طراحي سيستمهاي بازيابي اطلاعات بهينه در نرمافزارهاي كتابخانه‌اي و اطلاع‌رساني

فصلنامه اطلاع رساني. دوره ١٦، شماره ١و ٢

نوشته: علي گزني
كارشناس ارشد كتابداري در كتابخانة منطقه‌اي علوم و تكنولوژي شيراز

كليدواژه‌ها
سيستمهاي بازيابي اطلاعات، سازماندهي اطلاعات، بازيابي اطلاعات، نمايش اطلاعات

چكيده
بررسي فني نرمافزارهاي مختلف كتابخانه‌اي نشان مي‌دهد كه سرويس بازيابي اطلاعات در اين نرمافزارها به صورت بهينه طراحي و پياده‌سازي نشده‌اند. هدف پژوهش حاضر اين بوده است كه يك سيستم بازيابي اطلاعات بهينه را تعريف كند، راهنمايي براي طراحان سيستمهاي بازيابي اطلاعات، محققان و دانش‌پژوهان اين سيستمها را ارائه دهد و ارزيابي سيستمهاي بازيابي اطلاعات را ممكن سازد.
در اين مقاله به موضوعات زير پرداخته شده است:
امكانات و قابليتهاي سيستمهاي بازيابي اطلاعات ، بازخورد در سيستمهاي بازيابي اطلاعات، نحوة نمايش اطلاعات خروجي،ارزيابي سيستم بازيابي اطلاعات،و نمايش اطلاعات.

مقدمه
مسئلة بازيابي اطلاعات از زماني آغاز شد كه بشر سعي كرد محيط پيرامون خود را كنترل كند يا حداقل از فشارهاي خارجي كه باعث نابودي او مي‌شدند، جلوگيري كند. بشر براي ايجاد محيطي مطلوب براي ادامة بقاء به اتخاذ تصميمات سريع، صحيح و دقيق نياز داشت. كيفيت اين تصميمات به توانايي تصميم گيرنده در حل مسائل وابسته بود، ولي قبل از آن به ميزان ارتباط و كيفيت اطلاعاتي وابسته بود كه تصميم گيرنده براي حل مشكل فراهم آورده بود. بتدريج و در طول تاريخ، جمع‌آوري، سازماندهي و نگهداري امري متداول و مرسوم شد.
در چند دهة اخير، بازيابي رايانه‌اي مورد توجه خاصي قرار گرفته كه علت آن را مي‌توان چهار عامل دانست: افزايش اهميت زمان، تغيير در كميت و كيفيت اطلاعات، تغيير در ماهيت نيازهاي اطلاعاتي، تغيير در اهميت منابع اطلاعاتي. اين عوامل ، تغيير و توسعه در سيستمهاي بازيابي اطلاعات را به دنبال داشتند و موجب مطرح شدن بحث‌هايي از جمله لزوم تغيير در ابزارهاي اطلاعاتي، ايجاد و سيستمهاي ارتباطي جديد و استفاده از آن‌ها، تغيير معاني و اشاعه اطلاعات شده‌اند.
امروزه رشد تصاعدي و اهميت حياتي اطلاعات، لزوم صرفه‌جويي در وقت و هزينة جستجوكنندگان، جوان بودن شبكه‌هاي گسترده بخصوص از لحاظ سرعت، لزوم دستيابي سريع، جامع و مانع به اطلاعات خاص مورد نياز، از جمله مسائل مهمي هستند كه اهميت يك فرآيند بازيابي حساب شده، كنترل شده و كامل را آشكار مي‌سازند.

تعريف مسئله
در حال حاضر در سطح كشور نرمافزارهاي كتابخانه‌اي مختلفي طراحي شده‌اند، از جمله مي‌توان از نرمافزارهاي پارس آذرخش، نوسا، كاوش، بهمن برنا، گنجينه، ارغون و كتيبه نام برد. بررسي فني اين سيستمها نشان مي‌دهد كه سرويس بازيابي اطلاعات در اين نرمافزارها به صورت بهينه طراحي و پياده‌سازي نشده و كارآيي مطلوبي ندارد. با توجه به استفاده تعداد زيادي از مراكز كتابخانه‌اي و اطلاع‌رساني از اين نرمافزارها و نقشي كه در بازيابي و نيز گردش اطلاعات در جامعه دارند، مي‌توانند در مجموعه اطلاع‌رساني كشور موثر باشند. بدين لحاظ ارائة مباني صحيح طراحي و ارزيابي سيستمهاي بازيابي اطلاعات داراي اهميت مي‌باشد.

تعاريف عملياتي
سيستمهاي بازيابي اطلاعات: سيستمهايي كه به منظور بازيابي و پردازش داده‌هاي ساختار نيافته طراحي شده‌اند و به لحاظ نوع سازماندهي، ساختار پايگاه‌ها، راهبرد بازيابي، گروه‌هاي خدمت گيرنده، و فرآيندي كه در طي آن به درخواست‌هاي اطلاعاتي كاربران پاسخ مي‌دهند، از ديگر
سيستمهاي موجود متمايز مي‌شوند.
سازماندهي اطلاعات: تجزيه و تحليل داده‌هاي موجود در يك سيستم اطلاعاتي به منظور آماده‌سازي آن‌ها براي انجام بازيابي اطلاعات.
بازيابي اطلاعات: چرخه‌اي كه با فرمول‌بندي نياز اطلاعاتي شروع مي‌شود، با جستجو ادامه مي‌يابد، و با بازبيني و انتخاب اطلاعات مناسب خاتمه مي‌يابد.
نمايش اطلاعات: محاسبه و انتقال علائم به شكل هندسي به صورتي قابل درك و مشاهده توسط انسان، به منظور فهم و كشف روابط پنهان بين عناصر مختلف داده‌ها.

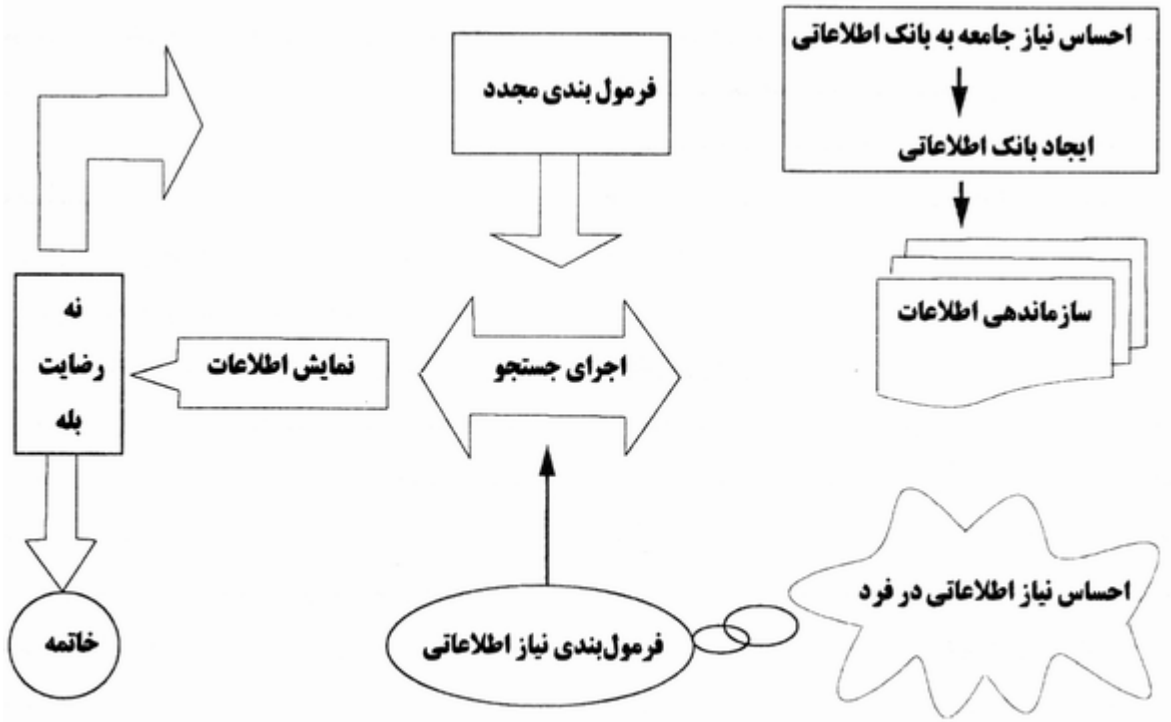
پرسش‌هاي اساسي پژوهش
نحوة سازماندهي اطلاعات، ساخت فايل‌هاي شاخص و مقلوب و ساير فايل‌هاي كمكي در سيستمهاي بازيابي

اطلاعات بهينه به چه صورت مي‌باشد؟
چه امكانات و قابليت‌هايي در سيستم‌هاي بازيابي اطلاعات بهينه لازم است تا كاربران بتوانند به بهترين شكل ممكن به بيان نيازهاي اطلاعاتي خود بپردازند؟ به بيان ديگر، چه قابليت‌ها و امكاناتي بايد در سيستم وجود داشته باشد تا كاربران بتوانند يك راهبرد بازيابي صحيح و دقيق را ايجاد كنند؟
بازخورد در سيستم‌هاي بازيابي اطلاعات بهينه بايد به چه صورت باشد؟ يا به بيان ديگر، سيستم بازيابي اطلاعات چگونه مي‌تواند كاربران خود را به مرتبطترين اطلاعات، در رابطه با نياز اطلاعاتيشان هدايت كند؟
نحوة نمايش اطلاعات در خروجي‌هاي حاصل از نتايج جستجو در سيستم‌هاي بازيابي اطلاعات بهينه چگونه كاربر را به انتخاب اطلاعات مرتبط هدايت مي‌كند؟ يا به بيان ديگر، شيوه‌هاي بهينه نمايش اطلاعات در خروجيهاي حاصل از نتايج جستجو كدامند؟
چگونه مي‌توان يك سيستم بازيابي اطلاعات را مورد ارزيابي قرار داد؟

روش پژوهش
پژوهش حاضر يك پژوهش از نوع تحليلي كاربردي مي‌باشد. بخشي از پژوهش به صورت بررسي كتابخانه‌اي، بر روي اجزاي
تشكيل دهنده سيستم‌هاي بازيابي اطلاعات صورت گرفته است. در اين بخش به مراحل زير توجه شده و اطلاعات لازم گردآوري شده است: ساختار داده‌ها، ساختار فايل‌هاي اطلاعاتي، ساختار فايل‌هاي نمايه، نمايه‌سازي خودكار، پردازش در سيستم‌هاي بازيابي اطلاعات، قابليت‌هاي سيستم‌هاي بازيابي اطلاعات،دسته‌بندي واژه‌ها، راهبرد بازيابي، الگوريتم‌هاي بازيابي، نمايش اطلاعات، ارزيابي سيستم‌هاي بازيابي اطلاعات. در بخش ديگري از پژوهش به بررسي چند نرم‌افزار داخلي و خارجي پرداخته شده است. انجام اين بررسي با توجه به اهداف پژوهش و به منظور تشريح وضعيت كنوني سيستم‌هاي بازيابي اطلاعات داخلي و خارجي مي‌باشد كه به همين منظور نمونه‌هايي انتخاب و تشريح مي‌شوند. از ميان نمونه‌هاي انتخاب شده در ايران نرم‌افزار كتابخانه‌اي پارس آذرخش ـ به لحاظ قدمت كاركرد، تعداد مراكز استفاده‌كننده، توانايي انجام جستجو در مقايسه با ساير همتاهاي خود ــ و نرم‌افزار كتابخانه‌اي پويا ــ به دليل جديد بودن و توانمندي سيستم فوق ـ مورد بررسي قرار خواهند گرفت. در ميان نرم‌افزارهاي خارجي نرم‌افزار «دايالوگ» (۱) به لحاظ قدمت كاركرد و اهميت آن مورد بررسي قرار خواهد گرفت و چند بستة نرم‌افزاري ديگر از جمله «اوآك» (۲) و «داوكواست» (۳) نيز به لحاظ دارا بودن موتورهاي جستجوي مبتني بر پردازش به زبان طبيعي و محيط گرافيكي، به صورت ضمني مورد بررسي قرار خواهند گرفت.
به لحاظ كاربردي و عملي بودن موضوع پژوهش، فهم بهتر مطالب و مفاهيم، كار عملي با زبان‌هاي برنامه‌نويسي++ C/C ، فاكس‌پرو و زبان پرس و جوي ساختاري (Structure Query language) صورت گرفت و يك سيستم بازيابي اطلاعات نمونه طراحي شد. در نهايت براساس مطالعات نظري و عملي، عوامل موثر در يك نظام بازيابي بهينه شناسايي و معرفي شدند.

يافته‌هاي پژوهش
براساس يافته‌هاي پژوهش و با بهره‌گيري از مدل‌هاي موجود مانند مدل «بوك‌استين» (۱۹۸۳) و «مدل بويس» و «كرافت» (۱۹۸۵)، يك مدل كه نشان‌دهندة ساختار يك سيستم بازيابي اطلاعات مي‌باشد ارائه شد. اين مدل در شكل ۱ مشاهده مي‌شود.
با توجه به مدل ارائه شده، بحث بر سه محور سازماندهي، بازيابي و نمايش اطلاعات متمركز شده كه در ادامه به بررسي يافته‌ها در اين زمينه‌ها مي‌پردازيم:

پژوهشگاه علوم انسانی و مطالعات فرهنگی
پرتال جامع علوم انسانی



شکل (۱): مدلی برای سیستم‌های بازیابی اطلاعات

شكل(١): مدلي براي سيستمهاي بازيابي اطلاعات

١ـ سازماندهي اطلاعات

هر سيستم اطلاعاتي داراي يك مبناي خاص براي تجزيه و تحليل اطلاعات مي‌باشد كه سيستم براساس آن به تفسير اطلاعات و مطابقت بين اقلام و درخواستهاي اطلاعاتي مي‌پردازد و بدين ترتيب بازيابي صورت مي‌گيرد. اين تجزيه و تحليل، سازماندهي ناميده مي‌شود. به لحاظ اهميت سازماندهي اطلاعات، سيستمهاي بازيابي اطلاعات به دو دستة هوشمند و غيرهوشمند تقسيم مي‌شوند. سيستمهاي بازيابي اطلاعات هوشمند براي اولين بار بين سال‌هاي ١٩٦٢ تا ١٩٦٥ در دانشگاه هاروارد طراحي و به كار گرفته شدند («سالتون»، ١٩٦٤؛ «سالتون»، ١٩٦٥). سيستمهاي هوشمند سيستمهايي مي‌باشند كه در آن‌ها تمام پردازش‌ها بر روي متن به صورت خودكار انجام مي‌شود.

هنگام پردازش متن‌هايي با زبان طبيعي، مسئلة پيچيدگي زبان و بيقاعدگي‌هايي كه در حوزة نحوي و معنايي وجود دارند خود را نشان خواهند داد كه بايد اين مشكلات را مد نظر قرار داد. اين مشكلات همان طوري كه «سالتون» (١٩٦٤) نيز به آن اشاره دارد هنگام اختصاص خودكار توصيفگرها به متن مشاهده مي‌شوند. اين مشكلات عبارت‌اند از واژه‌هاي مترادف، عملكردهاي نحوي، تركيبات نحوي، ارجاعات غيرمستقيم، متفاوت بودن معاني متناسب با هر متن، ارتباط معاني واژه با متن‌هاي قبلي، و تغيير معناي واژه‌ها در طول زمان.

در اين ميان واژه‌نامه‌ها قادر نيستند ابهامات و پيچيدگي‌هاي موجود در زبان را به صورت كامل برطرف كنند، اما اين بيقاعدگي‌ها را كاهش مي‌دهند. انواع مهمتر واژه‌نامه‌هايي كه در سيستمهاي بازيابي به كار مي‌روند عبارت‌اند از:

١) سياههٔ بازدارنده (٤) : شامل واژه‌هايي است كه نبايد مورد استفاده قرار گيرند.

٢) واژه‌نامهٔ ريشهٔ لغات و سياههٔ پسوندها: ريشهٔ لغات و پسوندهاي آن‌ها به صورت جداگانه در هر مدخل نگهداري مي‌شوند. يكي از روش‌هاي متداول در ريشه‌يابي، استفاده از روش «لاوينز» مي‌باشد.

٣) واژه‌نامهٔ عبارات: واژه‌ها براساس رخداد همزمان و ميزان اين رخداد، به صورت يك عبارت در نظر گرفته مي‌شوند. با واژه‌نامه كيفيت و كارآيي يك تحليل محتوايي بالا مي‌رود. با تعيين عبارات، حالات مبهم معنايي در متن بمراتب كاهش خواهد يافت. به عنوان مثال واژهٔ «برنامه‌نويسي» و «زبان» هر يك به تنهايي بيان‌كنندهٔ حالات مختلفي هستند، اما قطعاً «زبان برنامه‌نويسي» يك معناي مشخصي را القا مي‌كند.

٤) واژه‌نامهٔ مفاهيم (مترادف‌ها): در هر مدخل تعدادي از واژه‌هاي مترادف يا مفاهيم هم‌طبقه آورده شده است. گسترش دامنهٔ مفاهيم و اصطلاحات در فرمول جستجو از طريق جايگزين كردن واژه‌هاي مبهم با واژه‌هاي هم‌طبقه، از طريق اين واژه‌نامه صورت مي‌گيرد.

٥) سلسله مراتب مفاهيم: مشابه طرح‌هاي رده‌بندي كتابخانه‌اي مي‌باشد؛ با حركت به بالاي رده به قسمت‌هاي اعم و با حركت به قسمت‌هاي پايين به قسمت‌هاي اخص خواهيم رسيد.

به منظور سازماندهي واژه‌ها بايد دامنهٔ شمول واژه‌ها و نوع دسته‌بندي مفاهيم را مدنظر قرار داد. واژه‌هاي موجود در متون به سه دسته واژگان عمومي، واژگان با رخداد بالا، و واژگان با رخداد پايين قابل تقسيم مي‌باشند كه بايد در مورد شمول هر يك، تصميم‌گيري صورت گيرد. نوع طبقه‌بندي واژه‌ها و مفاهيم نيز از مسائلي هستند كه در ضرايب بازيابي و دقت، تاثير بسزايي خواهند داشت. بنابراين بايد در مورد اعم يا اخص بودن طبقات ، تصميم‌گيري صورت گيرد. پاسخگويي به مسائل فوق، وابسته به سياستگذاري سيستم در رابطه با ضرايب بازيابي و دقت مي‌باشد. اگر بخواهيم ضريب بازيابي را بالا ببريم بايد تمام واژه‌هايي كه به نحوي در بازيابي موثر باشند، به واژه‌نامه اضافه شوند. در مقابل اگر بخواهيم ضريب دقت بالاتري داشته باشيم و اقلام مرتبطتري مورد جستجو قرار گيرند، بايد به سمت

واژه‌نامه‌هاي تخصصي‌تر حركت كنيم.
ساخت واژه‌نامه‌ها به صورت خودكار با دو روش تمام خودكار و نيمه خودكار قابل اجرا مي‌باشد. به دليل گستردگي كار، اين كار به صورت كميته‌اي و گروهي انجام مي‌شود. اين كميته تعدادي سؤال را تعيين و مطرح مي‌نمايد و در نهايت، استانداردهاي لازم را براي تدوين واژه‌نامه پيشنهاد مي‌كند (دوول، ١٩٦٥). اغلب روش‌هاي مورد استفاده در ساخت واژه‌نامه‌ها با روش تمام خودكار، به واژه‌هاي موجود در مجموعة مدارك وابسته مي‌باشند
(دوول، ١٩٦٥؛ دنيس، سالتون، ١٩٦٦). عملي بودن روش خودكار بيش‌تر از آن كه به يك شاخه موضوعي خاص وابسته باشد، به مجموعة مدارك مورد پردازش بستگي دارد. بنابراين، بايد در رويه‌هاي پردازشي تغييراتي اعمال شود و از قضاوت انسان‌ها نيز استفاده كنيم. روش‌هاي مورد استفاده براساس اهداف سيستم‌ها متفاوت‌اند (اسپارك جونز، ١٩٦٥؛ لِوري، آبراهام، ١٩٦٥؛ رايزنر، ١٩٦٥).
به منظور بالا بردن كارايي سيستم‌هاي بازيابي اطلاعات از روش‌هاي محاسبة همبستگي و خوشه‌بندي اطلاعات نيز استفاده مي‌گردد. محاسبة همبستگي براساس دفعات رخداد همزمان واژه‌ها، به عنوان معياري براي سنجش ميزان مشابهت و همبستگي بين دو متن به كار مي‌رود. ميانگين حاصل از اين تعداد رخداد، بنحو شايسته‌تري بيان‌كنندة اين همبستگي خواهد بود. اگر بخواهيم اين همبستگي مقدار بالاتر و معنادارتري را نشان دهد بايد اين ارزيابي براساس لغات مشترك و غيرمشترك بين دو متن صورت گيرد. به منظور انجام اين محاسبه، فرمول‌هاي مختلفي پيشنهاد شده كه در پژوهش حاضر به بررسي فرمول «سالتون» و «مك‌گيل» (١٩٨٣) پرداخته شد كه به قرار زير است:

$$SIM(j,k)=\frac{\sum_{i=1}^{n} t_{ij}.t_{ik}}{\sum_{i=1}^{n} (t_{ij})^2+\sum_{i=1}^{n} (t_{ik})^2-\sum_{i=1}^{n} t_{ij}.t_{ik}}$$

در اين فرمول tij نشان‌دهندة ضريب وزني (دفعات تكرار) واژة i در متن j ، و n تعداد كل واژه‌هاي موجود در واژه‌نامه است كه مي‌تواند مجموع واژه‌هاي موجود در دو مدرك باشد. نكتة مهم اين است كه اگر واژة در i در مدرك j موجود نباشد، tij برابر صفر خواهد بود و اگر واژة i هيچ يك از دو مدرك وجود نداشته باشد در آن صورت نيز tij.tik برابر صفر خواهد بود.
خوشه‌بندي براي دسته‌بندي ركوردهاي اطلاعاتي مشابه با هم مورد استفاده قرار مي‌گيرد. خوشه‌بندي از طريق فرمول‌هاي مختلفي قابل اجرا است. يكي از ساده‌ترين فرمول‌هاي موجود كه‌توسط «بُنر» ارائه شده در پژوهش حاضر مورد بررسي قرار گرفته كه به قرار زير است:

$$S_{ij}=\frac{T_{ij}}{T_{ii}+T_{jj}-T_{ij}}$$

در اين فرمول Tij بيانگر تعداد واژه‌هاي مشتركي است كه در دو ركورد j و i وجود دارند. Tii و Tjj به ترتيب نشان‌دهندة كل واژه‌هاي موجود در ركوردهاي i و j مي‌باشند. اولين قدم در ايجاد يك خوشه ، تشكيل يك ماتريس از مشخصه‌هاي ركوردها مي‌باشد. به هر ركورد به صورت مجزا يك سطر اختصاص داده مي‌شود و هر واژه يا مشخصه به صورت يك ستون نمايش داده مي‌شود.

٢ـ بازيابي اطلاعات
مطالعات انجام شده توسط «پني‌من» (١٩٧٥) و «چاپمن » (١٩٨١) در موضوع جستجوي اطلاعات نشان داد كه جستجوگران بانك‌هاي متني و كتابشناختي از يك چرخ ة دستورها به صورت مشترك استفاده مي‌كنند. در اين مطالعات تاكيد بر دستورهاي پراستفاده‌تر بود تا اين كه بخواهند سلسله مراتب استفاده از اين دستورها را مورد توجه قرار دهند. چرخ ة دستورها شامل ٧ مرحله زير است:

١) انتخاب يك بانك اطلاعاتي براي انجام جستجو
بر مبناي بانك‌هاي موجود و نياز اطلاعاتي كاربر، يك بانك اطلاعاتي‌توسط او انتخاب مي‌شود. آنگاه سيستم خلاصه‌اي از اطلاعات، مانند محدودة تاريخي ركوردها، تعداد ركوردها، قيمت و... را در اختيار كاربر قرار مي‌دهد.

٢) جستجو براي واژه‌هاي مورد نظر در بانك واژگان
قبل از انجام جستجو و براي فرمول‌بندي جستجو، كاربر مي‌تواند هر يك از واژه‌نامه‌هاي موجود در سيستم را كه در بخش سازماندهي اطلاعات به آن اشاره شد مورد استفاده قرار دهد. واژه‌نامه‌ها نقش بسيار مهمي در بازيابي اطلاعات ايفا مي‌كنند، حداقل استفاده‌اي كه از بانك واژگان به عمل مي‌آيد، اين است كه كاربر خواهد فهميد كه آيا واژه‌هاي مورد نظر او در بانك اطلاعاتي وجود دارند؟ شكل صحيح آن‌ها به چه صورت مي‌باشد؟ واژه در چند ركورد، در كدام فيلدها و به چه ميزان تكرار شده است؟ اگر واژة درخواستي كاربر، اشتباه تايپ شده باشد يا در بانك واژگان وجود نداشته باشد، مقدار عددي صفر براي آن نمايش داده مي‌شود. كاربر با استفاده از اين فايل‌ها مي‌تواند به واژه‌هاي مرتبط با واژه‌هاي مورد نظر خود دست يابد و به اخص يا اعمتر كردن فرمول جستجو بپردازد.

٣) ايجاد فرمول جستجو و انجام جستجو
فرمول‌بندي صحيح و دقيق نيازهاي اطلاعاتي كاربران وابسته به امكانات بازيابي اطلاعات براي فرمول‌بندي مي‌باشد . هر چه قابليت‌ها و امكانات بهتري فراهم آورده شود، كاربر راحت‌تر و آسان‌تر مي‌تواند به بيان نيازهاي خود بپردازد. امكانات لازم براي فرمول‌بندي جستجو در يك سيستم بهينه شامل امكان استفاده از عملگرهاي بولي، عملگرهاي مجاورت، جستجوي املايي، كوته‌سازي واژگان، تعيين محدوده در داده‌هاي كمّي، گسترش مفاهيم و اصطلاحات،

استفاده از ضريب وزني، سياهة واژه‌ها مي‌باشد. بعد از انجام جستجو، نمره‌گذاري مجموعه‌هاي بولي و اندازه‌گيري مشابهت، در بازيابي بهتر اطلاعات كاربران بسيار مفيد است.

۴) نمايش و بازبيني ركوردها
بازخورد خوب سيستم در اين قسمت نقش مهمي در هدايت كاربر براي رسيدن به اطلاعات مورد نظرش دارد. در اين قسمت سيستم بايد بتواند به سؤالات زير پاسخ دهد:
●چه ركوردهايي با نياز اطلاعاتي كاربر مطابقت دارند؟
● آيا اطلاعاتي وجود دارد تا بتوان واژه‌هايي را كه حوزه‌هاي موضوعي وسيعي را پوشش داده‌اند تشخيص داد؟
● آيا اطلاعاتي وجود دارد كه بتوان واژه‌هايي را كه موجب اخص‌تر شدن دامنة جستجو مي‌شوند، تشخيص داد؟
● آيا مجموعة ركوردهاي بازيابي شده با نيازهاي اطلاعاتي كاربر مطابقت دارد؟
● آيا اطلاعاتي وجود دارد كه از طريق آن كاربر بتواند به تصحيح منطق بولي فرمول جستجو بپردازد؟

۵) سفارش مدرك

۶) درخواست براي اطلاعاتي دربارة سيستم بازيابي اطلاعات

۷) برقراري پارامترهاي نمايشي و ارتباطي
در مجموعة بازخورد سيستم بايد چهار مقوله به عنوان
مهم‌ترين اهداف، مد نظر طراحان سيستم‌هاي بازيابي اطلاعات قرار گيرند:
◊ بهينه‌سازي انتخاب واژگان جستجو توسط كاربر
◊ بهينه‌سازي فرمول جستجوي كاربر
◊ بهينه‌سازي و تنظيم تعداد ركوردهاي بازيابي شده
◊ بهينه‌سازي ضريب دقت و بازيابي يا بهينه‌سازي كل بازيابي اطلاعات
در نهايت بايد اين نكته را خاطرنشان كرد كه كه جستجوگران باتجربه و متخصصان بازيابي اين نكته را مي‌دانند كه انجام بازيابي مستلزم صرف وقت و هزينه است و در اين مسير بايد با اصلاح و بازنويسي و تكرار چرخة جستجو به نتايج دلخواه دست يافت.

۳ـ نمايش اطلاعات
امروزه در طراحي اغلب نرم‌افزارهاي حرفه‌اي، «نمايش اطلاعات» (۵) از جايگاه مهمي برخوردار است. امكانات نرم‌افزاري و سخت‌افزاري موجود، تحول مهمي در نمايش اطلاعات به وجود آورده اند. نمايش اطلاعات راه‌هاي آسان‌تر و موثرتري را براي تعامل با سيستم‌هاي رايانه‌اي، بخصوص براي جستجوگران سيستم‌هاي اطلاعاتي فراهم آورده است.
تعريف علمي نمايش گرافيكي اطلاعات عبارت است از «محاسبه و انتقال علائم به اشكال هندسي به صورتي قابل درك و مشاهده توسط انسان، به منظور فهم و كشف روابط پنهان موجود بين عناصر مختلف داده‌ها» (رُز، ۱۹۹۶).
نمايش گرافيكي اطلاعات به عنوان يك شيوة علمي و زيرشاخه‌اي از مبحث «تعامل انسان و رايانه» و با استفاده از تواناي‌هاي گرافيكي رايانه‌ها، اهداف زير را دنبال مي‌كند:
۱ـ بالا بردن سرعت فهم و پردازش اطلاعات توسط انسان در طي فرآيند ادراك و كم كردن درگيري‌هاي ذهني او ،
۲ـ ايجاد و برقراري ارتباط بين اجزاي مختلف اطلاعات ،
۳ـ انجام عمليات‌هاي پيچيده با اعمالي بسيار ساده ،
در نمايش گرافيكي اطلاعات جنبه‌هاي مختلفي مورد توجه قرار مي‌گيرند، كه مهم‌ترين آن‌ها عبارت‌اند از: جهت‌دار كردن اطلاعات نمايشي، برجسته كردن اطلاعات، رنگ‌ها، ابعاد و اندازه‌ها، پيكره‌بندي تصاوير، تكرار در يك فضا. در نمايش اطلاعات و ارائة فنون مختلف، همواره بايد چهار عامل را مورد توجه قرار داد. اين چهار عامل عبارت‌اند از:
الف ـ فضا: در نمايش ساختارهاي سلسله مراتبي، بايد از
روش‌هاي موثري براي نمايش اطلاعات استفاده گردد، زيرا باافزايش پيچيدگي، كار مشكل‌تر خواهد شد. به هر حال با هر حجمي از اطلاعات، بايد بتوان به نحوي به بازبيني اين ساختارها پرداخت كه ديد كلي و جزيي، با هم و به صورت همزمان، در اختيار كاربر قرار گيرند. مشكلاتي كه همواره در اين مورد وجود خواهند داشت عبارت‌اند از:
۱) محدوديت‌هاي موجود در تواناي ادراكي كاربران ،
۲) محدوديت موجود در فضاي نمايش رايانه‌ها .
ب ـ زمان: در ارائة فنون و راهكارهاي متفاوت براي نمايش اطلاعات، هميشه بايد سخت‌افزار و زمان لازم براي بارگذاري و روزآمد كردن صفحات يا مناطق نمايشي را در نظر گرفت. گاهي اوقات با يك تغيير كوچك، بايد كل منطقة نمايشي از نو ترسيم گردد. اين خود نيازمند زمان است و بسته به تواناي‌هاي سخت‌افزاري، از يك سيستم به سيستم ديگر متفاوت مي‌باشد.
ج ـ ثبات: ثبات نيز از جنبه‌هاي مهم در نمايش اطلاعات مي‌باشد كه بايد مورد توجه قرار گيرد. نبودن ثبات موجب بالا رفتن بار فكري و انحراف فكري كاربران مي‌گردد.
د ـ بازبيني: يكي از جنبه‌هاي مهم در نمايش اطلاعات، فراهم آوردن بازبيني بهينة اطلاعات مي‌باشد. بازبيني بهينة اطلاعات نيازمند توجه به سه عامل فضا، زمان و ثبات مي‌باشد.

۱ـ روش‌هاي متداول در نمايش اطلاعات
منظور از روش‌هاي متداول در نمايش اطلاعات، روش‌هايي هستند كه به صورت مرسوم در بسياري از سيستم‌هاي نرم‌افزاري موجود وجود دارند. اين روش، سه فن عمده را در برمي‌گيرد كه در ادامه به بررسي مختصر آن‌ها مي‌پردازيم:

۱/۱ـ پیمایش و تنظیم اندازه (٦)
منظور از پیمایش، حرکت در سطح صفحات نمایشي اطلاعات به منظور بازبیني قسمت‌هاي مختلف صفحه مي‌باشد. در این حالت گاهي براي نمایش بهتر اطلاعات، نیاز به تغییر اندازهٔ صفحات اطلاعات خواهیم داشت.

۱/۲ـ روش چند پنجره‌اي
در این روش از چند پنجره به صورت همزمان براي نمایش اطلاعات استفاده مي‌گردد. استفاده از چند پنجره باقابلیت‌هاي تنظیم و پیمایش، نتایج بهتري را به همراه خواهد داشت. هر پنجره داراي قابلیت تنظیم و پیمایش مخصوص به خود مي‌باشد، بنابراین کاربر مي‌تواند از چند پنجره براي مشاهده و مقایسه جنبه‌هاي مختلف یك سلسله مراتب استفاده کند.
۱/۳- راهبرد بازبیني نقشه‌اي(۷)
در این روش، از دو پنجره به صورت همزمان براي نمایش اطلاعات استفاده مي‌شود (برد، ۱۹۹۰). در پنجرهٔ اول، کل سلسله مراتب در تصویري کلي با مقیاسي کوچك‌تر قابل مشاهده مي‌باشد. همزمان در پنجرهٔ دوم، جزئیات سلسله مراتب در مقیاسي بزرگ‌تر مشاهده مي‌شود.

۲ـ فنون تاکید بر محتوا (۸)
با افزایش تعداد و عمق گره‌ها در ساختارهاي سلسله مراتبي، ضعف فنون متداول در مهار نمایشي کلي و جزیي اطلاعات و کاهش درگیري‌هاي ذهني کاربر، بیش‌تر نمایان مي‌شود. به دلایل فوق، تحقیقات مختلفي در این زمینه صورت گرفت. در این مسیر بتدریج راهبردهایي تدوین شدند که سعي در نمایش اطلاعات مبتني بر محتواي اطلاعات داشتند. از بهترین فنون مورد استفاده در این زمینه، فنون تاکید بر محتوا مي‌باشند. در این فنون با استفاده از روش‌هاي فشرده‌سازي، کلیات و جزئیات اطلاعات به صورت همزمان در اختیار کاربر قرار مي‌گیرد. در این روش از تعدادي از توابع براي تعیین محل استقرار و اندازهٔ هر یك از گره‌ها در کل ساختار استفاده مي‌شود. و گسترش فضاي اطلاعاتي از یك کانون مرکزي که به صورت پیش‌فرض در صفحه وجود دارد آغاز، متناسب با واکنش‌ها و تعاملات کاربر گسترش پیدا مي‌کند. در ادامه به بررسي مختصر چهار فن مورد استفاده در این روش مي‌پردازیم:
۲/۱ـ تورقگرهاي هایپربولیك (۹)
این تورقگرها براي نمایش اطلاعات از قوانین علم هندسه سود مي‌جویند. در ابتدا فضاي اطلاعاتي بر روي یك سطح هایپربولیك قرار داده مي‌شود و سپس بر روي یك منطقه‌دایره‌اي شکل ترسیم مي‌گردد و نمایش داده مي‌شود.
۲/۲- درخت‌هاي مخروطي (۱۰)
یکي از راه‌هاي نمایش دو بعدي و سه بعدي ساختارهاي سلسله مراتبي، استفاده از درخت‌هاي مخروطي مي‌باشد (کارپر،
۱۹۹۵). از لحاظ مکاني، ریشه یا بالاترین نقطه در ساختار سلسله مراتبي در رأس مخروط قرار مي‌گیرد و زیرشاخه‌هاي فرعي ساختار در فضاي رأس پراکنده مي‌شوند. هر کدام از این زیرشاخه‌ها به نوبهٔ خود یك مخروط دیگر را تشکیل مي‌دهند که زیرشاخه‌هاي دیگري در اطراف آن پراکنده مي‌شوند و این تقسیم‌بندي به همین صورت ادامه پیدا مي‌کند.
۲/۳ـ تنظیم هوشمند (۱۱)
این فن را مي‌توان در ارتباطات موجود از نوع دو طرفه (دو بعدي) مورد استفاده قرار داد (بارترام، ۱۹۹٤). در این مدل، در ابتدا اطلاعات را به روش مشخصي خوشه‌بندي مي‌کنند و سپس به هر خوشه یك نماد گرافیکي اختصاص داده مي‌شود و نمادهاي گرافیکي با نظم خاصي بر روي صفحه نمایش قرار مي‌گیرند، در حالي که خطوط ارتباطي بین هر کدام بر روي صفحهٔ نمایش قابل مشاهده مي‌باشند.
۲/۴ـ عدسي جادویي (۱۲)
عدسي جادویي از جمله فنوني است که در آزمایشگاه «زیراکس پارك» (۱۳) گسترش یافته و منطقه خاصي از صفحه نمایش را با حالتي متفاوت از سایر نقاط نمایش مي‌دهد (استون، ۱۹۹٤). اطلاعات اولیه به صورت کلي نمایش داده مي‌شوند و کاربر مي‌تواند عدسي را بر روي هر قسمت از این منطقه نمایشي اطلاعات قرار دهد ؛ که در این صورت اطلاعات جزیي مربوط به آن منطقهٔ خاص نمایش داده مي‌شود. بنابراین با استفاده از این عدسي‌ها مي‌توان به مطالعه و مشاهدهٔ جزئیات اطلاعات پرداخت و در همان حال، اطلاعات کلي در خارج از محدودهٔ عدسي قابل مشاهده مي‌باشند.

۳ـ روش «چه‌اوپز» (۱٤)
روش «چه‌اوپز» امکان مانور بهتر بر روي ساختارهاي سلسله مراتبي بزرگ را فراهم مي‌آورد. در این روش، کلیات اطلاعات در ساختارهاي سلسله مراتبي از بین نمي‌رود و در همان زمان امکان دسترسي آسان و سریع به جزئیات نیز وجوددارد. در این مدل نیاز به خوشه‌بندي اطلاعات یا عملیات‌هاي گرافیکي و محاسباتي پیچیده براي نمایش اطلاعات نمي‌باشد.

بحث و نتیجه‌گیري
در این جا پنج سؤال مطرح شده در بخش سؤالات پژوهش به ترتیب مطرح مي‌شوند و پاسخ آن‌ها براساس نتایج به دست آمده مورد بحث قرار مي‌گیرند. بعد از آن، نتیجه‌گیري و پیشنهادها مطرح خواهند شد.

سؤال۱. نحوهٔ سازماندهي اطلاعات، ساخت فایل‌هاي شاخص و مقلوب و سایر فایل‌هاي کمکي در سیستم‌هاي بازیابي اطلاعات بهینه به چه صورت مي‌باشد؟
امروزه با توجه به حجم انبوه اطلاعات و نگهداري اطلاعات تمام متن، ابزارها و شیوه‌هاي سنتي براي کنترل اطلاعات کافي نیستند. استفاده از روش‌هاي غیرخودکار در این زمینه غیراقتصادي است و با توجه به هوشمندتر شدن سیستم‌ها مقرون به صرفه نیست. هر چند که سازماندهي خودکار اطلاعات در سیستم‌هاي خارجي بازیابي اطلاعات

متداول است، اما در سيستم‌هايي كه در داخل كشور طراحي شده‌اند، چيزي تحت عنوان سازماندهي اطلاعات وجود ندارد.
سازماندهي اطلاعات هستة مركزي و ستون فقرات يك سيستم بازيابي اطلاعات را تشكيل مي‌دهد. هر سيستم اطلاعاتي داراي يك مبناي خاص براي تجزيه و تحليل اطلاعات مي‌باشد، كه سيستم براساس آن به تفسير اطلاعات و مطابقت بين اقلام و درخواست‌هاي اطلاعاتي مي‌پردازد. به لحاظ اهميت سازماندهي اطلاعات، سيستم‌هاي بازيابي اطلاعات به دو دسته هوشمند و غيرهوشمند تقسيم مي‌شوند.
هنگام پردازش متن‌هايي با زبان طبيعي، مسئلة پيچيدگي زبان و بي قاعدگي‌هايي كه در حوزة نحوي و معنايي وجود دارد خود را نشان خواهد داد كه بايد اين مشكلات را مدنظر قرار داد. در اين ميان واژه‌نامه‌ها قادر نيستند به صورت كامل، ابهامات و
پيچيدگي‌هاي موجود در زبان را برطرف كنند؛ اما مي‌توانند اين بي‌قاعدگي‌ها را كاهش دهند. اين واژه‌نامه‌ها شامل سياه ة بازدارنده، واژه‌نامة ريشة لغات و سياهة پيوندها، واژه‌نامة عبارات،واژه‌نامة مفاهيم، برقراري روابط سلسله مراتبي بين مفاهيم، و ريشه‌يابي واژگان مي‌شود.
به منظور سازماندهي واژه‌ها بايد دامنة شمول واژه‌ها، نوع دسته‌بندي مفاهيم و مكان واژه‌ها را مدنظر قرار داد. پاسخگويي به مسائل فوق وابسته به سياستگذاري سيستم در رابطه با ضرايب بازيابي و دقت مي‌باشد. اگر بخواهيم ضريب بازيابي را بالا ببريم بايد تمام واژه‌هايي كه مي‌توانند بنحوي در بازيابي مدارك موثر باشند، به واژه‌نامه اضافه شوند. هر چه واژه‌نامه عمومي‌تر باشد، ضريب بازيابي افزايش و ضريب دقت كاهش خواهد يافت. در مقابل اگر بخواهيم ضريب دقت بالاتري داشته باشيم و اقلام مرتبطتري مورد جستجو قرار گيرند، بايد به سمت واژه‌نامه‌هاي تخصصي‌تر حركت كنيم.
ساخت واژه‌نامه به صورت خودكار با دو روش تمام خودكار و نيمه‌خودكار قابل اجرا است، به دليل اين كه عملي بودن روش تمام خودكار بيش‌تر از آن كه به يك شاخة موضوعي خاص وابسته باشد به مجموعة مدارك مورد پردازش بستگي دارد ؛ بنابراين در رويه‌هاي پردازشي تغييراتي اعمال مي‌گردد و از قضاوت انسان‌ها استفاده مي‌شود.
به منظور بالا بردن كارايي سيستم‌هاي بازيابي اطلاعات، از روش‌هاي محاسبة همبستگي و خوشه‌بندي اطلاعات نيز استفاده مي‌گردد كه محاسبة همبستگي به عنوان معياري براي سنجش ميزان مشابهت و همبستگي بين دو متن و خوشه‌بندي براي دسته‌بندي ركوردهاي اطلاعاتي مشابه با هم، مورد استفاده قرار مي‌گيرد.

سؤال۲. چه امكانات و قابليت‌هايي در سيستم‌هاي بازيابي اطلاعات بهينه لازم است تا كاربران بتوانند به بهترين شكل ممكن به بيان نيازهاي اطلاعاتي خود بپردازند؟ به بيان ديگر، چه قابليتها و امكاناتي بايد در سيستم وجود داشته باشد تا كاربران بتوانند يك راهبرد بازيابي صحيح و دقيق داشته باشند؟
فرمول‌بندي صحيح و دقيق نيازهاي اطلاعاتي كاربران وابسته به امكانات بازيابي اطلاعات براي فرمول‌بندي مي‌باشد. هر چه قابليت‌ها و امكانات بهتري فراهم آورده شود، كاربر راحت‌تر و آسان‌تر مي‌تواند به بيان نيازهاي خود بپردازد. در ادامه به بررسي‌مختصر اين امكانات مي‌پردازيم:

۱ـ امكان استفاده از عملگرهاي بولي: با استفاده از اين كاربر عملگرها تصورات (AND, OR, NOT, XOR) مي‌تواند بين و واژه‌هايي كه براي بيان درخواست اطلاعاتي خود برگزيده است يك ارتباط منطقي برقرار كند.
۲ـ عملگرهاي مجاورت: با تركيب عملگرهاي منطق بولي و عملگرهاي مجاورت مي‌توانيم مجموعه‌هاي اطلاعاتي را مورد ارزيابي قرار دهيم
۳ـ جستجوي املايي: اين روش هم در جستجوي واژه‌ها و هم در تجزيه و تحليل ابتدايي مجموعه‌هاي اطلاعاتي به منظور ايجاد فايل‌هاي كمكي در سيستم، كاربرد دارد. به عنوان مثال، كاربر واژه Computer را براي جستجو وارد مي‌كند و آنگاه واژه‌هاي Conputer, Commpiter, Compute مورد جستجو قرار مي‌گيرند. استفاده از اين روش باعث افزايش ضريب فراخواني و كاهش ضريب دقت مي‌شود.
۴ـ كوته‌سازي واژگان: كوته‌سازي اين امكان را فراهم مي‌آورد كه در هنگام جستجو، بدون در نظر گرفتن حرف يا حرف‌هايي از يك واژه يا يك واحد اطلاعاتي، به بازيابي بپردازيم. بدليل ماهيت گوناگون اطلاعات و درخواست‌هاي اطلاعاتي، كوته‌سازي، بخصوص در سيستم‌هايي كه ريشه‌يابي واژه‌ها (در مرحله تجزيه و تحليل داده‌ها) و يكدست‌سازي اطلاعات بخوبي صورت نمي‌گيرد، اهميت بسيار مي‌يابد.
۵ـ تعيين محدوده در داده‌هاي كمي: بنابر ماهيت اين نوع داده‌ها نياز خواهد بود كه علاوه بر يك مورد خاص، در محدوده‌هاي خاصي نيز جستجو صورت گيرد.
۶ـ گسترش مفاهيم و اصطلاحات: سيستم‌هاي بازيابي اطلاعات با استفاده از ابزارها و شيوه‌هاي مختلف به گسترش مفاهيم موجود در درخواست اطلاعاتي كاربر مي‌پردازند و براي اين منظور به استفاده از اصطلاحنامه‌ها و واژه‌نامه‌هاي موجود يا ايجاد آن‌ها بر پاية اطلاعات موجود در سيستم مي‌پردازند.
۷ـ استفاده از ضريب وزني: منظور از ضريب وزني اين است كه كاربر به هر عضو دلخواه از فرمول جستجو، يك عدد صحيح كه مشخص‌كنندة اهميت عضو در ميان اعضا مي‌باشد اختصاص مي‌دهد و سيستم بازيابي اطلاعات با در نظر گرفتن اين‌ضريب ، بازيابي اطلاعات را انجام مي‌دهد.
۸ـ سياهة واژه‌ها: در سياهه فوق بين هر جفت واژه، عملگر OR فرض مي‌شود و هر كدام از آن‌ها داراي ضرايب وزني يكساني مي‌باشند. در نهايت، ركوردها براساس بالاترين تعداد واژه‌هاي فوق كه در ركوردها ظاهر مي‌شوند، و همچنين براساس تعداد تكرار آن‌ها در ركوردها نمره‌گذاري مي‌شوند و براساس همين نمره‌گذاري مرتب مي‌شوند.
بعد از انجام جستجو، نمره‌گذاري مجموعه‌هاي بولي و اندازه‌گيري مشابهت مي‌تواند در بازيابي بهتر اطلاعات كاربران بسيار مفيد باشد :
۱ـ نمره‌گذاري مجموعه‌هاي بولي: در اين روش نيز براساس تعداد تكرار اعضاي فرمول جستجو در ركوردهاي بازيابي شده، به هر كدام از ركوردها يك نمره اختصاص مي‌يابد و در نهايت، ركوردها برحسب همين نمره در خروجي مرتب مي‌شوند. اين نمره نشان‌دهنده درصد ارتباط ركورد بازيابي شدة حاضر، با درخواست اطلاعاتي كاربر مي‌باشد.

۲ـ اندازه‌گيري مشابهت: دسترسي آسان كاربر به حداقل يك ركورد كه بالاترين ميزان ارتباط را با درخواست اطلاعاتي او داشته باشد، بسيار اهميت دارد. اولين ركوردي كه به اين صورت توسط سيستم مورد بازيابي قرار گيرد، مي‌تواند مبناي جستجوهاي بعدي كار باشد.

سئوال۳. بازخورد در سيستم‌هاي بازيابي اطلاعات بهينه به چه صورت مي‌باشد؟ به بيان ديگر، سيستم بازيابي اطلاعات چگونه مي‌تواند كاربران خود را به مرتبطترين اطلاعات، در رابطه با نياز اطلاعاتيشان هدايت كند؟
بر مبناي چرخة دستورها و مطالعات انجام شده و با توجه به دستورهاي انتخابي كاربر، مي‌توان پيش‌بيني كرد كه جستجوگران در چه مرحله‌اي از جستجو قرار دارند و هدف آن‌ها چيست. با گروهبندي و با توجه به ترتيب استفاده از آن‌ها و پارامترهاي استفاده شده مي‌توان به دسته‌بندي كاربران اطلاعاتي پرداخت و بازخوردمتناسبي‌نشان‌داد.چرخةدستورهشامل مراحل زير است: ۱) انتخاب يك بانك اطلاعاتي براي انجام جستجو ،
۲) جستجو براي واژه‌هاي مورد نظر در بانك واژگان ،
۳) ايجاد فرمول جستجو و انجام جستجو ،
۴) نمايش و بازبيني ركوردها ،
۵) سفارش مدارك ،
۶) درخواست براي اطلاعاتي دربارة سيستم بازيابي اطلاعات ،
۷) برقراري پارامترهاي نمايشي و ارتباطي
با توجه به چرخة فوق و آنچه كه در بخش دستورها به آن اشاره شد، بازخورد مناسب سيستم بايد در چهار مقولة زير برنامه‌ريزي و دنبال گردد. اين چهار مقوله عبارت‌اند از:
۱ـ بهينه‌سازي انتخاب واژگان جستجو توسط كاربر ،
۲ـ بهينه‌سازي فرمول جستجوي كاربر ،
۳ـ بهينه‌سازي و تنظيم تعداد ركوردهاي بازيابي شده ،
۴ـ بهينه‌سازي ضريب دقت و بازيابي يا بهينه‌سازي كل بازيابي اطلاعات .

سؤال٤. نحوة نمايش اطلاعات در خروجي‌هاي حاصل از نتايج جستجو در سيستم‌هاي بازيابي اطلاعات بهينه چگونه مي‌تواند كاربر را در انتخاب اطلاعات مرتبط هدايت كند؟ به بيان ديگر، شيوه‌هاي بهينه نمايش اطلاعات در خروجي‌هاي حاصل از نتايج جستجو كدام‌اند؟
امروزه با بالا رفتن قدرت سيستم‌هاي سخت‌افزاري و نرم‌افزاري و راحت‌تر شدن برنامه‌نويسي براي ايجاد محيط‌هاي گرافيكي قوي، ديگر عذري براي عدم استفاده از شيوه‌هاي مختلف نمايش اطلاعات باقي نمي‌ماند. استفاده از جنبه‌هاي مختلف در نمايش گرافيكي اطلاعات مانند جهت‌دار كردن اطلاعات نمايشي، برجسته كردن اطلاعات، رنگ‌ها، ابعاد و اندازه‌ها، پيكره‌بندي تصاوير و تكرار در يك فضا به منظور نمايش بهينه را بايد مورد توجه قرار داد. استفاده از شيوه‌هاي پيمايش و تنظيم اندازه، چند پنجره‌اي، راهبرد بازپيني نقشه‌اي شيوه‌هاي تاكيد بر محتوا، عدسي جادويي و روش «چه‌اوپز»، تنظيم هوشمند، درخت‌هاي مخروطي و تورقگرهاي هايپربوليك را نيز مي‌توان به منظور نمايش بهينه اطلاعات مورد توجه قرار داد.
اگر رابطها در سيستم‌هاي نرم‌افزاري بر مبناي دو هدف تسريع در فرآيند ادراك و تسهيل در فرآيند تفكر طراحي شوند، قاعدتاً پيشرفت قابل ملاحظه‌اي در انتقال اطلاعات از رايانه‌ها به‌انسان صورت خواهد گرفت. حتماً شما هم اين جملة معروف را شنيده‌ايد كه «يك تصوير برابر با هزار واژه مي‌باشد» .
راه‌هاي مختلفي براي به تصوير كشيدن اطلاعات در يك محيط گرافيكي وجود دارند، فهم دقيق فرآيند ادراك انسان، اين بصيرت را براي طراحان سيستم‌ها ايجاد خواهد كرد كه بدانند با كدام شيوه نمايش قادر خواهند بود فهم و انتقال اطلاعات را به حداكثر برسانند. در اين جا بايد اين نكته را متذكر شد كه براي اين پرسش كه بهترين روش براي به تصوير كشيدن اطلاعات كدام است، پاسخ واحدي وجود ندارد. پاسخ به اين سؤال با توجه به اهداف نمايش اطلاعات و چهار متغير فضا، زمان، ثبات، و بازبيني متفاوت مي‌باشد. فنون متداول نمايش اطلاعات همانند پيمايش و تنظيم اندازه، روش چند پنجره‌اي، و راهبرد بازبيني نقشه‌اي، امكان نمايش ساختارهاي اطلاعاتي كوچك را فراهم مي‌آورند و با بزرگ‌تر شدن ساختارها، بازبيني اطلاعات با استفاده از روش‌ها كارآمد نخواهد بود. در فنون تاكيد بر محتوا مانند تورّقگرهاي هايپربوليك، درخت‌هاي مخروطي، تنظيم هوشمند و عدسي جادويي، امكان نمايش ساختارهاي سلسله مراتبي بزرگ‌تر وجود دارد. توزيع مناسب ساختار در فضاي نمايشي و ثبات، بازبيني مناسبي از اطلاعات را به وجود مي‌آورد. اما نمايش اطلاعات در اين فنون نيازمند محاسبات زيادي مي‌باشد. در اين صورت اگر سيستم از امكانات سخت‌افزاري پاييني برخوردار باشد، انجام محاسبات مدت زمان بيش‌تري را به خود اختصاص مي‌دهد. در روش «چه‌اوپز» امكان نمايش ساختارهاي سلسله مراتبي بزرگ با در نظر گرفتن فضا، زمان و ثبات براي يك بازبيني‌خوب وجود دارد، اما مشكل ذاتي كه در اين روش وجود دارد اين است كه در هر لحظه فقط يك شاخه از اطلاعات قابل مشاهده مي‌باشد.
در استفاده از اين فنون و راهبردها هميشه يك خطر وجود دارد و آن اين كه اين استفاده به فهم مشاهده‌كننده، پيش‌زمينه‌هاي قبلي او و محتواي اطلاعات وابسته خواهد بود. ذهن انسان هميشه از آخرين اطلاعات براي ادراك اطلاعات جديد استفاده مي‌كند. اگر مجموعه‌اي از اطلاعات به عنوان يك شيء نشان داده شوند، ذهن تمايل دارد كه اطلاعات دروندادي را همانند آن شكل تفسير كند. تجارب قبلي ما بر قضاوت ما از
تصاوير اثر خواهند گذاشت. اين نكته را بايد هميشه مد نظر قرار داد كه: يك كاربر ممكن است يك تصوير را براساس تجارب قبلي خود تفسير كند نه براساس آنچه طراح در نظر داشته است.

سؤال٥. چگونه مي‌توان يك سيستم بازيابي اطلاعات را مورد ارزيابي قرار داد؟
آنچه در هنگام ارزيابي سيستم‌هاي بازيابي اطلاعات در سطح دنيا مورد توجه قرار مي‌گيرد، ضريب بازيابي و ضريب دقت مي‌باشد. اما اين زماني صحيح است كه سيستم‌ها از يك حداقل امكانات ممكن برخوردار باشند.
در ادامه، عناصري كه لازم است در يك سيستم بازيابي اطلاعات بهينه وجود داشته باشد در جدول ۱ خلاصه مي‌شوند

و وضعيت نرم‌افزارهاي خارجي و داخلي در رابطه با وجود يا نبود اين عناصر نمايانده مي‌شود.

جدول۱. امكانات سيستم‌هاي داخلي بازيابي اطلاعات

پژوهشگاه علوم انسانی و مطالعات فرهنگی
پرتال جامع علوم انسانی

مقوله	زیرشاخه	نرم‌افزارهای ایرانی	نرم‌افزارهای خارجی
سازماندهی اطلاعات	سیاههٔ بازدارنده	وجود دارد	وجود دارد
	واژه‌نامهٔ ریشهٔ لغات و سیاههٔ پسوندها	وجود ندارد	وجود دارد
	واژه‌نامهٔ عبارات	وجود ندارد	وجود دارد
	واژه‌نامهٔ مفاهیم	وجود ندارد	وجود دارد
	برقراری روابط سلسله مراتبی بین مفاهیم	وجود ندارد	وجود دارد
	ریشه‌یابی واژگان	وجود ندارد	وجود دارد
	امکان استفاده از روش‌های خودکار	وجود ندارد	وجود دارد
	امکان استفاده از روش‌های نیمه‌خودکار	وجود ندارد	وجود دارد
	محاسبهٔ همبستگی	وجود ندارد	وجود دارد
	خوشه‌بندی اطلاعات	وجود ندارد	وجود دارد
فرمول جستجو			
	امکان استفاده از عملگرهای بولی		
	AND	وجود دارد	وجود دارد
	OR	وجود دارد	وجود دارد
	NOT	وجود دارد	وجود دارد
	XOR	وجود ندارد	وجود دارد
	عملگرهای مجاورت		
	ADJ	وجود دارد	وجود دارد
	W	وجود ندارد	وجود دارد
	N	وجود ندارد	وجود دارد
	P	وجود ندارد	وجود دارد
	S	وجود ندارد	وجود دارد
	F	وجود ندارد	وجود دارد
	جستجوی املایی	وجود ندارد	وجود دارد
	کوته‌سازی	وجود دارد	وجود دارد
	?	وجود دارد	وجود دارد
	*	وجود دارد	وجود دارد
	#	وجود ندارد	وجود دارد
	تعیین محدوده در داده‌های کمی	وجود دارد	وجود دارد
	گسترش مفاهیم و اصطلاحات	وجود ندارد	وجود دارد
	استفاده از ضریب وزنی	وجود ندارد	وجود دارد
	سیاهه‌واژه‌ها	وجود ندارد	وجود دارد
	مجموعه‌های بولی نمره‌گذاری شده	وجود ندارد	وجود دارد
	اندازه‌گیری مشابهت	وجود ندارد	وجود دارد
بازخورد سیستم			
	بهینه‌سازی انتخاب واژگان جستجو توسط کاربر	وجود ندارد	وجود دارد
	بهینه‌سازی فرمول جستجوی کاربر	وجود ندارد	وجود دارد
	بهینه‌سازی و تنظیم تعداد رکوردهای بازیابی شده	وجود ندارد وجود دارد	
	بهینه‌سازی ضریب دقت و بازیابی یا بهینه‌سازی کل بازیابی اطلاعات	وجود ندارد	وجود دارد
نمایش اطلاعات			
	استفاده از جنبه‌های مختلف نمایش اطلاعات	وجود ندارد	وجود دارد
	استفاده از شیوه‌های مختلف نمایش اطلاعات	وجود ندارد	وجود دارد

پژوهشگاه علوم انسانی و مطالعات فرهنگی
پرتال جامع علوم انسانی

پي‌نوشت‌ها:

1. Dialog
2. Oak
3. Dowquest
4. stop word list
5. information visualization
6. pan & zoom
7. map view strategy
8. focus+context
9. Hyperbolic Browser
10. Cone-Trees
11. intelligent zoom
12. Magic Lens
13. Xerox PARC
14. Cheops Apporach

منابع

1. Abraham. C.T. 1965. Techniques for Thesaurus Organization and Evaluation. Information Science 4(4).
2. Bartram, L. 1994. Contextual Assistance in User interfaces to Complex, Time-Critical Systems: The Intelligent Zoom. In: Graphics Interface. Banff, Albert: 216-224.
3. Beard, D.V. and Walker, J.Q. 1990. Navigating Techniques to Improve the Display of Large Two-Dimensional Spaces. Behavior and Information Technology. 9(6): 451-466.
4. Bonner, R.E. 1964. On Some Clustering Techniques. IBM Journal of Research 22-32.
5. Bookstein, A. 1983. Outline of a General Probabilistic Retrieval Model. Journal of Documentation. 39 (2): 63-72.
6. Boyce, B.R. and Kraft, D.H. 1985. Principles and Theories in Information Science. In: Annual Review of Information Science and Technology. Vol. 20. White Plains, Newyork: Knowledge Industry Publications.
7. Carrierre, J. and Kazman, R. 1995. Research Report Interacting With Huge Hierarchies: Beyond Cone Trees. Proceedings of Information Visualization. Allanta: GA.
8. Chapman, J.L. 1981. A State Transition Analysis of On-line Information-Seeking Behavior. Journal of the American Society for Information Science 35(2): 107-116.
9. Dennis, S.F. 1965. The Construction of a Thesaurus Automatically from a Sample of Text. Symposium on Statistical Association Methods for Mechanized Documentation. Natl. U.S. Misc. Publ. 369 (December).
10. Dovel, J.A. and Heald J.H. 1965. Project Lex Status. Proceedings of ADI Meeting. Santa Monica, Calif.
11. Levery, F.1966. Organization et Consultation d'un Thesaurus. 1965 FID Congress, Spartan Books. Washington, D.C.
12. Lovins, J.B. 1968. Developing of Stemming Algorithm. Mechanical Translation and Computational Linguistics 11(1).
13. Penniman, W.D. 1975. A Stochastic Process Analysis of On-line Behavior. Proceeding of the American Society for Information Science 38th Annual Meeting. Washington D.C: ASIS.
14. Resiner, P.1965. Semantic, Diversity and a Growing Mac-Machine Thesaurus. Information Science 4(4).
15. Rose, R.1996. P1000 Science and technology and Strategy for Information Visulization. Version 2.
16. Salton, G.1964. A Document Retrival System For Mac-Machine Interaction. Proceeding 19th ACM Natl. Philadelphia, Pa.
17. Salton, G. 1965. Progress in automatic Information Retrieval. IEEE Spectrum 2(8).
18. Salton, G. 1966. Data Manipulation and Programming Problems in Automatic Information Retrieval. Commun . ACM 9(3).
19. Salton, G. and McGill, M. 1983. Introduction to Modern Information Retrieval . New York: McGraw-Hill.
20. Sparck-Jones, K. 1965. Experiments in Semantic Classification. Mech. Transl 8(4).
21. Stone, M.C. and Fishkin, K. 1994. The Moveable filter as a User Interface Tool. Proceedings of CHI CONFERENCE on Human Factors in Computing Systems. Boston: ACM.