

بررسی میزان تطابق زبان نمایه‌سازان، نویسندگان و برچسب‌گذاران در پایگاه اطلاعاتی اریک و مندلی

مریم قنواتی

کارشناسی ارشد مدیریت اطلاعات؛ گروه علم اطلاعات و دانش‌شناسی؛
دانشکده مدیریت؛ دانشگاه تهران؛ maryamghanavati@ut.ac.ir

علیرضا نوروزی *

دکتری علم اطلاعات و دانش‌شناسی؛ استادیار؛ گروه علم اطلاعات و
دانش‌شناسی؛ دانشکده مدیریت؛ دانشگاه تهران؛ noruzi@ut.ac.ir

مریم ناخدا

دکتری علم اطلاعات و دانش‌شناسی؛ استادیار؛ گروه علم اطلاعات و
دانش‌شناسی؛ دانشکده مدیریت؛ دانشگاه تهران؛ mnakhoda@ut.ac.ir

اشکان خطیر

دانشجوی دکتری مهندسی فناوری اطلاعات؛ پژوهشگاه علوم و فناوری
اطلاعات ایران (ایرانداک)؛ khatir@students.irandoc.ac.ir

پذیرش: ۹۶/۰۴/۰۳

دریافت: ۹۵/۰۹/۰۸

فصلنامه علمی پژوهشی
پژوهشگاه علوم و فناوری اطلاعات ایران
شاپا (چاپی) ۸۲۲۳-۲۲۵۱
شاپا (الکترونیکی) ۸۲۳۱-۲۲۵۱
نمایه در SCOPUS، LISA و ISC
<http://jlist.irandoc.ac.ir>
دوره XX | شماره X | صص XX-XX
۱۳XX X

نوع مقاله: پژوهشی

به این مقاله به شکل زیر استناد کنید:

دورن متن:

(قنواتی، زودآیند)

در فهرست منابع:

قنواتی، مریم. زودآیند. بررسی میزان تطابق زبان
نمایه‌سازان، نویسندگان و برچسب‌گذاران در پایگاه
اطلاعاتی اریک و مندلی پژوهشنامه پردازش و
مدیریت اطلاعات.

<http://jipm.irandoc.ac.ir> (دسترسی در

روز/ماه/سال)

چکیده: هدف این پژوهش شناسایی میزان تطابق زبان نمایه‌سازان، نویسندگان و برچسب‌گذاران در پایگاه اطلاعاتی اریک و مندلی است. این پژوهش از نوع کاربردی است که برای انجام آن از روش تحلیل محتوا استفاده شده است. جامعه مورد مطالعه در این پژوهش شامل ۴۹۹ مقاله از ۲۰ مجله حوزه علوم تربیتی است که در سال ۲۰۱۴ در پایگاه اطلاعاتی اریک نمایه‌سازی شده بودند. برچسب‌های تخصیص داده شده در فاصله زمانی از اول ژانویه سال ۲۰۱۴ تا سوم اوت سال ۲۰۱۶ از وبگاه مندلی استخراج گردید. کلیدواژه‌های تخصیص داده شده به مقاله‌ها توسط نویسندگان‌شان، از خود مقاله‌ها استخراج گردید. داده‌های گردآوری شده برای بررسی میزان تطابق این سه زبان با استفاده از برنامه‌ای به زبان برنامه‌نویسی شی‌گرا سی شارپ مورد تجزیه و تحلیل قرار گرفت. همچنین به منظور مشخص شدن میانگین‌ها و فراوانی‌ها از آمار توصیفی استفاده شد و برای تحلیل و ارزیابی داده‌ها نرم‌افزار اکسل ۲۰۱۳ به کار گرفته شد. یافته‌ها نشان می‌دهد که میزان تطابق کلیدواژه‌های تخصیص داده شده توسط نویسندگان مدارک با برچسب‌های تخصیص داده شده توسط برچسب‌گذاران به همان

مدارک در وبگاه مندلی ۱۵ درصد، میزان تطابق توصیفگرهای تخصیص داده شده توسط نمایه‌سازان به مدارک در پایگاه اریک با برچسب‌های تخصیص داده شده توسط برچسب‌گذاران به همان مدارک در وبگاه مندلی ۳ درصد، میزان تطابق توصیفگرهای تخصیص داده شده توسط نمایه‌سازان به مدارک در پایگاه اریک با کلیدواژه‌های تخصیص داده شده توسط نویسندگان به همان مدارک ۴ درصد، و در نهایت، میزان تطابق هر سه زبان مورد مطالعه ۱/۱ درصد بود. همچنین میزان حضور توصیفگرهای تخصیص داده شده توسط نمایه‌سازان در اصطلاحنامه اریک ۳۴ درصد بود که از دو گروه نویسنده و برچسب‌گذار بیشتر است. یافته‌ها نشان داد که تطابق بیشتری بین واژگان و زبان نویسنده و برچسب‌گذار نسبت به تطابق بین واژگان و زبان نمایه‌ساز و برچسب‌گذار و همچنین تطابق بین واژگان و زبان نویسنده و نمایه‌ساز وجود دارد. نتایج این پژوهش نشان داد که برچسب‌گذاران واژه‌هایی متفاوت از توصیفگرهای نمایه‌سازان و کلیدواژه‌های نویسندگان استفاده کرده‌اند، و این نشان دهنده عدم آشنایی سه گروه «نمایه‌ساز، نویسنده و برچسب‌گذار» از زبان و واژگان مورد استفاده همدیگر است. پیشنهاد می‌شود که به منظور جستجوی سودمند و بازیابی اطلاعات در نظام‌های سنتی نمایه‌سازی در هنگام کنترل واژه‌ها و همارایی واژه‌ها (پیش‌همارا و پس‌همارا) زبان کاربر و نویسنده مورد توجه قرار گیرد.

کلیدواژه‌ها: نمایه‌سازی، برچسب‌گذاری، بازنمایی اطلاعات، سازماندهی اجتماعی اطلاعات، کلیدواژه‌های نویسنده، اریک، مندلی.

*علیرضا نوروزی noruzi@ut.ac.ir

۱. مقدمه و بیان مسئله

هدف نظام‌های بازیابی اطلاعات، تسهیل در شناسایی مدارک یا منابع اطلاعاتی است که متناسب با نیازهای اطلاعاتی جستجوگران باشد. استفاده‌کنندگان اصطلاح‌ها نیز با فرهنگ‌ها و سطوح دانش متفاوت، گاهی با کمترین اطلاع از وسعت و دامنه اصطلاح‌های موضوعی، برای بازیابی مدارک به جستجو می‌پردازند. در این میان زبان نمایه‌سازی برای ایجاد این هم‌زبانی و عمل انطباق بین مدارک ذخیره شده و پرسش مطرح شده نقش تعیین‌کننده‌ای دارد. بنابراین، انطباق میان توصیفگرهای نمایه‌ای با پرسش یا کلیدواژه‌های جستجوی کاربران، نقش مهم و تعیین‌کننده‌ای در امر بازیابی اطلاعات دارد. به عبارت دیگر، نحوه استفاده از زبان توسط نمایه‌سازان و جستجوگران بسیار متفاوت است (نوروزی و ولایتی، ۱۳۸۹، ص. ۴۳).

نظام‌های بازیابی اطلاعات اغلب مبتنی بر زبان طبیعی یا زبان کنترل شده هستند. جستجو در یک نظام بازیابی اطلاعات چه با استفاده از زبان طبیعی و چه با استفاده از اصطلاح‌های کنترل شده فقط در صورتی نتایجی را به همراه خواهد داشت که کل یا بخشی از عبارت جستجوی کاربر با کل یا بخشی از اصطلاح‌های نمایه‌ای در نظام بازیابی اطلاعات همخوان باشد. به همین دلیل است که نمایه‌سازی با جستجو رابطه‌ای تنگاتنگ دارند (کلیولند و کلیولند، ۱۳۸۵؛ تیلور، ۱۳۸۵). از سوی دیگر، ابزار اصلی نمایه‌سازی به زبان کنترل شده اصطلاحنامه است و در حال حاضر به دلیل مشکلاتی نظیر عدم دسترسی به اصطلاحنامه‌ها، عدم آشنایی با اصطلاحنامه‌ها و روش استفاده از آن‌ها بسیار کم مورد استفاده کاربران و جستجوگران اطلاعات قرار می‌گیرند. در نتیجه، در یک نظام ذخیره و بازیابی اطلاعات، کاربر به طور دقیق نمی‌داند که چه می‌خواهد، یا نمی‌تواند آنچه را که می‌خواهد به طور صریح مطرح کند. بنابراین، منجر به ایجاد فاصله بین زبان کاربر و زبان نویسنده (زبان مدرک) می‌شود. این در حالی است که زبان نظام ذخیره و بازیابی اطلاعات و یا به عبارتی، زبان نمایه‌سازی در پایگاه اطلاعاتی، ضمن اینکه خود برگرفته از زبان کاربر و زبان نویسنده است، با این دو بسیار متفاوت است.

امروزه پایگاه‌های اطلاعاتی با کمبود بودجه و نیروی انسانی نمایه‌ساز مواجه هستند. در نتیجه، در پایگاه‌های اطلاعاتی از کلیدواژه‌های تخصیص داده شده توسط نویسندگان^۱ مدارک (اعم از پایان‌نامه‌ها، مقاله‌ها، و طرح‌های پژوهشی) استفاده می‌کنند. به این معنا که نویسندگان به مدارک خود کلیدواژه‌هایی تخصیص می‌دهند که نشان‌دهنده و نمایانگر موضوع یا موضوع‌های آن مدارک است. پایگاه‌های اطلاعاتی نیز تمام یا بخشی از محتوای موضوعی یا همان کلیدواژه‌های هر مدرک را در فیلدهای مربوط به آن مدرک نمایش می‌دهند. از طرف دیگر، با توجه به اینکه نظام‌های رده‌بندی کتابخانه‌ای و نظام‌های سنتی نمایه‌سازی که مبتنی بر پشتوانه انتشاراتی^۲ هستند، پاسخ‌گوی نیازهای اطلاعاتی کاربران اینترنتی نیستند و استفاده از رده‌بندی‌های کتابخانه‌ای نیازمند آشنایی کامل با قواعد فهرست‌نویسی و رده‌بندی است. در نتیجه، نظام رده‌بندی اجتماعی یا نمایه‌سازی اجتماعی (فوکسونومی^۳) که مبتنی بر فناوری‌های نوینی چون برچسب‌گذاری مشارکتی^۴ و همکاری، نرم‌افزارهای اجتماعی و نشانه‌گذاری اجتماعی^۵ است و پشتوانه کاربری^۶

1. Author-driven keywords

2. Literary warrant

3. Folksonomy

4. Collaborative tagging

5. Social bookmarking

6. User warrant

دارد، با کمک سایر فناوری‌های نسل دوم وب راه‌حلی نو برای خودسازمان‌دهی^۱ منابع اطلاعاتی وبی در هر جای جهان و با هر زبانی است. این نظام سازمان‌دهی اجتماعی دانش^۲ فرصت مناسبی را برای کاربران فراهم نموده تا منابع دلخواه خود را به صورت برخط از طریق اینترنت دسترس‌پذیر نمایند؛ و همچنین امکان شناسایی برچسب‌های^۳ پرسامد را برای کاربران و کتابداران (نمایه‌سازان) فراهم نموده و اجازه می‌دهد تا برچسب‌های داده شده به یک اثر توسط سایر کاربران شناسایی شود (نوروزی، ۱۳۹۵).

به دیگر سخن، سازمان‌دهی اجتماعی اطلاعات در محیط وب دارای سه ضلع است که عبارتند از: نویسندگان مدارک، نمایه‌سازان مدارک و استفاده‌کنندگان (برچسب‌گذاران) اطلاعات. در این راستا پژوهش حاضر به بررسی میزان اشتراک و انطباق زبان نمایه‌سازان (توصیفگرهای نمایه‌سازی انتساب داده شده به هر مدرک)، کلیدواژه‌های نویسندگان مدارک، و برچسب‌های تخصیص داده شده توسط برچسب‌گذاران و استفاده‌کنندگان اطلاعات می‌پردازد.

قابل ذکر است که یکی از مهمترین و قدیمی‌ترین پایگاه‌های اطلاعاتی جهان، پایگاه اطلاعاتی علوم تربیتی (اریک^۴) است. این پایگاه به لحاظ گردآوری، سازمان‌دهی و نمایه‌سازی اطلاعات، به ویژه مجله‌ها و منابع الکترونیکی، نسبت به سایر پایگاه‌ها ساختارمندتر و نظام‌مندتر بوده و دارای اصطلاحنامه تخصصی است. همچنین پایگاه اطلاعاتی مندلی یکی از پایگاه‌ها و خدمات اطلاعاتی مبتنی بر رده‌بندی و نمایه‌سازی اجتماعی است که برای سازمان‌دهی اجتماعی، به اشتراک‌گذاری مقاله‌های علمی، یافتن داده‌های پژوهشی و همکاری برخط از آن استفاده می‌شود. عدم تطابق ساختار زبانی نویسنده و نمایه‌ساز (براث^۵، ۲۰۰۸) ما را بر آن داشت تا ساختار زبان کاربر (برچسب‌گذار^۶) را با دو ساختار پیش‌گفته مورد مطالعه قرار دهیم. بنابراین، مسئله اصلی پژوهش حاضر این است که آیا میان سه نوع ساختار زبانی مهار شده اصطلاحنامه‌محور، مهار نشده کاربرمدار و کلیدواژه‌های تخصیص داده شده به مدارک توسط نویسندگان تطابق و همخوانی وجود دارد؟ برای پاسخ به این پرسش، پژوهش حاضر به بررسی میزان تطابق زبان نمایه‌سازان، نویسندگان و برچسب‌گذاران در پایگاه اطلاعاتی اریک و وبگاه مندلی^۷ می‌پردازد.

1. Self-organization of information
2. Social organization of knowledge
3. Tag
4. ERIC
5. Bruce
6. Tagger
7. Mendeley

۲. پرسش‌های اساسی پژوهش

این پژوهش در صدد است تا به پرسش‌های اساسی زیر پاسخ دهد:

۱. میزان تطابق کلیدواژه‌های تخصیص داده شده توسط نویسندگان مدارک با برچسب‌های تخصیص داده شده توسط برچسب‌گذاران به همان مدارک در وبگاه مندلی چگونه است؟
۲. میزان تطابق توصیفگرهای تخصیص داده شده توسط نمایه‌سازان به مدارک در پایگاه اریک با برچسب‌های تخصیص داده شده توسط برچسب‌گذاران به همان مدارک در وبگاه مندلی تا چه اندازه است؟
۳. میزان تطابق توصیفگرهای تخصیص داده شده توسط نمایه‌سازان به مدارک در پایگاه اریک با کلیدواژه‌های تخصیص داده شده توسط نویسندگان مدارک به همان مدارک چگونه است؟
۴. میزان تطابق کلیدواژه‌های تخصیص داده شده توسط نویسندگان مدارک و توصیفگرهای تخصیص داده شده توسط نمایه‌سازان به مدارک در پایگاه اریک و برچسب‌های تخصیص داده شده توسط برچسب‌گذاران به همان مدارک در وبگاه مندلی چگونه است؟
۵. میزان حضور کلیدواژه‌های تخصیص داده شده توسط نویسندگان مدارک و توصیفگرهای تخصیص داده شده توسط برچسب‌گذاران در اصطلاحنامه اریک چگونه است؟
۶. میزان فراوانی کلیدواژه‌های به کار رفته توسط سه گروه نویسنده، نمایه‌ساز و برچسب‌گذار چگونه است؟

۳. پیشینه پژوهش

در زمینه همخوانی و تطابق بین توصیفگرهای نمایه‌سازان، کلیدواژه‌های نویسندگان و برچسب‌های برچسب‌گذاران در داخل و خارج از کشور پژوهش‌های اندکی انجام شده است که برخی از آن‌ها به شرح زیر است (جدول ۱).

محمدپور (۱۳۹۰) به «بررسی میزان همپوشانی میان توصیفگرهای مقاله‌های علم اطلاعات و دانش‌شناسی (کتابداری) در پایگاه لیستا^۱ و برچسب‌های اختصاص یافته به همان مقاله‌ها در وبگاه سایت یولایک^۲» پرداخت. نتایج حاصل از این پژوهش نشان داد که کمتر از ۲ درصد میان

^۱. LISTA

^۲. CiteUlike

برچسب‌های اختصاص یافته به مقاله‌های کتابداری و توصیف‌گرهای اختصاص یافته به همان مقاله‌ها همپوشانی وجود دارد.

بهادری و بنی‌اقبال (۱۳۹۲) به بررسی «میزان انطباق توصیف‌گرهای نمایه‌سازی و کلیدواژه‌های پایان‌نامه‌ای رشته پزشکی دانشگاه علوم پزشکی تهران با اصطلاحنامه پزشکی فارسی و Mesh» پرداختند. نتایج نشان داد که میزان انطباق در توصیف‌گرهای فارسی ۵۰/۱ درصد همخوانی و در توصیف‌گرهای انگلیسی ۹۹/۷ درصد همخوانی، در کلیدواژه‌های فارسی ۱۸ درصد همخوانی و در کلیدواژه‌های انگلیسی ۲۶/۸ درصد همخوانی وجود داشته است.

کیپ^۱ (۲۰۱۱) در پژوهش خود به بررسی «کاربر، نویسنده و نمایه‌سازان حرفه‌ای در محتوا: پژوهشی در عملکرد برچسب‌گذاری در سایت یو لایک» پرداخت. میزان انطباق بین کلیدواژه نویسنده و برچسب‌های برچسب‌گذاران ۳۳ درصد، میزان انطباق بین توصیف‌گرها و برچسب‌های برچسب‌گذاران ۱۶ درصد و میزان انطباق بین کلیدواژه نویسنده و توصیف‌گرها ۱۹ درصد بود. در زمینه همخوانی و تطابق بین توصیف‌گرهای نمایه‌سازان، سرعنوان‌های موضوعی تخصیص داده شده توسط فهرست‌نویسان و کلیدواژه‌های نویسندگان در داخل و خارج از کشور پژوهش‌هایی انجام شده است که برخی از آن‌ها در زیر معرفی می‌گردد؛ هرچند که خارج از حوزه موضوعی پژوهش حاضر است.

نقنه اصفهانی، چشمه سهرابی، بنی‌اقبال (۱۳۹۱) به بررسی «مطالعه تطبیقی کلیدواژه‌های فارسی و انگلیسی پایان‌نامه‌های دانشگاه علوم پزشکی اصفهان با اصطلاحنامه‌ها و سرعنوان‌های موضوعی پزشکی فارسی و انگلیسی» پرداخت. یافته‌ها نشان داد که بین میزان تطابق کلیدواژه‌های فارسی و انگلیسی با پزشکی فارسی و MeSH رابطه‌ی معناداری وجود دارد در این میان ساختار کلیدواژه‌ها بیشتر به سمت ساختار ترکیبی و تک‌واژه‌ای و بیشتر کلیدواژه‌ها نیز از چکیده انتخاب شدند. همچنین بیشتر کلیدواژه‌های انتخابی انگلیسی بودند.

رجبیان (۱۳۹۱) به «بررسی میزان همخوانی عبارت‌های جستجوی کاربران با اصطلاحات پیشنهادی مقاله‌های پایگاه‌های اطلاعاتی لاتین و تحلیل ساختار واژگانی این اصطلاحات بر اساس روابط اصطلاحنامه‌ای» پرداخت. یافته‌ها نشان داد که در پایگاه ابسکو، بیشترین میزان همخوانی کلیدواژه‌های نویسندگان، همخوانی کامل (۶۰/۲ درصد) است. در پایگاه آی‌تریپل‌ای، بیشترین میزان همخوانی در اصطلاحات آی‌تریپل‌ای، همخوانی نسبی (۵۱/۵ درصد)، در اصطلاحات کنترل نشده پایگاه اطلاعاتی اینسپیک، همخوانی کامل (۵۵/۹ درصد) و در کلیدواژه‌های

^۱. Kipp

نویسندگان، همخوانی نسبی (۴۸/۲ درصد) می‌باشد. بین میزان همخوانی کلیدواژه‌های جستجوی کاربران با کلیدواژه‌های نویسندگان، در پایگاه ابسکو و آی تریپل ای، تفاوت معناداری مشاهده شد.

لی^۱ (۲۰۱۵) به «مقایسه تحلیلی اصطلاح‌های نمایه‌ای و برجسب‌های اجتماعی: سرعنوان موضوعی پزشکی در مقایسه با بیبسونومی^۲ و دلیشز^۳» پرداخت. یافته‌ها نشان داد که میزان تطابق مش با برجسب‌های داده شده کاربران در کل ۳۴ درصد و میزان تطابق با بیبسونومی ۲۱/۷۷ درصد و میزان تطابق با دلیشز ۸/۲۲ درصد است.

مینت و دیگران^۴ (۲۰۱۵) «کیفیت سرعنوان‌های موضوعی پزشکی (مش)^۵ در حوزه داروسازی و تطابق آن با مجله‌های داروسازی که در مدلاین^۶ نمایه‌سازی شده‌اند» مورد بررسی قرار دادند. نتایج این پژوهش نشان داد که در مجموع ۶۹۸۹ مقاله از ۱۰ مجله داروسازی که ۳۲۸ (۴/۷ درصد) مقاله‌ها به طور کامل نمایه نشده بودند و در نتیجه، هیچ اصطلاح مش را شامل نشد. از میان ۶۶۶۱ مقاله‌ای که کاملاً نمایه‌سازی شده بودند ۱۰/۱ درصد آن از مش در نظر گرفته شده است. و در مجموع، ۳۹۸۴ مقاله از ۶۶۶۱ مقاله که به طور خاص در مورد داروسازی است برای نمایه‌سازی آن از مش استفاده شده که حدود ۰/۹ درصد تطابق دارد.

جدول ۱. جمع‌بندی پیشینه‌ها

مشخصات منبع	عنوان/هدف	جامعه	روش	یافته‌ها
محمدپور (۱۳۹۰)	بررسی میزان همپوشانی میان توصیفگرهای مقاله‌های کتابداری در پایگاه لیستا و برجسب‌های اختصاص یافته به همان مقاله‌ها در سایت یولایک	مقاله‌های کتابداری منتشر شده در پایگاه لیستا و پایگاه فولکسونومی سایت یولایک	توصیفی-پیمایشی و تطبیقی مبتنی برمشاهده	کمتر از ۲ درصد میان برجسب‌های اختصاص یافته به مقاله‌های کتابداری و توصیفگرهای اختصاص یافته به همان مقاله‌ها همپوشانی وجود دارد.

1. Danielle H. Lee
2. BibSonomy
3. Delicious
4. Minguet et al
5. MeSH
6. Medline

توصیف‌گرهای فارسی ۵۰/۱ درصد همخوانی کامل، ۱۲/۹ درصد عدم حضور در اصطلاحنامه، کلیدواژه‌های فارسی ۱۸ درصد همخوانی کامل، ۴۴/۴ درصد عدم حضور در اصطلاحنامه وجود داشت.	پیمایشی به شیوه کتابسنجی مقایسه‌ای و کتابخانه‌ای است و ابزار گردآوری داده‌ها چک لیستی	کلیدواژه‌های پایان‌نامه‌های رشته پزشکی از سال ۱۳۸۵-۱۳۸۹	میزان انطباق توصیف‌گرهای نمایه‌سازی و کلیدواژه‌های پایان‌ای رشته پزشکی دانشگاه علوم پزشکی تهران با اصطلاحنامه پزشکی فارسی و MESH	بهادری و بنی‌اقبال (۱۳۹۲)
میزان انطباق بین کلیدواژه نویسنده و برجسب‌های برجسب‌گذاران ۳۳ درصد، میزان انطباق بین توصیف‌گرها و برجسب‌های برجسب‌گذاران ۱۶ درصد و میزان انطباق بین کلیدواژه نویسنده و توصیف‌گرها ۱۹ درصد بود.	آمار توصیفی - روش‌های اطلاع‌سنجی	کتابداری و علم اطلاع‌رسانی	کاربر، نویسنده و نمایه‌سازان حرفه‌ای در محتوا: پژوهشی در عملکرد برجسب‌گذاری در سایت یولایک	مارگارت ای. کیپ (۲۰۱۱)

۴. روش پژوهش

روش این پژوهش از لحاظ هدف پژوهش، از نوع کاربردی است؛ و می‌توان گفت که از نوع مطالعات پیمایشی است که برای انجام آن از روش تحلیل محتوا استفاده شده است. جامعه مورد مطالعه در این پژوهش مجله‌هایی را در بر دارد که در سال ۲۰۱۴ هم در پایگاه اطلاعاتی اریک نمایه‌سازی شده بودند و هم در پایگاه مندلی برجسب‌گذاری شده و هم دارای کلیدواژه‌های نویسنده بودند؛ و در محدوده ضریب تأثیر ۳/۸۹۷ (بالاترین رقم ضریب تأثیر) تا ۱/۵۸۶ در پایگاه گزارش استنادی مجله‌ها در پایگاه اطلاعاتی آی.اس.آی. وب آف ساینس را قرار گرفته بودند. چون انتظار می‌رود که چنین مجله‌های علمی سطح بالایی هم در پایگاه اریک نمایه‌سازی شده باشند؛ هم در پایگاه مندلی حضور داشته باشند و هم اینکه دارای کلیدواژه‌های نویسنده باشند. در هنگام انتخاب مجله‌ها از بین ۳۰ مجله که بر اساس ضریب تأثیر اولویت‌بندی شده بودند ۱۰ مجله در سال ۲۰۱۴ دیگر در پایگاه اطلاعاتی اریک نمایه‌سازی نمی‌شدند؛ و فقط شماره‌های مربوط به سال ۲۰۱۳ و قبل از آن موجود بود؛ که در نهایت، ۲۰ مجله در سال ۲۰۱۴ انتخاب شد که هم در پایگاه اریک بود و هم در پایگاه مندلی برجسب‌گذاری شده بودند، از بین ۲۰ مجله مورد نظر که دارای ۴۹۹ مقاله بودند؛ ۱۴۶ مقاله در پایگاه مندلی در سال ۲۰۱۴ اصلاً برجسب‌گذاری نشده بودند و یا حتی خوانده نشده بودند. از این ۱۴۶ مقاله، ۱۲۷ مقاله در پایگاه مندلی فقط خوانده شده بودند و برجسب‌گذاران برجسبی به مقاله‌ها نداده بودند. در این بررسی ۱۹ مقاله موجود در پایگاه اریک در پایگاه مندلی یافت نشد. در نهایت، ۳۵۳ مقاله از ۴۹۹ مقاله موجود در پایگاه اریک در پایگاه مندلی برجسب‌گذاری شده بودند. در ضمن استخراج کلیدواژه‌های نویسندگان از مقاله‌ها، تعداد ۷۹ مقاله موجود فاقد کلیدواژه نویسنده بود. در واقع،

در مبحث مجموعه‌ها در علم ریاضی پژوهش حاضر شامل سه مجموعه است. این سه مجموعه عبارتند از:

$I =$ کلیدواژه‌های مجموعه مقاله‌های دارای توصیفگرهای نمایه‌سازی که N تعداد کلیدواژه دارد؛

$A =$ کلیدواژه‌های مجموعه مقاله‌های دارای کلیدواژه نویسنده که M تعداد کلیدواژه دارد؛ و

$T =$ کلیدواژه‌های مجموعه مقاله‌های دارای برچسب‌های برچسب‌گذار که Z تعداد کلیدواژه دارد.

در این پژوهش با توجه به محدودیت‌های شرح داده شده در بالا، باید گفت که در هر مجموعه تعداد مقاله‌ها ثابت و یکسان است و مقاله‌ای حذف نمی‌شود؛ اما در خود مجموعه تعداد کلیدواژه‌ها تغییر می‌کند. ممکن است مقاله‌هایی باشند که در مجموعه I دارای تعداد کلیدواژه توصیفگر باشند؛ ولی در مجموعه A دارای تعداد کلیدواژه نویسنده نباشد و همچنین در مجموعه T دارای کلیدواژه برچسب هم نباشد یا ممکن است هم باشد.

برچسب‌های تخصیص داده شده به مقاله‌های مورد مطالعه توسط برچسب‌گذاران در فاصله زمانی اول ژانویه سال ۲۰۱۴ تا سوم اوت سال ۲۰۱۶ به آن مقاله‌ها، از وبگاه مندلی استخراج گردید. کلیدواژه‌های تخصیص داده شده به مقاله‌ها توسط نویسندگان، از خود مقاله‌ها استخراج گردید. توصیفگرهای انتساب داده شده به هر مقاله توسط نمایه‌سازان هم از پایگاه اطلاعاتی اریک گردآوری شد.

میزان تطابق این سه زبان با استفاده از برنامه‌ای مبتنی بر زبان برنامه‌نویسی شیء گرا سی شارپ مورد تجزیه و تحلیل قرار گرفت. برای یکپارچه‌سازی داده‌ها و برای اینکه داده‌ها به شکل مرتب و منظم در یک قالب و فرمت یکپارچه شوند بایستی از زبان شیء گرا استفاده کرد. علاوه بر این، باید از زبانی استفاده شود که قابل فهم و قابل ذخیره در رایانه باشد. همچنین از آمار توصیفی به منظور مشخص شدن میانگین‌ها و فراوانی‌ها و نرم‌افزار اکسل ۲۰۱۳ به منظور تحلیل و ارزیابی داده‌ها استفاده شد (جدول ۲).

جدول ۲. نمونه چگونه بررسی میزان تطابق سه زبان

عنوان مقاله	کلیدواژه نویسنده	توصیفگرهای نمایه ساز	برچسب های برچسب گذار
-------------	------------------	----------------------	----------------------

<ul style="list-style-type: none"> - Articles - Assessment - MS -Simulation-Based Education -Simulation-Based Medical Educat - Systematic Review - Validation_ Studies - Validity - Validity Evidence 	<ul style="list-style-type: none"> - Health Personnel - Professional Education - Validity - Evidence - Simulation - Evaluation - Predictor Variables 	<ul style="list-style-type: none"> - Assessment - Educational technology - Evaluation - Medical education - Methods quantitative - Reliability - Reporting quality - Simulation - Validity 	<p>What Counts as Validity Evidence? Examples and Prevalence in a Systematic Review of Simulation-Based Assessment</p>
<ul style="list-style-type: none"> - CSCL - Dual Eye-Tracking - eye-tracking - Gaze Visualization - Network Analysis 	<ul style="list-style-type: none"> - Cooperative Learning - Network Analysis - Eye Movements - Learning Activities - Graphs - Students - Group Dynamics - Prediction - Correlation 	<ul style="list-style-type: none"> - Collaborative learning - Dual eye-tracking - Network analysis 	<p>Toward collaboration sensing</p>
<ul style="list-style-type: none"> - Case study - Chapter 1 - Education - Guided play - Joy - Learning - Literature review - Middle school - Mise en place - Play 	<ul style="list-style-type: none"> - Teacher Effectiveness - Middle School Teachers - Middle School Students - Teaching Methods - Play - Social Studies - Student Motivation - Student Development - Context Effect - Adolescent Development - Imagination - Creativity - Role Playing - Class Activities - Independent Study - Humor - Classroom Environment - Interviews - Student Surveys - Observation 	<ul style="list-style-type: none"> - Beginning teachers - Middle grades teaching - Play - Social studies teaching 	<p>Toward More Joyful Learning: Integrating Play into Frameworks of Middle Grades Teaching</p>

جدول ۲. نمونه چگونه بررسی میزان تطابق سه زبان (دنباله)

عنوان مقاله	اشتراک کلیدواژه نویسنده و توصیفگرهای نمایه‌ساز و برچسب‌های برچسب‌گذار	اشتراک توصیفگرهای نمایه‌ساز و برچسب‌های برچسب‌گذار	اشتراک کلیدواژه نویسنده و برچسب‌های برچسب‌گذار	اشتراک کلیدواژه نویسنده و توصیفگرهای نمایه‌ساز
Validity	Validity	Validity	Assessment Validity	Evaluation Simulation Validity
Network Analysis	Network Analysis	Dual Eye-	Network Analysis	Toward collaboration sensing

		Tracking Network Analysis		
Play	Play	Play	Play	Toward More Joyful Learning: Integrating Play into Frameworks of Middle Grades Teaching

۵. تجزیه و تحلیل یافته‌ها

الف. پرسش اول. میزان تطابق کلیدواژه‌های تخصیص داده شده توسط نویسندگان مدارک با برچسب‌های تخصیص داده شده توسط برچسب‌گذاران به همان مدارک در وبگاه مندلی چگونه است؟

برای پاسخ به این پرسش با توجه به جدول ۳ تعداد واژه‌های مشترک بین تعداد کلیدواژه‌های نویسنده و برچسب‌های برچسب‌گذاران ۴۰۸ واژه بوده است که این تعداد بر تعداد کل واژه‌ها تقسیم شده و نسبت اشتراک آن‌ها به صورت (۱۵ درصد) به دست آمده است.

جدول ۳. تطابق بین کلیدواژه نویسنده و برچسب‌های برچسب‌گذاران

تعداد کلیدواژه نویسنده	تعداد برچسب	جمع کل واژه‌ها	تعداد واژه‌های مشترک	نسبت اشتراک واژه‌ها به درصد
۱۴۶۸	۱۶۷۱	۲۷۳۱	۴۰۸	۱۵

ب. پرسش دوم. میزان تطابق توصیفگرهای تخصیص داده شده توسط نمایه‌سازان به مدارک در پایگاه اریک با برچسب‌های تخصیص داده شده توسط برچسب‌گذاران به همان مدارک در وبگاه مندلی تا چه اندازه است؟

برای پاسخ به این پرسش با توجه به جدول ۴ تعداد واژه‌های مشترک بین تعداد توصیفگرهای نمایه‌سازان و برچسب‌های برچسب‌گذاران ۱۹۷ واژه بوده است که این تعداد بر تعداد کل واژه‌ها تقسیم شده و نسبت اشتراک آن‌ها به صورت (۳ درصد) بیان گردیده است.

جدول ۴. تطابق بین توصیفگرهای نمایه‌سازان و برچسب‌های برچسب‌گذاران

تعداد توصیفگر	تعداد برچسب	جمع کل واژه‌ها	تعداد واژه‌های مشترک	نسبت اشتراک واژه‌ها به درصد
۴۷۷۶	۱۹۷۸	۶۵۵۷	۱۹۷	۳

ج. پرسش سوم. میزان تطابق توصیفگرهای تخصیص داده شده توسط نمایه‌سازان به مدارک در پایگاه اریک با کلیدواژه‌های تخصیص داده شده توسط نویسندگان مدارک به همان مدارک چگونه است؟

برای پاسخ به این پرسش با توجه به جدول ۵ تعداد واژه‌های مشترک بین تعداد توصیفگرهای نمایه‌سازان و برچسب‌های برچسب‌گذاران ۳۱۵ واژه بوده است که این تعداد بر تعداد کل واژه تقسیم شده و نسبت اشتراک آن‌ها به صورت (۴ درصد) به دست آمده است.

جدول ۵. تطابق بین کلید واژه نویسنده و توصیفگرهای نمایه سازان

تعداد کلیدواژه نویسنده	تعداد توصیفگر	جمع کل واژه‌ها	تعداد واژه‌های مشترک	نسبت اشتراک واژه‌ها به درصد
۲۱۶۷	۶۰۵۷	۷۹۰۹	۳۱۵	۴

د. پرسش چهارم. میزان تطابق کلیدواژه‌های تخصیص داده شده توسط نویسندگان مدارک و توصیفگرهای تخصیص داده شده توسط نمایه‌سازان به مدارک در پایگاه اریک و برچسب‌های تخصیص داده شده توسط برچسب‌گذاران به همان مدارک در وبگاه مندلی چگونه است؟ برای پاسخ به این پرسش با توجه به جدول ۶ تعداد واژه‌های مشترک بین تعداد کلیدواژه‌های نویسنده و توصیفگرهای نمایه‌سازان و برچسب‌های برچسب‌گذاران ۷۴ واژه بوده است که این تعداد بر تعداد کل واژه‌ها تقسیم شده و نسبت اشتراک آن‌ها به صورت (۱/۱ درصد) تعیین گردیده است.

جدول ۶. تطابق بین کلیدواژه‌های نویسنده و توصیفگرهای نمایه‌سازان و برچسب‌های برچسب‌گذاران

تعداد کلیدواژه نویسنده	تعداد برچسب	تعداد توصیفگر	جمع کل واژه‌ها	تعداد واژه‌های مشترک	نسبت اشتراک واژه‌ها به درصد
۱۴۶۸	۱۶۷۱	۳۹۶۰	۷۰۲۵	۷۴	۱/۱

ه. پرسش پنجم. میزان حضور کلیدواژه‌های تخصیص داده شده توسط نویسندگان مدارک و توصیفگرهای تخصیص داده شده توسط نمایه‌سازان و برچسب‌های تخصیص داده شده توسط برچسب‌گذاران در اصطلاحنامه اریک چگونه است؟ برای پاسخ به این پرسش با توجه به جدول ۷ میزان حضور توصیفگرهای تخصیص داده شده توسط نمایه‌سازان ۳۴ درصد است. که تعداد کل توصیفگرهای مقاله‌های بررسی شده این پژوهش ۱۷۰۶ واژه یکتا بوده است که تعداد ۵۸۶ واژه آن در اصطلاحنامه اریک حضور داشته است.

جدول ۷. میزان حضور واژه‌ها در اصطلاحنامه اریک به تفکیک

تعداد کل واژه‌ها	حضور واژه‌ها در اصطلاحنامه اریک	نسبت حضور در اصطلاحنامه اریک به درصد

نویسنده	۱۵۳۱	۳۷۹	۲۵
نمایه‌ساز	۱۷۰۶	۵۸۶	۳۴
برچسب‌گذار	۱۴۲۸	۳۵۲	۲۵

جدول ۸ اشتراک بین تعداد کلیدواژه‌های نویسنده و تعداد توصیفگرهای نمایه‌سازان و تعداد برچسب‌های برچسب‌گذاران به تفکیک مجله‌ها را نشان می‌دهد.

جدول ۸. اشتراک بین نویسنده و نمایه‌ساز و برچسب‌گذار به تفکیک مجله

درصد اشتراک	تعداد کل واژه‌ها	تعداد واژه‌های مشترک	تعداد مقاله‌هایی که اشتراک دارند	عنوان مجله	اشتراک بین نویسنده و برچسب‌گذار
۳۴/۵	۱۱۹	۴۱	۱۳	American Educational Research Journal	اشتراک بین نویسنده و برچسب‌گذار
۲۷/۳	۷۷	۲۱	۴	Reading Research Quarterly	اشتراک بین نمایه‌ساز و برچسب‌گذار
۱۰/۸۴	۸۳	۹	۶	Metacognition and Learning	اشتراک بین نویسنده و نمایه‌ساز
۷/۷	۱۵۶	۱۲	۶	Educational Researcher	اشتراک بین نویسنده و نمایه‌ساز و برچسب‌گذار

و. پرسش ششم. میزان فراوانی کلیدواژه‌های به کار رفته توسط سه گروه نویسنده، نمایه‌ساز و برچسب‌گذار چگونه است؟

برای پاسخ به این پرسش بسیاری از کلیدواژه‌های به کار رفته توسط نویسندگان و توصیفگرهای مورد استفاده نمایه‌سازان و برچسب‌های داده شده به مدارک توسط برچسب‌گذاران که در این پژوهش گردآوری شده است دارای فراوانی است. با توجه به جدول ۹ کلیدواژه «Educational Policy» دارای بیشترین فراوانی (۲۰ مرتبه) بوده که نویسندگان مدارک به کار برده‌اند. همچنین در جدول ۱۰ نمایه‌سازان از تخصیص توصیفگر «Correlation» که دارای فراوانی (۸۸ مرتبه) بوده، به مدارک در پایگاه اطلاعاتی اریک استفاده کرده‌اند. جدول ۱۱ نشان می‌دهد که برچسب‌گذاران از برچسب «Education» که دارای فراوانی (۲۸ بار) است، در وبگاه مندلی به مدارک اختصاص داده‌اند.

جدول ۹. بیشترین کلیدواژه به کار برده شده نویسندگان مدارک

کلیدواژه‌های نویسنده	فراوانی
Educational Policy	۲۰
Assessment	۱۵
Higher Education	۱۵
College Students	۱۳
Professional Development	۱۲
Equity	۱۲

Research Methodology	۱۱
Science Education	۱۰

جدول ۱۰. بیشترین توصیف‌های مورد استفاده نمایه‌سازان به مدارک در پایگاه اریک

توصیفگر	فراوانی
Correlation	۸۸
Foreign Countries	۷۱
Comparative Analysis	۶۵
Teaching Methods	۵۸
College Students	۵۷
Statistical Analysis	۵۷
Intervention	۵۵
Elementary School Students	۵۳

جدول ۱۱. بیشترین برچسب‌های مورد استفاده برچسب‌گذاران به مدارک در وبگاه مندلی

برچسب	فراوانی
Education	۲۸
Folder - Recent Educational Journal Articles	۱۶
Professional Development	۱۴
K	۱۱
Assessment	۱۰
Higher Education	۹
Proficiency	۹

جدول ۱۲ ترتیب کلیدواژه‌هایی را نشان می‌دهد که به مدارک حوزه علوم تربیتی در پایگاه اطلاعاتی اریک و وبگاه مندلی هر سه گروه «نویسندگان، نمایه‌سازان و برچسب‌گذاران» مورد

استفاده قرار گرفته است. در این بررسی کلیدواژه «Correlation» با فراوانی (۸۸ مرتبه) بیشتر مورد استفاده قرار گرفته است.

جدول ۱۲. بیشترین کلیدواژه‌های به کار رفته مدارک در پایگاه اریک و وبگاه مندلی

کلیدواژه نویسنده و توصیفگر و برچسب	فراوانی
Correlation	۸۸
College Students	۷۵
Foreign Countries	۷۲
Educational Policy	۶۹
Comparative Analysis	۶۵
Intervention	۶۲
Teaching Methods	۵۹
Statistical Analysis	۵۷
Academic Achievement	۵۴

۶. بحث

یافته‌های پژوهش حاضر نشان داد که میزان انطباق و اشتراک کلیدواژه‌های نویسنده و برچسب‌های برچسب‌گذار به میزان ۱۵ درصد وجود دارد. به عبارت دیگر، بین کلیدواژه‌های نویسنده و برچسب‌های برچسب‌گذار همخوانی و انطباق ضعیفی وجود دارد.

کیپ (۲۰۱۱) در مطالعه‌اش نشان داد که انطباق بین کلیدواژه‌های نویسنده و برچسب‌های برچسب‌گذاران به میزان ۳۳ درصد وجود دارد. نتایج این پژوهش از آن جهت که بین کلیدواژه‌های نویسنده و برچسب‌های برچسب‌گذاران انطباق وجود دارد و مثبت است، با پژوهش حاضر همخوانی دارد.

یافته‌های پژوهش حاضر در خصوص انطباق زبان نمایه‌سازی و برچسب‌گذار نشان داد که انطباق و اشتراک زبان نمایه‌ساز و برچسب‌گذار به میزان ۳ درصد وجود دارد. در مطالعه‌ای که محمدپور (۱۳۹۰) انجام داده بود، نشان داد که کمتر از ۲ درصد میان برچسب‌های اختصاص یافته به مقاله‌های علم اطلاعات و دانش‌شناسی (کتابداری) و توصیفگرهای تخصیص یافته به همان مقاله‌ها انطباق و اشتراک وجود دارد. نتایج این پژوهش با پژوهش حاضر همخوانی دارد.

کیپ (۲۰۱۱) در پژوهش خود نشان داد که انطباق بین توصیفگرها و برچسب‌های برچسب‌گذاران به میزان ۱۶ درصد وجود دارد. نتایج پژوهش وی با پژوهش حاضر همخوانی

دارد. یافته‌های پژوهش حاضر در خصوص انطباق توصیفگرهای نمایه‌سازی با کلیدواژه‌های نویسندگان مشخص نمود که انطباق و اشتراک زبان نمایه‌ساز و زبان نویسنده به میزان ۴ درصد وجود دارد. در مطالعه کیپ (۲۰۱۱) نشان داده شد که انطباق بین کلیدواژه‌های نویسنده و توصیفگرهای نمایه‌سازی به میزان ۱۹ درصد وجود دارد. نتایج پژوهش وی با پژوهش حاضر همخوانی دارد.

جالب توجه است که در مطالعاتی که در خصوص میزان تطابق و همخوانی کلیدواژه‌های پایان‌نامه‌ها و توصیفگرهای تخصیص داده شده به آن‌ها توسط نمایه‌سازان انجام شده است، نتایجی بر خلاف نتایج پژوهش‌ها در خصوص میزان تطابق و همخوانی کلیدواژه‌های مقاله‌های مجله‌ها و توصیفگرهای اختصاص داده شده به آن‌ها توسط نمایه‌سازان وجود دارد. برای نمونه، در مطالعه‌ای که بنی‌اقبال، خسروی و پیرهادی (۱۳۸۹) با عنوان «مقایسه واژه‌های عنوان و چکیده پایان‌نامه‌ها با توصیفگرهای تعیین شده در نمایه‌سازان اسناد و کتابخانه ملی ایران» انجام دادند، نتایج نشان داد که میزان همخوانی واژه‌های موجود در چکیده با توصیفگرها ۵۳/۵ درصد بوده است. نتایج این پژوهش با پژوهش حاضر متفاوت است. در تفسیر این نتایج می‌توان گفت که «نوع مدرک» ممکن است که در تطابق و همخوانی زبان نویسنده و نمایه‌ساز تأثیرگذار باشد.

مقایسه سه زبان نویسنده و نمایه‌ساز و برچسب‌گذار نشان داد که میزان انطباق زبان نویسنده و نمایه‌ساز و برچسب‌گذار به میزان ۱/۱ درصد وجود دارد. همچنین به منظور مشخص نمودن میزان حضور واژه‌های تخصیص داده شده مقاله‌ها توسط سه گروه نویسنده، نمایه‌ساز و برچسب‌گذار پس از بررسی مشخص شد که میزان حضور توصیفگرهای تخصیص داده شده توسط نمایه‌سازان در اصطلاحنامه اریک ۳۴ درصد می‌باشد که از دو گروه نویسنده و برچسب‌گذار بیشتر است. فراوانی برچسب‌های برچسب‌گذاران امکان شناسایی برچسب‌های پرسامد را برای کاربران و کتابداران (نمایه‌سازان) فراهم نموده و اجازه می‌دهد تا برچسب‌های داده شده به یک اثر توسط سایر کاربران نیز شناسایی شود که این می‌تواند نقش مهم و تعیین کننده‌ای در امر بازیابی اطلاعات داشته باشد.

۷. نتیجه‌گیری

بررسی یافته‌ها نشان داد که تطابق بیشتری بین زبان نویسنده و برچسب‌گذار نسبت به تطابق بین زبان نمایه‌ساز و برچسب‌گذار و همچنین تطابق بین زبان نویسنده و نمایه‌ساز وجود دارد. نتایج این پژوهش نشان می‌دهد که برچسب‌گذاران واژه‌هایی متفاوت از توصیفگرها و کلیدواژه‌های نویسندگان استفاده می‌کنند. همچنین کلیدواژه‌های تخصیص داده شده به مدارک توسط

نویسندگان به همان مدارک متفاوت از توصیفگرهایی است که نمایه‌سازان در پایگاه اطلاعاتی اریک به همان مدارک اختصاص می‌دهند. این نشان‌دهنده عدم آشنایی سه گروه «نمایه‌ساز، نویسنده و برجسب‌گذار» از زبان مورد استفاده همدیگر و به تبع آن ناهمخوانی در بازنمایی اطلاعات است. از دیگر اصطلاح‌هایی که تطابق نداشتند می‌توان در بسط اصطلاحنامه به عنوان ابزار نمایه‌سازی استفاده کرد تا این سه زبان را بیشتر بهم نزدیک نمود. همچنین از این اصطلاح‌ها برای تسهیل جستجو و پژوهش در مقاله‌های پایگاه‌ها استفاده نمود؛ زیرا در برجسب‌گذاری اجتماعی به برجسب‌گذاران اجازه داده می‌شود که عموماً و آزادانه با اختصاص کلیدواژه یا هر کلمه‌ای به منابع برجسب داده و همچنین نظرات شخصی خود درباره همین منابع برجسب‌گذاری شده را به اشتراک بگذارند.

در هر صورت در پایگاه‌ها برای افزایش مانعیت و کاهش جامعیت، اصطلاحنامه باید به شکل الکترونیکی در اختیار تمام کاربران قرار گیرد تا امکان همارایی و ترکیب واژه‌ها در مرحله نمایه‌سازی اجتماعی توسط کاربران و در مرحله جستجو و بازیابی اطلاعات فراهم شود. همارایی جزء قوی‌ترین ابزار مانعیت است. با افزایش تعداد اصطلاح‌ها در ترکیب با یکدیگر در نمایه‌سازی مفاهیم مورد نیاز با دقت بیشتری بازیابی می‌شود و اطلاعات اضافی حذف می‌گردد؛ و بدین‌سان از بازیابی مدارک نامربوط جلوگیری می‌شود. بنابراین، جستجوی سودمند و بازیابی اطلاعات در نظام‌های سنتی نمایه‌سازی در هنگام کنترل واژه‌ها، همارایی واژه‌ها (پیش‌همارا و پس‌همارا) بدون آشنایی با زبان کاربر و نویسنده، با مشکل مواجه می‌شود. ترکیب و تعامل بین زبان‌های کنترل شده و طبیعی می‌تواند فقدان اخص بودن در زبان کنترل شده را جبران کند و مانعیت را بهبود بخشد. در حالی که زبان کنترل شده می‌تواند جامعیت را از طریق کنترل مترادف‌ها و نمایش رابطه‌ها بهبود بخشد.

فهرست منابع

- بنی‌اقبال، ناهید و خسروی، فریبرز و پیرهادی، صدیقه (۱۳۸۹). مقایسه واژه‌های عنوان و چکیده پایان‌نامه‌ها با توصیفگرهای تعیین شده در نمایه سازمان اسناد و کتابخانه ملی ایران. فصلنامه مطالعات ملی کتابداری و سازماندهی اطلاعات، شماره ۸۶، ۱۳۴-۱۴۶.
- بهادری، لیلیا و بنی‌اقبال، ناهید (۱۳۹۲). میزان انطباق توصیفگرهای نمایه‌سازی و کلیدواژه‌های پایانی رشته پزشکی دانشگاه علوم پزشکی تهران با اصطلاحنامه پزشکی فارسی و MESH. فصلنامه دانش‌شناسی (علوم کتابداری و اطلاع‌رسانی و فناوری اطلاعات) سال هفتم، شماره ۲۴.

- حرّی، عباس (۱۳۹۱). *تقریرات درس روش تحقیق*. تهران: دانشگاه تهران، دانشکده علوم اطلاعات و دانش‌شناسی، نیمسال دوم ۹۲-۱۳۹۱.
- کیولند، دونالد بی، و کیولند، آنا دی. (۱۳۸۵). *درآمدی بر نمایه‌سازی و چکیده‌نویسی*. ترجمه مهدی حسینی. تهران: چاپار.
- محمدپور، معصومه (۱۳۹۰). بررسی میزان همپوشانی میان برجسب‌های مقاله‌های علم اطلاعات و دانش‌شناسی (کتابداری) سایت فولکسونومی *CiteUlike* و توصیف‌گرهای اختصاص یافته به همان مقاله‌ها در پایگاه *LISTA* پایان‌نامه کارشناسی ارشد کتابداری و اطلاع‌رسانی. دانشگاه علامه طباطبائی، تهران.
- نوروزی، علیرضا، منصوری، علی و حسینی، مهدی (۱۳۸۶). *رده‌بندی مردمی (فولکسونومی): سازماندهی دانش بر اساس خرد جمعی*. اطلاع‌شناسی، سال پنجم، شماره ۱ و ۲ (پیاپی ۱۷ و ۱۸)، ۱۵۱-۱۶۶.
- نوروزی، علیرضا، و ولایتی، خالد (۱۳۸۹). *نمایه‌سازی موضوعی: نمایه‌سازی مفهومی*. تهران: چاپار.
- نوروزی، علیرضا (۱۳۹۵). *برجسب و برجسب‌گذاری اجتماعی: مطالعات کتابداری و علم اطلاعات*، ۲۳ (۱۸).
- Bruce, R. (2008). Descriptor and folksonomy concurrence in education related scholarly research. *Webology*, 5(3), Article 59. Retrieved August 25, 2016, from <http://www.webology.ir/2008/v5n3/a59.html>
- ERIC (2016). ERIC. Retrieved August 25, 2016, from <https://www.eric.ed.gov>
- Kipp, M. E. (2011). User, Author and Professional Indexing in Context: An Exploration of Tagging Practices on CiteULike. *Canadian Journal of Information and Library Science*, 35(1), 17-48.
- Lee, D. H. (2015). *Comparative Analysis of Index Terms and Social Tags*. *한국문헌정보학회지*, *Journal of the Korean Library and Information Science Society*, 49(2), 291-311.
- Mendeley (2016). Mendeley. Retrieved August 25, 2016, from <https://www.mendeley.com>.
- Minguet, F., Salgado, T. M., van den Boogerd, L., & Fernandez-Llimos, F. (2015). Quality of pharmacy-specific Medical Subject Headings (MeSH) assignment in pharmacy journals indexed in MEDLINE. *Research in Social and Administrative Pharmacy*, 1(5), 686-695.
- Wikipedia Encyclopedia (2015). *Social network*. *Wikipedia, Free Encyclopedia*. Retrieved August 25, 2016, from http://en.m.wikipedia.org/wiki/Social_network.

consistency between descriptors, author-supported keywords and tags in the ERIC and Mendeley databases

Maryam Ghanavati

MA Student in Information Management; Department of Knowledge and Information Science; Faculty of Management; University of Tehran; Tehran, Iran; maryamghanavati@ut.ac.ir

Alireza Noruzi

PhD in Knowledge and Information Science; Assistant Professor; Department of Knowledge and Information Science; Faculty of Management; University of Tehran; Tehran, Iran; Corresponding Author: noruzi@ut.ac.ir

Maryam Nakhoda

PhD in Knowledge and Information Science; Assistant Professor; Department of Knowledge and Information Science; Faculty of Management; University of Tehran; Tehran, Iran; mnakhoda@ut.ac.ir

Ashkan Khatir

PhD Candidate in Information Technology; Iranian Research Institute for Information Science and Technology (IRANDOC); khatir@students.irandoc.ac.ir

Abstract

Objective: The purpose of this study was to identify the language consistency between indexers, authors and taggers in the ERIC and Mendeley databases.

Methodology: This survey was conducted using content analysis methods and techniques to evaluate the language consistency between indexers, authors and taggers in the ERIC and Mendeley databases and also to determine common keywords. The sample for this study was comprised of top twenty journals in the field of Educational Research based on the Journal Citation Reports (JCR) of Web of Science, indexed in the ERIC database in 2014. Finally 499 articles published in the above-mentioned journals in 2014 were chosen as the sample base for the dataset. Note that only articles with author-supported keywords, indexed in the ERIC database and also tagged in

the Mendeley database from January 2014 to August 2016 were eligible to be assessed. Descriptors assigned to the articles on the ERIC database and tags associated to the articles on the Mendeley database for the period from January 2014 to August 2016 were extracted. Also author-assigned keywords assigned to all 499 articles were collected. Finally we created a software based on object-oriented programming (OOP) in C++ to analyses the search results. Descriptive statistics and measures, and thesaural term comparison show that there are important differences in the context of keywords from the three groups.

Finding: This study demonstrated that there were differences between the tagger, author and professional indexer views of the words used as tags, descriptors, or author-assigned keywords. The results showed that the consistency between the author-supported keywords and user tags of the 499 articles in the Mendeley was 15 percent; while the consistency between descriptors designated to the articles in the ERIC database and user tags associated to the articles on the Mendeley were three percent. On the other hand, the consistency between descriptors assigned to the articles in the ERIC database and the author-assigned keywords were 4 percent. Finally, the language consistency between the three above-mentioned groups was 1.1 percent. Also note that the presence of descriptors in the ERIC thesaurus was 34 percent, which were more than the author-supported keywords and tags.

Conclusions: The findings showed that the consistency between the keywords used by authors and taggers were more than the keywords chosen by indexers and authors, and by indexers and taggers. This means that three sides of the information representation triangle, i.e., indexer, author and tagger are unfamiliar with each other's language. It is worth noting that tags are useful supplements to controlled vocabularies, since the former provide a means for social organization of knowledge outside the framework of the latter. The low consistency between tags and descriptors in this research indicates that Mendeley users do not use the same terminology as subject specialists who maintain descriptors in the ERIC thesaurus. Further research involving semantic analysis of Mendeley tags may reveal an emerging vocabulary suitable for inclusion in the ERIC thesaurus as a controlled vocabulary.

Keywords: Collaborative tagging; Social indexing; Social classification; Knowledge organization; ERIC; Mendeley; Folksonomies; Controlled vocabulary.

بررسی میزان تطابق زبان نمایه‌سازان، نویسندگان و برجسب‌گذاران در پایگاه اطلاعاتی اریک و مندلی | قنوتی و دیگران



مریم قنوتی: دارای مدرک کارشناس ارشد مدیریت اطلاعات از دانشگاه تهران است سازماندهی دانش، نمایه سازی، علم سنجی از جمله علایق پژوهشی ایشان است.



علیرضا نوروزی: متولد سال ۱۳۵۵، دارای مدرک تحصیلی دکتری در رشته علوم اطلاعات و ارتباطات از دانشگاه اکس ماری (پاول سزان) فرانسه است. ایشان هم اکنون استادیار گروه علوم اطلاعات و دانش شناسی دانشگاه تهران است. بازنمایی اطلاعات، نمایه سازی، ابر داده، علم سنجی، و وب شناسی از جمله علائق پژوهشی وی است.



مریم ناخدا: دارای مدرک دکتری در رشته علم اطلاعات و دانش شناسی از دانشگاه تهران است. وی هم اکنون به عنوان عضو هیات علمی و استادیار در گروه علم اطلاعات و دانش شناسی دانشگاه تهران فعالیت دارد. سازماندهی اطلاعات علمی تخصصی، مدیریت مراکز و خدمات اطلاع رسانی و مدیریت تغییر از جمله علایق پژوهشی وی است.



اشکان خطیر: متولد ۱۳۶۴ دانشجوی دکتری مهندسی فناوری اطلاعات در پژوهشگاه علوم و فناوری اطلاعات ایران. پردازش متن، داده کاوی و علم سنجی از جمله علایق پژوهشی ایشان است.

