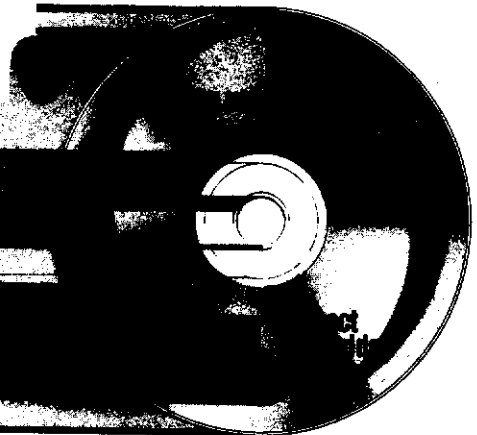




استراتژی کاوش

فاطمه نوشین فرد



مقدمه

رشد روز افزون اطلاعات و پیشرفت تکنولوژی موجب تحولی شگرف در زمینه اطلاع رسانی گردیده است. بدست آوردن اطلاعات از منابع چاهی مستلزم صرف وقت بسیار بوده و محدودیتهای زمانی و مکانی را به دنبال داشته است، لیکن با گسترش تکنولوژی، کامپیوتر جایگزین شیوه های دستی بازیابی اطلاعات شده و به دست آوردن اطلاعات با شیوه ای متفاوت از طریق کامپیوتر صورت می گیرد. پایگاههای اطلاعاتی پیوسته یکی از مهمترین ابزارهای تکنولوژی موجود در زمینه اطلاع رسانی است.

امروزه کاوش پیوسته و پرسش مستقیم از نظامهای کامپیوتری فعالیتی است که به طور فزاینده ای همه گیر شده است و به دلیل سرعت، دقت، قابلیت انعطاف، توان ترکیب منطقی و روز آمد بودن و سایر دلایل، تعداد استفاده کنندگان از آن روز به روز افزایش می یابد. لیکن کاوش پیوسته مستلزم صرف هزینه فراوان است و مدت زمان کاوش یکی از عوامل محاسبه کاوش پیوسته است، به طوری که هر چه زمان کاوش

افزایش می یابد بر هزینه نیز افزوده می شود. در کشور ما، علاوه بر هزینه استفاده از پایگاههای اطلاعاتی خارج از کشور، به دلیل بعد مسافت هزینه های سنگین مخابرات نیز افزوده می گردد. از این رو عامل زمان اهمیت بیشتری پیدا می کند و کاوشگر لازم است قبل از اقدام به کاوش پیوسته دارای طرح جامعی از کاوش باشد و استراتژی مناسب را بر آن اساس تدارک ببیند. این مقاله با توجه به مسائل و مشکلات فوق به تحلیل فرایند کاوش در نظامهای پیوسته می پردازد.

فرایند کاوش

کاوش بخش کامل کننده مصاحبه مرجع است و می تواند به همان اندازه پیچیده و غامض باشد. علی رغم کوششهایی که به دفعات برای تحلیل، نمایش و آموزش کاوش انجام شده است، همانند مصاحبه مرجع فرایندی ذهنی و مشهودی است. هدف نهایی کاوش مانند مصاحبه مرجع دستیابی به نتیجه است. مقدار زمانی که برای کاوش صرف می شود نیز عامل مهمی است. هر کاوشی که محل اطلاعات مورد نظر را مشخص

۳. طرح‌های ترکیبی بساتشخیص ترکیب مجموعه‌ها با استفاده از عملکردهای منطقی بول (NOT, OR, AND)؛

۴. تنظیم تصمیمات به طوری که واژه و ترکیباتی که تصور می‌رود از احتمال موفقیت بیشتری برخوردار باشند در ابتدا آزمایش گردند؛

۵. کنترل تصمیمات به طوری که نتایج حاصله منجر به تغییر استراتژی یا خاتمه کاوش گردد؛

اکنون پس از مراحل کاوش به تحلیل فرایند کاوش می‌پردازیم.

تحلیل پرسش

استفاده کنندگان نهایی نظامهای بازبایی اطلاعات نیازهای خود را دقیقاً مشخص نمی‌کنند. لکنکستر اهمیت تعریف دقیق آنچه را که مورد نیاز است، چنین بیان می‌دارد:

متقاضی ممکن است براساس نیاز به غلط تعریف شده‌اش خود در منابع به طور موفقیت‌آمیزی کاوش کند، ولی امکان ندارد بتوان استراتژی موفقیت‌آمیزی را بر همان مبنا تهیه کرد. برای انجام یک کاوش ماشینی با نتایج موفقیت‌آمیز حتی‌الامکان می‌بایست بیانات متقاضی را که صراحتاً نیاز اطلاعاتی وی را توصیف می‌کند بدست آورد. هر قدر شکاف بین درخواست بیان شده و نیاز اطلاعاتی بیشتر باشد احتمال موفقیت کاوشگر کمتر است.

کاوش پیوسته معمولاً توسط یک میانجی انجام می‌پذیرد. درخواست جهت کاوش پیوسته می‌تواند از طرق مختلفی چون مراجعه شخصی، تلفن، تلکس، نامه، فرمهای مخصوص درخواست کاوش پیوسته، پست الکترونیکی یا سایر رسانه‌های الکترونیکی دریافت گردد. هر طریقی که مورد استفاده قرار سرگیرد برای میانجی لازم است که اطلاعات هرچه بیشتر درباره پرسش کسب کند و اغلب مستلزم مصاحبه حضوری با استفاده کننده نهایی است و گاهی اوقات حضور وی در هنگام کاوش سودمند است. ترکیبی از دانش موضوعی استفاده کننده و دانش نسبت به سیستم معمولاً بهترین نتایج را به بار می‌آورد. چنانچه استفاده کننده قبل از کاوش پرسش را به خوبی مشخص نساخته باشد، لازم است میانجی در مشخص ساختن و تحلیل پرسش کمک نماید.

سازد. کاوشی مؤثر است. با این همه، کاوش می‌تواند مطلوب مکانی باشد که محل اطلاعات مورد نظر را با کمترین کار و زمان ارائه کند. به دلیل اینکه کاوش ساختار یافته‌تر از مصاحبه است تحلیل دقیق آن نیز امکان پذیر است. "بنسون" و "مالونی" مدل مختصر شده فرایند کاوش را به شرح زیر ارائه کرده‌اند:

در هر کاوش دو عامل وجود دارد. سیستم و پرسش، هر گونه ارائه سازمان یافته اطلاعات، یک سیستم است و هرگونه درخواست اطلاعات یک پرسش است ... مجموعه سستی یک کتابخانه یک سیستم اطلاعاتی است. برای نزدیکی و انطباق پرسش و سیستم ... کاوشگر بایستی ... شکافهای بین سیستم و پرسش را از میان بردارد، یا در واقع میان آن دو پل بزند... چنین روندی را اغلب فرایند کاوش می‌نامند.

مراحل کاوش

کاوش بازبایی اطلاعات از یک سیستم شامل مراحل زیر است:

۱. تحلیل پرسش مورد کاوش؛
 ۲. ترجمه مفاهیمی که در پرسش وجود دارد به زبان سیستم (زبان نمایه سازی پایگاه اطلاعاتی)؛
 ۳. فرمولبندی گزاره کاوش یعنی ترتیبی که واژه‌های انتخاب شده به کامپیوتر داده می‌شوند، ارتباط بین واژه‌ها از طریق منطق بول و دستورهای مورد استفاده.
- لنکستر در سال ۱۹۷۹ شش مرحله را در فرایند کاوش مشخص ساخت که شامل نیاز اطلاعاتی، بیان پرسش، انتخاب پایگاه اطلاعاتی، استراتژی کاوش (فرمول بندی کاوش)، کاوش در پایگاه اطلاعاتی و نمایش برون داد است. مراحل نخستین توسط فرایند باز خورد پالایش می‌شوند.

فرمول بندی کاوش

چهارمین مرحله در فرمت لنکستر فرمول بندی است که قلب یک استراتژی محسوب می‌شود و می‌تواند به تصمیمات زیر تقسیم گردد:

۱. برگرداندن سؤال به واژه‌هایی مناسب کاوش؛
۲. گسترش پرسش تا جایی که واژه‌های وابسته و جایگزین شونده را برای استفاده در کاوش دربرگیرد؛



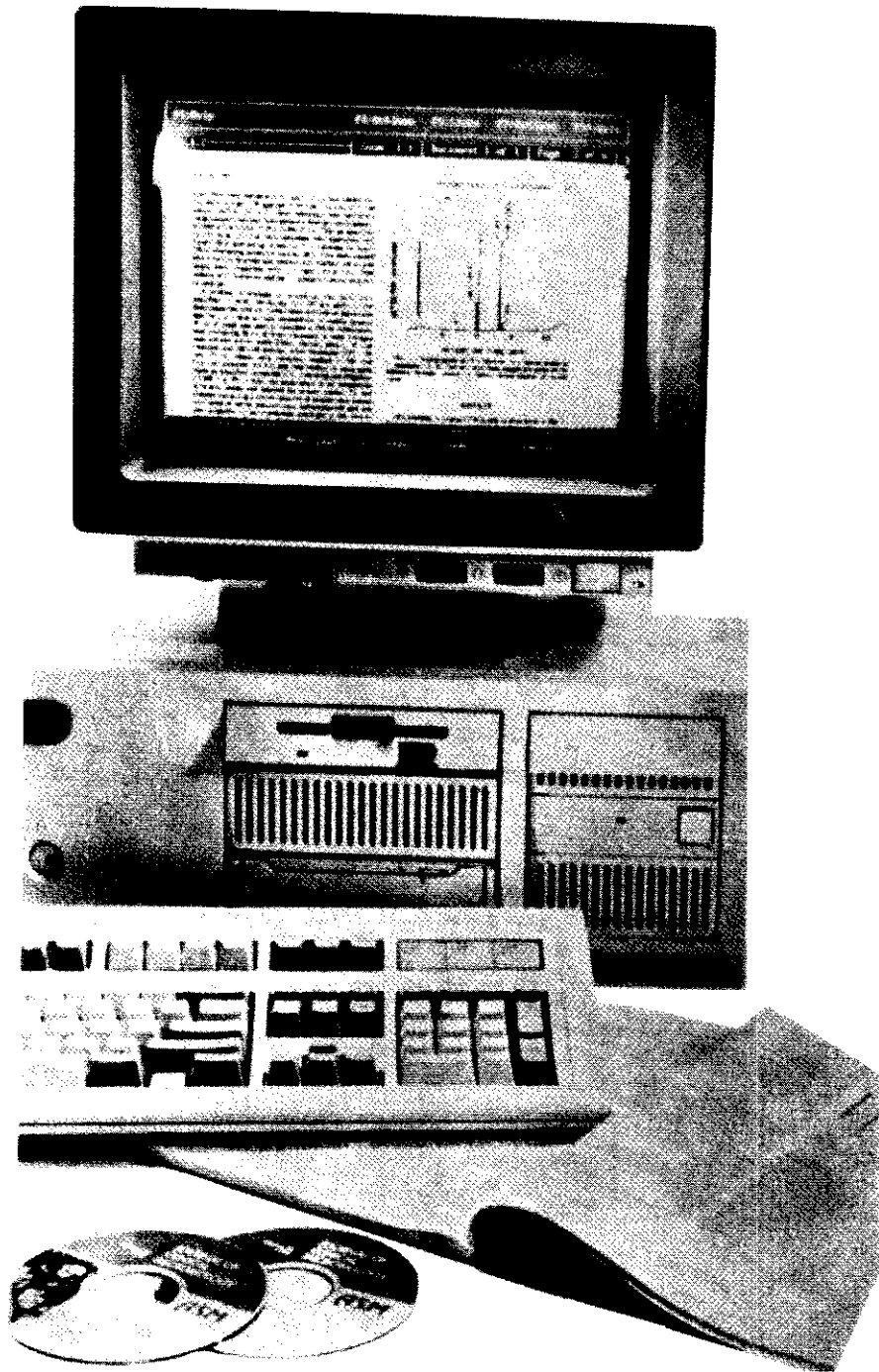
که مورد نیاز است می‌تواند توسط فرم درخواست کاوش پیوسته دقیقاً مشخص می‌گردد. این فرم می‌بایست طوری طراحی گردد که واسطه بتواند مفیدترین اطلاعات را برای تحلیل پرسش و انجام کاوش از آن استخراج نماید. استفاده کننده نهایی می‌بایست موضوع اخص یا واژه‌های مترادف آنچه را که نیاز دارد - گاه آنچه را نیاز ندارد - مشخص سازد. استفاده کننده اغلب موضوع را قبلاً به صورت دستی کاوش می‌کند، و کلید واژه‌هایی که در حین کاوش در چکیده‌ها و نمایه‌ها به کار برده هنگام کاوش پیوسته نیز مفید است. منابعی که قبلاً از سایر منابع بازایی شده‌اند کلید واژه‌های دیگری را انتقاء می‌نمایند که می‌توانند در گزاره کاوش بیابند. هرگونه محدودیتی در کاوش از قبیل زبان، تاریخ انتشار، تعداد منابع مورد نیاز (کلیه رکوردها یا دستیابی به جدیدترین رکوردها) و هزینه می‌بایست تا آنجا که امکان دارد قبل از شروع کاوش مشخص گردد.

حتی با فراهم بودن تمام این اطلاعات هنوز چیزهایی وجود دارد که قبل از شروع کاوش می‌بایست انجام شود. باید به این سؤال پاسخ داده شود که آیا پرسش برای کاوش پیوسته مناسب است. استفاده کننده ممکن است متقاضی کاوش پیوسته باشد، لیکن الزاماً از آنچه توسط کاوش پیوسته یا اشکال دیگر کاوش در کتابخانه و مرکز اطلاعات کسب می‌شود آگاه نیست. میانجی می‌بایست منابعی را تعیین و برای پاسخ به پرسش به بهترین نحو از آنها استفاده کند.

گاه از کاوش ناپیوسته، گاهی اوقات از کاوش پیوسته و غالباً از ترکیبی از هر دو استفاده می‌شود. هر منبعی که مورد استفاده قرار می‌گیرد میانجی می‌بایست دانش کاملی از موضوع آن خصوصاً هنگام مصاحبه حضوری داشته باشد. بعضی میانجی‌ها متخصص موضوعی هستند و فقط در یک رشته موضوعی مثل شیمی یا درحیطه وسیعتری مثل علوم و برخی دیگر در بسیاری از رشته‌ها کاوش می‌کنند که درگروه اخیر امکان ندارد واسطه متخصص موضوعی در همه حوزه‌ها باشد، لیکن می‌تواند با مطالعه زمینه‌ای و کمک استفاده کننده نهایی از استفاده از واژه‌های نامفهوم یا بازایی رکوردی نامربوط اجتناب نماید.

مرحله بعدی در فرایند کاوش معمولاً تعیین مفاهیم مربوط به کاوش است. برای مثال، عبارت گروه‌های خونی مصریان

بسیار از خدمات کاوش پیوسته در فاصله بعد از استفاده کننده نهایی فعالیت می‌کنند و همیشه امکان انجام مصاحبه رودرو وجود ندارد. جایی که امکان تماس شخصی وجود ندارد، استفاده کننده نهایی می‌بایست به هر طریقی اطلاعات دقیقی درباره پرسش خود به میانجی ارائه نماید. نوع اطلاعاتی



متممركز است، مانند Toxline يا Enviroline جنبه تبليغي دارد مانند Nasa؟ يا چندين رشته رامي پوشاند، مانند Social Seisearch .Scisearch

بامراجعه به مداركي كه توسط كامپيوتر ميزبان پيوسته تهيه مي گردد، منابعي كه در بخش قبل ذكر گرديد يا اطلاعاتي كه توليدكنندگان پايگاههاي اطلاعاتي دارند مي توان دامنه پايگاه اطلاعاتي رامشخص نمود. اين اطلاعات به صورت راهنما جهت استفاده كنندگان يافهرستي از اصطلاحات كنترل شده يا تركيبی از هر دو است. نشریاتی چون در «Database» و سایر نشریات پیوسته نیز مقالاتی را درباره پایگاههای خاص ارائه می نمایند.

نمایه ای که کامپیوتر میزبان پیوسته از پایگاههای اطلاعاتی خود ارائه می نماید صحیح ترین راه برای کسب اطلاع است که پایگاه اطلاعاتی چه موضوع خاصی را می پوشاند.

۲. فراگیری^{۱۱}

پوشش موضوعی پایگاه تاجه اندازه فراگیر است؟ اغلب بخش جداگانه ای از راهنماهای پایگاههای اطلاعاتی که توسط میزبان پیوسته تهیه می گردد به لیستی از مجلات تحت پوشش اختصاص می یابد. آیا تمام مقالات نشریات نمایه شده اند یا فقط بعضی از آنها؟ برای مثال در بعضی نشریات تحت پوشش مدلاین^{۱۲} تمام مقالات نمایه می شوند، ولی در سایر نشریات فقط مقالات منتخب نمایه می گردند. اندازه پایگاه اطلاعاتی چه قدر است؟ هرچند کمیت، کیفیت را تعیین نمی کند ولی می تواند در تعیین پایگاه اطلاعاتی پربارتر مفید باشد. هنگام کاوش لازم است بدانیم که پوشش پایگاه اطلاعاتی از چه زمانی شروع می شود، مثلاً پوشش CA Search از سال ۱۹۶۷ و Biosis و Inspec از سال ۱۹۶۹ است. بسیاری از پایگاههای اطلاعاتی که فاقد نسخه چاپی هستند فقط سالهای معدودی را تحت پوشش قرار می دهند.

۳. روز آمد بودن

پایگاه اطلاعاتی تا چه اندازه روز آمد است؟ بیشتر پایگاهها ماهانه یا زودتر روز آمد می شوند و بعضی هرسه ماه یک بار یا دیرتر روز آمد می گردند. با افزایش تولید اطلاعات مربوط به تجارت یعنی اطلاعات شرکتها نمایه انتشارات تجاری شامل

مومیایی شده دو مفهوم اساسی دربردارد: گروههای خونی- اجساد مومیایی شده و مفهوم جغرافیایی مصر.

کاوش به ندرت با مفهومی واحد انجام می گیرد. چنین پرسشهایی براحتی از طریق چکیده یا نمایه چاپی پاسخ داده می شوند و موضوع کاوش پیوسته نیستند. هرچند در موقعیتهایی که کاوش کننده به کلیه مدارک موجود درباره یک مفهوم نیاز دارد این نوع پرسش می تواند در سیستم پیوسته توسط زبان، تاریخ نشر و یا جایی که واژه مفهوم اصلی را دربرگیرد، محدود گردد.

انتخاب پایگاه اطلاعاتی

سؤال بعدی در فرایند کاوش این است که از کدام پایگاه اطلاعاتی استفاده گردد. امروزه پاسخ به این سؤال چندان آسان نیست، زیرا تعداد پایگاههای اطلاعاتی موجود بسیار زیاد است. بسیاری از این پایگاهها از طریق دو یاسه سیستم میزبان پیوسته قابل دستیابی هستند. حیطه موضوعی تحت پوشش این پایگاهها بسیار وسیع است و موضوع هر پایگاه اطلاعاتی اغلب شامل موضوعات وابسته نیز هست. هر میزبان پیوسته فهرستی از پایگاههای اطلاعاتی خود را تهیه می کند، اطلاعات جامع تر در مورد هر پایگاه معمولاً در یک راهنمای چاپی و اغلب به صورت پیوسته نیز در دسترس است. بسیاری از کامپیوترهای میزبان پیوسته نمایه ای از پایگاههای اطلاعاتی خود به صورت پیوسته دارند که تمام نمایه های اصلی از کلیه پایگاههای اطلاعاتی در یک نمایه ادغام گردیده است و کاوشگر می تواند با کاوش از طریق کلیه واژه برای تعیین پایگاه اطلاعاتی خود استفاده کند. نمونه ای از این نمایه خدمات پایگاههای اطلاعاتی یک یا چند میزبان پیوسته را ارائه می دهند که شامل پایگاههای اطلاعاتی در موضوعی خاص است. این منابع به صورت چاپی یا پیوسته در دسترس هستند.

ضوابطی که برای انتخاب پایگاه اطلاعاتی رعایت می گردد عبارت است از:

۱. پوشش موضوعی^{۱۰}

پایگاه اطلاعاتی چند موضوع را می پوشاند؟ آیا بر یک رشته علمی متممركز است. مثل CA Search یا Biosis؟ بر حل مسئله



ERIC وجود دارد. حدود ۳۷ درصد همپوشانی بین مجلات تحت پوشش Medlin و Embase وجود دارد. از ۶۰۴۰ عنوان موجود در هر پایگاه، Embase، ۴۳۴۹ رکورد و Medline ۳۸۸۱ رکورد را با یک همپوشانی ۲۲۰۶ در بر می‌گیرد.

۵. محتوای رکورد

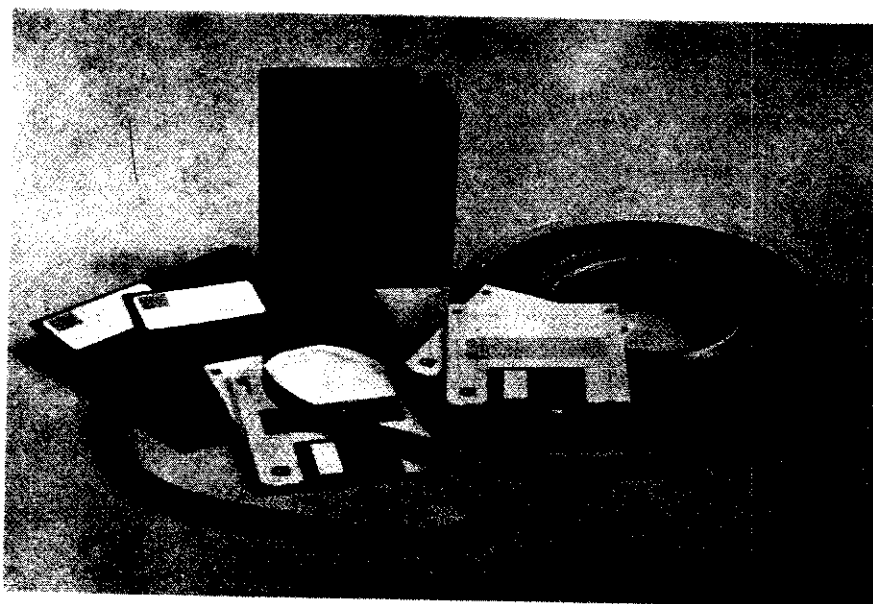
چقدر اطلاعات در هر رکورد پایگاه اطلاعاتی وجود دارد؟ راهنماهای کامپیوترهای میزبان، پیوسته این اطلاعات را ارائه می‌نمایند. هر بخش از رکورد فیلد، جزء طبقه یا پاراگراف می‌نامند. کلیه رکوردهای کتابشناختی اطلاعاتی درباره نویسنده، عنوان و منبع (عنوان نشریه، جلد شماره و تاریخ) را ارائه می‌دهند. به استثنای این اطلاعات تفاوت‌های بسیاری در اطلاعات ارائه شده وجود دارد. بیشتر پایگاه‌های اطلاعاتی حداقل یک فیلد برای اصطلاحات نمایه اختصاص می‌دهند و بسیاری دارای فیلد چکیده هستند. این پایگاه‌ها نه فقط به جهت اینکه لغات چکیده به عنوان یک واژه قابل کاوش هستند مفید می‌باشند، بلکه به استفاده کننده نهایی کمک بیشتری می‌نمایند تا تصمیم بگیرد آیا اصل مقاله را بخواند یا خیر. گاهی اوقات چکیده‌ها دارای اطلاعات کافی برای استفاده کننده‌ها هستند و دستیابی به اصل مقاله لزومی ندارد. بعضی پایگاه‌های اطلاعاتی دارای فیلدهای زیادی در هر رکورد هستند. مدولاین حداکثر دارای ۳۸ فیلد است. LCMARE^{۱۷} و BNB^{۱۸} بیش از حد ممکن دارای فیلد هستند، Social Scisearch تمام استثنائاتی که نویسنده در پایان مقاله‌اش ذکر می‌کند در رکورد سیستم پیوسته وجود دارد.

۶. نمایه سازی

نمایه سازی پایگاه چقدر فراگیری و شمول دارد؟ در Social Scisearch فقط لغات عنوان قابل کاوش است و هیچ گونه فیلدی برای واژه نمایه وجود ندارد. بیشتر پایگاه‌های اطلاعاتی نظیر Inspec، Eric و Medline یک فیلد برای واژه نمایه دارند که معمولاً حاوی واژه‌های کنترل شده است و اغلب فیلد دیگری برای واژه نمایه دارند که واژه‌های تکمیلی^{۱۹} توصیفگرها^{۲۰}، واژه‌های کنترل نشده^{۲۱}، واژه‌های تکمیلی آزاد^{۲۲} یا لغات متن^{۲۳} نامیده می‌شود معمولاً شامل

روزنامه‌ها و غیره ارائه می‌گردد که به روزکردن این اطلاعات در بعضی موارد به دقایق رسیده است. برای مثال، Dialog Business Wires و رویتر^{۱۳} هرپانزده دقیقه یک بار روزآمد می‌شوند.

یک پایگاه اطلاعاتی امکان دارد ماهانه روزآمد گردد، ولی از زمان انتشار مقالات چقدر گذشته است. برخی پایگاه‌های اطلاعاتی رکوردهای خود را هفته‌ها قبل از آنکه نسخه‌های چاپی راهی قفسه‌های کتابخانه‌ها شوند به صورت پیوسته ارائه می‌دهند. برای مثال کتابخانه ملی پزشکی آمریکا^{۱۴} از آوریل ۱۹۸۸ براساس تجربه، مدلاین را به صورت دو بار در ماه ارائه نمود. قبلاً مدلاین در آغاز هرماه روزآمد می‌شد و رکوردهایی را که در اندیس مدیکوس^{۱۵} چاپی پیدا نمی‌شد تا ماه بعد نداشت. امکان دارد به روز رسانی هر پایگاه اطلاعاتی در پایگاه میزبان پیوسته متفاوت باشد. رکوردهای مدلاین



چهار الی پنج هفته پس از چاپ مقاله به صورت پیوسته قابل دستیابی هستند.

۴. همپوشانی موضوع^{۱۶}

چقدر همپوشانی بین پایگاه‌های اطلاعاتی وجود دارد؟ بسیاری از رکوردهای مربوط به سرطان در سر پایگاه Cance line، Embase، Medline وجود دارد. در صد و پنجاه درصد از اطلاعات Exceptional Child Education Resources در

بعدی است. توصیه می‌شود ابتدا پایگاه اطلاعاتی ارزانتر مورد کاوش قرار گیرد (درونداد اولیه و اصلاح گزارش کاوش در نتیجه رکوردهای بازیابی شده، چاپ رکوردها و ضبط و جستجو برای استفاده در سایر پایگاههای اطلاعاتی در اولین پایگاه انجام می‌شود، سپس کاوش ضبط شده می‌تواند فراخوانی شود و در پایگاههای اطلاعاتی دیگر در زمان بسیار کوتاهتری انجام نمی‌گیرد).

میزبانهای پیوسته و تولید کنندگان پایگاههای اطلاعاتی قیمت پایگاههای خود را ارائه می‌دهند، همچنین انتشاراتی نظیر Clover Coparative Cost Chart قیمت پایگاههای اطلاعاتی کامپیوترهای میزبان پیوسته را ارائه می‌دهند و بنابراین مقایسه سریعی بین قیمت پایگاههای مختلف می‌تواند صورت گیرد.

تعدادی از کامپیوترهای میزبان پیوسته با امکان کاوش همزمان در چندین پایگاه، تصمیم‌گیری برای انتخاب پایگاههای مختلف را آسانتر ساخته‌اند. برای مثال در Esa_Irs کاوشگر می‌تواند تا ۸ پایگاه را با هم خوشه^{۲۶} کرده و به طور همزمان کاوش نماید. بعضی از خوشه‌ها در مورد موضوع یعنی از پیش تعیین شده‌اند. Dialog یک سرویس One Search ارائه کرده است که تا ۲۰ پایگاه می‌تواند به صورت همزمان مورد کاوش قرار گیرد و نرم افزار محاسبه Dialog زمان صرف شده در هر فایل را مشخص کرده و هزینه‌ها را بر اساس آن محاسبه می‌کند و هیچ گونه هزینه اضافی وجود ندارد.

انتخاب کامپیوتر میزبان پیوسته^{۲۷}

سازمانی امکان دارد از یک یا چند کامپیوتر میزبان پیوسته استفاده کند و تصمیم‌گیری برای پاسخ به پرسش محدود به میزبانهای پیوسته‌ای است که سازمان از آنها استفاده می‌نماید. گاهی اوقات یک پایگاه اطلاعاتی اغلب مورد نیاز فقط از طریق یک کامپیوتر میزبان قابل دستیابی است و بنابراین مشکلی در انتخاب کامپیوتر میزبان پیوسته وجود ندارد. هر چند یک پایگاه اطلاعاتی از طریق میزبانهای مختلفی قابل دستیابی است. تعیین کامپیوتر میزبان به طور مناسب شامل ضوابط زیر است:

اسامی خاص، واژه‌های جغرافیایی و واژه‌های موضوعی مهم کنترل نشده می‌باشد. سایر پایگاههای اطلاعاتی مانند Cancer line Biosis, CA Search از زبانهای طبیعی کنترل نشده استفاده می‌کنند. Medlin بیشتر مقالات حداقل با دوازده نمایه از زبان کنترل شده سرعنوانهای موضوع پزشکی^{۲۴} به طور عمیقی نمایه می‌شوند، گرچه بعضی از مقالات کاملاً^{۲۵} نمایه نمی‌شوند لیکن از پنج واژه استفاده می‌کنند.

آیا پایگاه اطلاعاتی از بعضی اشکال طرح رده بندی برای جستجوی مفاهیم گسترده‌تر استفاده می‌کند (برای مثال ساختار درختی Mesh)^{۲۵}؟

۷. نوع منابع موجود در پایگاه اطلاعاتی

پایگاه اطلاعاتی چه نوع انتشاراتی را در بر دارد؟ گاهی اوقات شامل یک نوع انتشارات مانند پروانه ثبت اختراع در World Potent Index یا استانداردها در BSI Standardline یا مشخصات و گزارشات کنفرانسها در نمایه مقالات کنفرانسها و نمایه گزارشات کنفرانسهاست. پایگاه اطلاعات کتابشناختی فقط شامل مقالات مجلات است؛ یا سایر انواع انتشارات از قبیل تک نگاشتها، رساله‌ها، پروانه‌های ثبت اختراعات، و روزنامه‌ها را هم در بر می‌گیرد. پایگاه اطلاعاتی اخبار تجاری فقط شامل روزنامه‌هاست و یا مجلات تجاری را نیز در بر می‌گیرد.

۸. هزینه

هزینه استفاده از پایگاههای اطلاعاتی مختلف می‌بایست بررسی گردد. در صورتی که لیست کوتاه پایگاههای اطلاعاتی شما با استفاده از ضوابطی که در بالا شرح داده شد محدود به دو پایگاهی باشد که هر یک از آنها می‌توانند به طور رضایت بخشی به پرسش پاسخ دهند، آنگاه هزینه می‌تواند در انتخاب اولین پایگاه برای کاوش نقش مهمی داشته باشد. در صورتی که احتمال کاوش یک پرسش در Eric به بهای هر ساعت ۳۰ دلار وجود داشته باشد و پاسخهای مورد انتظار از رأس هر دو پایگاه مشابه باشند، آیا عاقلانه است که Social Scisearch را به بهای ۱۲۰ دلار رساعت کاوش کنیم؟ در عمل امکان دارد برای اطمینان از پوشش و فراگیری، هر دو پایگاه را کاوش کنیم ولی چون زمان کاوش در پایگاه اول معمولاً طولانی تر از پایگاه



۱. پرسش گذشته نگر^{۲۸}

آیا کامپیوتر میزبان پیوسته تمام پایگاه اطلاعاتی و یا فقط بخشی از آن را دارد؟ در سیستمهای مختلف، پایگاه اطلاعاتی چند سال را می پوشاند؟

۲. روزآمد بودن

در صورتی که اطلاعات بسیار جدید مورد نیاز است، روز، ماه و دقیقه می تواند اهمیت داشته باشد.

۳. فایل های جداگانه در پایگاههای اطلاعاتی بزرگ

آیا میزبانهای پیوسته پایگاههای اطلاعاتی بزرگ از قبیل Biosis و Medline, CA Search را در یک فایل کامل یا چندین فایل کوچکتر و یا هر دو را ارائه می دهند. از حیث کاوش، کاوش فایل های کوچکتر سریعتر صورت می گیرد. خصوصاً فایل جاری زمانی که فقط رکوردهای جدید مورد نیاز باشد. امتیاز عمده کاوش فایل کامل این است که لزومی ندارد کاوشها را ضبط کرده و فایلها را عوض کنیم و جایی که روز آمد بودن مطرح نباشد پوشش جامعی دارد لیکن کاوش در آن از کاوش فایل های کوچکتر و جداگانه تر کندتر است.

۴. محتوای رکورد

آیا تمام کامپیوترهای میزبان پیوسته تمام اجزای رکوردهای پایگاههای اطلاعاتی خود را قابل دستیابی ساخته اند؟ برای مثال CA Search چکیده ها فقط از طریق CAS پیوسته و STN بین المللی قابل دستیابی است. توافق تولید کننده با سایر میزبانهای پیوسته، دستیابی چکیده ها از تمام پایگاههای اطلاعاتی را مجاز نمی سازد.

۵. خصیصه های کاوش^{۲۹}

خصیصه های کاوش که پیوسته توسط کامپیوترهای میزبان پیوسته متفاوت ارائه می شود تعیین می کند که کدام میزبان مورد استفاده قرار گیرد. در صورتی که کاوش خاصی مستلزم استفاده از کاوش همجواری یا رشته ای^{۳۱} یا هر دو باشد کدام میزبان پیوسته این خصیصه ها را فراهم می سازد. آیا کاوش در پایگاه اطلاعاتی خاصی در یک میزبان پیوسته می تواند به

زبان معینی محدود گردد، در حالیکه در میزبان دیگر این قابلیت وجود ندارد.

۶. جستجوی غایل متقابل^{۳۲}

این عمل می تواند به شکلهای متفاوتی باشد. ضبط کاوش یک پایگاه اطلاعاتی و کاوش در پایگاهها یا پایگاههای اطلاعاتی دیگر؛ ضبط کاوش روی یک پایگاه اطلاعاتی و انتقال اجزای رکوردها به پایگاه دیگر سپس با استفاده از این اجزای روی یک پایگاه اطلاعاتی جدید کاوش دیگری را انجام دهد؛ و سرانجام کاوش چندین پایگاه اطلاعاتی به صورت همزمان.

۷. قیمت

قیمت پایگاههای اطلاعاتی در میزبانهای متفاوت چه قدر است؟ این مسئله پیچیده ای است و به آسانی نمی تواند با انتخاب میزبانی که پایین ترین هزینه را برای زمان اتصال جهت کاوش در پایگاه اطلاعاتی خاصی عرضه می نماید حل گردد. سایر عوامل نیز باید در نظر گرفته شود:

(الف) هزینه های مخابرات برای دسترسی به یک کامپیوتر میزبان پیوسته: یک کاوشگر ممکن است در کالیفرنیا دستیابی محلی Toll-Hvee به یک میزبان پیوسته داشته باشد، ولی مجبور است برای دستیابی به سیستمی در بریتانیای کبیر نرخهای مخابرات بین الملل را پردازد، حتی اگر هزینه زمان اتصال برای دستیابی به آن سیستم از سیستم موجود در کالیفرنیا کمتر باشد.

(ب) آیا هزینه چاپ رکوردهای بازیابی شده به صورت پیوسته یا ناپیوسته و برای انتقال به یک پایگاه اطلاعاتی در کامپیوترهای میزبان پیوسته متفاوت است؟

(ج) اکثر کامپیوترهای میزبان پیوسته تخفیفهایی را برای استفاده بیشتر از پایگاههای اطلاعاتی خود ارائه می دهند. در صورتی که مبلغ کل کاوش در یک سال برای تخفیف کافی باشد، و با این فرض که تمام پایگاههای مورد نیاز در یک سیستم میزبان باشند، اهمیت دارد که تمام کاوشها روی یک سیستم میزبان متمرکز شود تا بتوان از تخفیف بیشتری بهره برد.

(د) در صورتی که کاوشگر با سیستم میزبان خاصی آشناتر و راحت تر باشد، مقرون به صرفه است که از این سیستم استفاده

استراتژی کاوش

پس از تعیین مفاهیمی که در پرسش هستند و انتخاب پایگاه اطلاعاتی در کامپیوتر میزبان پیوسته، کاوشگر باید تصمیم بگیرد، کدام استراتژی را آغاز نماید و چگونه استراتژی مورد نظر را تدارک ببیند. منظور از استراتژی کاوش در کاوش پیوسته مجموع کل تصمیم گیرها و اقداماتی است که در طول اجرای کاوش انجام می شود. تصمیماتی که به واسطه رکوردهای بازیابی شده و بازیابی نشده بر نتیجه کاوش تأثیر می گذارند، گاهی اوقات ممکن است لازم باشد در حین فرآیند کاوش پیوسته استراتژی اصلاح گردد. در واقع استراتژی چهار هدف زیر را در بر می گیرد:

۱. جور کردن تعداد رکوردهای مرتبط مطلوب؛
۲. اجتناب از بازیاب رکوردهای نامرتب؛
۳. احتراز از تعداد زیاد رکوردها و مجموعه های بسیار بزرگ؛
۴. احتراز از رکوردهای بسیار کم و مجموعه های بسیار کوچک یا نهی.

طرح ریزی می بایست قبل از ارتباط پیوسته صورت گیرد و همواره باید به خاطر داشت که کاوشگر با یک سیستم تعادلی کار می کند و اغلب مجبور است در حین کاوش به واسطه آنچه توسط گزاره اولیه کاوش بازیابی شده کاوش را اصلاح کرده و تاکتیکها را تغییر دهد. در واقع ماهیت تعاملی بودن سیستمهای پیوسته یکی از امتیازات اصلی آنهاست.

شود، حتی اگر پایگاه مورد نظر در سیستم میزبان دیگر هزینه کمتری داشته باشد.

ه) زمان پاسخ ممکن است در یک سیستم کندتر از دیگری باشد، بویژه هنگامی که گزاره های پیچیده کاوش توسط جستجوگر وارد می شود. بررسی مداوم قیمت یک پایگاه در سیستمهای میزبان متفاوت وقت گیر است، انتشاراتی نظیر Clover Comparative Cast Chart نام این اطلاعات را یکجا ارائه می نماید.

ضوابط بالا برای انتخاب یک میزبان پیوسته جهت پایگاه اطلاعاتی خاصی است با این فرض که جستجوگر با سیستمهای تمام کامپیوترهای میزبان پیوسته که آن پایگاه را ارائه می کنند به طور مساوی آشنایی داشته باشد. اغلب جستجوگر با یک سیستم آشناتر و راحت تر است. این مسئله به ویژه زمانی که در کاوش از نظر زمان تحت فشار هستیم، اهمیت دارد.

آشنایی با یک سیستم میزبان پیوسته خاص، کاوش سریعتر و مؤثرتری را فراهم می آورد که بر تفاوت قیمت با میزبانهای پیوسته دیگر که کاوشگر با سیستم آنها آشنایی کمتری دارد غلبه می نماید ولی لزوماً بهترین نتیجه را به بار نمی آورد، زیرا کاوشگر از بعضی امتیازات که در سایر سیستمهای میزبان پیوسته ارائه می شود (از قبیل آنچه در بند "الف" تا "د" شرح داده شد) محروم است.





استراتژی طرح جامعی برای تمام کاوش است و با انتخاب واژه‌های اخص و استفاده از فرمانها به کار گرفته می‌شود. معمول‌ترین کاوش شامل توجه پرسش به زبان پایگاه، اطلاعاتی مورد کاوش و استفاده از خصیصه‌هایی است که سیستم در اختیار می‌گذارد، استفاده از عملگرهای منطبق بول برای نشان دادن ارتباط بین واژه‌های مختلف در استراتژی کاوش و داشتن چند طرح برای اصلاح احتمالی کاوش پس از دیدن نتایج حاصل از اولین گزاره کاوش است. مطالب فوق با مثال زیر بوضوح نشان داده می‌شود: با جستجو در Fuel Abstracts مدارکی درباره "Oil Spillage in Atlantic Ocean" مورد نیاز است.

استراتژی طراحی شده برای این کاوش مستلزم بازیابی مدارکی است که واژه‌های Atlantic, Spillage, Oil را در بر داشته باشد (از واژه‌هایی استفاده شده که در خواست اصلی وجود داشته باشد) به همراه Spills و Leaks که مورد درخواست نبوده ولی مترادف واژه Spillage هستند (از واژه‌های دیگر نیز می‌توان برای جامعیت کاوش استفاده کرد، ولی برای این مثال این پنج واژه کافی است). واژه Spillage در گزاره نیامده است، زیرا Ocean به خوبی مفهوم Atlantic Ocean را می‌رساند.

با دسته بندی واژه‌های که یک مفهوم را می‌رساند این سه گروه ایجاد می‌شود:

Oil	Spillage	Atlanti
	Spills	
	Leaks	

این واژه‌ها می‌توانند با استفاده از عملکردهای بول به هم پیوندند. واژه‌هایی که در یک ستون هستند توسط عملگر OR به هم وصل می‌شوند، زیرا نقش جایگزینی دارند. واژه‌هایی که در ستونهای مختلف هستند توسط عملگر AND به هم وصل می‌شوند. زیرا آنها مفاهیم مختلفی را ارائه می‌نمایند و هر یک از آنها می‌بایست در رکوردهای بازیابی شده وجود داشته باشند.

گرچه بسیاری از خدمات کاوش پیوسته اتصال واژه‌ها را با استفاده از بیش از یک نوع عملگر در یک گزاره کاوش مجاز می‌سازند، لیکن برای نیل به نتیجه مطلوب بایست دقت کافی به عمل آید. بنابراین، به کاوشگر مبتدی توصیه می‌شود که

واژه‌ها را در یک گزاره کاوش فقط با استفاده از یک نوع عملگر به هم متصل نماید. با پیروی از این توصیه مرحله اول کاوش درباره "Oil Spillage in Atlantic Ocean" یافتن مدارکی که در پایگاه اطلاعاتی خواهد بود که واژه Leaks, Spills, Spillage را در بر داشته باشد. این واژه‌ها توسط اپراتور بول به هم وصل خواهند شد.

همان گونه که برای انتخاب پایگاه اطلاعاتی قبل از شماره پایگاه اطلاعاتی دستوری مورد نیاز است یک دستور نیز قبل از واژه‌های مورد کاوش لازم است تا به سیستم اعلام نماید این واژه‌ها باید در پایگاه اطلاعاتی مورد کاوش قرار گیرند. این دستور در خدمات کاوش تجسمی Find است. مرحله اول کاوش به قرار زیر است:

Find Spillage or spills or leaks?

Set	1	324	spillage
Set	2	208	spills
Set	3	159	leaks
Set	4	570	Spillage or s p i l l s or Leaks

سه واژه در جستجو به دنبال دستور توسط عملگر منطقی OR مجزا شده‌اند. رسم الخط نیز به اندازه واژه اهمیت دارد. یک فاصله می‌بایست در هر طرف عملگر گذاشته شود تا سیستم بتواند آنها را به عنوان عملگر بشناسد. اگر فاصله‌ای گذاشته نشود سیستم آن را به عنوان یک واژه در نظر خواهد گرفت.

دو واژه باقیمانده "Oil" و "Atlantic" با استفاده از دستور Find داده می‌شوند ولی این بار با این عملگر AND به هم می‌پیوندند. این عمل برای هر یک از واژه‌ها یک مجموعه تولید می‌کند و در مجموعه سوم آن دو واژه را با هم ترکیب نموده است. مجموعه‌های حاصل کاوش شماره‌های متوالی دارند و به دنبال مجموعه شماره ۴ ادامه پیدا می‌کنند.

Find Oil AND Atlantic

Set	5	725	Oil
Set	6	53	Atlantic
Set	7	26	Oil AND Atlantic

مرحله نهایی در کاوش، ترکیب مجموعه‌های نهایی هر مرحله است. یعنی نتیجه‌ای که با استفاده از عملگر باید OR در واژه‌های Leaks, Spills, Spillage (Set 4) با نتیجه استفاده از

است که به عنوان توصیفگر در این پایگاه اطلاعاتی به کار می رود و در یک تزاروس پیدا می شود. چنین اصطلاحی "Foot and mouth Disease" است. بنابراین استفاده از واژه های زبان کنترل شده اثر محدود کننده زیر را در بر خواهد داشت.

S3	1060	Foot	AND	Disease
S4	488	"Foot	AND	Mouth Disease"

قرار دادن عبارات توصیفگر در این مرحله باعث انتخاب عبارت به جای واژه های جداگانه Foot و Mouth Disease می شود که توسط اپراتور AND به هم متصل شده اند.

جایگزینی واژه ها و مفاهیم و استفاده از واژه ها و مفاهیم دیگر نیز در محدود کردن کاوش اهمیت دارد. عملکرد بول

اپراتور AND و ترکیب دو واژه Oil و Atlantic (Set 7) اسم اعدام شود.

Find Set 4 and Set 7

set 8 6 Set 4 and Set 7

شیوه های که جستجوگر درخواست را در سیستم وارد می کند بستگی به کامپیوتر میزبان پیوسته دارد. واژه های انتخاب شده خواص از زبان کنترل شده با آزاد باشند فقط بخشی از استراتژی کاوش شمرده می شوند. گاهی اوقات یک مفهوم نمی تواند به طور صحیح توسط یک یا چند واژه توصیف گردد. مفاهیمی که حیطه وسیعی را می پوشانند ممکن است به شکلی از کدهای رده بندی نیاز داشته باشند. برای مثال CA Search دارای ۸۰ بخش است که هر بخش دارای کد خاصی است. شکل مشابه آن در پایگاه مدلاین است. فهرست واژه های کنترل شده در پایگاه اطلاعاتی سرعنوانهای موضوعی پزشکی به صورت ساختار سلسله مراتبی تنظیم شده است و می توان با استفاده از یکی از موضوعات سرعنوان رکوردهایی را که نه تنها به آن سرعنوان بلکه به واژه های خاص تر در پایین سلسله مراتب نیز مربوط است، بازیابی نمود. برای مثال با قرار دادن پیشوند eaploc قبل از Peptic ulcer کاوشگر نه تنها می تواند رکوردهای مربوط به Peptic ulcer را به عنوان یک سرعنوان Mesh بلکه واژه های خاص تر نظیر Peptic duodenal ulcer و Esophagitis/Peptic و غیره را بازیابی نماید.

شیوه های محدود کردن کاوش

یکی از مشکلاتی که در فرایند کاوش پیش می آید زیاد بودن تعداد رکوردهای بازیابی شده است. در مثالهای زیر روشهای گوناگونی برای محدود کردن کاوش با استفاده از پایگاه اطلاعاتی کشاورزی CAB Dialog آمده است:

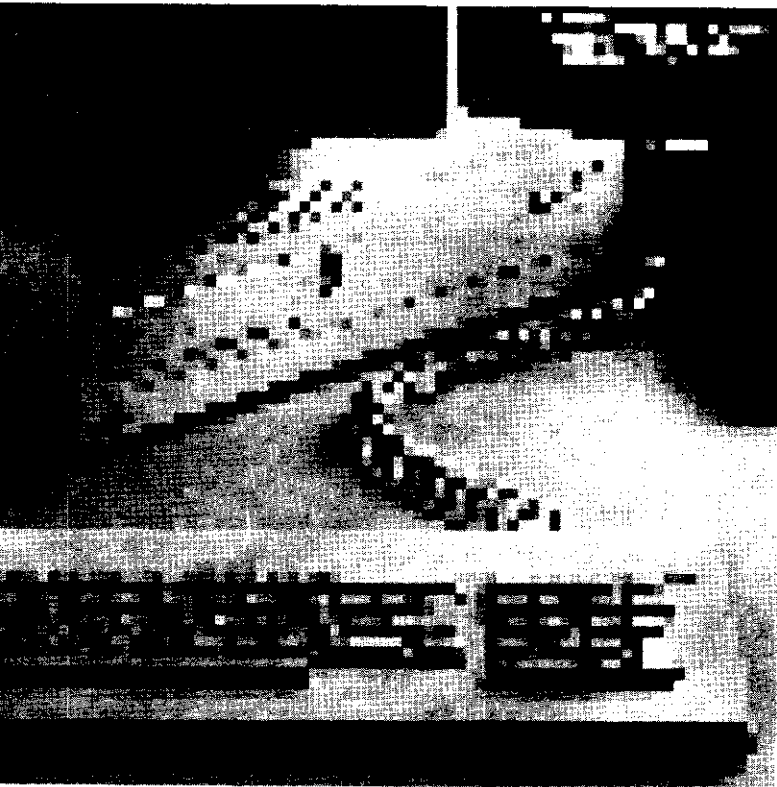
در مرحله اول مفهوم اصلی "Foot Disease" است که در سه مجموعه اولیه این مفهوم بازیابی می گردد.

SS	Foot	AND	Disease
S1	9191	Foot	
S2	35507	Disease	
S3	1060	Foot AND Disease	

یکی از روشهای محدود ساختن نتیجه کاوش مفهومی

AND اغلب یک ناکتیک مؤثر و سریع برای محدود کردن است. در مثال زیر کاوشگر یک اصطلاح جدید در استراتژی وارد می نماید و به اثری محدود کننده و مطلوب دست می یابد.

گرچه از Cattle، ۴۳۱۰۴ بار در پایگاه استفاده شده است ولی همواره Foot AND Mouth Disease رخ نداده است، با این





وجود ۲۲۲ رکورد با این تاکتیک از بین می‌رود.

استفاده از عملگر NOT نیز مؤثر است. مفهوم "Calves" در این مثال به صورت زیر استفاده شده است.

S "Foot AND Mouth Disease" A N D
Cattle NOT Calves

S4 488 "Foot AND Mouth Disease
43104 Cattle
8466 Calves

S6 244 ("Foot AND Mouth Disease" AND
Cattle) NOT Calves

تأثیر آن در این مورد خاص قابل توجه نیست و فقط ۲۲ مدرک کنار گذاشته می‌شود.

عملگر NOT می‌بایست به دقت در کاوش به کار برده شود، زیرا می‌تواند بعضی از رکوردهایی را که قابل قبول هستند به آسانی از بین ببرد.

تاکتیک دیگر محدود کردن واژه‌ها نسبت به رخدادشان در فیلدهای خاصی از رکورد است. این فیلدها ممکن است فیلدهای زبان طبیعی مانند عنوان و چکیده و یا فیلدهای زبان کنترل شده و توصیفگرها باشند. محدود کردن اصطلاح Cattle به عنوان یک توصیفگر در مثال زیر نشان داده شده است.

S "Foot AND Mouth Disease" A N D
Cattle/De
488 "Foot AND Mouth Disease"
37044 Cattle/De
219 "Foot AND Mouth Disease"
AND Cattle/De

امکان محدود کننده دیگر وارد کردن عبارت و اصطلاح کاوش با مقیاسهای همجواری است.

S8 1060 Foot AND Disease
S9 31 Foot (1W) Disease
S10 875 Foot(2W) Disease

در مجموعه شماره "۹" مجموعه "۸" به شدت محدود می‌گردد و با قبول یک لغت فاصله بین Foot و Disease تعداد رکوردهای بازایی شده به شدت کاسته می‌گردد. در مجموعه شماره "۱۰" محدودیت کمتر شده و حداکثر ۲ کلمه فاصله را مجاز می‌سازد.

استفاده از همجواری مشکل است و اگر خیلی محدود باشد ممکن است منجر به از دست دادن رکوردهای مربوط گردد و تأثیر چندانی در نتیجه مطلوب کاوش نداشته باشد.

تاکتیک محدود کردن با بازایی رکوردهای جدید در بسیاری از سیستمها امکان پذیر است و بنابراین استفاده از محدودیت سال نشر می‌تواند نتیجه مفیدی به بار آورد.

S7 219 "Foot AND Mouth Disease"
AND Cattle/De

S11 40 S7 AND PY=1987

محدود ساختن رکوردهای بازایی شده بر حسب زبان اصلی مدرک نیز می‌تواند به صورت مشابهی انجام شود و رکوردهایی که به زبان انگلیسی هستند بازایی می‌شوند.

روشهای گسترش جستجو

بازایی رکوردهای بسیار کم همانند بازایی تعداد زیاد رکوردها مشکل آفرین است. در صورتی که یک کاوش بسیار محدود باشد تاکتیکهای گسترش دهنده عبارتند از:

۱. واژه‌های کمتری توسط اپراتور AND به هم متصل گردند.

۲. جایگزینی واژه‌هایی که هم معنا هستند ولی رخداد بیشتری در پایگاه اطلاعاتی دارند.

۳. به جای کاوش در یک فیلد خاص واژه‌ها در چندین فیلد یا تمام فیلدها کاوش گردند.

۴. هرگونه محدودیتی در تاریخ نشر و زبان برداشته شود.

ضمناً در صورتی که واژه‌های توصیفگر رکوردهای محدودی را بازایی نمایند و امکان استفاده از زبان طبیعی در پایگاه وجود داشته باشد، می‌توان با استفاده از واژه‌های زبان طبیعی نتیجه کاوش را گسترش داد و رکوردهای بیشتری را بازایی نمود. هنگام استفاده از زبان طبیعی باید از تمام اشکال مختلف یک واژه استفاده شود. برای این کار برش واژه‌ها می‌تواند به بازایی رکوردهای بیشتری بیانجامد. مثلاً با استفاده از ریشه کلمه Fish و علائم برش می‌توان واژه‌هایی نظیر Fishes, Fishery, Fishing و غیره را بازایی نمود.

به حداکثر رساندن بازده بازایی

تاکتیکهای تحدید و گسترش فقط ابزاری برای تغییر یک کاوش

مفاهیم پرسش توسط واژه‌های مترادف و معادل با استفاده از عملگر OR توسعه می‌یابد و سپس تمام مفاهیم با استفاده از عملگر AND پاسخ نهایی را می‌زند. برای مثال:

S1 nuclear OR radioactive

S2 fallout OR pollution

S3 sheep OR lamb OR lambs

S4 S1 AND S2 AND S3

در کاوش فوق ۸ رکورد بازایی شده که سبک منطقی آن Building Blocks بوده است. گرچه این سبک در کاوش منطقی است ولی مستلزم صرف وقت و دانش کاملی از واژه‌های مناسب است. همچنین اگر تعداد رکوردهای بازایی شده کم یا زیاد باشد اصلاح استراتژی آسان نیست. این استراتژی غالباً تنها شیوه اجرای یک کاوش جامع درباره موضوع است، خصوصاً اگر میزان با جامعیت بالایی مورد نیاز باشد.

Successive Fractions در این سبک کاوشگر از عام به خاص حرکت می‌کند و با کل پایگاه اطلاعاتی شروع کرده و بتدریج حاصل کاوش را به بخشهای کوچکتر خرد می‌کند تا جایی که تمديد کاوش منجر به حذف رکوردهای مربوط نگردد. در این روش از عملگر منطقی OR و AND استفاده می‌گردد. برای مثال:

S1 318 swimming AND (women OR female)

S2 29 S1 AND fit ?

S3 8 S2 NOT advanced

در صورتی که تعداد رکوردهای بازایی شده بسیار زیاد باشد این سبک پس از استراتژی اصلی مورد استفاده قرار می‌گیرد. تمديد رکوردها از نظر تاریخ انتشار و زبان نیز برای نیل به نتیجه مطلوب در این سبک مؤثر است.

citationn pearl growing این سبک با یک رکورد کلیدی که قبلاً شناخته شده و در پایگاه اطلاعاتی بازایی گردیده است شروع می‌کند و سپس کاوش از راههای مختلفی گسترش می‌یابد. از کلید واژه‌هایی که برای توصیف رکورد کلیدی به کار رفته برای بازایی سایر رکوردهایی که همان کلید واژه‌ها را دربردارند استفاده می‌گردد، از نویسنده رکورد کلیدی برای بازایی سایر رکوردهایی که آن نویسنده یا نویسندگان درباره موضوعات مشابه نوشته‌اند استفاده می‌گردد یا از طریق فیلد منبع تالگان رکوردهای دیگر در موضوعات مشابه را که

از آنچه مورد انتظارات به مسر صحیح هستند. به حداکثر رساندن بازایی رکوردهای مربوط به موضوع به حداقل رسانیدن بازایی رکوردهای نامربوط اغلب به صورت به حداکثر رساندن جامعیت و مانعیت توصیف می‌گردد که از مقیاسهای مورد استفاده جهت ارزیابی آزمایش گرفته شده است.

میزان جامعیت = $\frac{\text{رکوردهای مرتبط بازایی شده}}{\text{تعداد کل رکوردهای مرتبط در پایگاه}}$ ^{۳۶}

میزان مانعیت = $\frac{\text{رکوردهای مرتبط بازایی شده}}{\text{تعداد کل رکوردها بازایی شده}}$ ^{۳۷}

برای مثال در یک آزمون ورودی توسط white, Mecain

و Griffith ^{۳۸} در سال ۱۹۸۷ از ۱۱ کاوش پیوسته در پایگاه اطلاعاتی مدلاین به طور متوسط ۶۴ رکورد بازایی شد که از آن تعداد ۳۶ رکورد مرتبط و ۲۸ رکورد نامربوط با موضوع بودند. بنابراین میانگین میزان مانعیت $\frac{۳۶}{۲۴}$ یا ۶۴ درصد بوده است. هرگاه تمام رکوردهای بازایی شده مرتبط با موضوع بودند، میزان دقت ۱۰۰ درصد می‌بود.

محاسبه میزان با مانعیت در عمل مشکل است، زیرا تعیین تعداد کل رکوردهای مرتبط در پایگاه اطلاعاتی توسط بررسی مستقیم مشکل است ولی باید دانست که میزان جامعیت و مانعیت همواره نسبت معکوس دارند.

۳۹ سبکهای استراتژی کاوش

چهار سبک عمده‌ای که تاکنون برای استراتژی کاوش مشخص شده عبارتند از: Brief Search (جستجوی مختصر)، Building Blocks (کنده ساخت، بناکنده)، Successive fraction (کسرهای متوالی) و citation pearl growing است در عمل امکان دارد بیش از یک سبک در جریان کاوش به کار رود.

Brief Search اساس این سبک استفاده از عملگر AND برای بازایی سریع بدون صرف وقت برای ردیابی و استفاده از واژه‌های مترادف دیگر است. امکان دارد کاوش با جامعیت کم نتیجه مطلوبی داشته باشد و فقط معدودی رکوردهای مرتبط بازایی گردد که می‌تواند استفاده کننده را قانع ساخته یا کاوش با استراتژی جامع‌تری ادامه یابد.

Building Blocks در هیافت Building Blocks همه



توسط آن سازمان (تالگان) و نویسندگان متفاوت انتشار یافته است بازایی می‌نماید.

خروجی نتایج کاوش

هنگامی که کاوشگر به مجموعه‌ای از رکوردهای مربوط به هم دست می‌یابد، چگونه این رکوردها به صورت فرمت مناسب در اختیار استفاده کننده نهایی قرار بگیرد؟ امروزه بیشتر کاوشهای پیوسته با استفاده از صفحه نمایش اطلاعات محدودی را در یک زمان در معرض دید قرار می‌دهند. فرمهای دیگری برای نمایش اطلاعات بسیار مورد نیاز است که عبارتند از:

الف) چاپ تمام کاوش یا بخشهایی از آن (بخشهای مورد نیاز) توسط چاپگری که به پایانه وصل شده است در زمانی که کاوش پیوسته انجام می‌پذیرد. استفاده از سرعتهای انتقال بالاتری که در شبکه‌های عمومی انتقال داده‌ها در سالهای اخیر ارائه شده (۱۲۰۰ باود و اخیراً ۲۴۰۰ باود)^{۴۱} نسبت به سابق باصرفه‌تر است هرچند چاپ هر رکورد به صورت پیوسته هزینه‌ای دربردارد که در پایگاههای اطلاعاتی مختلف متفاوت است.

ب) چاپ مجموعه رکوردها به صورت ناپیوسته^{۴۲} (یعنی با ورود یک دستور چاپ رکوردهایی از سیستم میزبان پیوسته تقاضا می‌گردد که این رکوردها بعداً برای کاوشگری استفاده کننده نهایی ارسال می‌گردد).

دو مزیت این شیوه هزینه و ارائه حاصل چاپی است. گر چه هزینه‌ای برای چاپ هر رکورد یا هر صفحه از اطلاعات منظور می‌گردد، ولی در مقایسه با هزینه اتصال پیوسته مقرون به صرفه‌تر است و ارائه رکوردها نیز بهتر از بسیاری چاپگرهای پیوسته است.

بعضی از کامپیوترهای میزبان پیوسته شیوه‌های دیگری را ارائه نموده‌اند که از نظر هزینه مزایای چاپ ناپیوسته را دارد و ضمناً فاصله زمانی بین سفارش چاپ ناپیوسته و دریافت کاوشگر یا استفاده کننده نهایی را کاهش می‌دهد. Dialog یک سرویس پست الکترونیکی به نام Dialmail در سال ۱۹۸۵ ارائه داده است. امکان درخواست انتقال چاپهای ناپیوسته به Dialmail وجود دارد که در صبح روز بعد برای چاپ به صورت پیوسته با هزینه اتصال بسیار کمتری آماده خواهند بود.

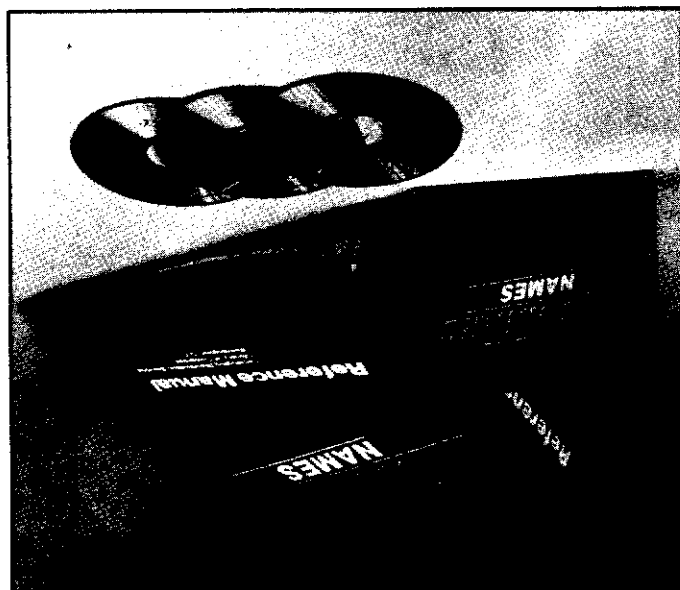
ج) شیوه دیگر برای کاوشگر این است که رکورد را به شکل الکترونیکی به منظور دوباره پردازی در ترمینال ذخیره سازد. رفتارهای قانونی در انتقال و ضبط اطلاعات وجود دارد. تولید کنندگان پایگاههای اطلاعاتی و کامپیوترهای میزبان پیوسته نگرانند که چگونه اطلاعات منتقل شده مورد استفاده قرار می‌گیرد. قوانین حق مؤلف در اینجا دخالت دارد. Nicolls انتقال بدون مجوز را در رابطه با قوانین جنایی بررسی کرد و به تفصیل درباره مفهوم مالکیت اطلاعات بحث کرد و نتیجه گرفت که اعطای حق مالکیت انحصاری اطلاعات



یک شبکه محلی یا خدمات پست الکترونیکی دارد. استفاده کننده نهایی با دریافت اطلاعات بدین صورت، می تواند آن را به شکل کامل یا گزینشی چاپ نماید یا آن را به شکل الکترونیکی به طور موقت یا دائمی ذخیره نماید.

معمولاً تقاضای اطلاعات کمتر هزینه کمتری را هم در بر خواهد داشت برای مثال هزینه چاپ یک رکورد کامل شامل تمام فیلدها در پایگاه اطلاعاتی Inspec در ISA-IRS به صورت پیوسته ۲۵/۲۰ پوند و به صورت ناپیوسته ۳۵/۳۰ پوند است^{۴۲} در صورتیکه هزینه چاپ اطلاعات کتابشناختی شامل نویسنده، عنوان و منبع در هر دو حالت کمتر است. بعضی از کامپیوترهای میزبان پیوسته دارای فرستهای از پیش تعیین شده هستند و در برخی، علاوه بر آن، امکان انتخاب فیلدهای خاصی براساس نیاز استفاده کننده وجود دارد. تصمیم گیری درمورد استفاده از فرست معمولاً در مرحله آماده سازی کاوش انجام می گیرد، زیرا کاوش فرست های مناسب در زمان ارتباط پیوسته پر هزینه خواهد بود.

چنانچه کاوش توسط واسطه انجام گیرد ارائه نتایج کاوش به استفاده کننده نهایی مهم است. بهتر است استفاده کننده نهایی بداند چه زمانی کاوش انجام شده است، از کدام پایگاه اطلاعاتی استفاده گردیده است و غیره. این اطلاعات می تواند به صورت یک گزارش مختصر به استفاده کننده نهایی ارائه گردد. این گزارش می تواند به صورت یک فرم همراه نتایج کاوش به استفاده کننده تحویل گردد.



گرفتاریهای غیرقابل قبول اجسامی و قانونی به دنبال خواهد داشت. بعضی ناشران حتی ادعا کردند که صنعت الکترونیکی اطلاعات با چکیده کردن مقالاتشان آنها را از درآمد و حق امتیاز محروم می کند.

Davenport سؤال می کند که آیا بحث انتقال و ضبط اطلاعات به همان اندازه اهمیت دارد که بعضی تولیدکنندگان و استفاده کنندگان آن را مهم تلقی کرده اند. وی انتقال و ضبط اطلاعات را تمهیدی برای تهیه کنندگان پایگاههای اطلاعاتی و کامپیوترهای میزبان پیوسته نمی داند زیرا بعضی از کامپیوترهای میزبان پیوسته دستورات خاصی را برای انتقال الکترونیکی ارائه کرده اند و هزینه انتقال و ضبط اطلاعات برای هر رکورد بیش از هزینه چاپ به صورت ناپیوسته است.

Sakida ابداع EsA-IRS را برای دستور انتقال و ضبط توصیف می کند که آن را حقی قانونی برای استفاده کننده می داند، زیرا وی حق استفاده دوباره از اطلاعات را دارد لیکن استفاده کننده حق انتقال، مبادله نسخه برداری و فروش دوباره هیچ يك از اطلاعاتی را که با انتقال و ضبط دریافت کرده است به شخص ثالث، بدون اجازه قبلی ISA-IRS نخواهد داشت و از اطلاعات منتقل شده برای ترکیب با سایر اطلاعات جهت ایجاد یک فایل ترکیبی جدید به صورت پیوسته یا هر شکل دیگری برای ارائه به شخص ثالث استفاده نخواهد کرد.

انتقال اطلاعات به شکل الکترونیکی برای ذخیره و ویرایش (مثلاً از بین بردن رکوردهای نامربط) نتایج یک کاوش مختصر قبل از تحویل به استفاده کننده نهایی مفید است، لیکن مستلزم امکانات سخت افزاری مناسب و حافظه RAM مناسب است. اگر کاوشگر به RAM بیشتری در یک میکروکامپیوتر، معنی کامپیوتر یا مین فریم دسترسی داشته باشد امکان پردازش دوباره اطلاعات منتقل شده افزایش می یابد. اطلاعات می تواند در یک پایگاه اطلاعاتی داخلی (درون سازمانی)^{۴۳} ذخیره گردد و در فواصل معینی با پایگاه اطلاعاتی محلی ادغام شود، بدین منظور که این پایگاه بتواند در محل بدون هیچ گونه هزینه اضافی مورد کاوش قرار گیرد. در بند الف و ب، اطلاعات به صورت چاپی یعنی لیستی چاپی از رکورها به استفاده کننده نهایی ارائه می گردد. با انتقال اطلاعات به شکل الکترونیکی کاوشگر قابلیت ارسال اطلاعات به استفاده کننده نهایی به شکل الکترونیکی از طریق



ارسال مدارک

هنگامی که استفاده کننده نهایی مدارک مورد نیاز خود را برای مطالعه کامل تعیین می کند، این مدارک می تواند از طریق امانت بین کتابخانه ای که با سیستم پیوسته ارتباطی ندارد سفارش داده شده و ممکن است از یک منبع محلی، منطقه ای، ملی یا بین المللی استفاده گردد. بسیاری از کامپیوترهای میزبان پیوسته امکان سفارشات به صورت پیوسته را ارائه می دهند که توسط آن استفاده کننده می تواند مدارک خاصی را به تهیه کنندگان مدارک سفارش دهند.

Dialog در سال ۱۹۷۹ خدمات سفارش پیوسته خود را به ۱۴ تهیه کننده ارائه داد و اکنون حدود ۷۵ تهیه کننده برای سفارش مدارک به صورت پیوسته در اختیار دارد. سفارش به صورت پیوسته^{۴۵} بویژه برای استفاده کنندگان نهایی که از منابع محلی یا ملی غنی برخوردار نیستند مفید است.

مسئله مهم سرعت ارسال مدارک است. بسیاری از تهیه کنندگان نرخهای متفاوتی را برای سفارشات اولویت دار و بدون اولویت ارائه می دهند. Colbect در یک بررسی مختصر از ده سال ارسال مدارک چهار دلیل عمده را برای افزایش سرعت ارسال و کاهش نسبی در هزینه ذکر می کند که عبارتند



از: خدمات تحویل شبانه،^{۴۶} پایگاههای تمام متن،^{۴۷} انتقال از طریق فاکس^{۴۸} و سفارش از طریق سیستم پیوسته.

ناشران و مرکز ارسال مدارک کوشیده اند تا از تکنولوژی دیسکهای نوری بهره جویند و مدارک را با صرفه بیشتری مهیا می سازد. در پروژه Adonis محتویات حدود ۲۲۰ نشریه پایه بیوشیمی روی CD-ROM^{۴۹} نگاهداری شده و درخواست برای مقالات این نشریات مستقیماً^{۴۹} از روی دیسک فشرده چاپ می گردد.

واسطه

شرایط و قابلیت هایی در میانجی - شخصی که کاوش از طریق سیستم پیوسته را به نفع استفاده کننده نهایی انجام می دهد - مطلوب است. این شرایط و قابلیت ها به دو دسته تقسیم می گردند: مهارتها و دانش که اکتسابی هستند و قابلیت ها که بیشتر ذاتی هستند و به شخصیت جستجوگر بستگی دارند.

دانش و مهارتهای اکتسابی

۱. تخصص موضوعی

هنگامی که کاوشگر در یک رشته موضوعی خاص بویژه علوم کار می کند و تعداد محدودی پایگاههای مربوط به آن موضوع مورد استفاده قرار خواهد گرفت تخصص موضوعی قطعاً^{۵۰} مطلوب ترین مهارت است. کاوشگر می تواند از کمک استفاده کننده نهایی براساس دانش و واژگان مشترک بهره جوید. مفاهیم مورد کاوش براساس زبان مشترک کاوشگر استفاده کننده نهایی بهتر تحلیل می گردند. در علوم اجتماعی و انسانی دارا بودن تخصصی موضوعی تا این حد حیاتی نیست زیرا این حیطه های موضوعی برای اشخاص غیر متخصص قابل فهم هستند و احتمال بیشتری در آشنایی با مفاهیم مورد کاوش وجود دارد. برای کاوشگرهایی که خدمات مرجع را در کتابخانه های عمومی و دانشگاهی ارائه می دهند زمینه آموزشی وسیع تر همراه با فنون سستی کاوش منابع بهترین اساس برای کاوش موفقیت آمیز است.

۲. شناخت پایگاههای مورد کاوش

آگاهی از ویژگیهای هر پایگاه خاص: فیلدهایی که می توانند مستقیماً^{۵۱} کاوش گردند، آشنایی با زبانهای نمایه سازی، اطلاع

قدرت تحلیل برای مشخص ساختن ماهیت اصلی پرسش مورد بررسی باشد. کلیه گزاره‌های کاوش پیوسته توسط زبان فرمان مانند ارتباطات منطقی بدل بیان می‌شوند، کاوش می‌بایست تحلیل پرسش را مطابق این علامت قراردادی انجام دهد.

۲. کنجکاوی و ثبات

جستجوگر می‌بایست دارای طبیعت کنجکاو باشد تا بتواند هیافت تحلیلی را که در فرایند کاوش اساسی است تعقیب نماید. کنجکاوی باید با ثبات همراه باشد تا کاوشگر را وادار سازد تمام رهیافتهای ممکن را نسبت به پرسش در نظر بگیرد.

۳. اعتماد به نفس

کاوش پیوسته مستلزم تصمیم‌گیری در حین ارتباط پیوسته است. برای اجرای کاوش مقرون به صرفه، این تصمیمات می‌بایست به سرعت اتخاذ گردند. کاوشگر می‌بایست از اعتماد به نفس کافی برخوردار باشد تا هنگامی که تحت فشار ارتباط پیوسته است و گاهی اوقات با حضور استفاده‌کننده نهایی، بتواند تصمیم‌گیری نماید.

۴. مهارت در ارتباطات

کاوشگر می‌بایست دارای مهارت در ارتباط باشد و بتواند با استفاده‌کننده نهایی ارتباط مفیدی ایجاد کند. می‌بایست اعتماد استفاده‌کننده نهایی را هنگام مصاحبه و در حین کاوش جلب نماید. باید در حین کاوش بتواند پیشرفت آن را تفسیر نموده و به اطلاع استفاده‌کننده برساند و سپس در صورت لزوم اطلاعات لازم را برای هرگونه اصلاح در کاوش از وی کسب نماید.

یادداشتها:

Borgman, Christine L.; Moghadam, Dineb and Corbett, Patrick. *Effective*

Online Searching. New York: Marcel Dekker Inc., 1984.

Convey, John. *Online information retrieval*. London: Clive Bingley, 1989.

Hartley, R. J., et al. *Online Searching: Principles and Practice*. London:

Bowker-Saur, 1990.

از اینکه حاصل کاوش چگونه می‌تواند محدود گردد و خصیصه‌های مشابه جستجوگر را قادر می‌سازد که سریعاً قبل از کاوش و در حین کاوش تصمیم‌گیری نماید. کمک از راهنماهای پایگاههای اطلاعاتی در حین ارتباط پیوسته مقرون به صرفه نیست. برای کاوشگر راحت تر است که در زمینه موضوعی خاص کار کند، زیرا تعداد پایگاههایی که مورد استفاده قرار می‌گیرند کمتر است. کاوشگرهایی که در حیطه‌های موضوعی وسیعی کاوش می‌کنند می‌بایست پایگاههای بیشتری را مورد کاوش قرار دهند.

۳. آگاهی کامل از زبانهای فرمان^{۵۱}

دانش نسبت به زبانهای فرمان سیستم‌های مورد استفاده، کاوشگر را قادر می‌سازد تا از امکانات هر سیستم استفاده نماید. به دلیل آنکه زبان فرمان پایه (اساس) ارتباط با سیستم است، در صورتی که کاوشگر دانش کاملی از فرمانهایی که احتمالاً مورد نیاز هستند نداشته باشد کاوش تعاملی در سیستم پیوسته به بهترین نحو اجرا نخواهد بود. آموزش ابتدایی در هر سیستم با حضور در یک دوره مقدماتی که توسط کامپیوتر میزبان پیوسته یا سایر سازمانها برگزار می‌شود صورت می‌گیرد، لیکن تسلط کامل بر سیستم فقط از طریق استفاده مداوم و تحت شرایط و کاوش‌های واقعی کسب می‌گردد.

۴. مهارت تایپ

بر اهمیت مهارت تایپ باید تأکید گردد. هرچند امروزه کمتر نیاز است کاوشگر در سیستم پیوسته تایپ کند زیرا استفاده از menu ها و اطلاعاتی که به صورت ناپیوسته تهیه شده و به سیستم انتقال می‌یابد این ضرورت را کمتر کرده است. هنگامی که در نظر است کاوشگر از طریق صفحه کلید پایانه با کامپیوتر میزبان پیوسته ارتباط برقرار نماید مشاهده اینکه چگونه وی می‌تواند در هر رهیانت به طور مؤثری عمل کند بدون فراگیری مهارت کافی در تایپ مشکل است.

قابلیت‌های ذاتی

۱. قدرت تفکر و تحلیل منطقی

لازم است کاوشگر بتواند به صورت منطقی فکر کند و دارای



26. Medical Subject Headings (MESH)

27. Tree Structure

28. Cluster

۲۹. در این متن Online Host معادل کامپیوتر میزبان پیوسته است لیکن در بعضی موارد نویسنده در متن اصلی عبارت میزبان پیوسته را به کار برده است که منظور همان Online Host می باشد.

30. Retrospective Coverage

31. Search Features

32. Proximity Searching

33. String searching

34. Cross-hill Searching

35. Affine

36. Interactive

37. R.J. Hartley, et al. *Online Searching: Principles and Practice*. (London: Bowker-Saur, 1990), p.87-90.

38. R.J. Hartley, et al. *Online Searching: Principles and Practice*. (London: Bowker-Saur, 1990), p.154-158.

39. Recall ratio

40. Precision ratio

41. R.J. Hartley, et al. *Online Searching: Principles and Practice* (London: Bowker-Saur, 1990), p.164.

42. Search Strategy Styles

43. R.J. Hartley, et al. *Online Searching: Principles and Practice*. (London: Bowker-Saur, 1990), p.170-171.

44. Baud

45. Affine Print

46. In-house database

47. John Convey. *Online Information retrieval* (London: Clive Bingley, 1989), p.77.

48. Online Ordering

49. Overnight Delivery Services

50. Full-text Databases

51. Facsimile Transmission

52. Compact Disc Read Only Memory

53. Subject Expertise

54. Command Languages

Katz, William A., *Introduction to Reference Work*. Vol. II. New York :

McGraw-Hill, 1978.

داورپناه، محمدرضا. "استراتژی جستجو" پیام کتابخانه، سال دوم، شماره چهارم، ص ۶۰ - ۵۷.

توضیحات:

۱. در اینجا روشهای دستی کاوش مدنظر هستند لیکن بسیاری از کاوشگران مجرب که از کامپیوتر استفاده می کنند اذعان دارند که فرایند کاوش چیزی فراتر از آموختن فرامین کامپیوتری است. بنابراین، مطلبی که در اینجا ارائه می شود در مورد کاوش کامپیوتری نیز صادق است.

2. William A. Katz, *Introduction to Reference Work* (New York: McGraw-Hill, 1978), p.92-3

3. John Convey, *Online Information Retrieval* (London: Clive Bingley, 1989) p.46

4. Formulation

5. Output Screening

6. Boolean Logical Operators

7. End-Users

8. Online Search

9. Non-Online Sources

10. Database

11. Online Host

12. Subject Coverage

13. Comprehensiveness

14. Medline (MEDLARS Online)

15. Reuter

16. National Library of Medicine

17. Index Medicus

18. Subject Overlap

19. Library of Congress Machine-Readable Catalog

20. British National Bibliography

21. Supplementary Terms

22. Identifiers

23. Uncontrolled Terms

24. Free Terms

25. Text words