

چگونگی نمایه‌سازی در پایگاه‌های اطلاعاتی مراکز اطلاع‌رسانی تهران

سیدحسین میرجلیلی^۱

چکیده: پژوهش حاضر به بررسی چگونگی سیاست‌گذاری و به‌کارگیری زبان نمایه‌سازی در ۳۲ پایگاه اطلاعاتی طراحی شده توسط ۱۳ مرکز اطلاع‌رسانی شهر تهران پرداخته است. روش پژوهش پیمایشی است و اطلاعات مورد نیاز از طریق پرسشنامه، مصاحبه، مشاهده، و کار بر روی تک‌تک پایگاه‌های اطلاعاتی گردآوری شده است. یافته‌های تحقیق نشان می‌دهد از مجموع پایگاه‌های مورد بررسی، ۱۹ پایگاه اطلاعاتی از زبان کنترل شده، ۷ پایگاه از زبان طبیعی، و ۶ پایگاه نیز از ترکیبی از زبان طبیعی و کنترل شده استفاده می‌کنند. نمایه‌سازی ۴ پایگاه اطلاعاتی، رایانه‌ای و نمایه‌سازی ۲۸ پایگاه به شیوه دستی است. همچنین در هیچ یک از مراکز مورد بررسی تلاشی برای نمایه‌سازی رایانه‌ای با به‌کارگیری زبان کنترل شده صورت نگرفته است. نتایج حاصل از کار بر روی پایگاه‌ها حاکی از آن است که در اکثریت پایگاه‌های اطلاعاتی که از زبان کنترل شده استفاده می‌کنند ناهماهنگی‌های آشکاری در گزینش توصیفگرها به چشم می‌خورد که نشان از عدم دقت و توجه نمایه‌سازان و عدم پیروی از یک سیاست واحد دارد.

کلیدواژه‌ها: نمایه‌سازی، سیاست‌گذاری، پایگاه اطلاعاتی، مراکز اطلاع‌رسانی، تهران

مقدمه

نمایه‌سازی به عنوان عملی برای توصیف موضوع و محتوای مدارک در قالب واژه‌ها و به منظور جست‌وجوی مدارک و به دست آوردن اطلاعات از جای جای متون توسعه یافت. شیوه‌ها و فنون متفاوتی برای شناسایی و توصیف محتوای مدارک به وجود آمد و ماشین‌ها برای دسته‌بندی، سازماندهی، و پردازش اطلاعات به کار گرفته شدند. رشته جدیدی نیز برای مطالعه فنون و نظریات نمایه‌سازی ایجاد گردید، و ملاک‌هایی برای ارزیابی و کارآیی نمایه‌ها تدوین شد (۱۲: ۲۹۴).

بعد از جنگ جهانی دوم با اشباع شدن کتابخانه‌ها و مراکز تحقیقاتی از گزارش‌های علمی و فنی، روش مطلوبی برای دستیابی موضوعی به اطلاعات دقیق به منظور رفع نیازهای اطلاعاتی محققان وجود نداشت. افزایش ناگهانی اطلاعات همراه با کاهش زمان مجاز برای پردازش اطلاعات و اشباع آن، گرایش به سمت منابعی غیر از کتاب نظیر اسناد، مدارک و مقالات نشریات را به وجود آورد. عدم کارآیی روش‌های سنتی در سازماندهی این منابع سبب تمرکز تلاش‌ها بر روش‌های جدید سازماندهی اطلاعات شد. بنابراین

زبان طبیعی در محیط‌های رایانه‌ای رایج و متداول است.

در حال حاضر از روش‌های گوناگونی برای ذخیره و بازیابی اطلاعات استفاده می‌شود. بعضی از این روش‌ها ساده و برخی پیچیده‌اند. تعدادی از زبان مفید و واژه‌های کنترل شده برای نمایه‌سازی بهره می‌گیرند و برخی از زبان طبیعی مدرک استفاده می‌کنند. برخی افراد با استفاده از رایانه و با سرعت زیاد و صرف وقت کم این کار را انجام می‌دهند و برخی دیگر با روش‌های دستی به ذخیره و بازیابی اطلاعات می‌پردازند. بعضی شدیداً به مهارت‌های نمایه‌ساز در مرحله درونداد تکیه دارند در حالی که گروهی دیگر مهارت‌های تجسسی بیشتری را طلب می‌کنند. درباره امتیازهای نسبی زبان طبیعی و کنترل شده بحث‌های متفاوتی وجود دارد. یک دیدگاه نمایه‌سازی کنترل شده را تنها راه صحیح نمایه‌سازی می‌دانند در حالی که دیگران، ترجیح می‌دهند از فرصت‌هایی که به وسیله رایانه ارائه می‌شود به‌طور کامل استفاده کنند، حتی اگر نمایه حاصله یک نمایه کامل به حساب نیاید (۱۵۶:۶).

با توجه به نقش غیرقابل انکار پایگاه‌های اطلاعاتی در اطلاع‌رسانی و جایگاهی که زبان اطلاعات دارد و همچنین با عنایت به رشد پایگاه‌های اطلاعاتی داخلی و تلاش برای ایجاد پایگاه‌های اطلاعاتی در مراکز اطلاع‌رسانی، در این پژوهش سعی شده است، چگونگی سیاست‌گذاری و به‌کارگیری زبان نمایه‌ای در پایگاه‌های اطلاعاتی که توسط مراکز اطلاع‌رسانی دولتی شهر تهران طراحی و ایجاد شده‌اند، مورد بررسی قرار گیرد.

تعریف مسئله

یکی از عوامل اساسی که هنگام طراحی و ایجاد پایگاه‌های اطلاعاتی باید مد نظر قرار گیرد توجه به زبان

یکی از مفاهیمی که به هنگام نمایه‌سازی و بازیابی به شکل دستی یا ماشینی باید مورد توجه قرار گیرد و در واقع بخشی از سیاست‌های اصلی نمایه‌سازی به حساب می‌آید، زبان نمایه‌سازی است. زبان نمایه‌سازی زبانی است که برای توصیف موضوع یا سایر جنبه‌های اطلاعاتی مدارک در به کار می‌رود (۹۲:۶). به عبارت دیگر منظور از زبان نمایه مجموعه کامل اصطلاحات به کار رفته در نمایه‌سازی برای توصیف محتوای مدارک است (۸۵:۷). زبان نمایه‌سازی یا زبان نمایه‌ای برای ایجاد هم‌زمانی بیشتر میان تولیدکننده و استفاده‌کننده اطلاعات نقش تعیین‌کننده‌ای دارد و با همین زبان است که موضوع و محتوای مدارک بیان می‌شود.

زبان‌های نمایه‌سازی را براساس معیارهای متفاوت به‌گونه‌های مختلف دسته‌بندی کرده‌اند، ولی به‌طور کلی می‌توان گفت این زبان‌ها بر دو نوعند:

۱. زبان‌های مفید یا کنترل شده: در این زبان اصطلاحات از پیش برای نمایه‌سازی مدارک انتخاب شده‌اند. مانند طرح‌های رده‌بندی، فهرست‌های قراردادی، سرعنوان‌های موضوعی، و اصطلاحنامه‌ها (۶۱:۴). در این زبان نمایه‌ساز ابتدا واژه‌ها را انتخاب می‌کند و سپس مطابق با آن اصطلاحاتی را به مدارک اختصاص می‌دهند. زبان کنترل شده معمولاً در نظام‌هایی که واژه موضوعی به وسیله انسان تخصیص می‌یابد به کار می‌رود، در این شیوه مقدر است سیاهه واژه‌هایی را که باید به وسیله رایانه تخصیص یابد (در صورت وجود در یک مدرک) مشخص کنند (۹۲:۹).

۲. زبان طبیعی و آزاد: در این شیوه نمایه‌ساز آزادانه دست به انتخاب اصطلاحات می‌زند و در استفاده از هر اصطلاحی که صلاح بداند و به ذهنش خطور کند، مجاز است. اصطلاحات ممکن است در متن مدرک بوده یا منتخب نمایه‌ساز باشد (۱۵:۱). واژه‌ها ممکن است به وسیله انسان یا ماشین تخصیص یابد. نمایه‌سازی به

دقت در پیروی از یک رویهٔ واحد با توجه به سیاست‌های در نظر گرفته شده است که با کار بر روی پایگاه‌های اطلاعاتی و با در نظر گرفتن سیاست‌های هر پایگاه انجام می‌شود.

پایگاه‌های اطلاعاتی: پایگاه مجموعه‌ای مرتبط به هم و نظام یافته از داده‌های گردآوری شده در حوزه‌های موضوعی مشخص است که پوشش فراسازمانی داشته، در چارچوب یک نظام مدیریت پایگاه‌های اطلاعاتی ذخیره شده، و بازایی آنها نیز امکان‌پذیر باشد. در این تحقیق پایگاه‌های دارای اطلاعات کتابشناختی و موجود در کشور بررسی شده‌اند.

مراکز اطلاع‌رسانی: این مراکز شامل مراکز اطلاع‌رسانی دولتی است که خدمات اطلاع‌رسانی فراسازمانی را ارائه می‌دهند و تولیدکنندهٔ همه یا تعدادی از پایگاه‌های خود هستند.

پرسش‌های اساسی

۱. نمایه‌سازی پایگاه‌های اطلاعاتی داخلی مراکز اطلاع‌رسانی شهر تهران براساس کدام یک از دو زبان طبیعی و کنترل شده انجام می‌شود؟
۲. در پایگاه‌هایی که از نمایه‌سازی رایانه‌ای استفاده می‌کنند، کدام یک از زبان‌های نمایه‌سازی کاربرد دارد؟
۳. پایگاه‌هایی که زبان طبیعی را مورد استفاده قرار می‌دهند، از چه سیاست‌هایی پیروی می‌کنند؟
۴. نمایه‌سازی با زبان کنترل شده در پایگاه‌های اطلاعاتی براساس چه سیاست‌هایی صورت می‌گیرد؟
۵. سیاست نمایه‌سازی پایگاه‌های اطلاعاتی از لحاظ هم‌ارزی چیست؟
۶. سیاست‌های پیش‌بینی شده در استفاده از زبان نمایه‌ای در هر یک از پایگاه‌های اطلاعاتی تا چه حد در عمل به کار گرفته شده است؟

نمایه‌سازی آنهاست. به‌کارگیری هر یک از زبان‌های کنترل شده و یا طبیعی در نتایج بازایی اطلاعات تأثیر مستقیم دارد. در سال‌های اخیر پایگاه‌های اطلاعاتی متعددی در کتابخانه‌ها و مراکز اطلاع‌رسانی ایران به‌خصوص در شهر تهران طراحی و ایجاد شده‌اند. هر مرکز با توجه به اهداف، امکانات و دانش کارکنان خود شیوه‌هایی را برای نمایه‌سازی مدارک و اطلاعات پایگاه‌های اطلاعاتی خود به‌کار گرفته‌اند، لیکن اطلاعات مدونی در مورد سیاست‌ها، کیفیت، و عملکرد آنها در دسترس نیست. نظر به نقش و اهمیت زبان نمایه‌ای در بازایی پایگاه‌های اطلاعاتی، در این تحقیق، سیاست‌ها و خط‌مشی‌های طراحان و دست‌اندرکاران نمایه‌سازی پایگاه‌های اطلاعاتی داخلی مراکز اطلاع‌رسانی دولتی شهر تهران مورد بررسی قرار می‌گیرد تا معین شود زبان نمایه‌سازی هر پایگاه اطلاعاتی چیست و چه سیاست‌هایی برای استفاده از هر زبان نمایه‌ای در نظر گرفته شده است. همچنین رایانه چه جایگاهی در نمایه‌سازی هر پایگاه دارد و با کار بر روی هر پایگاه اطلاعاتی بررسی می‌شود که پایگاه‌های اطلاعاتی در به‌کارگیری زبان نمایه‌ای تا چه اندازه از سیاست‌های یکسان و رویهٔ واحدی پیروی کرده‌اند.

تعریف‌های به‌کار رفته در پژوهش

زبان نمایه‌ای: منظور از زبان نمایه‌ای مجموعهٔ کامل اصطلاحات به‌کار رفته در نمایه‌سازی برای توصیف محتوای مدارک است. در این تحقیق بر زبان‌های نمایه‌سازی الفبایی تأکید، و رده‌بندی به‌طور کامل نادیده گرفته شده است.

سیاست‌گذاری: این اصطلاح به مجموعهٔ تدابیر و خط‌مشی‌هایی اطلاق می‌شود که برای استفاده از زبان طبیعی یا آزاد و زبان کنترل شده در نظر گرفته شده است. به‌کارگیری زبان نمایه‌ای عبارت از تعیین میزان

هدف و فایده پژوهش

هدف از این تحقیق بررسی سیاست‌ها و خط‌مشی‌های استفاده از زبان نمایه‌سازی در پایگاه‌های اطلاعاتی داخلی مراکز اطلاع‌رسانی شهر تهران و بررسی میزان دقت در به‌کارگیری این سیاست‌ها در پایگاه‌های اطلاعاتی هر مرکز است.

فایده این تحقیق آن است که علاوه بر آن که زبان نمایه‌سازی پایگاه‌های اطلاعاتی و سیاست‌های حاکم بر آنها و وضعیت نمایه‌سازی در مراکز اطلاع‌رسانی شهر تهران مشخص می‌شود می‌تواند مورد استفاده مراکز اطلاع‌رسانی که به نحوی درگیر تولید پایگاه‌ها هستند، قرار گیرد تا به سابقه و تجربه مراکز اطلاع‌رسانی در نمایه‌سازی پایگاه‌های اطلاعاتی پی ببرند.

روش پژوهش و جامعه تحقیق

در این پژوهش، روش تحقیق پیمایشی است. اطلاعات مورد نظر از طریق پرسشنامه، مشاهده، مصاحبه، و کار بر روی تک‌تک پایگاه‌ها و مطالعه اسناد و مدارک جمع‌آوری شده است. جامعه آماری شامل پایگاه‌های اطلاعاتی داخلی فراسازمانی مراکز اطلاع‌رسانی دولتی شهر تهران و نمایه‌سازان یا مدیران پایگاه‌های اطلاعاتی این مراکز است. جمعاً ۳۲ پایگاه اطلاعاتی از ۱۳ مرکز اطلاع‌رسانی در این تحقیق مورد بررسی قرار گرفته‌اند که مشخصات آنها در جدول ۱ آمده است.

پیشینه پژوهش

معرفزاده (۱۳۶۶) در پایان‌نامه خود که به روش پیمایشی انجام گرفته، به بررسی وضعیت ابزارهای بازبایی اطلاعات پزشکی به منظور تهیه نمایه پزشکی پرداخته است. در این تحقیق انواع نمایه‌های عنوان، مورد بررسی قرار گرفته و نمایه‌سازی کوئیک^۱ و کوک^۲

بر روی مدارک جمع‌آوری شده اجرا شده است. نتایج نشان می‌دهد که به وضوح لازم است که کلیدواژه‌هایی از متن به عنوان مدارک فارسی افزوده شود. در پایان، نمایه‌سازی گردان نوع افزوده به عنوان، برای تهیه ابزارهای بازبایی اطلاعات پزشکی تحت عنوان "نمایه پزشکی ایران" پیشنهاد شده است (۸). بابایی (۱۳۷۴) یا استفاده از روش پیمایشی به بررسی ساختار و وضعیت ذخیره و بازبایی اطلاعات در پایگاه‌های اطلاعاتی مرکز اطلاعات و مدارک علمی و سازمان پژوهش‌های علمی و صنعتی پرداخته است. نتایج تحقیق حاکی از آن است که پایگاه‌های اطلاعاتی مرکز اطلاعات و مدارک، علی‌رغم موفقیت نسبی در گردآوری و اشاعه اطلاعات، در نظام نمایه‌سازی ناهماهنگی‌هایی دارد. نبودن توصیفگرهای انتخابی، استفاده از زبان نمایه‌سازی ترکیبی طبیعی و کنترل شده، عدم کنترل دقیق مترادف‌ها، مفرد و جمع، از جمله این ناهماهنگی‌هاست. این امر ضرورت ویرایش توصیفگرها را ایجاد می‌کند (۲). خسرویانی (۱۳۷۴) به بررسی چگونگی سیاست‌گذاری و نمایه‌سازی پایگاه‌های اطلاعاتی معاجم لفظی و موضوعی مرکز

تحقیقات رایانه‌ای علوم اسلامی به روش پیمایشی، پرداخته است. نتایج تحقیق نشان می‌دهد عمده‌ترین مشکل پایین بودن ضریب دقت در اطلاعات بازبایی شده است (۵). چشمه سهرابی (۱۳۷۸) به بررسی تأثیر استفاده از اصطلاحنامه بر میزان جامعیت و مانعیت و مدت زمان جست‌وجوی اطلاعات پرداخته است. در این تحقیق هر کلیدواژه یک‌بار از طریق اصطلاحنامه و بار دیگر با استفاده از زبان طبیعی مورد کاوش قرار گرفته است. نتایج تحقیق بیان می‌کند که استفاده از

1. Kwic: Keyword in context

2. Kwoc: Keyword out of context

جدول ۱. مشخصات پایگاه‌های اطلاعاتی مورد بررسی

ردیف	پایگاه اطلاعاتی	مرکز اطلاع‌رسانی	تاریخ ایجاد	تعداد رکورد	نوع اطلاعات	نرم افزار
۱	اطلاعات دریای خزر	مرکز اطلاعات و مدارک علمی ایران	۱۳۷۰	۱۰۰۳۷	عنوان	CDS-ISIS
۲	چکیده پایان‌نامه‌های فارغ‌التحصیلان ایرانی خارج از کشور	مرکز اطلاعات و مدارک علمی ایران	۱۳۷۰	۱۱۲۱۸	چکیده	CDS-ISIS
۳	چکیده پایان‌نامه‌های ایران (دکتری و کارشناسی ارشد)	مرکز اطلاعات و مدارک علمی ایران	۱۳۷۲	۶۶۴۷۸	چکیده	CDS-ISIS
۴	چکیده گزارش‌های دولتی	مرکز اطلاعات و مدارک علمی ایران	۱۳۷۲	۱۴۸۷۰	چکیده	CDS-ISIS
۵	چکیده صح‌های پژوهشی	مرکز اطلاعات و مدارک علمی ایران	۱۳۶۸	۴۹۴۱۸	چکیده	CDS-ISIS
۶	عرب‌پژوهانه‌های همایش‌ها، اجلاس‌های ایران	مرکز اطلاعات و مدارک علمی ایران	۱۳۷۲	۴۳۴۹۲	عنوان	CDS-ISIS
۷	چکیده مقالات علمی و فنی	مرکز اطلاعات و مدارک علمی ایران	۱۳۶۹	۵۰۰۸۰	چکیده	CDS-ISIS
۸	اطلاعات علوم تربیتی	مرکز اطلاعات و مدارک علمی ایران	۱۳۷۰	۹۵۰۰	عنوان	CDS-ISIS
۹	میراث و معماری	سازمان میراث فرهنگی، مرکز اطلاع‌رسانی	۱۳۷۴	۵۲۱۲	چکیده	CDS-ISIS
۱۰	بایگ‌های علمی	سازمان میراث فرهنگی، مرکز اطلاع‌رسانی	۱۳۷۶	۴۰۸	چکیده	CDS-ISIS
۱۱	میراث‌شناسی	سازمان میراث فرهنگی، مرکز اطلاع‌رسانی	۱۳۷۶	۸۰۰	چکیده	CDS-ISIS
۱۲	پایگاه اطلاعات آرشین ملی	مرکز اسناد ملی ایران، مرکز اطلاع‌رسانی اسنادی	۱۳۷۰	۱۰۹۹	عنوان	unix
۱۳	پایگاه اطلاعات مدیریت	مرکز مدیریت دولتی، مرکز توسعه اطلاعات و ارتباطات مدیریت	۱۳۷۵	۳۲۲۹	متن	Exar
۱۴	مقالات و مدارک بازرگانی و اقتصاد	مرکز پژوهش‌های بازرگانی	۱۳۷۳	۹۸۰۰	عنوان	پایس آذرخش
۱۵	پایگاه اطلاعات منابع	مرکز اطلاع‌رسانی منابع	۱۳۷۳	۵۵۰۰۰	متن	بهن برنا
۱۶	پایگاه اطلاعات مقالات فارسی	سازمان مدیریت و برنامه‌ریزی کشور، مرکز اطلاع‌رسانی	۱۳۶۸	۶۰۰۰۰	عنوان	فاکس پرو
۱۷	پایگاه اطلاعات مقالات غیرفارسی	سازمان مدیریت و برنامه‌ریزی کشور، مرکز اطلاع‌رسانی	۱۳۶۵	۹۳۰۰۰	عنوان	فاکس پرو
۱۸	صح‌های تحقیقاتی	سازمان مدیریت و برنامه‌ریزی کشور، مرکز اطلاع‌رسانی	۱۳۷۹	۴۸۰۰	چکیده	CDS-ISIS
۱۹	پایگاه اطلاعات کنفرانس‌های کشاورزی ایران	مرکز مطالعات برنامه‌ریزی و اقتصاد کشاورزی	۱۳۷۱	۴۴۴۰	عنوان	CDS-ISIS
۲۰	گزارش‌های کشاورزی	مرکز مطالعات برنامه‌ریزی و اقتصاد کشاورزی	۱۳۷۲	۴۵۰۰	عنوان	پایس آذرخش
۲۱	اسناد تاریخی	مؤسسه مطالعات تاریخ معاصر ایران	۱۳۷۲	۲۵۰۰۰	عنوان	نیسا
۲۲	سند چاپ شده در کتاب‌ها و مجلات فارسی	مؤسسه مطالعات تاریخ معاصر ایران	۱۳۷۲	۱۷۰۰۰	عنوان	نیسا
۲۳	اسناد تاریخی	مؤسسه مطالعات تاریخ معاصر ایران	۱۳۷۲	۲۵۰۰۰	عنوان	نیسا
۲۴	مقالات تاریخی	مؤسسه مطالعات تاریخ معاصر ایران	۱۳۷۶	۸۰۰۰	عنوان	نیسا
۲۵	مقالات فرهنگی فارسی	کتابخانه ملی جمهوری اسلامی ایران	۱۳۷۷	۱۵۸۰۰۰	عنوان	پایس آذرخش
۲۶	اعلامه‌ها	کتابخانه ملی جمهوری اسلامی ایران	۱۳۷۸	۶۶۰۰	عنوان	پایس آذرخش
۲۷	حکمت، روان‌پزشکی، فلسفه، ادبیات، تاریخ و علوم اسلامی	کتابخانه ملی جمهوری اسلامی ایران	۱۳۷۹	۷۷۶	چکیده	پایس آذرخش
۲۸	سجده‌های حصص	کتابخانه مجلس شورای اسلامی	۱۳۷۱	۲۷۱۱۳	عنوان	پایس آذرخش
۲۹	اسناد تاریخی مجلس	کتابخانه مجلس شورای اسلامی	۱۳۷۱	۱۷۶۶۲	عنوان	پایس آذرخش
۳۰	گزارش‌های فرهنگی	سازمان میراث و ارتباطات اسلامی	۱۳۷۴	۱۳۰۰۰	متن	حویا
۳۱	پایگاه اطلاعات نیرو افراسی	کتابخانه و مرکز اطلاع‌رسانی وزارت نیرو	۱۳۷۵	۸۰۰۰	عنوان	نیسا
۳۲	پایگاه اطلاعات نیرو افراسی	کتابخانه و مرکز اطلاع‌رسانی وزارت نیرو	۱۳۷۵	۲۱۹۵	عنوان	نیسا

موضوعی آزمایش کند. یکی از دو بانک مورد بررسی، از زبان کنترل شده و دیگری از زبان طبیعی استفاده می‌کند. نتایج تحقیق نشان داد که بانک اطلاعاتی که از زبان کنترل شده در نمایه‌سازی استفاده می‌کرد عملکرد و کارایی بهتری نسبت به بانک اطلاعاتی کنترل نشده دارد ولی بانک‌ها به صورت معناداری مکمل یکدیگرند (۱۰: ۲۷-۵۴).

تجزیه و تحلیل یافته‌ها

جدول ۱ مشخصات پایگاه‌های اطلاعاتی جامعه مورد پژوهش را نشان می‌دهد. تحلیل یافته‌های این جدول حاکی از آن است که به جز پایگاه اطلاعات مقالات غیرفارسی که در سال ۱۳۶۵ شکل گرفته، سایر پایگاه‌ها حاصل تلاش‌هایی بوده که از سال ۱۳۶۸ آغاز شده است. از مجموع پایگاه‌های مورد بررسی ۱۹ پایگاه حاوی عنوان، ۱۰ پایگاه حاوی چکیده، و ۳ پایگاه اطلاعاتی حاوی متن کامل مدارک هستند. همچنین نرم‌افزار ۱۳ پایگاه، CDS-ISIS ۷ پایگاه، پارس آدرخش و ۶ پایگاه، نوساست و نرم‌افزار ۶ پایگاه نیز توسط سازمان‌های آنها طراحی شده است.

از مجموع ۳۲ پایگاه اطلاعاتی مورد بررسی، نمایه‌سازی ۱۹ پایگاه با زبان کنترل شده، ۷ پایگاه با زبان طبیعی، و ۶ پایگاه براساس ترکیبی از زبان طبیعی و کنترل شده انجام می‌شود که پایگاه‌هایی که از زبان ترکیبی استفاده می‌کنند همگی وابسته به مرکز اطلاعات و مدارک هستند. نمایه‌سازی ۴ پایگاه اطلاعاتی مقالات غیرفارسی، طرح‌های تحقیقاتی، مقالات تاریخی، و اسناد تاریخی مجلس ماشینی و زبان نمایه‌ای آنها نیز طبیعی بوده و هیچ سیاست خاصی نیز برای آنها اعمال نشده است. نمایه‌سازی سایر پایگاه‌ها به شیوه دستی

اصطلاحنامه باعث کاهش میزان جامعیت و افزایش مانعیت جست‌وجو می‌شود و استفاده از اصطلاحنامه مدت زمان جست‌وجوی اطلاعات را کاهش می‌دهد (۳).

در خارج از کشور رگازی^۱ (۱۹۷۸) به تشریح چگونگی به‌کارگیری زبان نمایه‌سازی در طراحی یک نظام ماشینی خوان می‌پردازد. وی اشاره می‌کند که در آغاز زبان طبیعی به دلیل هزینه کمتر و اجرای سریع‌تر آن انتخاب گردید ولی به تدریج کاهش هزینه درونداد در زبان طبیعی با افزایش هزینه‌های جست‌وجو مزیت خود را از دست داد و کاوش‌گران متوجه نامطلوب بودن استفاده از زبان طبیعی به علت ابهام و عدم کنترل مترادف‌ها و دشواری انتقال راهبرد کاوش بر روی نظام شدند. بنابراین تصمیم گرفته شد یک نمایه کنترل شده ایجاد شود و به اجرا درآید (۱۱). ولیچ (۱۹۹۴) سیاست نمایه‌سازان، زبان نمایه‌ای، عمق نمایه‌سازی، و شکل فیزیکی نمایه‌های پایگاه‌های اطلاعاتی را مورد بررسی قرار داده است. وی نتیجه می‌گیرد که بانک‌های اطلاعاتی کتابشناختی بزرگ به‌طور کامل نمایه‌سازی خودکار را مورد استفاده قرار نمی‌دهند و ترجیح می‌دهند به جای استفاده از نظام‌های خودکار، از نمایه‌سازی بهره بگیرند و از نظام خودکار فقط در برخی از حوزه‌ها، به‌طور مثال برای کنترل واژگان، کمک می‌گیرند (۱۴: ۶۲۰-۶۲۷). راولی^۲ (۱۹۹۴) به بازنگری بحث در مورد زبان نمایه‌سازی کنترل شده و طبیعی در دوره‌های مختلف می‌پردازد و نتیجه می‌گیرد که بازیابی اطلاعات در عمل مستلزم ترکیبی از زبان‌های طبیعی و کنترل شده است که هم اکنون نیز برای کاوش طیف وسیعی از انواع مختلف بانک‌های اطلاعاتی مورد استفاده قرار می‌گیرد (۱۳: ۱۰۸-۱۱۹). ریبیرو^۳ (۱۹۹۶) تحقیقی بر روی آرشیوهای شهر پورتو در پرتغال انجام داد تا ارزش نسبی واژگان کنترل شده را در نمایه‌سازی

1. Regazzi 2. Rowley 3. Reberro

بهره می‌گیرند. همگی پایگاه‌هایی که دست به کنترل واژگان می‌زنند بر کنترل شکل‌های مختلف اصطلاحات نظیر مفرد و جمع، ریشه‌های متفاوت یک اصطلاح، املاهای متفاوت، اختصارها و آغازه‌ها، شکل‌های فارسی و غیرفارسی، کنترل متشابهات، کنترل اسامی اشخاص، مناطق جغرافیایی و سازمان‌ها و... تأکید کرده‌اند و جز در موارد محدود، معیارهایی برای انتخاب اصطلاح مرجح از نام‌رجح بیان شده است که با دستورالعمل اصطلاحنامه‌نویسی نیز مطابقت دارد. ولی نتایج حاصل از کار بر روی تک‌تک پایگاه‌ها و مطابقت نظرات نمایه‌سازان با وضعیت موجود پایگاه نشان می‌دهد که آنچه نمایه‌سازان پایگاه‌ها برای انتخاب توصیفگرها بیان کرده‌اند در بسیاری از پایگاه‌ها در عمل به کار گرفته نشده است.

در این میان پایگاه اطلاعات کتابشناختی کشاورزی ایران و چکیده پایان‌نامه‌های کتابداری به علت تخصص طراحان آن و پایگاه اطلاعات آرشیو ملی و اسناد تاریخی مجلس به علت تهیه فهرست واژه‌های پذیرفته شده مرکز و ضرورت استفاده از این واژه‌ها، از وضعیت مناسبی برخوردارند. ولی در سایر پایگاه‌ها بازیابی شکل‌های مختلف یک اصطلاح منجر به بازیابی تعداد زیادی رکورد می‌شود که این نشان می‌دهد که عملاً سیاست واحدی در انتخاب کلیدواژه‌ها حاکم نیست. در این پایگاه‌ها بازیابی با املاهای متفاوت یک اصطلاح، شکل مفرد و جمع، ریشه‌های متفاوت یک اصطلاح، اصطلاح خارجی و معادل فارسی آن، و... منجر به بازیابی تعداد زیادی رکورد می‌شود، این مسئله به‌خصوص در پایگاه‌های اطلاعاتی مرکز اطلاعات و مدارک علمی به علت حجم کار، تعدد نمایه‌سازان، گستردگی موضوعات، و عدم استفاده از در میان پایگاه‌های مورد بررسی در پایگاه اطلاعات نمایه به علت انتخاب توصیفگرهای عام و کلی و ارجاع

انجام شده و سپس اطلاعات آنها به پایگاه منتقل می‌شود. علاوه بر چهار پایگاه فوق‌الذکر که از زبان طبیعی استفاده می‌کنند، زبان نمایه‌ای پایگاه اطلاعات مدیریت، اسناد چاپ شده در کتاب‌ها، مجلات فارسی، و لغات اسناد تاریخی نیز طبیعی است و نمایه‌سازان به‌صورت مکانیکی واژه‌هایی را از مدارک استخراج و به مدارک منتسب می‌کنند. این در حالی است که پایگاه اطلاعات مدیریت به‌صورت تمام متن است و شیوه قلمی نمایه‌سازی تأثیری در بازیابی آن ندارد و نرم‌افزار آن نیز قابلیت جست‌وجوهای ترکیبی با به‌کارگیری منطق بول را داراست.

یافته‌های حاصل از جدول ۲. نشان می‌دهد که سیاست همارایی پایگاه‌های اطلاعاتی مرمیت و معماری، باستان‌شناسی، نمایه، و چکیده پایان‌نامه‌های کتابداری و اطلاع‌رسانی از نوع پیش همارایی ساختگی و شیوه همارایی سایر پایگاه‌های اطلاعاتی پسرهماراست. یعنی ترکیب اصطلاحات در زمان جست‌وجو و توسط جست‌وجوگران انجام می‌گیرد. پژوهشگاه علوم انسانی و اطلاع‌رسانی

پایگاه‌های اطلاعاتی که از زبان کنترل شده یا ترکیبی از زبان کنترل شده و طبیعی استفاده می‌کنند برای گزینش و استاندارد کردن توصیفگرها از سیاست‌ها و خط‌مشی‌های متفاوتی پیروی می‌کنند. پایگاه‌هایی نظیر مقالات و مدارک بازرگانی و اقتصاد، پایگاه اطلاعات مقالات فارسی، گزارش‌های کشاورزی، چکیده پایان‌نامه‌های کتابداری و اطلاع‌رسانی، و نسخه‌های خطی، ملاک گزینش توصیفگرها را بر سرعنوان‌های موضوعی و سیاست‌های حاکم بر این منبع قرار داده‌اند. پایگاه‌هایی نظیر آرشیو ملی و اسناد تاریخی براساس فهرست واژه‌های پذیرفته شده در سازمان عمل می‌کنند. پایگاه‌های مرکز اطلاعات و مدارک علمی به علت چند موضوعی بودن از انواع مختلف منابع و ابزارها به‌خصوص فرهنگ‌های دو زبانه

جدول ۲. زبان نمایه‌ای و شیوه نمایه‌سازی پایگاه‌های اطلاعاتی

پایگاه اطلاعاتی	زبان نمایه‌ای			شیوه نمایه‌سازی		همارایی	
	طبیعی	کنترل شده	ترکیبی از صیغی و کنترل شده	دستی	ماشینی		
اطلاعات دریای حزر			x	x		x	
چکیده پایان‌نامه‌های خارج از کشور		x		x		x	
چکیده پایان‌نامه‌های ایران			x	x		x	
چکیده گزارش‌های دولتی			x	x		x	
چکیده طرح‌های پژوهشی			x	x		x	
عناوین مقالات همایش‌ها و اجلاس‌های ملی ایران			x	x		x	
چکیده مقالات علمی و فنی			x	x		x	
اطلاعات علوم تربیتی		x		x		x	
مرمت و معماری		x		x		x	
باستان‌شناسی		x		x		x	
مردم‌شناسی		x		x		x	
پایگاه اطلاعات آرشبو ملی		x		x		x	
پایگاه اطلاعات مدیریت			x	x		x	
مقالات و مدارک بازرگانی و اقتصاد		x		x		x	
پایگاه اطلاعات نمایه			x	x		x	
پایگاه اطلاعات مقالات فارسی			x	x		x	
پایگاه اطلاعات مقالات غیرفارسی	x				x		
طرح‌های تحقیقاتی	x				x		
پایگاه اطلاعات کنایشناختی کشاورزی ایران		x		x		x	
گزارش‌های کشاورزی		x		x		x	
اسناد تاریخی		x		x		x	
اسناد چاپ شده در کتاب‌ها و مجلات فارسی		x		x		x	
لغات اسناد تاریخی		x		x		x	
مقالات تاریخی		x		x		x	
مقالات فرهنگی فارسی		x		x		x	
اعلامیه‌ها		x		x		x	
چکیده پایان‌نامه‌های کنش‌اندازی و اطلاع‌رسانی		x		x		x	
نسخه‌های خطی		x		x		x	
اسناد تاریخی مجلس	x				x		
گزارش‌های فرهنگی		x		x		x	
پایگاه اطلاعات نیرو (فارسی)		x		x		x	
پایگاه اطلاعات نیرو (لاتین)		x		x		x	
	۷	۱۹	۶	۲۸	۴	۲۷	۵

دسترسی به اطلاعات از طریق نمایه موضوعی امکان‌پذیر نیست.

اصطلاحنامه‌های تخصصی بیشتر به چشم می‌خورد. موضوع‌های مختلف به توصیفگرهای عام،

نتیجه‌گیری و پیشنهادها

بررسی حاضر بر روی زبان نمایه‌ای پایگاه‌های اطلاعاتی شهر تهران صورت گرفته است و نشان می‌دهد که اکثریت پایگاه‌های اطلاعاتی از زبان کنترل شده برای نمایه‌سازی استفاده می‌کنند و نمایه‌سازی با زبان طبیعی در پایگاه‌های اطلاعاتی دارای سیاست و ضابطه خاصی نیست، یا ماشینی است و انسان در آن نقشی ندارد و یا اینکه بدون هیچ دخل و تصرف و کنترلی صرفاً واژه‌هایی از متن استخراج و به مدارک منتسب می‌شود. به‌طور کلی نمایه‌سازی در پایگاه‌های اطلاعاتی از وضعیت مناسبی برخوردار نیست. اگرچه سیاست‌ها و خط‌مشی‌های آنها در نمایه‌سازی به جز پایگاه اطلاعات مدیریت اسناد چاپ شده در کتاب‌ها و مجلات فارسی و پایگاه اطلاعات نمایه، مناسب و قابل قبول به نظر می‌رسد ولی در به‌کارگیری این سیاست‌ها هیچ کنترل و دقتی صورت نگرفته است. وضعیت نمایه‌سازی پایگاه‌ها تا حدود زیادی به تخصص نمایه‌سازان آن بستگی دارد و ادامه وضعیت این پایگاه‌ها به شکل کنونی، در دراز مدت، استحکام نمایه پایگاه‌ها را به خطر می‌اندازد. نتایج مبین آن است که در اکثریت مراکز، منابع و ابزارهای لازم برای کنترل واژگان وجود ندارد به همین دلیل سلیقه و نظرات نمایه‌سازان نقشی اساسی دارد. لذا ضروری است مراکز اطلاع‌رسانی به تهیه منابع و اصطلاحنامه‌هایی که قسمتی از پوشش موضوعی پایگاه‌ها را دربر می‌گیرند اقدام کنند و نمایه‌سازان و مدیران همه پایگاه‌های اطلاعاتی نیز باید با تدوین یک برنامه منسجم به ویرایش توصیفگرهای پایگاه‌های خود اقدام کنند.

در میان مراکز اطلاع‌رسانی مورد بررسی، مرکز اطلاعات و مدارک علمی به عنوان یک مرکز ملی با سطح خدمات وسیع، باید برای یکتواخت کردن و کنترل

توصیفگرهای پایگاه‌های خود اقدام کند زیرا ادامه روند فعلی استحکام نمایه پایگاه‌ها را به خطر می‌اندازد. مرکز مدیریت دولتی با توجه به قابلیت نرم‌افزاری پایگاه اطلاعات مدیریت باید از ادامه نمایه‌سازی به زبان طبیعی، بدون کنترل، و به شیوه دستی با توجه به تمام متن بودن پایگاه خودداری کند و یا با تدوین سیاست اصولی به کنترل توصیفگرها بپردازد.

مؤسسه مطالعات تاریخ معاصر نیز باید از ادامه نمایه‌سازی بدون کنترل و بی ضابطه پایگاه اطلاعات اسناد چاپ شده در کتاب‌ها و مجلات فارسی اجتناب نماید و به جای صرف هزینه زیاد به نمایه‌سازی دقیق اطلاعات و مدارک پایگاه بپردازد.

با توجه به کثرت پایگاه‌های اطلاعاتی طراحی شده در مراکز اطلاع‌رسانی شهر تهران و با توجه به این که اکثریت نمایه‌سازان فعلی پایگاه‌ها فاقد تخصص و آموزش کافی هستند پیشنهاد می‌شود کتابخانه ملی به عنوان متولی، با ایجاد دوره‌های کوتاه مدت نمایه‌سازی و اعلام آن به مراکز اطلاع‌رسانی، به آموزش نمایه‌سازان همت گمارد.

تاریخ دریافت: ۸۰/۵/۴

مآخذ

- اصطلاحنامه فرهنگ فارسی اصفا، تدوین شورای عالی اصطلاحنامه، قریب‌حسروی، [و دیگران] تهران: سازمان مدارک انقلاب اسلامی، ۱۳۷۴
- بابایی، محمد. بررسی ساختار و وضعیت ذخیره و بازیابی اطلاعات در پایگاه‌های اطلاعاتی ۱ مرکز اطلاعات و مدارک علمی؛ ۲ سازمان پژوهش‌های علمی و صنعتی ایران؛ پایادنامه کارشناسی ارشد کتابداری و اطلاع‌رسانی، دانشکده علوم تربیتی، دانشگاه تهران، ۱۳۷۴
- جسمه سهراسی، مطلق. تأثیر استفاده از اصطلاحنامه در بانک‌های اطلاعاتی کتابساختی بر میزان جامعیت و مانعیت و مدت زمان جست‌وجوی اطلاعات بازیابی شده؛ پایان‌نامه کارشناسی ارشد کتابداری و اطلاع‌رسانی،

- and for an Indexing of Policy Using Controlled Language. *Journal of the Society of Archivists*. Vol.17, No.1 (1996): 27-54.
11. Regazzi, John J. "Indexing and Online Searching of Multimedia- Purpose Textual Database: Conflict Confluence" presented at the Annual Meeting of the Special Libraries Association, KANSAS, Missouri (June 1978).
12. Rothman, John "Index, Indexer, Indexing". *Encyclopedia of Library and Information Science*. Vol. II. (1974): 286-294.
13. Rowley, J. "The Controlled Versus Natural Indexing Language Debate Revisited: a Perspective on Information Retrieval Practice and Research". *Journal of Information Science*. vol. 20, Issue2 (1994): 108-119.
14. Wellisch, H.H. "Book and Periodical indexing. *Journal of the American Society for Information Science*. Vol.45, Issue 8 (September 1994): 620- 627.
- دانشکده علوم انسانی، دانشگاه تربیت مدرس، ۱۳۷۸.
۴. حریری، مهرانگیر "اصطلاحنامه در دحیره و بازیابی اطلاعات و اطلاع‌رسانی" اطلاع‌رسانی، ۱۱ (۱۳۶۱): ۶۰-۷۶
۵. حسروبیانی، جعفر. بررسی وضعیت پایگاه‌های اطلاعاتی مهاجم لفظی و مفروضه‌ی محرک تحقیقات کامپیوتری علوم اسلامی. پایان‌نامه کارشناسی ارشد کتابداری و اطلاع‌رسانی، دانشکده روانشناسی و علوم تربیتی، دانشگاه تهران، ۱۳۷۴
۶. راوی، حسین. *نمایه‌سازی و چکیده‌نویسی*. ترجمه جعفر مهرداد تهران سازمان مدارک فرهنگی انقلاب اسلامی، ۱۳۷۴
۷. لکستر، فردریک ویلفرد *نظام‌های بازیابی اطلاعات، ویژگی‌ها، آزمون و ارزیابی*. ترجمه جعفر مهرداد، شیراز: نوید، ۱۳۷۹.
۸. معرف‌زاده، عبدالحمید "بررسی وضعیت ابزار بازیابی اطلاعات پزشکی و پیشنهادی برای تهیه نمایه پزشکی" پایان‌نامه کارشناسی ارشد کتابداری و اطلاع‌رسانی، دانشکده مدیریت و اطلاع‌رسانی، دانشگاه علوم پزشکی ایران، ۱۳۶۶.
۹. وسمان، هرمان. *سیستم‌ها و خدمات اطلاع‌رسانی رایانه*. ترجمه جعفر مهرداد. شیراز: نوید، ۱۳۷۹
10. Rebeiro, F. "Subject indexing and authority control in archives: the Need for Subject Indexing in Archives: