

نوشته: **جواد بشیری**

کارشناس مرکز اطلاع رسانی و خدمات علمی جهادسازندگی

کلیدواژه ها: آگاهی رسانی جاری / پردازش اطلاعات / اوس.سی.آر. / پویش (اسکن) / اشاعه اطلاعات / فهرست

مندرجات

چکیده

"پایگاه اطلاعاتی فهرست مندرجات جاری" یکی از فعالیت‌های آگاهی رسانی جاری است که توسط مرکز اطلاع رسانی و خدمات علمی جهاد و براساس اطلاعات صفحه فهرست مندرجات بیش از ۵۰۰ عنوان نشریه ادواری هسته در زمینه‌های کشاورزی، دامپروری، جنگل و مرتع، آب و خاک و آبخیزداری، مهندسی، زیست شناسی و... انجام می‌گردد. اطلاعات موجود در صفحه مندرجات هر نشریه پویش (Scan) می‌شود و با استفاده از سیستم تشخیص نوری کاراکترها (OCR) پردازش می‌گردد. محصول آماده علاوه بر کپی بر روی دیسکت، از طریق شبکه بی بی اس نیز در اختیار ۷۰ مرکز آموزشی و تحقیقاتی تحت پوشش در سراسر کشور قرار می‌گیرد. در آغاز برای تهیه این محصول به روش دستی عمل می‌شد، که مشکلات و نارسایی‌هایی به همراه داشت، از این رو سیستم الکترونیکی OCR جایگزین روش قبلی شد که مزایای قابل توجهی دارد.

مرکز اطلاع رسانی خدمات علمی جهاد وابسته به معاونت آموزش و تحقیقات وزارت جهاد سازندگی از سال ۱۹۹۳ با هدف گردآوری، سازماندهی، و اشاعه اطلاعات شروع به فعالیت کرد. جامعه استفاده کنندگان از خدمات مرکز اطلاع رسانی جهاد به دو دسته سازمانی و ملی تقسیم می‌شوند که از نظر سازمانی شامل محققین، پژوهشگران، استادان، دانشجویان و کارکنان وزارت جهاد می‌شود و از نظر ملی نیز محققین، استادان، دانشجویان و دانش پژوهش پژوهان دانشگاهها و مؤسسات دولتی و بخش خصوصی را پوشش می‌دهد.

معاونت آموزش و تحقیقات وزارت جهاد دارای ۷۰ مرکز و مؤسسه آموزشی و تحقیقاتی در استان‌ها و شهرهای مختلف سراسر کشور می‌باشد که در زمینه‌های منابع طبیعی، جنگل و مرتع، آب و خاک و آبخیزداری، دامپروری، دامپزشکی و... فعالیت می‌کنند و تهیه و تأمین اطلاعات مورد نیاز آنها بر عهده مرکز اطلاع رسانی جهاد می‌باشد. به همین جهت این مرکز برای نیل به اهداف خود در اولین گام اقدام به خرید متمرکز منابع اطلاعاتی مورد نیاز خود از جمله نشریات ادواری لاتین نمود. این اقدام، مسؤلیت‌هایی را برای مرکز اطلاع رسانی به دنبال داشت و آن، فراهم کردن امکان استفاده آسان و روزآمد از نشریات ادواری دریافتی برای کلیه استفاده آسان و روزآمد از نشریات ادواری دریافتی برای کلیه استفاده کنندگان بود. بنابراین گام دوم ارائه خدمات آگاهی رسانی جاری بود.

آگاهی رسانی جاری

خدمات "آگاهی رسانی جاری" (۱) خدماتی است مستمر و دائمی که بر اساس منابع اطلاعاتی مختلف از قبیل کتاب، نشریه، و... صورت می‌گیرد و به شکل‌های مختلف مانند "اشاعه اطلاعات گزیده" (۲)، "فهرست مندرجات جاری" (۳) تهیه کتابشناسی جاری، بولتن، خبرنامه و غیره در اختیار استفاده کنندگان قرار می‌گیرد. این خدمات به منظور روزآمد نگه داشتن اطلاعات جامعه استفاده کننده و صرفه جویی در وقت آنها ارائه می‌شود. هر چه جامعه استفاده کننده خاص‌تر باشد، اطلاعات نیز تخصصی‌تر می‌شود. آگاهی از فعالیت‌های پژوهشی و تحقیقاتی و علائق مطالعاتی استفاده کنندگان در موفقیت خدمات آگاهی رسانی جاری نقش حیاتی دارد. این خدمات باید پوشش مناسبی از منابع اطلاعاتی موجود به صورتی روزآمد و گزینشی داشته باشد تا بتوانند علاقه استفاده کنندگان را برانگیزد. یکی از مهم‌ترین و متداول‌ترین شکل های آگاهی رسانی جاری، خدمات فهرست مندرجات جاری می‌باشد.

خدمات فهرست مندرجات جاری

در مرحله نخست از طرح ایجاد پایگاه اطلاعاتی فهرست مندرجات جاری، ابتدا از بین ۷۰ مرکز آموزشی و تحقیقاتی پراکنده در سراسر کشور، ۵ مرکز که دارای منابع اطلاعاتی غنی و امکانات سرویس دهی بهتری بودند به عنوان مراکز مادر (سایت‌های اصلی) انتخاب شدند. سپس از طریق مذاکره با مسئولان اطلاع رسانی این مراکز و همچنین نظرخواهی از جامعه اطلاعاتی پراکنده در بقیه مراکز (۶۵ مرکز)، سیاهه نشریات مورد نیاز محققین و پژوهشگران تحت پوشش تهیه شد و برای اشتراک بیش از ۴۰۰ عنوان نشریه هسته برای اولین سال (۱۹۹۴) اقدام شد. تعداد مجلات در سال‌های ۱۹۹۵ به بعد بنا بر پیشنهاد استفاده کنندگان و مراکز تحت پوشش و کارشناسان مرکز اطلاع رسانی، افزایش

یافته و در حال حاضر به ۵۰۰ عنوان رسیده است.

پس از اشتراك نشریه ادواری سال ۱۹۹۴ و دریافت تدریجی آنها (۲۰۰ تا ۲۵۰ شماره در هر ماه)، مهم‌ترین مسئله، ارائه خدمات فهرست مندرجات جاری بود. بنابراین ابتدا از فهرست مندرجات نشریات، يك نسخه کپی بر روی کاغذ آ ۴ تهیه و اطلاعات کتابشناختی هر شماره در بالای فهرست مندرجات درج می‌گردید. سپس کپی‌های تهیه شده بر اساس عنوان نشریات و به ترتیب الفبایی مرتب می‌شد و حاصل کار، مجموعه‌های را تشکیل می‌داد که به ترتیب شماره گذاری می‌شد. به منظور سهولت در دستیابی به اطلاعات موجود در هر مجموعه، سیاهه عنوان نشریات موجود تهیه و در مقابل هر عنوان (در سیاهه عنوان)، شماره صفحه درج می‌گردید. پس از این مراحل، دستورالعمل مختصری مبنی بر چگونگی استفاده از مجموعه، فرم درخواست مقاله و سایر اطلاعات لازم به آن اضافه می‌شد. محصول به دست آمده يك مجلد ۳۰۰ تا ۴۰۰ صفحه‌های از صفحه فهرست مندرجات نشریات ادواری بود که به تعداد مراکز تحت پوشش (۷۰ نسخه) تکثیر و صحافی می‌شد و در اختیار آن‌ها قرار می‌گرفت. استفاده کنندگان می‌توانستند با مرور سیاهه عنوان فهرست مندرجات در آغاز هر جلد، عناوین نشریات مورد نظر خود را جستجو، و با مراجعه به صفحه مشخص شده، فهرست مندرجات نشریات مورد علاقه خود را بیابند. چنانچه مقاله یا مقالاتی را انتخاب می‌کردند، اطلاعات کتابشناختی کامل نشریه (شامل عنوان نشریه، عنوان مقاله، سال نشر، دوره، شماره، صفحات) را در فرم درخواست مقاله که در اختیار مراکز و مؤسسات تحقیقاتی بود درج و متن کامل مقاله یا مقالات را از مرکز اطلاع رسانی جهاد درخواست می‌کردند. در مدت ۱۸ ماه تهیه فهرست مندرجات به روش فوق، تعداد ۱۸ شماره (هر ماه يك شماره) از آن تهیه شد و در اختیار استفاده کنندگان قرار گرفت. پس از این مدت با نظر سنجی از استفاده کنندگان و بررسی‌های به عمل آمده توسط کارشناسان مرکز اطلاع رسانی مشخص شد که تهیه این محصول به روش دستی مشکلات و مسائلی را در بر دارد، از جمله:

- بالا بودن هزینه تولید (برای تهیه هر شماره از این مجموعه لازم بود که از ۳۰۰ تا ۴۰۰ صفحه فهرست مندرجات نشریاتی که در يك ماه دریافت می‌شد، به تعداد مورد نیاز (حداقل ۷۰ نسخه) تکثیر و صحافی شود)؛
 - صرف وقت زیاد برای تولید (برای تهیه کپی اولیه، کنترل و الفبایی کردن آن‌ها، تهیه سیاهه جداگانه مربوط به هر شماره، تکثیر به تعداد مورد نیاز و صحافی آن، زمان زیادی صرف می‌شد)؛
 - بالا بودن هزینه ارسال (با توجه به حجم زیاد مجموعه که در مجلدهای ۳۰۰ تا ۴۰۰ صفحه‌های و در قطع بزرگ تهیه می‌شد، ارسال آن از طریق پست، هزینه زیادی داشت)؛
 - دشواری در دستیابی به اطلاعات (این مجموعه براساس عنوان نشریات، الفبایی می‌شد و بنابراین جستجوی اطلاعات فقط از طریق عنوان امکان‌پذیر بود)؛
 - عدم امکان درهم‌کرد (با توجه به اینکه هر شماره به صورت جداگانه و بر روی کاغذ آ ۴ تهیه می‌شد، ایجاد درهم‌کرد امکان‌پذیر نبود).
- بنابراین با نظرسنجی از استفاده کنندگان و بررسی‌های انجام شده، مشخص شد که تهیه این محصول به روش دستی، با وجود هزینه بالا، کارایی زیادی ندارد؛ در نتیجه تصمیم گرفته شد که این خدمات به صورت الکترونیکی ارائه شود.

تشخیص نوری کاراکترها

افزایش روزافزون تولید اطلاعات که غالباً به صورت مواد چاپی به بازار عرضه می‌گردد و ضرورت دسترسی سریع و آسان به داده‌های موجود از يك سو، و حضور فراگیر رایانه در عرصه‌های مختلف زندگی از سوی دیگر، متخصصان سخت افزاری و نرم افزاری رایانه و کارشناسان و صاحب‌نظران اطلاع رسانی را بر آن داشته که همواره در پی یافتن راهکار مناسب برای گردآوری و پردازش اطلاعات به کمک رایانه باشند. از این رو مهم‌ترین مسئله، فرآیند تغییر شکل داده‌ها از حالت چاپی به صورت رقمی (ضبط بر روی حافظه رایانه که با عنوان ورود اطلاعات از آن نام برده می‌شود) می‌باشد. برای حل این مشکل، نخستین راه حلی که به ذهن خطور می‌کند انتقال داده‌ها از طریق صفحه کلید است که خود مستلزم مسائل و مشکلاتی است و برای عملی شدن آن، سال‌ها وقت، و تعداد بی‌شماری کاربر و تجهیزات مورد نیاز است. این روش، صرف‌نظر از این که هزینه بالایی دارد، به علت اشتباهاتی که در حین ورود اطلاعات پیش می‌آید، کاربرد چندانی نخواهد داشت. لذا این مسائل منجر به تولید و تکمیل دستگاه‌های گوناگون پویشگر (اسکنر) و نرم افزارهای مختلف تشخیص نوری کاراکترها (۴) گردیده است. با استفاده از سیستم الکترونیکی ورود اطلاعات، با وجود صرفه‌جویی در هزینه، سرعت ورود اطلاعات نیز به طور قابل ملاحظه‌ای افزایش می‌یابد. در این فرایند دو عمل اتفاق می‌افتد:

- پویش،

- پردازش.

در مرحله پویش (اسکن)، ابتدا متن به تصویر تبدیل می‌شود و سپس توسط نرم افزار OCR نقطه به نقطه تصویر (صفحه پویش شده) با مجموعه کاراکترهای تعریف شده در نرم افزار OCR که بر روی حافظه موقت رایانه می‌باشد، مطابقت می‌گردد. مجموعه کاراکترها یا حروفی که درست تشخیص داده شده و تشکیل دهنده تصاویر یا کلمات در متن اصلی بوده،

بر روی حافظه رایانه ذخیره می‌شود. کاراکترهایی که توسط نرم افزار قابل تشخیص نبوده‌اند به صورت علائم قراردادی مشخص می‌شود تا توسط کاربر اصلاح شود. از مشکلات قابل ذکر در این مرحله این که صفحه مندرجات اغلب مجلات از نظر شکل، رنگ و اندازه با یکدیگر متفاوت‌اند و غالباً از فونت‌های مختلف، عکس، تصاویر گرافیکی و حروف غیرانگلیسی استفاده می‌کنند و در برخی موارد، صفحات فهرست مندرجات در جاهای مختلف مجله می‌آید که مشکلاتی را برای کاربر به وجود می‌آورد. در مرحله پویس سعی می‌شود از اطلاعات توضیحی غیرضروری، تصاویر و حروف گرافیکی صرف نظر شود تا ضمن بالا رفتن سرعت عمل، از اشغال اضافی حافظه خودداری گردد. هر قدر کیفیت صفحه فهرست مندرجات بهتر و واضح‌تر باشد و نرم افزار مورد استفاده، قابلیت و انعطاف بیشتری داشته باشد، حاصل کار موفقیت‌آمیزتر خواهد بود. پس از پویس صفحات فهرست مندرجات ۲۵۰ تا ۳۰۰ شماره از نشریات دریافتی در یک ماه، داده‌های جمع‌آوری شده در یک پرونده (فایل) ذخیره می‌شود تا در مرحله بعدی پردازش شود. مرحله بعد از پویس صفحات فهرست مندرجات، پردازش داده‌ها است.

در این مرحله با فراخوانی یک به یک رکوردها، اطلاعات کتابشناختی مربوط به هر کدام شامل عنوان نشریه، سال نشر، دوره، شماره، و محل نگهداری در فیلد خاص خود وارد می‌شود [منظور از محل نگهداری، کتابخانه‌های مراکز مادر (که قبلاً ذکر شدند) می‌باشند]. پس از آن کلید واژه‌های مربوط به هر رکورد به طور جداگانه استخراج می‌شود. البته در نرم افزار مورد استفاده حروف اضافه، حروف ربط و غیره (مانند a, an, and, or, the) کلید واژه محسوب نمی‌شوند.

در پایان، تمام کلید واژه‌های به دست آمده به ترتیب الفبایی مرتب و در پنج فایل با نام‌های CDX.SERIAL, KEYS.DBF, SERIAL.FPT, SERIAL.DBF، KEYS.CDX ذخیره می‌شوند.

پس از این مراحل، محصول قابل استفاده، از طریق شبکه داخلی (۵) در اختیار مراکز و مؤسسات آموزشی و تحقیقاتی تحت پوشش قرار می‌گیرد. علاوه بر این، اطلاعات مورد نظر به صورت فشرده بر روی دیسکت ضبط و در اختیار استفاده کنندگان غیر جهادی قرار می‌گیرد.

نحوه استفاده از پایگاه اطلاعاتی فهرست مندرجات جاری

برای استفاده از این پایگاه اطلاعاتی، کارشناسان مرکز اطلاع رسانی نرم افزاری را طراحی کرده‌اند که همراه با اولین شماره فهرست مندرجات برای مراکز تحت پوشش ارسال شده است. کاربران با نصب برنامه بر روی سیستم خود می‌توانند با اجرای دستور JSIS.BAT وارد برنامه شوند و با انتخاب گزینه جستجو، به منوی اصلی راه یابند. در این قسمت کاربران می‌توانند براساس فیلدهای مختلف مانند عنوان نشریه، عنوان مقاله، موضوع نشریه، کلید واژه و محل نگهداری نشریات، کاوش کنند و با انتخاب مقاله یا مقالات مورد نیاز خود، اطلاعات کامل کتابشناختی آن را در فرم درخواست مقاله درج و برای مرکز اطلاع رسانی ارسال نمایند.

نتیجه

صرفه جویی ارزی در تهیه و تأمین منابع اطلاعاتی، روزآمد نگه داشتن اطلاعات استفاده کنندگان، و معرفی راهکاری نو در جهت ذخیره و اشاعه اطلاعات، از مسائل مهمی است که در این مقاله مورد بحث قرار گرفته است. در همین راستا، مرکز اطلاع رسانی و خدمات علمی جهاد برای عملی ساختن این سه عامل، سیاست‌ها و روش‌هایی را در پیش گرفته است که عبارت‌اند از: خرید متمرکز منابع اطلاعاتی، خدمات آگاهی رسانی جاری، و پویس و پردازش اطلاعات با استفاده از روش OCR.

پی نوشت‌ها

- 1.Current Awareness
- 2.SDI
- 3.Current Contents
- 4.Optical Characters Recognition (OCR)
- 5.Local Area Network

منابع

- ۱.ویونگ، ادموند. مراکز اطلاع رسانی و خدمات آگاهی رسانی جاری. ترجمه عباس گیلوری (در دست انتشار).
- ۲.هنروران، فرامرز. "اوس.آر. و کاربرد آن در کتابخانه‌ها و مراکز اطلاع رسانی". فصلنامه پیام کتابخانه، سال ششم، شماره اول: (۵۳ - ۵۰).