



دانشگاه علوم پزشکی کرمان

دانشکده مدیریت و اطلاع رسانی پزشکی

پایان نامه مقطع کارشناسی ارشد رشته کتابداری و اطلاع رسانی پزشکی

عنوان:

مقایسه پایگاه ها و موتورهای جستجوی منتخب از لحاظ ساخت و آژیه ای در

بازیابی اسناد پزشکی فارسی

توسط: محسن نوری

استاد راهنما: دکتر مریم اخوتی

استاد مشاور: جناب آقای علی طالبیان - سرکار خانم دکتر اعظم استاجی

سال تحصیلی: ۹۵-۱۳۹۴

چکیده

مقدمه: هدف نهایی هر سیستم بازیابی اطلاعات، بازیابی حداکثر اطلاعات مرتبط و حداقل اطلاعات نامرتبط در کمترین زمان ممکن می‌باشد. در بازیابی اطلاعات، ربط مدارک با سوالی خاص بر اساس شباهت واژه‌ها سنجیده می‌شود. زبان فارسی به دلیل ماهیت آن مانند سایر زبان‌ها در امر بازیابی اطلاعات دارای مشکلات گوناگونی است. یکی از این مشکلات مربوط به ساخت واژه می‌باشد، به طوری که اشکال مختلف نوشتاری موجب از دست رفتن اسناد مرتبط در بازیابی اطلاعات می‌شود. در این مطالعه به بررسی ۱۵ مورد از مشکلات ساخت واژه‌های زبان فارسی که در بازیابی اطلاعات و اسناد پزشکی مشکل ساز هستند پرداخته شده است.

روش اجرا: این مطالعه به روش پیمایش مقایسه‌ای انجام شده است. بدین صورت که مشکلات ساخت واژه‌های برگرفته از متون (شامل ۱۵ مشکل) با استفاده از مطالعات صورت گرفته انتخاب شده و با استفاده از اصطلاح نامه پزشکی فارسی برای هر مشکل ساخت واژه‌های یک مصداق شامل یک کلیدواژه انتخاب شده و یک کلیدواژه انتخاب نشده انتخاب شد. سپس کلید واژه‌ها در چهار پایگاه اطلاعاتی مگ ایران، ایران مدکس، مرکز منطقه‌ای شیراز، مدلب و موتورهای جستجوی فارسی پارسیک و پارسیجو مورد جستجو قرار گرفت و داده‌های توصیفی به صورت جدول به ثبت رسید. نتایج به دست آمده توسط نرم افزار SPSS و با استفاده از آزمون کای دو، ضریب فی و ضریب کرامر برای تعیین وجود رابطه و همچنین شدت آن مورد تجزیه و تحلیل قرار گرفتند.

نتایج: بر اساس نتایج به دست آمده بیشترین فراوانی بازیافت بین پایگاه‌های اطلاعاتی مربوط به پایگاه ایران مدکس می‌باشد. از میان دو موتور جستجو، موتور جستجوی پارسیجو، مدارک بیشتری را بازیابی نمود. مدت زمان بازیافت کلید واژه‌ها نیز در ابزارهای جستجو نزدیک به هم می‌باشد به جز پایگاه مرکز منطقه‌ای شیراز که بیشترین مدت زمان بازیابی نتایج را به خود اختصاص داد. همچنین نتایج آمار تحلیلی نشان داد که بین کلید واژه‌ها و ابزارهای جستجو از لحاظ فراوانی بازیافت رابطه معنی داری وجود دارد ولی از لحاظ مدت زمان بازیافت بین کلیدواژه‌ها و ابزارهای جستجو رابطه معنی داری مشاهده نشد.

بحث و نتیجه گیری:

با توجه به این که در ابزارهای جستجو بیشترین فراوانی مربوط به شکل انتخاب شده کلیدواژه‌ها (بر اساس اصطلاح نامه پزشکی) می‌باشد ولی شکل انتخاب نشده کلیدواژه‌ها نیز بازیابی شده است. همچنین مشاهده شد که به جز مشکل ساخت واژه‌های مربوط به تشدید که تمام ابزارهای جستجو به رفع آن در بازیابی اطلاعات پرداخته اند

هیچ‌کدام از موتورهای جستجو به طور کامل به رفع مشکلات ساخت واژه ای زبان فارسی نپرداخته اند. به عبارت دیگر بین کلیدواژه ها و ابزارهای جستجو رابطه معنی داری وجود دارد به طوری که جستجوی یک کلید واژه خاص در ابزارهای جستجو نتایج مختلفی را بازیابی خواهد کرد.

کلید واژه ها: ساخت واژه، زبان فارسی، بازیابی اطلاعات، اسناد فارسی، اطلاعات پزشکی، اصطلاحات پزشکی،

ابزارهای جستجو