

بررسی تأثیر شیوه‌های نمایه‌سازی
بر قابلیت بازیابی منابع مشخص و موجود
در سه نرم‌افزار کتابخانه‌ای فارسی^(۱)

غلامرضا سلگی^(۲)

چکیده

پژوهش حاضر به بررسی تأثیر شیوه‌های متفاوت نمایه‌سازی بر قابلیت بازیابی اطلاعات در نرم‌افزارهای کتابخانه‌ای فارسی پرداخته است. در این پژوهش با استفاده از روش تحقیق علی/مقایسه‌ای با محور قرار دادن یک فرضیه اصلی و سه فرضیه دیگر، وجود یا عدم وجود تفاوت بین متغیر مستقل "شیوه‌های نمایه‌سازی" و متغیر وابسته "قابلیت بازیابی یا میزان بازیابی" در سه نرم‌افزار فارسی پارس آذرخش، کاوش و نوسا و در سه فیلد عنوان، پدیدآور و موضوع را مورد بررسی قرار داده است. مشخص گردید که بین شیوه‌های نمایه‌سازی و قابلیت بازیابی منابع مشخص و موجود در سه نرم‌افزار مورد تحقیق تفاوت معنی‌داری وجود دارد. آزمون مجذور کای χ^2 در سطح ۰/۰۱ و درجه آزادی ۲، میزان این تفاوت را ۱۶/۵۹۵ محاسبه نموده است. تفاوت موردنظر عمدتاً ناشی از مسایل و مشکلاتی است که در حوزه‌های مهندسی نرم‌افزار (برنامه‌نویسی)، نمایه‌سازی، ورود اطلاعات و بازیابی اطلاعات وجود دارد.

مقدمه

۱- برگرفته از پایان‌نامه کارشناسی ارشد کتابداری و اطلاع‌رسانی به راهنمایی دکتر رحمت‌الله فتاحی

۲- کارشناس ارشد کتابداری و اطلاع‌رسانی

با ظهور رایانه و فناوری اطلاعات به عرصه ذخیره و بازیابی اطلاعات تحول شگرفی در این حوزه پدید آمد. از طرفی، حجم بسیار بالایی از اطلاعات را می‌توان در نرم‌افزارهای کتابخانه‌ای ذخیره نمود. از طرف دیگر، سرعت و قابلیت‌های بازیابی اطلاعات نیز افزایش چشمگیری یافته است. این مزایا باعث شده که انتظارات ما نیز از فناوری اطلاعات و سیستم‌های رایانه‌ای ذخیره و بازیابی اطلاعات افزایش یابد. اما در کنار این مزایا، این سیستمها باعث بروز برخی از مشکلات و نارسایی‌ها در امر ذخیره و بازیابی اطلاعات شده‌اند. از جمله مشکلات و اختلالات موجود حوزه بازیابی اطلاعات در نرم‌افزارهای رایانه‌ای، افت اطلاعات است. منظور از افت اطلاعات این است که رکوردهای مرتبط با موضوع یا منابع مشخص و موجود که در فایل قرار دارند به دلایلی بازیابی نمی‌شود.

علی‌رغم پیشرفتهایی که در زمینه فهرست‌های رایانه‌ای در دنیا صورت گرفته است، در ایران نظام‌های موجود هنوز با وضعیت مطلوب فاصله زیادی دارند. این امر از یکسو به خود نرم‌افزارها و مهندسی آنها مربوط می‌شود، و از سوی دیگر، به کاربری آنها. برخی از عوامل مربوط به مشکلات نرم‌افزارها عبارتند از توان و دانش فنی برنامه‌نویسان نظام‌های کتابخانه‌ای، میزان آگاهی آنان نسبت به نیازهای کاربران نظام‌ها، قابلیت‌هایی که جامعه استفاده‌کننده از آنها انتظار دارد (فرج‌پهلوی، ۱۳۷۹، ص ۴۳)، و ارتباط کمرنگ مهندسان برنامه‌نویس با کتابداران و اطلاع‌رسانان برنامه‌ساز.

سوالات پژوهش

- ۱ - آیا با به کارگیری شیوه‌های متفاوت نمایه سازی فیلدها در نرم افزارهای فارسی، نتایج بازیابی متفاوتی به دست می‌آید؟
- ۲ - تأثیر شیوه‌های نمایه سازی متفاوت بر قابلیت بازیابی اطلاعات در نرم افزارهای فارسی تا چه میزان است؟

هدف و فایده پژوهش

۱. تأثیر شیوه نمایه‌سازی بر قابلیت بازیابی منابع مشخص و موجود در سه نرم‌افزار فارسی و مشخص ساختن اینکه شیوه‌های نمایه‌سازی به کار گرفته شده در کدامیک از نرم‌افزارهای مورد بررسی از قابلیت بازیابی بیشتری برخوردار است.
۲. فراهم‌آوری بستر مناسب برای ارزیابی سایر نرم‌افزارهای کتابخانه‌ای با توجه به

یافته‌های این پژوهش.

فرضیه‌های پژوهش

فرضیه اصلی: بین قابلیت بازیابی متاثر از شیوه‌های نمایه‌سازی در سه نرم‌افزار مورد تحقیق تفاوت معنی‌داری وجود دارد.

فرضیه‌های فرعی

۱-۱- بین قابلیت بازیابی متاثر از شیوه‌های نمایه‌سازی فیلد عنوان در سه نرم‌افزار مورد تحقیق تفاوت معنی‌داری وجود دارد.

۲-۱- بین قابلیت بازیابی متاثر از شیوه‌های نمایه‌سازی فیلد پدیدآور در سه نرم‌افزار مورد تحقیق تفاوت معنی‌داری وجود دارد.

۳-۱- بین قابلیت بازیابی متاثر از شیوه‌های نمایه‌سازی فیلد موضوع در سه نرم‌افزار مورد تحقیق تفاوت معنی‌داری وجود دارد.

مروری بر پژوهش‌های انجام شده در خارج از کشور

اهمیتی که بازیابی اطلاعات و موضوعات مرتبط با این حوزه مانند نمایه‌سازی (شیوه‌ها، زبانها و اصطلاحنامه‌ها)، از دیدگاه کتابداران وجود دارد، موجب شده است که این موضوع در کشورهای پیشرفته نظیر ایالات متحده، کانادا و بریتانیا مورد توجه قرار گرفته و پژوهش‌های بسیاری در این زمینه صورت گیرد که هر یک به جنبه‌های مشخصی از موضوع پرداخته‌اند. در این قسمت چند اثر مهم از صاحب‌نظران کشورهای دیگر مورد بررسی و مرور قرار می‌گیرد.

پژوهش‌های سیریل کلوردن^(۱): به لحاظ تاریخی، نخستین تلاش عمده برای ارزشیابی و آزمایش نرم‌افزارها و پایگاه‌های اطلاعاتی، توسط سیریل کلوردن و در دو مجموعه آزمایش تجربی بنام کرینفیلد ۱ و ۲^(۲) انجام شد. مجموعه آزمایش‌های نخست در سال ۱۹۵۷ اجرا و در سال ۱۹۶۲ گزارش شد. هدف کرینفیلد (۱) مقایسه زبان نمایه‌سازی پس‌هما را با زبانهای نمایه‌سازی سنتی بود که با استفاده از ۱۸۰۰۰ مدرک و ۱۲۰۰ موضوع جستجو انجام گرفت. نکته قابل توجه آنکه، اخص بودن^(۳) و ازگان نمایه‌سازی و درجه جامعیت^(۴) نمایه‌سازی اهمیت

1- Cyril Cleversdon

2- Cranfield 1 & 2

3- Specificity

بیشتری نسبت به تفاوت‌های زبانهای نمایه‌سازی مورد استفاده داشت (پاٲو، ۱۳۷۹، ص ۴۱۸).
مجموعه مطالعات وی بصورت گزارشی در سال ۱۹۶۲ با مشخصات زیر انتشار یافت:

Report on the Testing and Analysis an Investigation into the Comparative Efficiency of Indexing Systems ".Aslib - Cranfield Report, Cranfield, England, October 1962.

نظام بازیابی اسمارت (SMART)^(۵) به مدیریت سالتون^(۶): نظام اسمارت در دانشگاه هاروارد بین سالهای ۱۹۶۱ تا ۱۹۶۴ و سپس با همکاری دانشگاه کرنل تا سالیان متمادی ادامه داشت. برای مدت بیست سال، سالتون (مسئول این نظام) با استفاده از طرحهای گوناگون ارزش‌گذاری اصطلاحها، راهبردهای جستجو، فرمولبندی خودکار جمله جستجو که بر مبنای بازخوردهای ربط انجام می‌شود، و بعد هم جستجوهای بسط یافته بولی، دست به آزمایشهای گسترده‌ای زد. در این آزمایشها از شیوه‌های پیشرفته تحلیل معنایی و نحوی استفاده شد. یکی از یافته‌های او به این نکته اشاره دارد که بازیابی و پردازش اطلاعات با استفاده از زبان طبیعی بسیار پیچیده است و بدون درک عمیقی از نظریه‌های زبانشناختی نمی‌توان نمودارهای معنایی واقع‌گرایانه و ساختارهای نحوی را برای بازیابی موثرتر پیاده کرد. در این مرحله، روشهای ساده‌تر نمایه‌سازی نتایج نسبتاً بهتری به همراه داشته است (پاٲو، ۱۳۷۹، ص ۴۱۹).

مطالعات لنکستر^(۷): لنکستر به عنوان نخستین فردی شناخته می‌شود که در زمینه سودمندی یکی از نظامهای موجود، یعنی مدلارز^(۸) مطالعه‌ای گسترده انجام داده است. همچنین وی پژوهشهای بسیاری در زمینه نظامهای بازیابی اطلاعات علوم زیستی و پزشکی انجام داد که نتایج بسیار خوبی را به همراه داشت.

پژوهشهای سیورتز و همکاران وی^(۹): سیورتز و همکاران وی مطالعات گسترده‌ای در مورد نظامهای بازیابی و نرم افزارهای موجود انجام دادند. وی طی سلسله مقالاتی از سال ۱۹۹۲ تا ۱۹۹۳ به بررسی قابلیت نرم‌افزارهای موردنظر پرداخت و با توجه به معیارهای ۱۱۰ گانه ارزشیابی عملکرد پایگاههای اطلاعاتی، ۸ برنامه نرم‌افزاری مختلف را مورد ارزیابی، سنجش و مقایسه قرار داده است. برخی از این معیارها عبارتند از: ورودی‌ها و ویژگی اصلی فایل‌ها،

4- Exhaustivity

5- Salton, 1983

6- Salton

7- Lancaster

8- MEDLARS

9- Seiverts

تجدید سازمان فایل‌ها، شیوه‌نمایه‌سازی، زبانهای نمایه‌سازی، خروجی‌ها و....

پژوهشهای انجام شده در ایران

بیشتر پژوهشهای حوزه نمایه‌سازی در پایگاههای اطلاعاتی در ایران از یک رویکرد عمومی در بررسی موضوع برخوردارند و کمتر به جنبه‌های تخصصی موضوع پرداخته‌اند. در پژوهشی که بابایی (۱۳۷۴) انجام داد سوابق و مقدماتی که منجر به ایجاد پایگاههای اطلاعاتی در مرکز اطلاعات و مدارک علمی ایران و سازمان پژوهشهای علمی و صنعتی ایران گردیده، مورد بررسی قرار گرفته است. محورهای مورد بحث آن پژوهش عبارتند از: اهداف، تعداد رکورد و پوشش زمانی اطلاعات پایگاهها، منابع و شیوه جمع‌آوری اطلاعات، گردش کار اطلاعات و مدارک، چگونگی آماده‌سازی مدارک و نمایه‌سازی آنها در هر یک از پایگاههای اطلاعاتی که بطور جداگانه مورد بررسی قرار گرفته و مسائل و مشکلات هر مقوله تجزیه و تحلیل شده است.

خسروانی (۱۳۷۴) نیز در پایان‌نامه خود پایگاههای اطلاعاتی معاجم لفظی و معاجم موضوعی مرکز تحقیقات کامپیوتری علوم اسلامی را مورد بررسی و تجزیه و تحلیل قرار داده است. هر پایگاه مستقلاً از نظر مشخصات کتاب، اهداف، تعداد رکوردها، برونداد پایگاه، نحوه آماده‌سازی اطلاعات، منابع و شیوه استخراج واژگان و موضوعات و قابلیت‌های هر پایگاه مورد بررسی قرار گرفته است. همچنین چگونگی انتخاب موضوعات و نمایه‌سازی که از دستورالعمل‌های مشخصی پیروی می‌کنند مورد ملاحظه قرار گرفته است. عمل بازیابی اطلاعات در پایگاههای معاجم لفظی از طریق عملگرهای متناسب شش‌گانه و نیز «فهرست منوی اصلی» صورت می‌گیرد، لیکن دایره جستجو در پایگاههای معاجم «لفظی و موضوعی» تنها به این دو روش محدود نبوده، بلکه از طریق کلیه کلید واژه‌ها، نمایه‌ها و موضوعات نیز مقدور است. عمده‌ترین مانع در بازیابی عدم استفاده از محیط گرافیکی است که کلیه متون را که به زبان عربی می‌باشند بدون علائم و حرکات به کاربر عرضه می‌نماید که در نتیجه کاهش ضریب دقت در بازیابی اجتناب‌ناپذیر است.

در پژوهشی دیگر (قریب، ۱۳۷۷) وضعیت نمایه‌سازی (شیوه و زبان نمایه‌سازی، تخصص، نیروی انسانی، بودجه تخصیص یافته برای این بخش و استفاده از رایانه) در کتابخانه‌ها و مراکز اطلاعات رسانی مستقر در شهر تهران را مورد بررسی قرار داده است. روش تحقیق پیمایشی است. برای گردآوری اطلاعات ابتدا نام و نشانی مراکز تهیه شده و سپس پرسشنامه به مراکز ارسال شده و پس از استخراج اطلاعات، مجدداً از اکثریت آنها بازدید بعمل

آمده و مصاحبه صورت گرفته است. جامعه مورد مطالعه شامل ۲۳ مرکز می باشد. نتایج بررسی نشان می دهد حدود ۲/۵ درصد مراکز از هر دو شیوه پس همارا و پیش همارا برای نمایه سازی منابع خود استفاده می کنند. بالاترین درصد انتخاب زبان نمایه سازی، زبان کنترل شده است که از سوی نزدیک به ۷۳ درصد مراکز مورد استفاده قرار گرفته است.

رضایی (۱۳۷۷) در پایان نامه خود، ارزشیابی قابلیت های ذخیره و بازیابی در برنامه های نرم افزاری مراکز اطلاع رسانی علوم قرآنی و معارف اسلامی را مورد بررسی قرار داده است. روش تحقیق توصیفی است. برای گردآوری اطلاعات به طور مستقیم به کتابخانه و مراکز اطلاع رسانی مراجعه شده است. جامعه مورد مطالعه شامل: نرم افزارهای علوم اسلامی موجود در مراکز علوم اسلامی در شهرهای تهران، قم و اصفهان است (۱۰ نرم افزار). به منظور ارزیابی نرم افزارها صدو یازده معیار با استفاده از پژوهش ای. جی سیورتنز^(۱) و دیگران و تحقیق غائبی در نظر گرفته شده است. برای اندازه گیری نرم افزار و تجزیه و تحلیل داده ها از روش میانگین حسابی وزنی استفاده شده است. در کل محقق در پژوهش خود به این نتیجه رسیده است که حدود ۷۵ درصد از نرم افزارهای موجود فاقد کارآیی، کیفیت و توانایی های لازم می باشند.

چشمه سهرابی (۱۳۷۸) در پایان نامه خود، تعیین میزان تاثیر استفاده از اصطلاحنامه در بانک های اطلاعات کتابشناختی بر میزان جامعیت، مانعیت، و مدت زمان جستجوی اطلاعات بازیابی شده را مورد بررسی قرار داده است. روش پژوهش تجربی، و ابزار گردآوری اطلاعات مشاهده مستقیم، و مصاحبه می باشد.

جامعه پژوهش متشکل از هفت بانک اطلاعاتی، ERIC, MEDLINE, INE, (AGRIS, INSPECT & WRA) CAB, است که بر اساس روش نمونه گیری تصادفی انتخاب شده اند. بر اساس یافته های تحقیق، استفاده از اصطلاحنامه به هنگام بازیابی باعث افزایش چشمگیر میزان مانعیت و کاهش میزان جامعیت و مدت زمان جست و جو می شود. حال آنکه استفاده از زبان طبیعی به هنگام ارزیابی، افزایش میزان جامعیت و مدت زمان جستجو و کاهش میزان مانعیت را بدنبال دارد.

موسوی (۱۳۷۹) در تحقیق خود به ارائه شاخصهایی برای ارزیابی کارآیی نرم افزارهای کتابخانه ای در زمینه فهرستنویسی پرداخته است. در این تحقیق با استفاده از روش پیمایشی سه نرم افزار نوسا، کاوش و پارس آدرخش را از دیدگاه فهرستنویسی مورد ارزیابی قرار داده است.

ارزیابی نرم افزارهای مورد نظر شامل ورود اطلاعات کتابشناختی، برگه آرایی، آرایش داده‌ها، امکانات جستجو در فهرستهای رایانه‌ای، و انتقال اطلاعات از CD-MARC می‌باشد.

یوسفی (۱۳۷۹) در مقاله‌ای تحقیقی به بررسی و تجزیه و تحلیل ریزش کاذب ناشی از ذخیره نادرست اطلاعات، خطا در ورود اطلاعات، بازیابی ناخواسته هم نویسه‌ها، ریزش کتاب در جستجوی سرنام‌ها، جابجایی کلمات مرکب و عبارات، عدم جستجوی بولی، قابل کاوش نمودن برخی از فیلدها، مبنای موتورهای جستجو برای تشخیص ابتدا و انتهای کلمه، ریزش کاذب ناشی از برنامه‌نویسی یا استفاده از واژه بر در هنگام جستجو، جستجو در فیلد آزاد، عدم رتبه بندی مدارک بازیابی شده، ارجاع خودکار نادرست در سرعنوانهای موضوعی فارسی، جستجوی حرف به حرف، و امثال آن در سه نرم‌افزار پارس آذرخش، نوسا و کاوش پرداخته است.

نوع پژوهش، جامعه آماری، حجم نمونه و روش نمونه‌گیری

پژوهش حاضر یک مطالعه کتابخانه‌ای و از نوع کاربردی و به روش علی - مقایسه‌ای در نظر گرفته شده است.

جامعه آماری از نوع جامعه محدود بوده و متشکل از سه نرم‌افزار پارس آذرخش، کاوش و نوسا است.

برای محاسبه نمونه آماری از فرمول کرجسی^(۱) و مورگان^(۲) (دیانی، ۱۳۷۸، ص ۶۶) استفاده شده است. بدین ترتیب با استفاده از این فرمول تعداد یا حجم نمونه ۲۷۸ رکورد تعیین شد. با توجه به حجم نمونه بدست آمده لازم بود که ۲۷۸ رکورد عنوان کتاب به صورت مشترک برای هر سه نرم‌افزار مورد نظر انتخاب و مورد آزمون قرار گیرند، اما برای استحکام، جامعیت بیشتر و عینی‌تر نمودن تحقیق، حجم نمونه به ۳۰۰ رکورد افزایش یافت. روش انجام کار نیز بدین گونه بود که کتابشناسی ملی بر روی دیسک فشرده (نسخه ۱۳۷۹) به عنوان پایگاه اطلاعاتی مبنا قرار گرفت. در مرحله اول با جستجو در این پایگاه و با استفاده از روش نمونه‌گیری منظم یا نظام یافته^(۳)، پس از تعیین فاصله رکوردها که برای آن عدد ۴۵۰ به دست آمد به صورت تصادفی عددی انتخاب و مبنای شروع جستجو و گزینش رکوردها قرار گرفت. سپس با در نظر گرفتن فاصله رکوردها، بعد از هر ۴۵۰ رکورد آخرین رکورد انتخاب و عنوان آن

در فرم مربوطه ثبت گردید. در این مرحله چون به طور طبیعی دستیابی به رکوردهای مشترک حجم نمونه مورد نظر ادامه پیدا نمود که در هر مرحله عدد جدیدی به صورت تصادفی انتخاب و مبنای دستیابی به حجم نمونه مورد نظر قرار گرفت. و به این ترتیب ۳۰۰ عنوان رکورد به عنوان حجم نمونه انتخاب و اطلاعات آنها در فرم عنوان ثبت گردید. پس از آن از ۳۰۰ عنوان رکورد انتخاب شده ۳۰۰ مورد نام پدیدآور و به همین تعداد نیز موضوع استخراج و در فرم‌های پدیدآور و موضوع ثبت گردید.

در جدول شماره (۱) نسبت توزیع شده نمونه آماری برحسب جامعه آماری برای سه نرم افزار مورد نظر و به نسبت مساوی نشان داده شده است.

جدول شماره ۱: نمونه آماری تحقیق در سه نرم افزار مورد بررسی تحقیق بر اساس فیلدهای سه گانه

عنوان، پدیدآورنده و موضوع (بر حسب تعداد دفعات جستجو)

| نمونه آماری | تعداد دفعات جستجو | | | جمع |
|-------------|-------------------|--------------|---------|------|
| | عنوانها | پدیدآورندگان | موضوعها | |
| الف | ۳۰۰ | ۳۰۰ | ۳۰۰ | ۹۰۰ |
| ب | ۳۰۰ | ۳۰۰ | ۳۰۰ | ۹۰۰ |
| ج | ۳۰۰ | ۳۰۰ | ۳۰۰ | ۹۰۰ |
| جمع | ۹۰۰ | ۹۰۰ | ۹۰۰ | ۲۷۰۰ |

ابزار گردآوری داده‌ها

در تحقیق حاضر از فرم‌های محقق ساخته برای گردآوری و ثبت داده‌ها استفاده شد. فرم‌های مورد نظر در قالب ۳ فرم الف: فرم بازیابی اطلاعات بر اساس فیلد عنوان ب: فرم بازیابی اطلاعات بر اساس فیلد پدیدآور ج: فرم بازیابی اطلاعات بر اساس فیلد موضوع، طراحی گردید. در هر یک از فرم‌های مورد نظر، شماره رکورد در پایگاه منبع مشترک و شیوه نمایه سازی (کلمه، عبارت، فیلد کامل) به صورت مشترک در سه نرم افزار مورد نظر جستجو و بازیابی گردید.

فرم‌های مورد نظر ابتدا در مورد تعدادی پیشینه، پیش آزمون گردید و مشکلات و مسایل احتمالی آن شناسایی و مرتفع گردید. نهایتاً این فرم‌ها با کمک فکری استادان راهنما و مشاور و

برخی از صاحب نظران اصلاح شد.

شیوه‌های تجزیه و تحلیل یافته‌ها

داده‌های ثبت شده در فرمهای موردنظر، در نرم افزار آماری SPSS وارد شده و بر اساس متغیرهای موجود در سوالها و فرضیه‌ها، و با استفاده از روش‌های آماری نظیر آزمون مجذور کای X^2 ، آزمون نسبت و جداول توافقی، فراوانی، درصد، تجزیه و تحلیل انجام گرفت.

تجزیه و تحلیل یافته‌ها بر اساس شیوه‌های نمایه‌سازی در فیلدها

بررسی نتایج بازیابی بر اساس شیوه‌های نمایه‌سازی در فیلد عنوان

در این مرحله مجموع قابلیت بازیابی (درصد بازیابی) هر یک از نرم افزارهای مورد تحقیق در سه مرحله‌ی قبل بر اساس شیوه نمایه‌سازی بصورت (کلمه، عبارت، فیلد کامل) در "فیلد عنوان" موردنظر تحقیق است. نتایج بدست آمده نشان می‌دهد که از مجموع ۳۰۰ رکورد منابع مشخص و مشترک موجود بین سه نرم‌افزار مورد تحقیق، نرم‌افزار "الف" با ۲۶۲ مورد بازیافت درست (۸۷/۳٪) نرم‌افزار "ب" با ۲۴۱ مورد (۸۰/۳٪)، و نرم‌افزار "ج" با ۲۵۸ مورد (۸۶٪) قابلیت بازیابی درست را دارا هستند.

جدول شماره ۲ - نسبت کلی بازیابی (درصد بازیابی) درست در هر یک از نرم افزارها با

توجه به سه شیوه نمایه سازی در فیلد عنوان

| جمع | نتیجه بازیابی | | | | نرم افزار |
|-----|---------------|-------|-----------|-------|-----------|
| | پیدا شده | | پیدا نشده | | |
| | تعداد | درصد | تعداد | درصد | |
| ۳۰۰ | ۲۶۲ | ۸۷/۳٪ | ۳۸ | ۱۲/۷٪ | الف |
| ۳۰۰ | ۲۴۱ | ۸۰/۳٪ | ۵۹ | ۱۷/۹٪ | ب |
| ۳۰۰ | ۲۵۸ | ۸۶٪ | ۴۲ | ۱۴٪ | ج |

بررسی نتایج بازیابی بر اساس شیوه‌های نمایه‌سازی در فیلد پدیدآور

در این مرحله مجموع قابلیت بازیابی (درصد بازیابی) هر یک از نرم افزارهای مورد تحقیق در سه مرحله قبل بر اساس شیوه نمایه سازی بصورت (کلمه، عبارت، فیلد کامل) در "فیلد پدیدآور" مورد نظر تحقیق است. نتایج بدست آمده نشان می دهد که از مجموع ۳۰۰ رکورد منابع مشخص و مشترک موجود بین سه نرم افزار مورد تحقیق، نرم افزار "الف" با ۲۵۴ مورد بازیافت درست (۸۴/۷٪)، نرم افزار "ب" با ۲۳۳ مورد (۷۷/۷٪)، و نرم افزار "ج" با ۲۴۲ مورد (۸۰/۷٪) قابلیت بازیابی درست را دارا هستند.

جدول شماره ۳ - نسبت کلی بازیابی (درصد بازیابی) درست در هر یک از نرم افزارهای با توجه به سه شیوه نمایه سازی در فیلد پدیدآور

| نرم افزار | نتیجه بازیابی | | | |
|-----------|---------------|-------|-----------|-------|
| | پیدا شده | | پیدا نشده | |
| | درصد | تعداد | درصد | تعداد |
| الف | ۱۵/۳٪ | ۴۶ | ۸۴/۷٪ | ۲۵۴ |
| ب | ۲۲/۳٪ | ۶۷ | ۷۷/۷٪ | ۲۳۳ |
| ج | ۱۹/۳٪ | ۵۸ | ۸۰/۷٪ | ۲۴۲ |

بررسی نتایج بازیابی بر اساس شیوه های نمایه سازی در فیلد موضوع

در این مرحله مجموع قابلیت بازیابی (درصد بازیابی) هر یک از نرم افزارهای مورد تحقیق در سه مرحله قبل بر اساس شیوه نمایه سازی بصورت (کلمه، عبارت، فیلد کامل) در "فیلد موضوع" مورد نظر تحقیق است. نتایج بدست آمده نشان می دهد که از مجموع ۳۰۰ رکورد منابع مشخص و مشترک موجود بین سه نرم افزار مورد تحقیق، نرم افزار "الف" با ۲۵۹ مورد بازیافت درست (۸۶/۳٪)، نرم افزار "ب" با ۲۳۷ مورد (۷۹٪)، و نرم افزار "ج" با ۲۵۳ مورد (۸۴/۳٪) قابلیت بازیابی درست را دارا هستند.

جدول شماره ۴ - نسبت کلی بازیابی (درصد بازیابی) درست در هر یک از نرم افزارهای با

توجه به سه شیوه نمایه‌سازی در فیلد موضوع

| جمع | نتیجه بازیابی | | | | نرم‌افزار |
|-----|---------------|-------|-----------|-------|-----------|
| | پیدا شده | | پیدا نشده | | |
| | تعداد | درصد | تعداد | درصد | |
| ۳۰۰ | ۲۵۹ | ۸۶/۳٪ | ۴۱ | ۱۳/۷٪ | الف |
| ۳۰۰ | ۲۳۷ | ۷۹٪ | ۶۳ | ۲۱٪ | ب |
| ۳۰۰ | ۲۵۳ | ۸۴/۳٪ | ۴۷ | ۱۵/۷٪ | ج |

بررسی نتایج بازیابی بر اساس شیوه‌های نمایه‌سازی کلمه، عبارت و کل فیلد در فیلدهای عنوان، پدآور و موضوع

در این مرحله مجموع قابلیت بازیابی (درصد بازیابی) هر یک از نرم‌افزارهای مورد تحقیق در مراحل قبل (فیلدهای عنوان، پدآور و موضوع) و بر اساس شیوه نمایه‌سازی بصورت (کلمه، عبارت، کل فیلد) در نظر تحقیق است. از مجموع ۹۰۰ مورد جستجو، ۳۰۰ مورد به جستجو و بازیابی منابع مشخص و مشترک موجود بین سه نرم‌افزار مورد تحقیق بر اساس عنوان، ۳۰۰ مورد جستجو بر اساس نام پدآور و ۳۰۰ مورد جستجو بر اساس موضوع (ها) اختصاص یافت. نتایج بازیابی به صورت زیر به دست آمد: نرم‌افزار "الف" با ۷۷۵ مورد بازیافت درست (۸۶/۱٪)، نرم‌افزار "ب" با ۷۱۱ مورد (۷۹٪)، و نرم‌افزار "ج" با ۷۵۳ مورد (۸۳/۷٪) قابلیت بازیابی درست را دارا هستند.

جدول شماره ۵ - نسبت کلی بازیابی (درصد بازیابی) درست در هر یک از نرم‌افزارهای با توجه به شیوه‌های نمایه‌سازی به صورت کلمه، عبارت و فیلد کامل در فیلدهای عنوان، پدآور و موضوع

| جمع | نتیجه بازیابی | | | | نرم‌افزار |
|-----|---------------|-------|-----------|-------|-----------|
| | پیدا شده | | پیدا نشده | | |
| | تعداد | درصد | تعداد | درصد | |
| ۳۰۰ | ۷۷۵ | ۸۶/۱٪ | ۱۲۵ | ۱۳/۹٪ | الف |
| ۳۰۰ | ۷۱۱ | ۷۹٪ | ۱۸۹ | ۲۱٪ | ب |

| | | | | | | | |
|--|--|--|--|--|--|--|-----------------------------|
| | | | | | | | در سه نرم افزار مورد نظر |
|--|--|--|--|--|--|--|-----------------------------|

آزمون مجذور کای (X^2) میزان تفاوت قابلیت ارزیابی متأثر از شیوه‌های نمایه سازی (کلمه، عبارت، فیلد کامل) در فیلدهای عنوان، پدیدآور و موضوع در سه نرم افزار مورد بررسی را ۱۶/۵۹۵ با درجه آزادی $df=2$ نشان می دهد.

چون مجذور کای محاسبه شده ۱۶/۵۹۵ در مقایسه با مقدار ۹/۲۱ جدول مجذور کای (X^2) در سطح ۰/۰۱ و درجه آزادی ۲، بزرگتر است، بنابراین تفاوت موجود بین قابلیت ارزیابی متأثر از شیوه‌های نمایه سازی (کلمه، عبارت، کل فیلد) در فیلدهای سه گانه عنوان، پدیدآور و موضوع در سه نرم افزار مورد بررسی تأیید گردیده و در واقع فرض تحقیق یا H_1 تأیید شده است.

نتایج بدست آمده بالا نشان می دهد که بین قابلیت ارزیابی متأثر از شیوه‌های نمایه سازی در سه نرم افزار مورد تحقیق تفاوت معنی داری وجود دارد. در این قسمت به ترتیب نرم افزار "الف" رتبه اول، نرم افزار "ب" رتبه سوم و نرم افزار "ج" رتبه دوم را کسب نموده اند.

فرضیه ۱

بین قابلیت ارزیابی متأثر از شیوه‌های نمایه سازی فیلد عنوان در سه نرم افزار مورد تحقیق تفاوت معنی داری وجود دارد.

این فرضیه با استفاده از داده‌های فرم الف (فیلد عنوان) مورد ارزیابی قرار گرفت. از تعداد ۳۰۰ رکورد که بطور یکسان در سه نرم افزار مورد جستجو قرار گرفت نتایج آن به شرح زیر بدست آمد:

برای دستیابی به نتایج این فرضیه، امتیازهای کسب شده نرم افزارها برای ۳۰۰ عنوان رکورد مورد محاسبه قرار گرفت که در مجموع نرم افزار "الف" با ۲۶۲ رکورد قابلیت ارزیابی (۳/۸۷٪)، نرم افزار "ب" با ۲۴۱ رکورد قابلیت ارزیابی (۳/۸۰٪) و نرم افزار "ج" با ۲۵۸ رکورد قابلیت ارزیابی (۸۶٪) را کسب کرده اند. یافته‌های بدست آمده در این قسمت با استفاده از آزمون مجذور کای (X^2) مورد ارزیابی قرار گرفت که نتایج حاصله در جدول شماره ۷ آمده است:

آزمون مجذور کای (X^2) میزان تفاوت تعداد رکوردهای ارزیابی شده بر اساس

شیوه‌های نمایه‌سازی (کلمه، عبارت، فیلد کامل) در فیلد عنوان در سه نرم‌افزار مورد بررسی را ۶/۳۴۷ با درجه آزادی $df=2$ نشان می‌دهد.

مقایسه مجذور کای محاسبه شده با مقدار مجذور کای - ۵/۹۹ نشان دهنده آن است که بین قابلیت بازیابی متأثر از شیوه‌های نمایه‌سازی فیلد عنوان در سه نرم‌افزار مورد بررسی در سطح ۰/۰۵ تفاوت معنی داری وجود دارد. بنابراین فرضیه تحقیق تأیید می‌شود.

جدول شماره ۷ - وجود تفاوت معنی دار بین قابلیت بازیابی متأثر از شیوه‌های نمایه‌سازی (کلمه،

عبارت، فیلد کامل) در فیلد عنوان در سه نرم‌افزار مورد بررسی

| نتیجه | سطح معناداری | درجه آزادی | χ^2 | تعداد رکورد بازیابی درست نرم افزارها | | | تعداد رکورد مورد جستجو | آماره متغیر |
|-------------------------|--------------|------------|----------|--------------------------------------|-----|-----|------------------------|---|
| | | | | الف | ب | ج | | |
| تأیید فرضیه تحقیق H_1 | ۰/۰۴۲ | ۲ | ۶/۲۴۷ | ۲۵۸ | ۲۴۱ | ۲۶۲ | ۳۰۰ | قابلیت بازیابی متأثر از شیوه‌های نمایه‌سازی در فیلدهای عنوان، پدیدآور و موضوع در سه نرم‌افزار موردنظر |

در این قسمت با توجه به نتایج بدست آمده و محاسبات آماری به ترتیب نرم‌افزار "الف" رتبه اول، نرم‌افزار "ج" رتبه دوم و نرم‌افزار "ب" رتبه سوم را کسب نموده‌اند.

فرضیه ۲

بین قابلیت بازیابی متأثر از شیوه‌های نمایه‌سازی فیلد پدیدآور در سه نرم‌افزار مورد تحقیق تفاوت معنی داری وجود دارد.

این فرضیه با استفاده از داده‌های فرم ب (فیلد پدیدآور) مورد ارزیابی قرار گرفت. تعداد ۳۰۰ رکورد مشترک در سه نرم‌افزار مورد بررسی بر اساس شیوه‌های نمایه‌سازی فیلد پدیدآور مورد جستجو قرار گرفت نتایج زیر بدست آمد:

محاسبات نشان‌دهنده این است که در مجموع نرم‌افزار "الف" با ۲۵۴ رکورد ۸۴/۷٪

درصد قابلیت بازیابی، نرم افزار "ب" با ۲۳۳ رکورد ۷/۷۷٪ درصد قابلیت بازیابی و نرم افزار "ج" با ۲۴۲ رکورد ۷/۸۰٪ درصد قابلیت بازیابی را کسب کرده‌اند. یافته‌های بدست آمده در این قسمت با استفاده از آزمون مجذور کای (X^2) مورد ارزیابی قرار گرفت که نتایج بدست آمده در جدول شماره ۸ آمده است:

جدول شماره ۸ - وجود تفاوت معنی دار بین قابلیت بازیابی متأثر از شیوه‌های نمایه سازی فیلد

پدیدآور در سه نرم افزار مورد بررسی تحقیق

| متغیر | آماره | تعداد رکورد | | | X^2 | درجه آزادی | سطح معناداری | نتیجه |
|--|-------|-------------|------------------|---------|-------|------------|-------------------|-------|
| | | مورد جستجو | درست نرم افزارها | الف ب ج | | | | |
| قابلیت بازیابی متأثر از شیوه‌های نمایه سازی در فیلدهای عنوان، پدیدآور و موضوع در سه نرم افزار مورد نظر | ۳۰۰ | ۲۵۴ | ۲۳۳ | ۲۴۲ | ۲ | ۰/۹۰ | تأیید فرضیه H_0 | |

با توجه به مجذور کای محاسبه شده و مقایسه آن با جدول مجذور کای (X^2) که میزان آن در سطح ۰/۰۵ و با درجه آزادی $df=2$ ، برابر است با عدد ۵/۹۹، می توان بیان داشت که به دلیل کوچکتر بودن عدد مجذور کای (۴/۸۰۸) از عدد جدول (۵/۹۹)، بین قابلیت بازیابی متأثر از شیوه‌های نمایه سازی (کلمه، عبارت، فیلد کامل) در فیلد پدیدآور در سه نرم افزار مورد بررسی تفاوت معنی داری وجود ندارد. بنابراین فرضیه صفر یا H_0 تایید و فرضیه تحقیق رد شده است. با توجه به نتایج به دست آمده به ترتیب نرم افزار "الف" رتبه اول، نرم افزار "ج" رتبه دوم و نرم افزار "ب" رتبه سوم را کسب کرده‌اند.

بین قابلیت بازیابی متأثر از شیوه‌های نمایه‌سازی فیلد موضوع در سه نرم‌افزار مورد تحقیق تفاوت معنی‌داری وجود دارد.

این فرضیه با استفاده از داده‌های فرم ج (فیلد موضوع) مورد ارزیابی قرار گرفت. تعداد ۳۰۰ رکورد مشترک در سه نرم‌افزار مورد بررسی بر اساس سه شیوه نمایه‌سازی در فیلد موضوع مورد جستجو قرار گرفت و نتایج زیر بدست آمد:

نتایج حاصله نشان دهنده این است که در مجموع نرم‌افزار "الف" با ۲۵۹ رکورد قابلیت بازیابی (۰/۸۶/۳)، نرم‌افزار "ب" با ۲۳۷ رکورد قابلیت بازیابی (۰/۷۹) و نرم‌افزار "ج" با ۲۵۳ رکورد قابلیت بازیابی (۰/۸۳/۷) از رکوردها را دارند.

یافته‌های بدست آمده در این قسمت با استفاده از آزمون مجذور کای (X^2) مورد ارزیابی قرار گرفت که نتایج بدست آمده در جدول شماره ۹ ارائه شده است:

جدول شماره ۹ - وجود تفاوت معنی دار بین قابلیت بازیابی متأثر از شیوه‌های نمایه‌سازی از فیلد پدیدآور در سه نرم‌افزار مورد بررسی تحقیق .

| متغیر | آماره | تعداد رکورد مورد جستجو | تعداد رکورد بازیابی درست نرم افزارها | | | X^2 | درجه آزادی | سطح معناداری | نتیجه |
|--|-------|------------------------|--------------------------------------|-----|-------|-------|------------|-------------------|-------|
| | | | الف | ب | ج | | | | |
| فصلیت بازیابی متأثر از شیوه‌های نمایه‌سازی در فیلدهای عنوان، پدیدآور و موضوع در سه نرم‌افزار موردنظر | ۳۰۰ | ۲۵۹ | ۲۳۷ | ۲۵۳ | ۶/۱۷۵ | ۲ | ۰/۴۶ | تأیید فرضیه تحقیق | |

آزمون مجذور کای (X^2) میزان تفاوت شیوه‌های نمایه‌سازی (کلمه، عبارت، فیلد کامل) در فیلد موضوع و در سه نرم‌افزار مورد بررسی را ۶/۱۷۵ با درجه آزادی $df=2$ نشان می‌دهد. چون مجذور کای محاسبه شده در مقایسه با مقدار ۵/۹۹ موجود در جدول مجذور کای (X^2) در سطح ۰/۰۵ و با درجه آزادی ۲، بزرگتر است، بنابراین تفاوت موجود بین قابلیت

بازیابی متأثر از شیوه‌های نمایه سازی (کلمه، عبارت، کل فیلد) در فیلد موضوع در سه نرم افزار مورد بررسی تایید گردیده و در واقع فرض تحقیق یا H_1 تایید شده است. در این قسمت با توجه به نتایج به دست آمده و محاسبات آماری به ترتیب نرم افزار "الف" رتبه اول، نرم افزار "ج" رتبه دوم و نرم افزار "ب" رتبه سوم را کسب نموده‌اند.

نتیجه‌گیری

در مجموع با توجه به داده‌ها و تجزیه و تحلیلی که در ارزیابی نرم افزارهای مورد بررسی در این تحقیق انجام گرفته است، نتایج هر کدام از قسمتهای مورد ارزیابی به شرح زیر بیان می‌گردد:

تفاوت موجود بین شیوه‌های نمایه‌سازی و قابلیت بازیابی منابع مشخص و موجود در سه نرم افزار مورد تحقیق ناشی از مسایل و مشکلاتی است که در حوزه‌های نمایه‌سازی و برنامه‌نویسی، ورود اطلاعات و جستجو و بازیابی اطلاعات وجود دارد. برخی از مسایل و مشکلات حوزه‌های نمایه‌سازی و برنامه‌نویسی را می‌توان به صورت زیر برشمرد:

عدم رعایت قواعد و ویژگیهای خاص زبان فارسی نظیر حد کلمات، واژه‌های مرکب، شیوه املاء، جایگاه الفبایی واژه‌ها، آواها و نظایر آن، ارجاع خودکار نادرست در سرعنوانهای موضوعی فارسی، عدم تحلیل مناسب پرسش از سوی سیستم، عدم شیوه نمایه‌سازی به صورت زیر فیلد در برخی از نرم افزارها، کنترل واژگان به صورت ناقص. عدم اعلان خطا در ورود اطلاعات، عدم وجود واژگان مبنا و پذیرفته شده از سوی همگان.

مسایل و مشکلات حوزه‌های ورود اطلاعات و جستجو و بازیابی اطلاعات با توجه به دلایل زیر به نظر می‌رسد عمدتاً ناشی از اشتباه در وارد کردن حروف توسط کارکنان و کاربران باشد.

- چون در حین نمایه سازی فیلدهای سه گانه در نرم افزارها عمل کنترل واژگان (Vocabulary Control) به صورت کامل انجام نمی‌پذیرد و اگر واژه مورد نظر به شکلهای دیگری نیز نوشته شده باشد، طبعاً در هنگام جستجو بازیابی نمی‌شوند، در این حالت افت اطلاعاتی رخ داده و جامعیت را از دست می‌دهیم.

- در طراحی فیلد عنوان باید جایگاه هر یک از اجزاء تشکیل دهنده اطلاعات کامل عنوان از قبیل عنوان کامل، عنوان فرعی، عنوان برابر و عنوان موازی در پایگاه اطلاعاتی را به صورتی دقیق تعریف شود. همچنین با توجه به ویژگی خاص عنوان قراردادی، لازم است فیلدی

جداگانه برای آن در نظر گرفته شود.

عدم تفاوت موجود بین شیوه نمایه سازی (کلمه، عبارت، کل فیلد) در فیلد پدیدآور در سه نرم افزار مورد بررسی به مسایلی نظیر تفاوت در ساختار فیلدهای تعریف شده، تعریف نوع الگوریتم و پرسش در برخورد با ویرگول، فاصله و نظایر آن برمی گردد. البته با توجه به عدد محاسبه شده عدم تفاوت موجود جزئی بوده، ولی به هر حال می تواند به عنوان الگویی برای سایر فیلدها مدنظر قرار گیرد.

البته شایان ذکر است که ارزش گذاری و رتبه بندی در این پژوهش فقط حاصل بررسی قابلیت بازیابی شیوه های نمایه سازی نرم افزارهای مورد نظر بوده و به هیچ عنوان نباید آن را به کل نرم افزار تعمیم داد. با توجه به اینکه هر یک از نرم افزارهای مورد بررسی علاوه بر مسایل و مشکلات در حوزه بازیابی اطلاعات، مسایل و مشکلات خاص خود را دارند و سایر نرم افزارها آن نوع مشکل را ندارد و به گونه ای مشکل سایر نرم افزارها را رفع نموده اند، به این نتیجه می رسیم که کلیه انواع مسایل و مشکلات خاص هر یک از نرم افزارها در این حوزه نیز قابل رفع می باشد.

پیشنهادها

با توجه به نتایج و یافته های این تحقیق، برای بهینه سازی و ارتقاء نرم افزارهای مورد بررسی در زمینه شیوه های نمایه سازی فیلدها و مواردی که مستقیم یا غیرمستقیم با آن در ارتباط می باشد، پیشنهادها و توصیه هایی به شرح زیر ارائه می شود:

- استفاده از فیلد متغیر به جای فیلد ثابت: برای فیلدهایی که دقیقاً نمی توان حجم اطلاعاتی که باید در آن درج شود را از قبل پیش بینی نمود، باید هنگام طراحی و برنامه نویسی نوع این گونه فیلدها را فیلد متغیر در نظر گرفت.

- انتقال اطلاعات هر فیلدی در فیلد خاص خود در هنگام ورود داده ها (به صورت دستی یا خودکار) تا به این وسیله از مفقود شدن داده ها و در نتیجه بروز افت اطلاعات جلوگیری شود.

- بررسی و مطالعه بیشتر بر روی جستجوی عبارت در نرم افزارها و عدم استفاده از عملگر در این نوع جستجو

- استفاده از جستجو بر اساس زیرفیلد یا بخشی از فیلد که هم اکنون در برخی نرم افزارهای این قابلیت وجود ندارد.

- در کلیه فیلدها حروف بزرگ و کوچک یکسان تلقی شوند.
- امکان جستجوی مجدد در درون نتایج حاصل از جستجو شده قبلی فراهم گردد.
- استفاده از نمایه‌های آماده و تهیه شده برای افزایش دقت کاربرد و جلوگیری از اشتباهات تایپی و آگاهی از تعداد پیشینه‌هایی که بازیابی می‌شود. این نمایه لازم است نه فقط در یک فیلد خاص، بلکه در تمام فیلدها تدارک دیده شود.
- برای جستجوی کارآمدتر، نقاط دسترسی متعدد و متنوعی پیش‌بینی شود به نحوی که با در اختیار داشتن هر یک از اجزای اطلاعاتی رکوردها - حتی اجزاء اطلاعات فرعی - به جستجو و بازیابی آنها پرداخت.
- اعلان پیام خطا در هنگام ورود و جستجوی غلط داده‌ها
- تهیه یک واژگان مبنا و استفاده از آن در هنگام ورود داده‌ها به طور خودکار و هوشمند از سوی هر یک از نرم‌افزارهای موجود، به گونه‌ای که نرم‌افزار و نمایه مزبور در پیوند با هم بوده و بر اساس این پیوند و ارتباط، نرم‌افزار موردنظر شکل صحیح واژه‌ها را تشخیص داده و در جای خود درج نماید.
- چون عمل کنترل واژگان (Vocabulary Control) هنگام نمایه‌سازی فیلدهای سه‌گانه (عنوان، پدیدآور، موضوع) در نرم‌افزارها به صورت کامل انجام نمی‌پذیرد و اگر واژه مورد نظر به شکلهای دیگری نیز نوشته شده باشد، در هنگام جستجو بازیابی نمی‌شود، در این حالت افت اطلاعاتی رخ داده و جامعیت را از دست می‌دهد بنابراین لازم است این فرآیند کنترلی در حد بالایی در نظر گرفته شود.
- با توجه به ویژگیهای خاص زبان فارسی در مورد کلمات، واژه‌های مرکب، شیوه املاء، جایگاه الفبایی واژه‌ها، آواها، واژه‌های دخیل و نظایر آن، همکاری گروه‌های برنامه‌نویس، کتابداران و اطلاع‌رسانان و همچنین متخصصین زبان و ادبیات فارسی در جهت اتخاذ خط‌مشی واحد و دستیابی به واژگان و عبارتهای دقیق و یکسان، امری ضروری و بدیهی به نظر می‌رسد.
- ارتباط بیش از پیش کتابداران و اطلاع‌رسانان برنامه ساز با مهندسان برنامه نویس جهت بهینه سازی نرم افزارهای کتابخانه‌ای که نتیجه این همکاری رشد طبیعی و منطقی اینگونه نرم افزارها خواهد بود.

پیشنهاد برای پژوهشهای آینده

با توجه به اینکه جستجو و بازیابی اطلاعات در پایگاه‌های اطلاعاتی یکی از عوامل

اساسی در نرم افزارهای کتابخانه‌ای به حساب می‌آید و از طرفی عملکرد نرم افزارهای کتابخانه‌ای موجود فارسی با همه محاسن و توانمندی‌هایی که دارا بوده، بی‌نیاز از ارزشیابی در همه زمینه‌های خاص آن نخواهند بود، بنابراین توصیه می‌شود همان‌گونه که در نیم قرن اخیر در کشورهای غربی - به ویژه انگلستان و آمریکا - پژوهش‌های گسترده‌ای در حوزه بازیابی اطلاعات و ارزشیابی نرم افزارهای کتابخانه‌ای در این حوزه انجام گرفته و دستاوردهای چشمگیری را نیز به همراه داشته است، پژوهشگران کتابداری و اطلاع‌رسانی و برنامه‌نویسان رایانه‌ای در ایران نیز به انجام مطالعه و تحقیقات عمیق‌تر در این حوزه پرداخته، در جهت بهینه‌سازی نرم افزارها گام برداشته و مهمتر اینکه به خلق و ارائه نظریات نو و بدیع مبادرت ورزند.

منابع

- بابایی، محمود. بررسی ساختار و وضعیت ذخیره و بازیابی اطلاعات در پایگاههای اطلاعاتی: ۱- مرکز اطلاعات و مدارک علمی ایران. ۲ - سازمان پژوهشهای علمی و صنعتی ایران. پایان نامه کارشناسی ارشد (کتابداری و اطلاع‌رسانی)، استاد راهنما: دکتر عباس حری. تهران: دانشگاه تهران، دانشکده روانشناسی و علوم تربیتی، ۱۳۷۴.
- پاڤو، میراندالی. مفاهیم بازیابی اطلاعات. ترجمه اسدالله آزاد، رحمت الله فتاحی، مشهد: دانشگاه فردوسی مشهد، دانشکده علوم تربیتی و روانشناسی، ۱۳۷۹.
- پولیت، استون. نظامهای ذخیره و بازیابی اطلاعات: خاستگاه، توسعه و کاربردها. ترجمه محمد حسین دیانی، جعفر مهرداد. شیراز: کتابخانه منطقه‌ای علوم و تکنولوژی شیراز، ۱۳۸۰.
- چشمه سهرابی، مظفر. تاثیر استفاده از اصطلاحنامه در بانک‌های اطلاعاتی کتابشناختی بر میزان جامعیت، و مدت زمان جستجوی اطلاعات بازیابی شده. به راهنمایی: دکتر عباس حری، پایان نامه کارشناسی ارشد (کتابداری و اطلاع رسانی). تهران: دانشگاه تربیت مدرس، دانشکده علوم انسانی، ۱۳۷۸.

دیانی، محمد حسین. «حجم نمونه در پژوهشهای پیمایشی». کتابداری و اطلاع رسانی. جلد دوم، شماره ۳ (پاییز ۱۳۷۸): ص ۵۹-۷۰.

رضایی، محمد جواد. ارزشیابی قابلیت‌های ذخیره و بازیابی در برنامه‌های نرم‌افزاری مراکز اطلاع‌رسانی علوم قرآنی و معارف اسلامی (در شهرهای تهران، قم و اصفهان). استاد راهنما: ماندانا صدیق بهزادی. استاد مشاور: حبیب‌الله کشتکار ملکی. پایان‌نامه کارشناسی ارشد (کتابداری و اطلاع رسانی). دانشگاه تربیت مدرس، دانشکده علوم انسانی، ۱۳۷۷.

شاپوری، سودابه. بررسی مشکلات جستجوی موضوعی استفاده‌کنندگان در فهرست رایانه‌ای کتابخانه مرکزی دانشگاه فردوسی مشهد. پایان‌نامه کارشناسی ارشد. مشهد: دانشگاه فردوسی مشهد، دانشکده علوم تربیتی و روان‌شناسی، ۱۳۷۹.

فرج‌پهلوی، عبدالحسین: احمد زرگر. «تقابل فرهنگ سنتی فهرست‌نویسی با خدمات مبتنی بر نظامهای خودکار کتابخانه‌ای». در: فهرست‌های رایانه‌ای: کاربرد و توسعه: مجموعه مقالات همایش کاربرد و توسعه فهرست‌های رایانه‌ای در کتابخانه‌های ایران، ۲۷ و ۲۸ آبان ۱۳۷۸. مشهد: دانشگاه فردوسی مشهد دانشکده علوم تربیتی و روانشناسی، ۱۳۷۹.

قریب، زهرا. بررسی وضعیت نمایه سازی در کتابخانه‌ها و مراکز اطلاع رسانی مستقر در شهر تهران. استاد راهنما: عباس حری. استاد مشاور: نورالله مرادی. پایان‌نامه کارشناسی ارشد (کتابداری و اطلاع رسانی)، دانشگاه آزاد اسلامی، ۱۳۷۷.

موسوی، افشین. بررسی تطبیقی کارایی نرم افزارهای نوسا، کاوش و پارس آذرخش از دیدگاه فهرست‌نویسی. پایان‌نامه کارشناسی ارشد، تهران: دانشگاه آزاد اسلامی، واحد شمال، دانشکده علوم انسانی، ۱۳۷۹.

یوسفی، احمد. «ریزش کاذب در نرم افزارهای کتابخانه‌ای نوسا، پارس آذرخش و کاوش». در: فهرست‌های رایانه‌ای: کاربرد و توسعه: مجموعه مقالات همایش کاربرد و توسعه فهرست‌های رایانه‌ای در کتابخانه‌های ایران، ۲۷ و ۲۸ آبان ۱۳۷۸. مشهد: دانشگاه فردوسی مشهد، دانشکده

علوم تربیتی و روانشناسی، ۱۳۷۹.

American Society of Indexers. **Indexing Evaluation Checklist**. Accessed 8 September 2000}. Available from:

<http://www.asindexing.org/checklist.html>

Borgman, hrisine L. «Why are Online Cataloges still Hard to Use?». **Journal of the American Society for Information Science (JASIS)**. Vol. 47.N.7.(1996):P496.

Lancaster, F.W. .**The Measurement and Evaluation of Linary Services**. Washington D.C: Information Resources Press, 1977.

Lancaster, F.W. "MEDLARS: Report on the Evaluation of Its Operating Efficiency." **American Documentation**, 20(2)(1969): 119-142.

Mulvany, Nancy C. «Software Tools for Indexing Revisited». **The Indexer**. Vol.21, No.4 (October 1999): p160-163.

Report on the Testing and Analysis of an Investigation into the Comparative Efficiency of Indexing Systems". Aslib - Cranfield Report, Cranfield, England, October 1962.

Salton, G; and M.J.McGill. **Introduction to Modern Information retrieval**. New York, McGraw-Hill, 1983.

Hood, William; Concepcion S. Wilson. **An Analysis of the Indexing Used in the LISA Database**. New South Wales: University of New South Wales, School of Information, Library and Archive Studies, 1992.