظم ارائه نتایج بازیابی شده براساس ربط می باشد که هر چه مدرکی دارای ربط بالاتری باشد در رتبه بالاتری قرار می گیرد. البته چنانچه قبلاً نیز توضیح داده شد، میزان ربط در پایگاه های اطلاعاتی عمدتاً براساس دفعات تکرار می باشد که گاهی نمی تواند ملاکی معتبر باشد. علاوه بر نظم براساس ربط، نظم براساس الفبای عنوان مدرک، نظم زمانی از قدیم به جدید و از جدید به قدیم رایج می باشد.

هنگامی که نظامی مدرک هایی را بازیابی و در اختیار کاربر قرار می دهد، این قانون همچنان حاکم است که با حداقل اطلاعات حداکثر اطلاعات را به دست آورد. این بدین معنا است که در رابطه با یک مدرک حداقل اطلاعات که اطلاعات کتابشناختی است به نمایش درآمده و کاربر بتواند از آن اطلاعات برای تصمیم گیری بهره برداری کند.

اطلاعاتی که در اکثر موارد و در مرتبه اول در یک نظام در ارتباط با مدرک های جست و جو شده بازیابی می شود اطلاعات کتابشناختی می باشد، که عبارتند از عنوان، اطلاعات کامل منبع و اطلاعات مربوط به مولف. در بسیاری از موارد رویت عنوان و مدرک برای تصمیم گیری کفایت می کند.

البته نوع دیگری از مرور نیز وجود دارد که از ابتدا قابل اجراء می باشد. در این نوع مرور معمولاً بخشی از نظام اطلاعاتی به طبقه بندی منابع از کل به جزء می پردازد. برای مثال در نظام های بازیابی اطلاعاتی مربوط به ادواری ها این کار از طریق رفتن بر روی نام نشریات و کلیک کردن بر روی آنها امکان پذیر است که ابتدا از لیست نمایش داده شده که ممکن است براساس الفبایی یا موضوعی یا زمانی باشد، نشریه مورد نظر انتخاب و سپس به سایت آن مجله در آن نظام وارد شده و حرکت ادامه می یابد تا کاربر به شماره مورد نظر خود از نشریه و در نهایت به مقاله مورد نظر خود دسترسی پیدا کند. اهمیت مرور از این طریق در این نهفته است که کاربر معمولاً می تواند به آنچه

که در یک نشریه یافته است دسترسی پیدا کند و یا آن مواردی که نمی توانسته برای آنها کلید واژه مناسب حدس بزند و یا به طور کل نمی دانسته است چطور آنها را جست و جو کند.

چگونگی انجام یک جستجوی کارآمد

۱- نظام اطلاعاتی متناسب با موضوع مورد نظر خود و یا کاربر را انتخاب نمایید. جستجوگران با تجربه به آسانی تشخیص می دهند برای هر نیاز به چه منابعی مراجعه نمایند به عنوان مثال اگر کاربر سوالی در رابطه با صنایع غذایی مطرح کند از منابع تخصصی مانند FSTA که منابع تخصصی تغذیه و صنایع غذایی را پوشش می دهد استفاده می کنند در این راستا ERI C برای جستجوی منابع آموزشی، CAB پایگاه کشاورزی، MEDLINE پزشکی و LI SA کتابداری می باشد.

۲- روش جستجو در نظام مورد نظر خود را از طریق منوی HELP فرا بگیرید. هر نظام یا پایگاه اطلاعاتی شیوه جستجوی خاص خود را دارد برای جستجوی موثر و بازیابی مرتبط تر بهتر است از منوی HELP شیوه ها و تکنیک های جستجو در آن مجموعه را آموخت. ۳- مفاهیم مهم را در جستجوی خود تشخیص دهید.

فهم سوال اصلی و شناسایی کلماتی که فاقد ارزش در بازیابی هستند بسیار مهم می باشد. در حین صحبت با کاربر مفاهیم مهم و کلیدی را مشخص نمایید.

۴- کلید واژه هایی که مفاهیم مورد نظر شمار را تعریف می کنند انتخاب نمایید. مفاهیم مهم و کلیدی به صورت کلید واژه اصلی در جستجو بکار می روند بنابراین فهم دقیق کلید واژه های اصلی پرسش گام اصلی در انجام یک جستجوی موثر می باشد.

۵- مترادف ها ، اسامی علمی، اصطلاحات مرتبط و یا سایر کلید واژه هایی که می تواند جست و جو را محدود و به بازیابی بهتر کمک کند در نظر داشته باشید.

در همه زبان ها از جمله فارسی و انگلیسی برای هم مفهوم بیش از یک کلید واژه وجود دارد. بسیاری از کاربران به این نکته توجه ندارند. عدم شناخت و بکارگیری مترادف ها و اصطلاحات معادل منجر به عدم بازیابی پیشینه های مرتبط با موضوع می شود. برای مثال دو کلمه Renal و Kidney هر دو به معنی کلیه می باشند و دو واژه Cancer و Neoplasm هر دو به معنی سرطان هستند.

- ۶- ویژگی های جست و جو که ممکن است کاربرد داشته باشند، از قبیل کوتاه سازی، عملگرهای مجاورت و عملگرهای بولین را در نظر داشته باشید.
 - ۷- کاراکترهای مشخص شده برای هر کدام از عملگرهای بولین را در هر نظام مشخص و بر اساس آنها جست و جو انجام دهید.
 - ۸- ارقام بازیابی شده را مورد ارزیابی قرار دهید.
 - ۹- در صورت نیاز و عدم بازیابی ارقام مناسب، راهبرد جستجو را تغییر دهید.
- اگر ارقام بازیابی شده مرتبط با نیاز اطلاعاتی نمی باشد جستجو را با شیوه ای دیگر تکرار نمایید.
- ۱۰- در صورت استفاده از موتورهای جستجو و عدم بازیابی ارقام مناسب، عمل جستج و جو را در موتور جستجوی دیگری تکرار نمایید.

خلاصه مطالب

اطلاع یابی یکی از بنیادی ترین نیازهای بشر برای توسعه و پیشرفت در تمدن های مختلف بوده است. ویژگی های اطلاعات موجود در شبکه جهانی وب شامل اطلاعات منحصر به فرد، روزآمد ، رایگان ،سهولت جستجو و بازیابی اطلاعات، دسترسی به اطلاعات چند رسانه ای سبب اهمیت این رسانه به عنوان یکی از محمل های قدرتمند اطلاعات گردیده است .حجم اطلاعات انبوه قابل دسترس از طریق شبکه جهانی وب نیاز به نوعی مهارت های اطلاع یابی و جستجو را حیاتی می نماید .جستجوهای که در نظام های اطلاعاتی الکترونیکی انجام می شود را می توان به اعتبار نوع اطلاعاتی که درخواست می شود و روشی که در هنگام جست و جو به کار گرفته می شود به جستجوهای معلوم و جستجوی نامعلوم یا موضوعی تقسیم می گردد .اتخاذ استراتژی مناسب جهت جستجوی ارقام اطلاعاتی مورد نیاز منجر به دسترسی به اطلاعات مناسب تر و مرتبط تر در زمان کمتر می گردد .استفاده از عملگرهای بولی ، جستجو از طریق عملگر نزدیک یابی یا مجاورتی، کوتاه سازی کلید واژه ها ، ریشه سازی، جستجوی عبارتی از جمله مهارت هایی هستند که متخصصین کتابداری و اطلاع رسانی باید به خوبی کسب نمایند تا بتوانند جستجوی دقیق تری را هدایت نمایند . رتبه بندی و پالایش نتایج جستجوی انجام شده منتهی به کارآمدی بازیابی اطلاعات مورد نظر می گردد که در متن به تفصیل و با مثال ارائه گردیده است.

نکات کلیدی

۱- مهارت های اطلاع یابی "داشتن دانش و مهارت های ویژه برای استفاده موثر از منابع و پایگاه های اطلاعاتی الکترونیکی می باشد. فرایند اطلاع یابی در محیط وب، می تواند از طریق " جستجو " یا " مرور " دنبال شود. این دو شیوه اطلاع یابی به ترتیب منجر به ظهور دو نوع ابزار کاوش یعنی "موتورهای کاوش " و " راهنماهای موضوعی وب " شده است.

۲- در یک تقسیم بندی کلی می توان جست و جوهایی که در نظام های اطلاعاتی صورت می گیرد را به دو دسته جست و جو برای " منابع معلوم " و جست و جو برای " منابع نامعلوم " تقسیم کرد که نوع اخیر را همچنین می توان "جست و جوی موضوعی " نام داد.

۳- فهم ویژگی های مشترک جستجو به لحاظ کاربرد آن در بیشتر نظام های بازیابی اطلاعات اهمیت دارد. به علاوه هر ابزار جستجو معمولا دارای دو گزینه می باشد که یکی به جستجوی ساده و دیگری به جستجوی پیشرفته اختصاص دارد.

۴- اهمیت اطلاعات شبکه جهانی وب به دلیل منحصر به فرد، روزآمد، رایگان، سهولت جستجو و بازیابی اطلاعات و دسترسی به اطلاعات چند رسانه ای می باشد.

۵- جورج بول، ریاضی دان انگلیسی، در سال ۱۸۴۷ منطق ریاضی را ارائه کرد که امروزه اساس بازیابی اطلاعات در بسیاری از پایگاه های اطلاعاتی و همچنین ابزارهای کاوش اینترنت است. از طریق منطق بول می توان مفاهیم مختلف (کلید واژه ها) را به منظور گسترش دادن یا محدود کردن دامنه جستجو با یکدیگر ترکیب کرد و به نتایج کاوش دقیق تر و مرتبط تری در یک پایگاه اطلاعاتی دست یافت. برای این منظور، از عملگرهای خاصی استفاده می شود که عبارتند از: AND،OR و

NOT.

۶- استفاده از عملگر AND ممکن است منجر به بازیابی کلیدواژه ها در " هرکجای متن " شود، بدون آنکه هیچ ارتباط مفهومی با یکدیگر داشته باشند (برای مثال یک کلیدواژه در سطر اول و دیگری در سطرانتهایی). این مشکل بویژه در نظام های بازیابی اطلاعات که فاقد نمایه سازی موضوعی هستند، بیشتر مشاهده می شود. برای جلوگیری از بازیابی اطلاعات غیر مرتبط یا به عبارت دیگر

"ریزش کاذب" در چنین مواردی، می توان از امکان کاوش دیگری تحت عنوان جستجوی نزدیک یابی استفاده کرد. استفاده از این امکان کاوش، بویژه در جستجوی آزاد در متن کاربرد فراوانی دارد. زیرا از طریق آن می توان کلید واژه ها و مفاهیم مورد نظر را در "کنار هم" بازیابی کرد. به تعبیر دیگر، از طریق جستجوی نزدیک یابی می توان ارتباط مکانی و جایگاه کلید واژه ها را در کنار هم تعیین کرد.

۷- کوتاه سازی عمل اضافه کردن علامتی به ابتدا یا انتهای ریشه یک کلمه برای بازیابی حالات مختلف آن کلمه می باشد. کوتاه سازی خصوصا برای بازیابی حالات مفرد و جمع یک کلمه مفید می باشد. عمل کوتاه سازی این امکان را فراهم می کند که با وارد کردن بخشی از کلمات، مشتقات مختلف آنها را نیز در فرایند کاوش بازیابی کنیم و به این ترتیب، دامنه جستجو را گسترش دهیم. اکثر ابزارهای کاوش اینترنت، امکان کوتاه سازی کلید واژه ها را از طریق علامت ستاره (*) در اختیار می دهند.

۸- استفاده از تکنیک ریشه سازی به کاربر اجازه می دهد تا کلماتی را بازیابی کند در یک یا دو حرف با یکدیگر فرق دارند.

۹- جستجوی عبارتی شامل کلید واژه هایی است که بایستی کنار یکدیگر بیایند. از این اپراتور هنگامی استفاده می شود که کلید واژه های مورد نظر کاربر بایستی حتما در کنار یکدیگر و همچنین به نظمی که وارد می کند پشت سرهم بیایند.

۱۰- قابلیت جستجوی پیوندهای فرامتنی در صفحات وب دیگر، به ما امکان می دهد که دریابیم کدام سایت ها، نشانی یک صفحه یا سایت وب مشخص را در صفحه خود گنجانده اند. به تعبیر دیگر، جستجوی پیوندهای فرامتنی (Hypertext Links) در سایر صفحات وب، نوعی تحلیل استنادی به شمار می آید.

۱۱- بازیابی مدارک موجود در یک نظام بازیابی اطلاعات از طریق جستجوی کلید واژه ای موثر می باشد ولی این تنها راه حل نمی باشد. هنگامی که نویسندگان آثاری از قبیل مقالات و کتاب ها به منابعی که استفاده کرده اند، چه در متن و چه در ماخذ، استناد می کنند این عمل جستجوی

منابع را برای آنها افزایش می دهد. قدرت جستجوی استنادی در این است که منابع مربوطی که تحقیقات جدید را تحت تاثیر قرار داده اند، را مشخص می کند. موسسه اطلاعات علمی سال ها است که این روش جستجو را به کار گرفته است. یوجین گارفیلد ایده تولید اطلاعات علمی با استفاده از این روش را معرفی نمود.