

سیستم جامع اطلاع رسانی سیمرغ

مفهوم شبکه کتابخانه‌ها در سیمرغ نوسا
جستجوی همزمان در مراکز اطلاع رسانی



نوسا 

شرکت نرم افزار و سخت افزار ایران
www.nosa.com

تهران، خیابان سپهبد قرنی، خیابان شاداب شرقی، پلاک ۶
تلفن/نمابر: ۸۸۸۲ ۲۹ ۷۹ - ۸۸۸۳ ۵۳ ۶۰
پست الکترونیک: info@nosa.com



پیدایش مفهوم جستجوی همزمان

کتابخانه ها و مراکز اسناد، مکانهایی هستند که در آنها انواع گوناگونی از اسناد و مدارک اطلاعاتی نگهداری می شود. اسنادی از قبیل کتاب، نشریه، پایان نامه، مقاله، گزارش، مواد سمعی بصری و ... که در مکانهای جدا از هم و با روشهای چیدمان و مرتب سازی متفاوت نگهداری می شوند. در طراحی پایگاههای اطلاعاتی برای کتابخانه ها، دو رویکرد به این اسناد وجود دارد: یک دیدگاه که متأثر از کتابخانه کنگره آمریکاست، معتقد است ماهیت کلی تمام اسناد یک چیز است لذا باید به یک شکل فهرست نویسی شده و یک بانک اطلاعاتی یکپارچه از همه آنها تشکیل شود. دیدگاه دیگر معتقد است هر نوع سند خاص، ماهیتاً با اسناد دیگر متفاوت است و بهتر است که بطور جداگانه فهرست نویسی و در پایگاه جداگانه ذخیره و نگهداری شود. روش اخیر روشی است که در شرکت نوسا از ابتدا تا به حال استفاده شده است. این روش، مزایا و معایبی نسبت به روش بانک اطلاعات یکپارچه دارد که اهم آن به شرح زیر است:

معایب :

- ۱) دشواری در طراحی و به روز کردن ساختار پایگاهها
- ۲) دشوار بودن نگهداری پایگاهها به علت تعدد آنها
- ۳) دشواری گرفتن لیست ها و گزارشهای یکپارچه از تمام پایگاهها بطور همزمان

مزایا :

۱) آزادی در طراحی هر پایگاه توسط هر مرکز

اگرچه در نوسا برای هر نوع سند یک پایگاه صفر (استاندارد) تعریف و به مشتریان ارائه می شود اما در عین حال، هر مرکز بنا به شرایط و نیازهای خود می تواند این پایگاه را به دلخواه خود تغییر داده و متناسب با نیازهای خود آن را اصلاح و تدوین (Modify) کند. به عنوان مثال می تواند فیلدی به آن اضافه یا از آن کم کند و یا مداخل جستجوی مورد نیاز خود را به آن اضافه کند. فراتر از آن، حتی یک مرکز می تواند بدون توجه به پایگاه صفر، ساختاری کاملاً دلخواه و متفاوت با آن را طراحی و استفاده کند. البته در این حالت، شخص باید کاملاً با امکانات تعریف (Define) سیمرغ در حد تخصصی آشنا بوده و توان نگهداری و رفع اشکال این گونه پایگاهها را داشته باشد.

۲) انعطاف پذیری بالا در طراحی تعاریف خاص آن پایگاه

هر سند خاص، بسته به اطلاعات ویژه خود، به کاربرگهای ورود اطلاعات، مداخل جستجو، گزارشهای چاپی و دیگر تعاریف مخصوص و مناسب خود نیازمند است. برای مثال برای پایان نامه ها نام استاد راهنما اهمیت دارد در حالیکه برای کتابها، این مورد موضوعیت ندارد. در این موارد اگر پایگاهها یکپارچه باشد، تعریف تعداد زیادی کاربرگ، مدخل، صفحات خلاصه نمایش و محاوره های جستجوی مناسب همه انواع اسناد باعث ایجاد شلوغی بیش از حد و دشواری در تعریف و اصلاح این گزارشها خواهد شد. اما با پایگاههای تفکیک شده این کار بسیار سهل تر خواهد بود و می توان در هر پایگاه هر تعداد گزارش، فهرست برگه، محاوره جستجو، کاربرگ ورود اطلاعات و ... دلخواه را تعریف و استفاده کرد.

۳) امکان تعریف پایگاههای خاص و پایگاههای غیر کتابشناختی

در بعضی از مراکز اطلاع رسانی منابع و اسنادی وجود دارد که کاملاً خاص آن مرکز می باشد. اساساً هر نوع سند ویژگیهای خاص مربوط به خود را داراست که در هر عمل یکپارچه سازی، قطعاً بعضی از این خصوصیات، به نفع یکپارچگی میان آنها حذف می شود. به عنوان مثال در یک کارخانه تولید اتومبیل، یک سری گزارشهای فنی یا کاتالوگهایی وجود دارد که توسط متخصصین و کارشناسان آن کارخانه استفاده می شود. محتوای اطلاعاتی این اسناد به شکلی است که هیچ کدام از پایگاههای قبلاً تعریف شده نمی تواند آنها را در بر گیرد. معمولاً در روش پایگاه یکپارچه تا جایی که امکان دارد سعی می شود اسناد جدید با ساختار موجود منطبق شوند و اطلاعات خاص را در فیلدهای مشابه آن وارد کنند. اما در صورتی که از هم منفک و جدا باشند می توان تا حد بسیار زیادی خاصترین و کوچک ترین موارد مربوط به هر نوع اسناد را در طراحی ساختار پایگاه مربوطه وارد کرد. این اسناد فیلدهای بخصوصی را برای ساختار خود می طلبد که باعث می شود تا یک پایگاه کاملاً اختصاصی برای آن طراحی شود. بدیهی است که در این حالت، تعریف یک پایگاه اختصاصی بسیار آسان تر و ایمن تر از تغییر و دست بردن در ساختار پایگاه یکپارچه می باشد.

نکته دیگر این که برای نگهداری اطلاعات غیر کتابشناختی ساده نیز می توان از این سیستم استفاده کرد. برای مثال برای نگهداری اطلاعات پرسنلی اعضای کتابخانه یا مشخصات اشیای موجود در یک موزه و ... می توان پایگاههای ساده ای را طراحی و اطلاعات مورد نیاز را در آن وارد کرد. البته لازم به توضیح است که سیمرغ مناسب



پایگاههای اطلاعاتی متنی و کتابشناختی است و لذا نباید توقع ایجاد پایگاههای غیر کتابشناختی بصورت حرفه ای را در آن داشت.

محدودیت کمتر از نظر حد تعاریف

در هر سیستم مدیریت پایگاههای اطلاعاتی، محدودیتهایی از نظر تعداد فیلد قابل تعریف در یک پایگاه، تعداد رکوردهای وارد شده، ظرفیت یک رکورد برای ورود اطلاعات، تعداد مداخل جستجو و ... وجود دارد. بدیهی است که در پایگاههای یکپارچه به دلیل گستردگی و ضرورت شمول همه انواع اسناد، این محدودیتهای خیلی سریعتر از پایگاههای تفکیک شده خود را نشان داده و مشکل ایجاد خواهد کرد. لذا این مشکل باعث خواهد شد تا از بعضی از جزئیات مربوط به اسناد چشم پوشی شود و جامعیت اطلاعاتی آن از بین برود. در حالیکه در پایگاههای مجزا بسیار کم اتفاق می افتد که به حد تعاریف برسیم.

امکان استفاده از دو نوع تعریف برای یک نوع سند بطور همزمان

بعضی از مراکز بنا به نیاز خود لازم می بینند که از یک نوع سند (مثلا کتاب) دو پایگاه مجزا از هم و با تعاریف متفاوت داشته باشند. برای مثال فرض کنید کتابخانه ای می خواهد کتابهای خود را پس از ثبت در دفتر، به نوعی وارد کامپیوتر کرده و به اطلاع رسانی این کتابها پردازد. بدیهی است که اگر کتابداران بخواهند تا زمان فهرست نویسی و آماده سازی کتاب صبر کنند ممکن است مدت زمان زیادی طول بکشد. لذا در این حالت می توانند یک پایگاه با تعاریف و ساختار ساده برای دفتر ثبت خود تعریف و رکوردهای موجود در آن را قابل جستجو کنند و زمانی که فهرست نویسی کتابها انجام شد آن را وارد پایگاه کتابها کرده و اطلاعات کامل آن را در معرض استفاده قرار دهند.

امکان استفاده از پایگاه بطور موقت

گاهی اوقات ضرورتی پیش می آید که لازم می شود اسناد و مدارکی بطور موقت وارد یک پایگاه شده و پس از مدتی این پایگاه حذف شود. برای مثال فرض کنید شخصی کتابخانه شخصی خود را به کتابخانه ای هدیه می کند. کتابداران مرکز برای اینکه کاربران را تا فهرست نویسی و سازماندهی کامل این کتابها در انتظار نگذارند می توانند یک پایگاه موقت به نام شخص اهدا کننده ایجاد و اطلاعات مختصر کتابها را در آن وارد کنند و پس از فهرست نویسی کامل، رکوردهای یاد شده را به پایگاه اصلی کتابهای خود منتقل کنند. برای مثال دیگر فرض کنید کتابخانه مرکزی یا هر مرجع دیگری وظیفه بررسی و اصلاح اطلاعات چند مرکز را به عهده دارد. این مرکز برای بررسی اطلاعات هر یک از مراکز دیگر می تواند یک پایگاه خاص آن مرکز ایجاد کرده و پس از هر گونه اصلاح و پردازش مورد نظر، آن پایگاه را حذف کند.

امکان استفاده از دو نوع پایگاه پدر در یک مرکز

لازم به یادآوری است که در سیمرغ امکان استفاده از نوع خاصی از پایگاههای رابطه ای (Relational) وجود دارد. در این نوع خاص، جهت ارتباط یک طرفه است به این شکل که رکوردهای یک سری از پایگاهها که حاوی اطلاعات مستند اند (مانند سرعنوانهای موضوعی یا فهرست مستند مشاهیر) با یک کلید (Key) به فیلدهای متناظر خود در پایگاههای دیگر (مثلا کتابها) مرتبط می باشند. پایگاههای نوع اول پایگاه پدر و پایگاههای نوع دوم پایگاه فرزند نامیده می شوند. حال با روش پایگاههای جدا از هم می توان برای یک پایگاه از یک مستند و برای پایگاه دیگر از یک مستند دیگر استفاده کرد. برای مثال برای گزارشها یا پایان نامه ها از پایگاه «توصیفگرها» که با روش اصطلاحنامه (Thesaurus) ای (پس همارا) تهیه شده و برای کتابها از پایگاه «سرعنوانهای موضوعی» که با روش سرعنوانی (پیش همارا) تهیه شده استفاده کرد.

امنیت بیشتر داده ها و امکان مدیریت بهتر پایگاههای اطلاعاتی

با تعدد پایگاهها امکانات بهتری برای مدیریت اختیارات و دسترسی کاربران به پایگاهها فراهم می شود. اگر تصمیم بر این است که اعضای خاصی اسناد خاصی را نبینند یا اینکه امکان اصلاح بعضی از اطلاعات (مثلا نشریات) را داشته و حق اصلاح در بعضی دیگر (مثلا کتابها) را نداشته باشند می توان این اختیار را به دلخواه به اعضا داد یا از آنها گرفت. اما در مورد پایگاههای یکپارچه این امکان یا وجود ندارد و یا اگر هم وجود داشته باشد بصورت کاملاً ابتدایی و ضعیف می باشد. باید دقت کرد که بهرحال چاره ای بجز باز کردن پایگاه اطلاعاتی برای تمام کاربران وجود ندارد و نقطه ضعف این طراحی درست در همینجاست. اگرچه طراحان نرم افزار در این موارد ممکن است با نوشتن لایه های نرم افزاری نوعی سطوح دسترسی را شبیه سازی کنند و کاربر تصور کند که به برخی از اسناد دسترسی ندارد اما واقعیت این است که این طراحی از اساس در این مورد ضعف دارد و تمام کاربران در مراحل مختلف پردازشهای اطلاعاتی به تمام اسناد دسترسی دارند و امنیت اطلاعات در این حالت در سطح پایین قرار دارد. نکته دیگر این است که اگر به فرض هم پایگاهی دچار خرابی یا ویروس شد پایگاههای دیگر آسیب نمی بینند و به سرعت می توان از گسترش خرابی و ویروس جلوگیری کرد.

سرعت بالا در انواع پردازشهای اطلاعاتی

در نتیجه ساده تر بودن پایگاهها، پردازشهایی همچون واژه سازی (Indexing)، مرتب سازی، انتقال اطلاعات و ... با سرعت خیلی بیشتر انجام می شود. در غیر این صورت برای هر مرحله پردازش اطلاعات، سیستم باید یک مرحله فیلترینگ براساس نوع سند را انجام داده و سپس به درخواست کاربر جواب دهد که این مرحله، باعث کند شدن عملیات



خواهد شد.

سوالی که در اینجا و در رابطه با این گونه طراحی پایگاهها پیش می آید این است که اگر یک جستجوکننده، بخواهد صرف نظر از نوع سند، جستجویی انجام دهد چه باید بکند؟ در واقعیت نیز عموماً همینطور است و محققى که به دنبال موضوعى خاص است برايش نوع سند اهمیتی ندارد. در این صورت آیا باید به تعداد پایگاهها جستجویش را تکرار کند؟ بدیهی است که قطعاً این روش خوبی نیست و انرژی و زمان زیادی را از کاربر خواهد گرفت. لذا برای حل این مشکل، امکانی طراحی شد به نام «جستجوی همزمان در پایگاهها». مساله ای که در اینجا وجود دارد این است که از چه طریقی و با چه واسطه (Interface) ای می توان این جستجوی همزمان را انجام داد؟

طراحی یک محاوره (Dialog) جستجوی عمومی

قبل از ادامه مطلب در اینجا لازم است که مفهوم محاوره جستجو توضیح داده شود. یک محاوره جستجو، واسطه (Interface) ای است میان کاربر و بانک اطلاعاتی که از طریق آن درخواست جستجو داده می شود. واضح است که این محاوره ها برای پایگاههای مختلف به اشکال مختلف می باشد. برای ایجاد یک محاوره (Dialog) عمومی که امکان تبادل اطلاعات با تمام انواع پایگاهها را داشته باشد باید:

ابتدا چند مداخل عمومی که موارد مشابه یا نزدیک به آن در اکثر پایگاههای اطلاعاتی موجود باشد تعریف شود. مداخلی که به نظر می آید در اکثر پایگاهها وجود داشته باشد عبارت است از: پدیدآور (نویسنده)، عنوان و موضوع زیرا هر سندی را به هر حال یک شخص یا گروهی از اشخاص (تئالگان) به وجود آورده و قطعاً عنوانی هم دارد و نیز می توان آن را در یک طبقه موضوعی جای داد. لذا محاوره ای به شکل زیر برای جستجوی عمومی (همزمان) در پایگاهها طراحی شده است:

• پدیدآور:

• عنوان:

• موضوع:

در مرحله بعد در تک تک پایگاهها، مداخل متناظر با این مداخل عمومی مشخص و تعیین شود. در واقع در این مرحله تعیین می شود که وقتی کاربر، عبارتی را در یکی از مداخل عمومی (مثلاً پدیدآور) وارد می کند در هر پایگاه در کدام مداخل جستجو خواهد شد. برای مثال در پایگاه کتاب در مداخل جستجوی پدیدآور و تئالگان، در پایگاه پایان نامه در مداخل نویسنده، استاد راهنما و استاد مشاور، در پایگاه مقالات در مداخل نویسنده مقاله و به همین ترتیب در پایگاههای دیگر. طبیعی است که برای بعضی از پایگاهها امکان انطباق کامل مداخل فوق وجود ندارد. برای مثال در پایگاه نشریات فیلد پدیدآور معنی ندارد. لازم به توضیح است که این کار اجباری و الزامی نبوده و مدیر مرکز اطلاع رسانی با صلاحدید خود ممکن است یک یا چند پایگاه را از لیست جستجوی همزمان خارج کند.

طراحی این واسط و محاوره تا حد زیادی مشکل جستجوی همزمان در پایگاهها را حل کرد. اما مشکل اساسی این بود که چگونه می توان سند خاصی را در چند کتابخانه جستجو کرد؟ به عبارت دیگر محقق می خواست بداند سند مورد نظر او در کدام کتابخانه ها موجود است. در بخش بعد ابتدا راه حلهای سنتی و ابتدایی که برای این کار پدیدار شد توضیح داده شده و در انتها راه حل شرکت نوسا برای حل این مشکل تشریح شده است.

فهرستگانها به عنوان اولین راه حل

برای پاسخ به این سوال و رفع این نیاز، ابتدا و زمانی که شبکه های مخابراتی و ارتباطی و زیرساختهای سخت افزاری و نرم افزاری ارتباطات رشد نیافته بود، راه حل تهیه فهرستهای مشترک که اصطلاحاً فهرستگان نامیده می شد پیشنهاد شد. در این روش، مجموعه رکوردهای بانکهای اطلاعاتی مختلف در سراسر کشور بصورت درهمکرد در یک پایگاه اطلاعاتی جمع آوری و تشکیل یک بانک بزرگ مشترک را می داد که هر رکورد دارای فیلدی بنام محل نگهداری بود که نشان می داد هر کتاب (سند) در کدام کتابخانه ها وجود دارد. این روش در آن زمان تنها راه حل بود و تعدادی از سازمانها در حیطه تخصصی خود اقدام به تهیه چنین پایگاههایی کردند. اما کم کم اشکالات این روش آشکار شد. اهم نواقص و ایرادات این روش به شرح زیر است:

مکانیزه نبودن بسیاری از کتابخانه ها

در بسیاری از کتابخانه ها نرم افزار مدیریت پایگاههای اطلاعاتی وجود نداشت. بنابراین عملاً روشی که انتخاب شد این بود که یک منبع جامع فهرست نویسی انتخاب می شد (مثل CD-MARC) و سپس با روشهایی مثل کپی گرفتن از رف برگه ها یا تهیه لیستهای چند ستونی، موجودی مراکز را تهیه و این موجودی ها تک تک در آن منبع جستجو شده و جوابها در یک بانک اطلاعاتی ریخته و ذخیره می شد. بدیهی است که به علت موجود نبودن یک منبع کامل و معتبر فهرست نویسی فارسی در آن زمان، این روش تنها برای کتابهای لاتین قابل اجرا بود و البته بسیاری از کتابها نیز بودند که در CD MARC یافت نمی شد و لذا از موجودی خارج می شد.



عدم تطبیق ساختار بسیاری از بانکهای اطلاعاتی با ساختار فهرستگان

هر مرکز و کتابخانه ای در کشور نرم افزار خاص خود را تهیه و اطلاعات خود را در آن وارد کرده بود. این پایگاهها که معمولاً توسط کارشناسان کامپیوتر مراکز تولید شده بود از استانداردهای نرم افزاری و کتابداری بسیار دور بود و لذا یکپارچه سازی این بانکها بدون تبدیل (Convert) اطلاعات تقریباً غیرممکن بود. علاوه بر آن در اکثر موارد بخش کوچکی از موجودی وارد سیستم شده بود و معمولاً این سیستمها پس از ورود اطلاعات به مشکلاتی برمی خورد و متوقف می شد.

کهنه شدن و تاریخ مصرف داشتن

از آنجایی که کتابخانه ها و مراکز اسناد دائماً در حال افزایش موجودی و احیاناً کاهش و حذف تعدادی از مجموعه خود هستند لذا فهرستگانها به علت Offline بودن به سرعت کهنه می شد و روزآمد کردن آنها مستلزم صرف انرژی، زمان و هزینه بسیار زیادی بود.

دشواری بدست آوردن اطلاعات از مراکز مختلف

مراکز مختلف به علل فرهنگی و دشواریهای سازمانی و اداری به سختی اطلاعات خود را عرضه می کنند. لذا یکی از دشوارترین مراحل تهیه فهرستگان همین مرحله بود. بخصوص در مراکزی که دارای نرم افزار کامپیوتری نبودند این مشکل شدیدتر بود. منبع اصلی موجودی در این کتابخانه ها، رف برگه (Shelf List) بود و معمولاً مدیران و کتابداران حساسیت خاصی به این رف برگه نشان می دادند و لذا بسیار به سختی قابل انتقال و دسترسی دیگران بود.

دشواری دسترسی به فهرستگان

به علت محدود بودن ظرفیت محیط های ذخیره اطلاعات، جمع آوری فهرستگان و سپس توزیع آن به سراسر کشور یکی دیگر از مشکلات این طرح بود. ضمن اینکه سخت افزارهای موجود در مراکز مختلف نیز معلوم نبود که بتواند جوابگوی این بانک عظیم باشد. موارد فوق باعث می شد که قابلیت اعتماد و اطمینان اطلاعات فهرستگان بسیار پایین باشد. و به همین خاطر هم بود که این راه حل به سرعت شکست خورد و کنار گذاشته شد.

گام دوم - دسترسی جداگانه به کتابخانه ها از طریق اینترنت

با گسترش امکانات ارتباطی و توسعه نرم افزارها، تولید کنندگان نرم افزارهای بانک اطلاعاتی به این فکر افتادند که بانکهای اطلاعاتی کتابخانه ها را در سایتهای وب قرار داده و از این طریق امکان دسترسی از راه دور را به کاربران بدهند. این روش اگر چه مزایای زیادی را نسبت به روش قبل داشت اما معایب خاص خود را داشت که اهم آنها به شرح زیر است:

دشواری در یافتن و نگهداری آدرس وب مراکز اطلاع رسانی

این روش کاملاً وابسته به نام و آدرس مراکز اطلاع رسانی در وب می باشد. به عبارت دیگر برای اتصال به سایت هر مرکز باید آدرس وب (URL) مرکز را در IE وارد کرد. برای مثال برای اتصال به سایت کتابخانه مرکزی دانشگاه صنعتی شریف باید عبارت www.sut.ac.ir/library را در IE وارد و به آن متصل شد. به همین ترتیب برای دیگر کتابخانه ها و مراکز اطلاع رسانی نیز کاربر باید تمام URL های مربوط به آن مراکز را یا در حافظه خود داشته باشد یا اینکه در جایی ذخیره کرده و در مواقع لزوم از آن استفاده کند. دشواری به خاطر سپاری این آدرس ها باعث شده تا موتورهای جستجویی نظیر google و yahoo در وب به وجود بیاید تا کاربران را در پیدا کردن آدرسهای وب یاری کند. لذا این مورد یکی از اساسی ترین مشکلات این روش می باشد و کاربر باید بخشی از انرژی خود را هر بار برای یافتن و وارد کردن آدرسهای متنوع مراکز مختلف کرده و دائم در جریان تغییرات احتمالی این نامها نیز باید باشد.

میزان صرف وقت و انرژی زمانی خود را نشان می دهد که تعداد مراکز قابل توجه باشد. برای مثال فرض کنید که تمام کتابخانه های تخصصی و دانشگاهی ایران حدود ۵۰۰ کتابخانه باشد و تمامی آنها دارای آدرس اینترنتی نیز باشند. حال اگر شخصی بخواهد در تمام این مراکز جستجو کند باید ۵۰۰ بار آدرس وب مراکز را در مرورگر اینترنتی خود وارد کرده و به سایتهای متصل گردد.

همزمان نبودن جستجو

در این روش جستجوی همزمان معنی ندارد. کاربر در هر لحظه تنها در یک کتابخانه جستجو می کند و برای جستجو در مرکز بعدی باید به مرکز بعدی متصل شود. حداکثر کاری که کاربر می تواند انجام دهد این است که چندین صفحه IE را همزمان باز کرده و در هر کدام به یک کتابخانه خاص متصل شود. بدیهی است که این مساله نمی تواند همزمانی



جستجو را ایجاد کند زیرا کاربر در هر لحظه می تواند تنها در یک سایت به تعامل و تبادل اطلاعات و جستجو بپردازد.

تنوع نرم افزارها در مراکز مختلف

هر مرکز نرم افزار خاص خود با رابطهای مخصوص به خود را دارد و لذا کاربر و محقق، باید طیف وسیعی از نرم افزارها را فرا بگیرد و با وارد شدن به هر مرکز باید امکانات آن نرم افزار خاص را به خاطر داشته باشد. این مساله زمانی خود را بیشتر نشان می دهد که هر نرم افزار نیازمندیهای نرم افزاری و سخت افزاری متفاوتی با دیگران داشته باشد. برای مثال فرض کنید یک نرم افزار تنها با ویندوز عربی و IE عربی کار کند و برعکس نرم افزار دیگر تنها با ویندوز انگلیسی کار کند. یا فرض کنید یک نرم افزار تنها با ویندوز ۲۰۰۰ کار کند و نرم افزارهای دیگر با ویندوز ۹۸. واضح است که در این موارد کاربر به راحتی نمی تواند به مراکز مختلف دسترسی داشته باشد و باید از تعدادی از مراکز صرف نظر کند.

دشواریهای ارتباطی و شبکه ای

با توجه به اینکه ظرفیت ارتباطی هر شبکه و سایت اطلاع رسانی محدود می باشد. لذا امکان اتصال به یک مرکز همواره و با شرایط دلخواه موجود نیست. لذا روش اتصال به سایت مناسبی نیست و بسته به وضعیت شبکه بسیار متغیر بوده و ممکن است با قطع ارتباط دیگر به سادگی نتوان ارتباط مجدد برقرار کرد. نکته دیگر این است که کاربر برای دسترسی به چند مرکز باید چند صفحه IE را همزمان باز نگهدارد اما ظرفیت و پهنای باند خطوط ارتباط اینترنتی اجازه باز کردن تعداد محدودی صفحه همزمان را در اینترنت می دهد و همین مساله باعث لختی و سنگینی در ارتباط می گردد.

گام سوم - ایجاد شبکه مراکز اطلاع رسانی (مدل سیمرغ نوسا)

همانطور که قبلاً اشاره شد در طرح اولیه جستجوی همزمان در چند پایگاه، مدلی بر اساس سه مدخل «پدیدآور»، «عنوان» و «موضوع» طراحی و پیاده سازی شد. واضح است که این مدل به دو دلیل ناقص می باشد: دلیل اول این است که فیلدهای دیگر بانکهای اطلاعاتی که حاوی اطلاعات بارزش می باشند از این مدل کنار گذاشته شده است. به عبارت دیگر در این مدل اطلاعاتی نظیر «ناشر»، «سال نشر»، «مشخصات فیزیکی»، «شماره رده بندی»، «شابک» و ... قابل بازیابی نیستند. و دلیل دوم محدودیت این روش به یک کتابخانه یا مرکز اطلاع رسانی بود. یعنی اینکه مفهوم همزمانی جستجو تنها بین پایگاههای مختلف در یک مرکز معنی داشت لذا برای حل این دو مشکل راه حلهایی در نوسا ارائه شده و در نسخه ۴ نرم افزار سیمرغ پیاده سازی شده است.

برای حل مشکل اول یعنی توسعه و گسترش مداخل جستجو به تمامی اطلاعات بارزش یک سند (رکورد) از راه حل پیشنهادی توافقی (Dublin Core) و برای حل مشکل دوم یعنی گسترش مفهوم همزمانی جستجو از همزمانی بین پایگاهها به همزمانی بین مراکز متعدد اطلاع رسانی، راه حل ایجاد شبکه کتابخانه های سیمرغ ارائه شده که موارد فوق در زیر توضیح داده شده است.

پشتیبانی مدل دوبلین (Dublin Core)

تنوع سلیقه ها در طراحان بانکهای اطلاعاتی باعث شده تا هر متخصص نامهای متنوع و متفاوتی را برای فیلدهای اطلاعاتی خود انتخاب کند و لذا این مساله باعث تشتت و چندگانگی پایگاههای اطلاعاتی شده و مانعی برای یکدستی و قابلیت جستجوی همزمان میان مراکز گشته است. بهر حال برای امکان همزمانی و همگرایی اطلاعاتی تنها راه حل، توافق بر یک مدل واحد و عمومی است که تلاشهای بین المللی برای طراحی مدلی تعمیم یافته و عمومی برای جستجو که حداکثر انطباق را با پایگاههای متنوع و متعدد اطلاعاتی داشته باشد سرانجام منجر به توافقی شده است که بنام توافقی دوبلین معروف است. شرکت نوسا نیز از نسخه ۴ به بعد، در مرورگر تحت ویندوز سیمرغ، این مدل جهانی را برای جستجو پشتیبانی می کند. در واقع این توافق، مدل توسعه یافته جستجوی عمومی نوسا با سه مدخل (پدیدآور، عنوان، موضوع) است که در بخش اول توضیح داده شد. مداخل این مدل پیاده سازی شده در نوسا به شکل زیر می باشد:

• عنوان (Title)

• پدیدآور اصلی (Author/Creator)

• سایر پدیدآوران (Contributor)

• کلیه پدیدآوران (Authors)

• موضوع (Subject)

• ناشر (Publisher)



- شماره بازیابی (Callnumber)

- شابک/شاپا (ISBN/ISSN)

- شرح (Description)

- فرم فیزیکی (Format)

- منبع اثر (Source)

- پوشش زمانی و مکانی اثر (Coverage)

با نگاهی دقیق به مداخل فوق مشخص می شود که طراحان این مدل سعی کرده اند تمام محتوا و اطلاعات ارزشمند رکوردهای اطلاعاتی و اسناد مختلف را قابل جستجو و بازیابی کنند. تولید کنندگان پایگاههای اطلاعاتی در تمام نقاط دنیا برای ارائه اطلاعات خود باید ساختار پایگاههای خود را به نحوی سازگار و منطبق با این مدل کنند. با این کار می توان با یک واسط (Interface) واحد دوبلین به بانکهای اطلاعاتی متنوع مربوط به اسناد متعدد در مراکز اطلاع رسانی دنیا دسترسی پیدا کرد. در اینجا لازم به توضیح است که لزومی ندارد ساختار پایگاههای اطلاعاتی یکسان باشد بلکه باید امکان دسترسی به رکوردها را از طریق مدل دوبلین داشته باشد. شاید بتوان برای روشن تر شدن قضیه بتوان گفت مدل دوبلین مانند پنجره ای است که به مناظر و مکانهای متفاوت باز شده و محتویات آنها را نشان می دهد. این مدل هم اکنون در طراحی پایگاههای اطلاعاتی نوسا وارد و تکمیل شده و مراکز دارای سیمرغ می توانند با تغییراتی در ساختار و تعاریف پایگاههای خود این امکان را پشتیبانی کنند.

نکته بسیار مهم در این مدل این است که جوابهای جستجو از پایگاههای مختلف، در یک لیست واحد و بصورت درهمکرد نمایش داده می شود و می توان از این لیست یک گزارش چاپی یکپارچه گرفت. به عبارت دیگر رکوردهای بازیابی شده از منابعی مانند کتاب، مقاله، پایان نامه و ... در یک لیست واحد نمایش داده شده و قابلیت گزارشگیری های مختلف از این لیست را به کاربر می دهد. نمونه ای از لیست درهمکرد را می توانید در صفحه بعد ببینید:

نوع	تاریخ	عنوان	پدیدآور اصلی
کتاب	۱۲۷۱	اقتصاد ایران	اقتصاد ایران
کتاب	۱۲۷۴	اقتصاد ایران	اقتصاد ایران
کتاب	۱۲۷۴	جنبه های کاربردی خصوصی سازی (بانضمام قوانین و مقررات ایران)	کیمیا
نشریه ادواری			
کتاب	۱۲۷۴	مبانی نظری حسابداری و گزارشگری مالی در ایران: متن پیشنهادی	کیمیا
نشریه ادواری			
کتاب	۱۲۷۵	نامه اتاق بازرگانی	احمدی، علی اصغر، ۱۲۲۴ -
کتاب	۱۲۷۵	تحلیلی تربیتی بر روابط دختر و پسر در ایران	اشرفی ساری بیگلو، صالح
پایان نامه	۱۲۷۱	بهبود سازی کلیماتورنوترونی برای دیفرانکومترنوترونی سازمان انرژی اتمی ایران	اصغری، جعفر
کتاب	۱۲۷۸	علفهای هرز مهم مزارع و مراتع ایران	بشیری نژاد، علی اکبر
کتاب	۱۲۷۵	گستره مالی و ساختار بودجه ای ایران والگوی منابع و مصارف	

از آنجایی که این مدل در یک توافق بین المللی صورت گرفته، پشتیبانی آن گامی در جهت گسترش امکانات جستجوی همزمان به فراتر از مرزهای ملی و در پهنه جهانی است.

شبکه مراکز اطلاع رسانی (شبکه سیمرغ)

اگرچه با انتخاب مدل دوبلین به عنوان مدل عمومی برای محاوره (واسط) جستجو، یک گام بزرگ برای نزدیک شدن به همزمانی جستجو برداشته شد اما گام بزرگ بعدی این بود که امکان ارتباط و تبادل داده ها بصورتی میان مراکز و کتابخانه ها برقرار گردد که یک کاربر کاملاً احساس همزمانی واقعی را داشته باشد. راه حلی که در نهایت، نوسا برای این مساله ارائه داد مدل شبکه Online برای کتابخانه ها بود. در این رویکرد، مفهومی بنام شبکه سیمرغ پدیدار می شود که از کتابخانه های دارای سیمرغ تشکیل می شود. در این روش، هر کتابخانه بعنوان یک گره (Node) از شبکه سیمرغ تعریف می شود و کاربر با اتصال به یک میزبان سیمرغ وارد شبکه On Line کتابخانه های سیمرغ می شود و در نتیجه این کار، در یک لحظه می تواند در چندین کتابخانه جستجو کند و از مراجعات مکرر به تک تک مراکز بی نیاز شود. چنین راه حلی مزایایی زیادی دارد که اهم آن به شرح زیر است:

در این روش دیگر نیازی به اتصال به تک تک کتابخانه ها نیست و کاربر با اتصال به یکی از مراکز میزبان سیمرغ، تمام مراکز میزبان سیمرغ را در اختیار خواهد داشت. تکنولوژی که سیمرغ برای این کار از آن بهره می برد، تکنولوژی XML برای انتقال اطلاعات است. هنگام اتصال به یک کتابخانه، نرم افزار در مدت زمان محدودی به دنبال مراکز زنده و فعال میزبان سیمرغ می گردد و سپس تمام مراکز فوق را در لیست مراکز اطلاع رسانی در اختیار کاربر قرار می دهد تا از میان آنها مواردی را که می خواهد انتخاب کند. نمونه



ای از مراکز موجود در شبکه سیمیرغ در صفحه بعد نشان داده شده است.



از این لحظه کاربر سیمیرغ وارد شبکه ای شده است که از مراکز و کتابخانه های مجهز به سیمیرغ تشکیل شده است. کاربر مختار است به یک یا چند مرکز و یا همه مراکز متصل شده و در آنها به جستجوی همزمان بپردازد. بدیهی است با انتخاب کلید «اتصال عمومی» می توان به تمام مراکز متصل شده و در آنها بطور همزمان به جستجوی اطلاعات پرداخت اما بهر حال هر کاربر با توجه به زمینه موضوعی خود می تواند تعدادی از کتابخانه های مربوط به موضوع خود را انتخاب و در آنها به شکل همزمان جستجو کند. در مثال بالا از میان مراکز اطلاع رسانی فعال در شبکه سیمیرغ (مراکز اطلاع رسانی موجود)، سه مرکز برای اتصال و جستجو انتخاب شده است (مراکز اطلاع رسانی متصل).

مزایای راه حل نوسا

راه حل نوسا که در دو محور اصلی بیان شد تا حد زیادی مشکلات و نقایص راه حل های قبلی را از میان برداشته و دارای مزایای زیادی است که اهم این مزایا به شرح زیر می باشد:

۱) همزمانی واقعی در جستجو

سیمیرغ برای شبیه سازی همزمانی جستجو، به این شکل عمل می کند که پس از دریافت درخواست کاربر، آن را به طور همزمان به تمام مراکز می فرستد و منتظر جواب می ماند. سپس جوابهای درخواست خود را که از مراکز فرستاده شده جمع آوری کرده و در انتها همه را نشان می دهد. نکته بسیار مهم در اینجا این است که این فرایند ارسال درخواست و گرفتن پاسخها، بصورت تک به تک نیست. به عبارت دیگر سیمیرغ منتظر مرکز خاصی برای پاسخ نمی ماند و در همان زمانی که مرکز فوق در حال پردازش و ارسال پاسخ است، درخواستها را به مراکز دیگر می فرستد. در واقع سیستم سیمیرغ از زمانهای بیکاری (Idle Time) و تاخیر مراکز استفاده بهینه می کند و نوعی شبیه سازی همزمانی به وجود می آورد. لذا به همین دلیل است که اگر تعداد مراکز چند برابر هم بشود در زمان انجام جستجو و گرفتن جوابها تغییر محسوسی رخ نخواهد داد و کارایی سیستم پایین نخواهد آمد.

۲) عدم تفاوت و تنوع نرم افزاری

از آنجایی که کاربر در تمام مدت جستجو تنها با یک نرم افزار روبروست لذا کافی است که با امکانات و ابزارهای همین نرم افزار (مرورگر تحت ویندوز سیمیرغ) آشنا باشد. اصولاً کاربر، فرایند ارسال درخواستها به مراکز دیگر و دریافت اطلاعات از آنها را مستقیماً مشاهده نمی کند و به همین دلیل از ابتدا تا انتهای کار با یک واسط (Interface) روبروست و لذا راحت تر می تواند با آن سازگار و منطبق شود.

۳) سازگاری بیشتر با امکانات ارتباطی

با نصب مرورگر تحت ویندوز سیمیرغ، کارایی ارتباطی بسیار بالاتر خواهد رفت. در این روش تنها درخواستهای کاربر و پاسخهای سیستم است که ردوبدل می شود و دیگر ظرفیت خط ارتباطی توسط اطلاعات مربوط به عناصر گرافیکی صفحه (تیترها، ستونها، کادرها و ...) اشغال نمی شود. بدیهی است که این مساله در کیفیت و کمیت جستجوها تاثیر بسیار زیادی خواهد داشت. حتی در حالتی که از مرورگر تحت ویندوز استفاده نشود و از طریق دسترسی وب (Web Access) ارتباط برقرار شود نیز وضعیت ارتباط از حالت دسترسی مجزا بهتر خواهد بود. دلیل این مساله در این است که کاربر تنها به یک مرکز متصل می شود و از مشکلات اتصال به مراکز چندگانه به دور است و ارتباط و تبادل اطلاعات میان مراکز اطلاع رسانی، در سطح پایین تر از اتصال، و توسط پروتکل XML انجام می شود. لذا این مساله باعث می شود که پهنای باند کمتری را اشغال کرده و کارایی سیستم بالا برود.