

جامعیت و مانعیت پایگاه‌های اطلاعاتی فارسی در بازیابی اسناد حیطة مصدومیت‌های شغلی

علی بهاری^۱، همایون صادقی بازرگانی^۲، فهیمه بختیاری^۳

مقاله پژوهشی

چکیده

مقدمه: استفاده از پایگاه‌های اطلاعاتی به منظور دستیابی به اسناد الکترونیکی، ضرورت بررسی عملکرد و پوشش پایگاه‌ها را دوچندان می‌سازد. هدف از انجام پژوهش حاضر، تعیین جامعیت و مانعیت پایگاه‌های اطلاعاتی فارسی در بازیابی اسناد مربوط به حیطة مصدومیت‌های شغلی بود.

روش بررسی: این مطالعه به روش توصیفی و با استفاده از مشاهده و مقایسه انجام شد. جامعه تحقیق شامل پایگاه‌های اطلاعاتی ملی Magiran، SafeLir، Irandoc و پایگاه اطلاعات علمی جهاد دانشگاهی (SID Scientific Information Database) و حجم نمونه شامل تعداد اسناد بازیابی شده طی ۲۰ سال اخیر از این پایگاه‌ها بود. کلید واژه‌ها بر اساس نظرات اعضای خبره تیم پژوهش و بررسی اسناد مربوط تعیین شد. در نهایت، ۴۰ کلید واژه انتخاب گردید و پایگاه‌های مذکور مورد جستجو قرار گرفت. تعداد اسناد بازیابی شده ثبت و شاخص‌ها و نمودارهای مرتبط با جامعیت و مانعیت تهیه شد.

یافته‌ها: تعداد اسناد بازیابی شده از پایگاه‌های SafeLir، SID، Magiran و Irandoc به ترتیب ۴۱۱، ۱۵۳، ۵۲۳ و ۱۴۹ مدرک بود. پایگاه SafeLir جامعیت نسبی بالاتر (۸۳/۷۰ درصد) و پایگاه SID مانعیت بالاتری (۸۹/۵۴ درصد) در بازیابی اسناد مربوط، نسبت به سایر پایگاه‌ها داشت. بیشترین همپوشانی بین پایگاه‌های SafeLir و Magiran مشاهده گردید (۲۶۹ سند).

نتیجه‌گیری: نتایج حاصل شده به محققان حیطة مصدومیت‌ها و رشته‌های مرتبط در شناسایی پایگاه مناسب (دارای جامعیت و مانعیت مناسب) جهت بازیابی اسناد در حوزه تخصصی آن‌ها کمک می‌کند. چنانچه جامعیت، مانعیت و یا هم مانعیت و هم جامعیت جستجو مهم باشد، می‌توان به ترتیب از پایگاه‌های SafeLir، SID و SafeLir استفاده نمود.

واژه‌های کلیدی: مصدومیت‌های شغلی؛ بهداشت حرفه‌ای؛ پایگاه داده؛ ذخیره و بازیابی اطلاعات

پیام کلیدی: نتایج مطالعه حاضر بیان کرد که جامعیت و مانعیت پایگاه‌های اطلاعاتی فارسی متفاوت هستند. بنابراین، برای انجام یک جستجوی مناسب، نباید فقط از یک پایگاه استفاده کرد. در حیطة مصدومیت‌ها اگر جامعیت و مانعیت یا هر دو ملاک انتخاب پایگاه قرار گیرد، می‌توان از یافته‌های این مطالعه بهره‌مند شد.

دریافت مقاله: ۱۳۹۸/۶/۸

پذیرش مقاله: ۱۳۹۸/۹/۳

تاریخ انتشار: ۱۳۹۸/۹/۱۵

ارجاع: بهاری علی، صادقی بازرگانی همایون، بختیاری فهیمه. جامعیت و مانعیت پایگاه‌های اطلاعاتی فارسی در بازیابی اسناد حیطة مصدومیت‌های شغلی. مدیریت اطلاعات سلامت ۱۳۹۸؛ ۱۶ (۵): ۲۳۶-۲۴۲

مقدمه

در چند دهه اخیر، بازیابی اطلاعات مورد توجه خاصی قرار گرفته است که علت آن را می‌توان ناشی از تغییر در کمیت و کیفیت اطلاعات، افزایش اهمیت زمان و تغییر در اهمیت منابع اطلاعاتی دانست. این عوامل باعث تغییر و توسعه در سیستم‌های بازیابی اطلاعات به خصوص در حیطة‌های مختلف علوم پزشکی شده است (۱).

در مطالعاتی که به منظور بررسی عملکرد ابزارهای جستجو انجام شده است، از تعداد نتایج بازیابی شده، به عنوان نمودی از قدرت بازیابی استفاده می‌شود (۲). معیار غالب جهت بررسی سیستم‌های بازیابی اطلاعات، ربط (Relevance) می‌باشد که با استفاده از دو ملاک جامعیت (Recall) و مانعیت (Precision) بررسی می‌شود (۳). مانعیت عبارت است از صحت یک جستجو و زمانی کامل می‌شود که هر نتیجه بازیابی شده مربوط به یک سؤال، حتماً به آن سؤال مرتبط باشد و جامعیت نیز مقیاسی از اثربخشی در بازیابی تمام اطلاعات مورد جستجو در یک پایگاه اطلاعاتی است (۴). در این زمینه، نتایج مطالعه صادقی و اخوتی که با هدف بررسی جامعیت و مانعیت پایگاه‌های اطلاعاتی

ایرانی در حوزه داروسازی انجام شد، نشان داد که اهمیت هر یک از این دو شاخص، تعیین‌کننده انتخاب پایگاه اطلاعاتی برای انجام تحقیقات می‌باشد (۵). همچنین، دمرچی‌لو و حاجی زین‌العابدینی با بررسی جامعیت نسبی پایگاه‌های اطلاعاتی ملی در حوزه دندان‌پزشکی، گزارش کردند که بیشترین مقدار این شاخص متعلق به پایگاه IranMedex است (۶).

مقاله حاصل تحقیق مستقل بدون حمایت مالی و سازمانی است.

۱- کارشناس ارشد، اپیدمیولوژی، مرکز تحقیقات پیشگیری از آسیب حوادث جاده‌ای و گروه آمار و اپیدمیولوژی، دانشکده بهداشت، دانشگاه علوم پزشکی تبریز، تبریز، ایران (نویسنده طرف مکاتبه)

Email: alibahari2012@gmail.com

۲- دانشیار، اپیدمیولوژی، مرکز تحقیقات پیشگیری از آسیب حوادث جاده‌ای و گروه آمار و اپیدمیولوژی، دانشکده بهداشت، دانشگاه علوم پزشکی تبریز، تبریز، ایران

۳- کارشناس ارشد، کتابداری و اطلاع‌رسانی پزشکی، مرکز تحقیقات علوم اعصاب، دانشگاه علوم پزشکی تبریز، تبریز، ایران

یک کلید واژه را ایجاد می‌کرد و جستجو می‌شد. بنابراین، ۴۰ کلید واژه به عنوان کلید واژه‌های اصلی در مطالعه حاضر مورد استفاده قرار گرفت. به طور مثال: مصدومیت محیط کار.

بازه زمانی ۲۰ سال اخیر برای جستجو در نظر گرفته شد. جستجو در هر چهار پایگاه مذکور به صورت جداگانه اما در یک روز انجام و اسناد بازیابی شده در یک فایل Endnote ذخیره گردید؛ بدین صورت که هر ترکیب واژگان در یکی از پایگاه‌ها جستجو می‌شد و بلافاصله همان ترکیب در پایگاه‌های دیگر مورد جستجو قرار می‌گرفت و پس از این که تعداد اسناد یادداشت می‌شد، با استفاده از Endnote ابتدا اسناد تکراری مشخص و حذف می‌شد و سپس با مقایسه فایل‌های Endnote هر یک از پایگاه‌ها، در صورتی که اسناد مشترکی بین پایگاه‌ها وجود داشت، مشخص می‌گردید. اعضای بهداشت حرفه‌ای تیم با استفاده از فایل Endnote اسناد، مرتبط بودن آن‌ها را بر اساس عناوین اسناد بررسی و یادداشت می‌کردند. تمام جستجوها در فیلد «عنوان» اسناد پایگاه‌ها انجام گرفت. بنابراین، جستجو آزاد نبود و علت آن نیز وجود تفاوت در ساختار و مکانیسم جستجوی پایگاه‌ها بود. با انجام جستجو در عناوین اسناد، اسناد بازیابی شده از پایگاه‌ها قابلیت مقایسه بیشتری را داشتند. به همین دلیل جستجو در عناوین اسناد صورت گرفت.

به منظور محاسبه جامعیت نسبی پایگاه‌ها از روش مذکور در تحقیق Shafi و Rather (۱۲) استفاده گردید. در این روش جامعیت نسبی به صورت رابطه ۱ تعریف می‌گردد.

رابطه ۱

$$\text{جامعیت نسبی} = \frac{\text{تعداد اسناد مرتبطی که توسط یک پایگاه‌های مورد مطالعه بازیابی شده}}{\text{تعداد کل اسناد مرتبطی که توسط تمامی پایگاه‌های مورد مطالعه بازیابی شده}}$$

اگر چهار پایگاه مورد نظر با A، B، C و D نشان داده شود، در این صورت تعداد اسناد بازیابی شده آن‌ها A₁، B₁، C₁ و D₁ خواهد بود. بر اساس روش فوق، دو حالت برای محاسبه جامعیت نسبی وجود دارد. حالت اول: اگر بین پایگاه‌ها همپوشانی وجود نداشته باشد، جامعیت نسبی حاصل تقسیم A₁ بر A₁ + B₁ + C₁ + D₁ خواهد بود و حالت دوم: اگر بین پایگاه‌ها همپوشانی وجود داشته باشد و به عنوان مثال همپوشانی پایگاه A و دیگر پایگاه‌ها مانند B به صورت B₂ نامگذاری شود، جامعیت نسبی پایگاه A برابر با A₁ خواهد بود و در صورت کسر و A₂ + B₂ + D₂ + A₁ در مخرج کسر قرار می‌گیرد. مانعیت هر یک از پایگاه‌های اطلاعاتی نیز با تقسیم تعداد اسناد مرتبط بازیابی شده هر پایگاه به کل اسناد بازیابی شده همان پایگاه محاسبه گردید.

یافته‌ها

تعداد اسناد بازیابی شده و تعداد اسناد مرتبط بازیابی شده برای هر یک از ۴۰ کلید واژه اصلی در هر یک از چهار پایگاه اطلاعاتی مورد نظر ثبت شد. تعداد اسناد بازیابی شده، در جدول ۱ آورده شده است. بر اساس یافته‌ها، بین تمام پایگاه‌های اطلاعاتی مورد بررسی، همپوشانی مشاهده گردید. تعداد اسناد بازیابی شده مشترک بین پایگاه‌های اطلاعاتی ثبت گردید (جدول ۲).

وجود اسناد بازیابی شده مشترک بین تمامی پایگاه‌های اطلاعاتی مورد بررسی باعث شد تا جامعیت نسبی آن‌ها با استفاده از حالت دوم روش مذکور در مطالعه Shafi و Rather (۱۲) بررسی گردد.

از طرف دیگر، اغلب مجلات ایرانی، فارسی هستند و بیشتر آن‌ها به دلایل مختلفی (مانند تحریم، زبان مجله و...) در پایگاه‌های اطلاعاتی بزرگ دنیا نمایه نمی‌شوند (۷). بنابراین، این اطلاعات از طریق پایگاه‌های مذکور قابل دسترسی نیست. زمانی که برای جستجوی منابع آنلاین صرف می‌گردد و همچنین، نتایج مرتبطی که از آن‌ها به دست می‌آید، همواره حایز اهمیت می‌باشد (۹، ۸). پژوهشگران ایرانی اغلب از یک پایگاه اطلاعاتی برای انجام تحقیقات استفاده می‌کنند (۱۰) و مهارت جستجوی اطلاعات در بیشتر آن‌ها (دانشجویان و افراد غیر متخصص) پایین است (۱۱). کاربران باید با مزایای هر یک از پایگاه‌های اطلاعاتی آشنا شوند تا بتوانند پایگاه متناسب با نیاز خود را انتخاب کنند. بنابراین، ارزیابی کارایی پایگاه‌هایی که به نمایه‌سازی مجلات این حوزه می‌پردازند، ضروری به نظر می‌رسد و از آن‌جایی که جامعیت و مانعیت دو معیار مهم برای ارزیابی نظام‌های بازیابی اطلاعات می‌باشد، پژوهش حاضر به بررسی جامعیت و مانعیت پایگاه‌های اطلاعاتی فارسی در حیطه مصدومیت‌های شغلی پرداخت.

روش بررسی

این مطالعه به روش توصیفی و با استفاده از مشاهده و مقایسه انجام گردید. جامعه تحقیق متشکل از پایگاه‌های اطلاعاتی ملی Magiran، SafeLir، IranDoc و پایگاه اطلاعات علمی جهاد دانشگاهی SID (Scientific Information Database) و حجم نمونه شامل تعداد اسناد بازیابی شده از این پایگاه‌ها بود. بنابراین، کل جامعه مقالات به عنوان نمونه انتخاب شد. تیمی متشکل از متخصصان اپیدمیولوژیست، کتابداری و اطلاع‌رسانی پزشکی و بهداشت حرفه‌ای ایجاد گردید. بر اساس ارزیابی‌های اولیه، مشخص شد که بازیابی اسناد فارسی در حوزه علوم پزشکی کشور از طریق پایگاه‌های SafeLir، ParsMedline، SID، IranDoc و Magiran صورت می‌گیرد که به دلیل مشکل فنی و عدم دسترسی به پایگاه ParsMedline و IranMedex، به ناچار این پایگاه‌ها از مطالعه کنار گذاشته شدند.

برای انتخاب کلید واژه‌ها، ابتدا از کلید واژه‌های اصطلاحنامه پزشکی فارسی استفاده شد، اما مشاهده گردید که اصطلاحات پذیرفته شده قابل توجهی در بین پژوهشگران حیطه مصدومیت‌های شغلی رایج است که در اصطلاحنامه پزشکی فارسی وجود ندارد و در بسیاری از موارد در صورتی که از واژگان اصطلاحنامه پزشکی فارسی استفاده گردد، هیچ نتیجه‌ای از پایگاه‌های SID و IranDoc بازیابی نمی‌شود. از طرف دیگر، به دلیل این که در نمایه‌سازی هر چهار پایگاه اطلاعاتی مذکور از اصطلاحنامه پزشکی فارسی و یا زبان کنترل شده استفاده نمی‌شد، تعداد اسناد بازیابی شده با استفاده از کلید واژه‌های رایج در بین پژوهشگران حیطه مذکور (که با بررسی متون مربوط به دست آمده بودند)، بیشتر از حالتی بود که از واژگان منتخب اصطلاحنامه پزشکی فارسی استفاده می‌شد. در نهایت، تصمیم بر این شد که با استفاده از عملگرهای بولی (Boolean operators) AND (یعنی «و») و OR (یعنی «یا»)، واژگان مورد نظر به صورت ترکیبی از دو گروه واژه باشند. گروه اول واژگان شامل «آسیب، مصدومیت، جراحت، صدمات و زخم» و گروه دوم واژگان شامل «شغلی، محل کار، محیط کار، ناشی از کار، کارکنان، کارگران، پرسنل و شاغلین» بود. بدین صورت که واژگان داخل هر گروه با یکدیگر «یا» و با واژگان گروه دیگر «و» می‌شدند؛ به این معنی که هر واژه از یک گروه با هر واژه از گروه دیگر با هم

جدول ۱: تعداد اسناد بازایی شده و اسناد مرتبط بازایی شده برای چهار پایگاه اطلاعاتی مورد بررسی به تفکیک کلید واژه‌ها

کلید واژه	تعداد اسناد بازایی شده از پایگاه SID		تعداد اسناد بازایی شده از پایگاه Magiran		تعداد اسناد بازایی شده از پایگاه SafeLir		تعداد اسناد بازایی شده از پایگاه Irandoc	
	کل اسناد	اسناد مرتبط	کل اسناد	اسناد مرتبط	کل اسناد	اسناد مرتبط	کل اسناد	اسناد مرتبط
آسیب شغلی	۲۸	۲۶	۵۱	۴۴	۳۹	۳۳	۱۳	۱۱
مصدومیت شغلی	۰	۰	۲	۲	۹	۷	۰	۰
جراحت شغلی	۰	۰	۰	۰	۳	۳	۰	۰
صدمات شغلی	۳	۳	۸	۷	۱۶	۱۵	۱	۱
زخم شغلی	۶	۶	۲۶	۱۹	۱۷	۱۵	۰	۰
آسیب محل کار	۰	۰	۱۷	۱۴	۱۹	۱۷	۸	۳
مصدومیت محل کار	۰	۰	۰	۰	۹	۷	۱	۱
جراحت محل کار	۳	۳	۲	۲	۱۴	۱۲	۷	۵
صدمات محل کار	۱۶	۱۲	۱۲	۱۰	۱۸	۱۳	۳	۳
زخم محل کار	۰	۰	۵۹	۴۱	۱۱	۱۰	۱۷	۳
آسیب پرسنل	۵	۴	۷	۵	۱۹	۱۶	۴	۲
مصدومیت پرسنل	۶	۵	۲	۲	۱۶	۱۶	۱	۱
جراحت پرسنل	۱	۱	۱	۱	۸	۶	۳	۲
صدمات پرسنل	۱	۱	۴	۳	۷	۵	۳	۱
زخم پرسنل	۰	۰	۱۷	۱۲	۱۱	۸	۲	۱
آسیب محیط کار	۸	۸	۱۲	۱۰	۲۴	۱۶	۱۲	۴
مصدومیت محیط کار	۰	۰	۰	۰	۲۴	۲۰	۸	۲
جراحت محیط کار	۲	۲	۲	۲	۲۰	۱۴	۰	۰
صدمات محیط کار	۴	۴	۴	۴	۲۰	۱۸	۲	۰
زخم محیط کار	۱۰	۶	۴۰	۲۸	۱۷	۱۲	۰	۰
آسیب کارگران	۵	۶	۸	۸	۱۱	۱۰	۲	۲
مصدومیت کارگران	۳	۳	۳	۳	۱۶	۹	۱	۱
جراحت کارگران	۱	۱	۲	۱	۹	۸	۵	۳
صدمات کارگران	۰	۰	۱	۰	۱۰	۷	۵	۲
زخم کارگران	۲	۲	۱۳	۴	۸	۵	۳	۱
آسیب ناشی از کار	۱۷	۱۶	۱۲	۱۰	۲۸	۲۴	۹	۷
مصدومیت ناشی از کار	۱	۱	۳	۲	۱۳	۹	۵	۳
جراحت ناشی از کار	۰	۰	۲	۰	۷	۳	۷	۵
صدمات ناشی از کار	۳	۳	۷	۴	۱۲	۸	۱	۰
زخم ناشی از کار	۰	۰	۱۰	۵	۱۱	۴	۳	۱
آسیب کارکنان	۲۱	۱۹	۴۷	۲۳	۱۴	۷	۱۲	۱۰
مصدومیت کارکنان	۱	۱	۱	۱	۷	۵	۳	۲
جراحت کارکنان	۱	۰	۱	۱	۳	۱	۰	۰
صدمات کارکنان	۲	۲	۸	۳	۱۵	۱۴	۳	۱
زخم کارکنان	۰	۰	۱۷	۳	۱۵	۵	۰	۰
آسیب شاغلین	۲	۲	۳	۱	۹	۹	۳	۲
مصدومیت شاغلین	۰	۰	۰	۰	۷	۳	۰	۰
جراحت شاغلین	۰	۰	۱	۰	۰	۰	۱	۰
صدمات شاغلین	۱	۰	۶	۶	۲	۲	۱	۱
زخم شاغلین	۰	۰	۰	۰	۵	۴	۰	۰
مجموع	۱۵۳	۱۳۷	۴۱۱	۲۸۱	۵۲۳	۴۰۰	۱۴۹	۸۱

SID: Scientific Information Database

جدول ۲: تعداد اسناد مشترک بین چهار پایگاه اطلاعاتی مورد بررسی

نام پایگاه	SID و Magiran	SID و SafeLir	SID و Irandoc	Magiran و SafeLir	Magiran و Irandoc	SafeLir و Irandoc
تعداد اسناد مشترک	۱۲۱	۱۳۲	۲۶۹	۷۷	۷۹	۸۱

به منظور پاسخگویی به سؤال تحقیق حاضر مبنی بر این که عملکرد کدام یک از پایگاه‌های ملی در بازیابی و پوشش اطلاعاتی حیطه مصدومیت‌های شغلی بهتر است، از روش پژوهش Shafi و Rather (۱۲) استفاده گردید. بر این اساس، عملکرد پایگاه اطلاعاتی SafeLir در بازیابی اسناد بیشتر از سایر پایگاه‌ها بود و اسناد مرتبط بیشتری را بازیابی نمود. از نظر موضوعی، پایگاه اطلاعاتی SafeLir فقط در زمینه صدمات و ایمنی فعالیت دارد و اسناد آن‌ها را پوشش می‌دهد، اما سایر پایگاه‌های اطلاعاتی در زمینه‌های مختلف پزشکی و غیر پزشکی حاوی اسناد انبوهی هستند. از طرف دیگر، موضوع مورد جستجو به موضوعات تحت پوشش پایگاه اطلاعاتی SafeLir نزدیک‌تر بود. بنابراین، جامعیت این پایگاه اطلاعاتی در حیطه مورد نظر بیشتر از سایر پایگاه‌های اطلاعاتی برآورد گردید. مطالعات دیگر نیز به این موضوع اشاره کرده‌اند که اگر حیطه‌های موضوعی تحت پوشش یک پایگاه بیشتر باشد، بازیابی اسناد یک حیطه معین از همان پایگاه دارای کارایی بالایی نخواهد بود (۱۵).

از جمله موارد قابل ذکر در مقایسه جامعیت نسبی پایگاه‌ها، می‌توان به عدم ورود پایگاه اطلاعاتی IranMedex به جامعه مورد بررسی اشاره کرد که این اتفاق به دلیل نقص فنی در این پایگاه بود. این موضوع از این نظر حایز اهمیت است که بررسی پژوهش‌های مشابه نشان داد که در ارزیابی جامعیت پایگاه‌های اطلاعاتی ملی، پایگاه اطلاعاتی IranMedex جامعیت بالاتری نسبت به سایر پایگاه‌های اطلاعاتی داشت (۱۳، ۵). بنابراین، این موضوع قابلیت مقایسه سایر پایگاه‌های اطلاعاتی ملی را تحت تأثیر قرار می‌دهد، اما چیزی که مسلم است و نتایج سایر مطالعات نیز نشان داده‌اند، بالاتر بودن جامعیت نسبی پایگاه Magiran نسبت به پایگاه SID می‌باشد (۱۶، ۱۳، ۵).

بر اساس بررسی متون انجام شده توسط تیم تحقیق، پژوهش حاضر اولین مطالعه‌ای است که عملکرد پایگاه اطلاعاتی SafeLir را در بازیابی اسناد با سایر پایگاه‌های اطلاعاتی ملی مورد بررسی هم‌زمان قرار می‌دهد. با توجه به نتایج به دست آمده، می‌توان گفت که جامعیت و عملکرد پایگاه اطلاعاتی نوپای SafeLir در بازیابی اسناد مربوط، در نوع خود قابل توجه می‌باشد و تحقیق حاضر توانسته است یک منبع مهم اطلاعاتی را به پژوهشگران و متخصصان حیطه مصدومیت‌ها معرفی نماید. به ویژه در شرایط کنونی که دسترسی به پایگاه اطلاعاتی IranMedex به عنوان پایگاه مناسب برای حیطه علوم پزشکی مقدور نیست.

تمام پایگاه‌های اطلاعاتی مورد بررسی در مطالعه حاضر با یکدیگر همپوشانی داشت. بیشترین همپوشانی مربوط به پایگاه‌های اطلاعاتی SafeLir و Magiran با ۲۶۹ سند مشترک بود. با توجه به این که تمام اسناد به زبان فارسی و اغلب مربوط به تحقیقات داخل کشور بود، مجلاتی که این مقالات را پذیرش و چاپ می‌کنند، در بیشتر پایگاه‌های اطلاعاتی فارسی‌زبان نمایه می‌شوند (۱۸، ۱۷). بنابراین، تعداد مجلات نمایه شده در پایگاه‌های اطلاعاتی مورد بررسی رابطه مستقیمی با جامعیت موضوع مورد نظر داشت و هر قدر این مجلات در پایگاه‌های مختلفی نمایه شوند، همپوشانی اطلاعاتی موضوع مورد نظر در پایگاه‌ها افزایش می‌یابد.

بر همین اساس، جامعیت نسبی پایگاه‌های اطلاعاتی محاسبه شد. مقادیر مربوط به این محاسبات در جدول ۳ آمده است. همان‌گونه که مشاهده می‌گردد، بیشترین مقدار جامعیت نسبی مربوط به پایگاه اطلاعاتی SafeLir و کمترین آن مربوط به پایگاه اطلاعاتی Irandoc می‌باشد.

جدول ۳: مقادیر جامعیت نسبی و مانعیت مربوط به پایگاه‌های اطلاعاتی

نام پایگاه	جامعیت نسبی (درصد)	مانعیت (درصد)
Magiran	۵۹/۷	۶۸/۳۶
SID	۴۱/۳	۸۹/۵۴
SafeLir	۸۳/۷	۷۶/۴۸
IranDoc	۳۴/۲	۵۴/۳۶

SID: Scientific Information Database

با استفاده از روش بیان شده در قسمت روش بررسی مطالعه حاضر، میزان مانعیت پایگاه‌ها محاسبه گردید. با وجود این که پایگاه SID سومین جایگاه جامعیت پایگاه‌های اطلاعاتی را به خود اختصاص داد، اما در بررسی مانعیت، این پایگاه رتبه اول را کسب کرد. اطلاعات مربوط به این شاخص در جدول ۳ آمده است.

بحث

نتایج پژوهش حاضر نشان داد که به نظر می‌رسد کلید واژگان اصطلاحنامه پزشکی فارسی، در این حوزه موضوعی نمی‌تواند ملاک مناسبی برای انتخاب کلید واژه اصلی در انجام جستجو باشد. نتایج برخی مطالعات در خصوص انجام جستجو در پایگاه‌های اطلاعاتی ملی گزارش کرده‌اند که استفاده از کلید واژگان اصطلاحنامه پزشکی فارسی اغلب منجر به بازیابی جامع اسناد نمی‌گردد (۱۳). با وجود این که موضوع مورد جستجو در بررسی حاضر متفاوت‌تر از موضوع مورد جستجو در سایر تحقیقات مشابه بود، این یافته قابل توجه است که واژگان رایج مورد استفاده در مقالات متخصصان و پژوهشگران حیطه‌های مختلف، به مراتب عملکرد مناسب‌تری نسبت به واژگان اصطلاحنامه پزشکی فارسی در بازیابی اسناد از پایگاه‌های اطلاعاتی دارند (۱۴).

انتخاب دو گروه از واژگان، ترکیب آن‌ها و استفاده از این ترکیب به عنوان کلید واژه‌های اصلی برای انجام جستجو در پایگاه‌های اطلاعاتی، از جمله تفاوت‌های مطالعه حاضر با سایر تحقیقات مشابه (۱۴-۱۲) بود. در پژوهش حاضر، انتخاب کلید واژه‌های اصلی بر اساس نظرات اعضای خبره گروه به همراه مرور کلید واژگان اسناد مربوط صورت گرفت. یافته‌ها نیز نشان داد که این نوع روش انتخاب کلید واژه‌ها عملکرد مناسب‌تری در بازیابی نتایج از پایگاه‌های اطلاعاتی دارد، اما به دلیل این که هدف مطالعه حاضر، مقایسه تأثیر استفاده از نوع کلید واژگان اصلی بر جامعیت بازیابی اسناد نبود، نتایج بررسی عملکرد کلید واژگان اصطلاحنامه پزشکی فارسی در یافته‌ها و بحث نیامده است.

نظریاتی در انتخاب کلید واژه‌های اصطلاح‌نامه پزشکی فارسی صورت گیرد و پایگاه‌های اطلاعاتی مربوط نیز از کلید واژگان کنترل شده استفاده نمایند. نتیجه تحقیق حاضر به پژوهشگران و دانشجویان حیطه صدمات کمک خواهد کرد تا پایگاهی را که برای بازایی اسناد فارسی در حوزه تخصصی آن‌ها جامعیت و مانعیت مناسبی دارد، شناسایی کنند.

پیشنهادهای

به نهادها، متصدیان تهیه منابع، کتابخانه‌ها و مراکز پژوهشی توصیه می‌شود که آشنایی و دسترسی به تمام پایگاه‌های اطلاعاتی ملی را فراهم نمایند تا محققان حوزه صدمات از پایگاهی که جامعیت بالا یا مانعیت بالا یا هر دو را دارد، استفاده نمایند.

تشکر و قدردانی

بدین وسیله از کلیه افرادی که در انجام این مطالعه همکاری نمودند، تشکر و قدردانی به عمل می‌آید.

تضاد منافع

در انجام مطالعه حاضر، نویسندگان هیچ‌گونه تعارض منافی نداشته‌اند.

نتایج پژوهش‌های مختلف نشان داده است، با وجود این که جامعیت پایگاه SID کمتر از سایر پایگاه‌های اطلاعاتی بود، اما مانعیت بالاتری را نسبت به سایر پایگاه‌ها دارد (۱۹، ۵). در مطالعه حاضر نیز پایگاه اطلاعاتی SID با مانعیت ۸۹/۵ درصد در مقایسه با پایگاه‌های اطلاعاتی SafeLir (۷۶/۵ درصد)، Magiran (۶۸ درصد) و Irandoc (۵۴ درصد)، دارای بیشترین مانعیت بازایی اسناد بود. با توجه به این که پایگاه SID در تحقیق حاضر جامعیت نسبی کمتری در مقایسه با پایگاه‌های اطلاعاتی SafeLir و Magiran داشت و اساساً دو مقیاس جامعیت و مانعیت در روند جستجوی اسناد نتیجه معکوسی دارند (یعنی هر قدر مانعیت به سطح مطلوب سوق داده شود، از جامعیت کاسته می‌شود)، قابل توجه باشد (۲۰). مقادیر هر دو شاخص جامعیت و مانعیت در پایگاه اطلاعاتی Irandoc کمترین بود. با توجه به این که پایگاه اطلاعاتی مذکور انبوهی از اسناد غیر علوم پزشکی، پروپوزال‌ها و پایان‌نامه‌ها را ثبت می‌نماید، شاید در کیفیت بازایی اسناد در این پایگاه اثر نامطلوبی بگذارد.

از محدودیت‌های پژوهش حاضر می‌توان به بررسی بازایی اسناد فارسی‌زبان و عدم بررسی بازایی اسناد انگلیسی‌زبان اشاره کرد و علت آن هم نبود قابلیت مقایسه پایگاه‌ها از این نظر بود.

نتیجه‌گیری

بر اساس نتایج به دست آمده از مطالعه حاضر، به نظر می‌رسد که باید تجدید

References

- Masic I, Pandza H, Toromanovic S, Masic F, Sivic S, Zunic L, et al. Information technologies (ITs) in medical education. *Acta Inform Med* 2011; 19(3): 161-7.
- Zandian F, Maseudinasrabadi A, Latifi M, Zadolkhajeh L. Study of comparative the performance of search engines and meta engines. *Journal of Academic Librarianship and Information Research* 2016; 49(4): 511-93.
- Shakeri S. Precision and recall in Farsi search Engines in regard with information retrieval in library and information science. *National Studies on Librarianship and Information Organization* 2008; 19(1): 177-99. [In Persian].
- Hariri N, Babalhavaeji F, Farzandipour M, Nadi Ravandi S. Evaluation criteria of information retrieval systems: What we know and what we do not know. *Iranian Journal of Information Processing and Management* 2014; 30(1): 199-221. [In Persian].
- Sadeghi H, Okhovati M. Recall and precision of Iranmedex, Magiran and SID (Scientific Information Database) databases for retrieval of scholarly information in the Field of Pharmacy. *Health Inf Manage* 2014; 11(3):435-42. [In Persian].
- Damerchiloo M, Haji Zeinolabedini M. A comparison of relative exhaustivity among three Persian language databases (Magiran, SID, Iranmedex) for retrieving information in the field of Dentistry. *Library and Information Science* 2010; 13(1): 53-72. [In Persian].
- Sadeghi-Bazargani H, Poursaghar F, Soori H, Bahari A, Bakhtyari F, Moghisi A. Developing a national database for Iranian injury and safety literature: SafeLir. *Medical Science* 2019; 23(99): 757-68.
- Larsen PO, von Ins M. The rate of growth in scientific publication and the decline in coverage provided by Science Citation Index. *Scientometrics* 2010; 84(3): 575-603.
- Pautasso M. Publication growth in biological sub-fields: Patterns, predictability and sustainability. *Sustainability* 2012; 4(12): 3234-47.
- Eslami A, Keshavarz H. Study of online electronic information search skills among Ph.D. students at faculty of geography, University of Tehran. *Iranian Journal of Information Processing and Management* 2007; 23(1): 45-60. [In Persian].
- Hashemian M, Janatikia M, Hashemian A. Information seeking skills in online databases of Iranian National Medical Digital Library: A study among residents of Isfahan University of Medical Sciences. *Health Inf Manage* 2013; 10(1): 1-8. [In Persian].
- Shafi SM, Rather R. Precision and recall of five search engines for retrieval of scholarly information in the field of biotechnology. *Webology* 2005; 2(2): 8-14.
- Damerchiloo M, Haji Zeinolabedini Mohs. A comparison of relative exhaustivity among three Persian language databases (Magiran, SID, IranMedex) for retrieving information in the field of dentistry. *Library and Information Science* 2010; 13(1): 53-72. [In Persian].
- Ghaebi A, Khosravi F, Malmir A. Evaluating the status of information retrieval in "Namayeh" and "Nama" databases and assessing the effectiveness of using controlled terms in indexing these two databases. *Research on Information Science and Public Libraries* 2011; 17(2): 247-62. [In Persian].

15. Gasparyan AY, Ayyvazyan L, Kitas GD. Multidisciplinary bibliographic databases. *J Korean Med Sci* 2013; 28(9): 1270-5.
16. Asadollahi Z, Nookarizi M. Evaluating the structure and content of electronic databases of iranian periodicals. *Library and Information Science* 2010; 13(2): 118-45. [In Persian].
17. Persian Subject Index. *Abstract Articles and Insurance Research* 2006; 10(1): 219-40. [In Persian].
18. Ebrahimi H. An examination of the degree of overlap in Namayeh and Nama databases in 2005. *National Studies on Librarianship and Information Organization* 2008; 18(4): 299-308. [In Persian].
19. Shiri AA. Assessing the comprehensiveness and determinance of information retrieval systems. *Research on Information Science and Public Libraries* 1994; 12(4): 38-45. [In Persian].
20. Buckland M, Gey F. The relationship between recall and precision. *J Am Soc Inf Sci* 1994; 45(1): 12-9.

Recall and Precision of Persian Databases for Retrieval of Scholarly Information in the Field of Occupational Injuries

Ali Bahari¹, Homayoun Sadeghi-Bazargani², Fahimeh Bakhtiari³

Original Article

Abstract

Introduction: Use of databases for accessing electronic documents has increased the necessity of reviewing the performance and coverage of these databases. The aim of this study was to assess the recall and precision of Persian databases for retrieval of scholarly information in the field of occupational injuries.

Methods: This descriptive study was carried out using observation and comparison approach. The study population included national databases of SafeLir, Scientific Information Database (SID), Magiran, and Irandoc, and the sample size equaled the number of documents retrieved from these databases in the last two decades. Keywords were determined based on the opinions of the members of the research team and review of the relevant documents. Finally, forty key words were chosen for searching the sites were searched, and the number of recovered documents was recorded. Indices and graphs related to recall and precision were then prepared.

Results: The number of recovered documents from Magiran, SID, SafeLir, and Irandoc were 411, 153, 523, and 149, respectively. SafeLir database had a higher relative recall (83.7%) and SID had a higher precision (89.54%) in retrieving relevant documents compared to other databases. The greatest overlap in results was observed between Magiran and SafeLir (269 documents).

Conclusion: The findings of this study will help researchers in the field of injuries and related fields to identify the appropriate database (with an appropriate recall and precision) for retrieval of documents in their specialized fields. If recall of the search or precision of the search or both of them are important. It is recommended to use SafeLir, SID, and SafeLir databases, respectively, in cases where precision, recall and both are important.

Keywords: Occupational Injuries; Occupational Health; Database; Information Storage and Retrieval

Received: 30 Aug., 2019

Accepted: 24 Nov., 2019

Published: 06 Dec., 2019

Citation: Bahari A, Sadeghi-Bazargani H, Bakhtiari F. **Recall and Precision of Persian Databases for Retrieval of Scholarly Information in the Field of Occupational Injuries.** Health Inf Manage 2019; 16(5): 236-42.

Article resulted from an independent research without financial support.

1- MSc, Epidemiology, Road Traffic Injury Research Center AND Department of Biostatistics and Epidemiology, School of Health, Tabriz University of Medical Sciences, Tabriz, Iran (Corresponding Author) Email: alibahari2012@gmail.com

2- Associate Professor, Epidemiology, Road Traffic Injury Research Center AND Department of Biostatistics and Epidemiology, School of Health, Tabriz University of Medical Sciences, Tabriz, Iran

3- MSc, Medical Library and Information, Neurosciences Research Center, Tabriz University of Medical Sciences, Tabriz, Iran