

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; Associate 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

Iranian Journal of
**Information
Processing and
Management**

Received: 28, Nov. 2016

Accepted: 24, Jun. 2017

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

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

Iranian Research Institute
for Information Science and Technology
(IranDoc)

ISSN 2251-8223

eISSN 2251-8231

Indexed by SCOPUS, ISC, & LISTA

Vol. 33 | No. 4 | pp. 1729-1750

Summer 2018



created a software based on object-oriented programming (OOP) in C++ to analyze 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.

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.

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 برچسب‌گذاران در پایگاه اطلاعاتی اریک و مندلی

مریم فنواقی

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

علیرضا نوروزی

دکتری علم اطلاعات و دانش‌شناسی؛ دانشیار؛
گروه علم اطلاعات و دانش‌شناسی؛ دانشکده مدیریت؛
دانشگاه تهران؛
پدیده‌آور رابط 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



مقاله برای اصلاح به مدت ۴۰ روز نزد پدیدآوران بوده است.

پذیرش: ۱۳۹۷/۰۴/۰۳

دریافت: ۱۳۹۵/۰۹/۰۸

فصلنامه | علمی پژوهشی

پژوهشگاه علوم و فناوری اطلاعات ایران
(ایرانداک)

شاپا (چاپی) ۸۲۲۳-۲۲۵۱

شاپا (الکترونیکی) ۸۲۳۱-۲۲۵۱

نمایه در SCOPUS، ISC، LISTA، و

jipm.irandoc.ac.ir

دوره ۳۳ | شماره ۴ | صص ۱۷۱۷-۱۷۳۸

تابستان ۱۳۹۷



چکیده: هدف این پژوهش شناسایی میزان تطابق زبان نمایه‌سازان، نویسندگان و برچسب‌گذاران در پایگاه‌های 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 نویسندگان آن‌ها، از خود مقاله‌ها استخراج گردید. داده‌های گردآوری شده برای بررسی میزان تطابق این سه زبان با استفاده از برنامه‌ای به زبان برنامه‌نویسی شیء‌گرا «سی‌شارپ» 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 داده‌شده توسط نویسندگان مدارک با برچسب‌های تخصیص داده‌شده توسط برچسب‌گذاران به همان مدارک در وبگاه «مدلی» ۱۵ درصد، میزان تطابق توصیفگرهای تخصیص داده‌شده توسط نمایه‌سازان به مدارک در پایگاه «اریک» با برچسب‌های تخصیص داده‌شده توسط برچسب‌گذاران به همان مدارک در وبگاه «مدلی» ۳ درصد، میزان تطابق توصیفگرهای تخصیص داده‌شده توسط نمایه‌سازان به مدارک در پایگاه «اریک» با کلیدواژه‌های تخصیص داده‌شده 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 داده‌شده 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 نشان‌دهنده عدم آشنایی سه گروه «نمایه‌ساز، نویسنده و برچسب‌گذار» به زبان و واژگان مورد استفاده همدیگر است. پیشنهاد می‌شود که به منظور جست‌وجوی 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، مدلی

۱. مقدمه و بیان مسئله

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

نظام‌های بازیابی اطلاعات اغلب مبتنی بر زبان طبیعی یا زبان کنترل‌شده هستند.

جست‌وجو در یک نظام بازبایی اطلاعات چه با استفاده از زبان طبیعی و چه با استفاده از اصطلاح‌های کنترل‌شده فقط در صورتی نتایجی را به همراه خواهد داشت که کل یا بخشی از عبارت جست‌وجوی کاربر با کل یا بخشی از اصطلاح‌های نمایه‌ای در نظام بازبایی اطلاعات همخوان باشد. به همین دلیل است که نمایه‌سازی با جست‌وجو رابطه‌ای تنگاتنگ (کلیولند و کلیولند، ۱۳۸۰). از سوی دیگر، ابزار اصلی نمایه‌سازی به زبان کنترل‌شده، 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 جست‌وجوگران اطلاعات قرار می‌گیرند. در نتیجه، در یک نظام ذخیره و بازبایی اطلاعات، کاربر به‌طور دقیق نمی‌داند که چه می‌خواهد، یا نمی‌تواند آنچه را که می‌خواهد، به‌طور 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 بازبایی اطلاعات و یا به‌عبارتی، زبان نمایه‌سازی در پایگاه اطلاعاتی، ضمن این که 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 داده‌شده 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 این که نظام‌های رده‌بندی کتابخانه‌ای و نظام‌های سنتی نمایه‌سازی که مبتنی بر پشتوانه انتشاراتی^۲ هستند، پاسخ‌گوی نیازهای اطلاعاتی کاربران اینترنتی نیستند، 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 برجسب‌گذاری مشارکتی^۴ و

1. author-driven keywords
2. literary warrant
3. folksonomy
4. collaborative tagging

همکاری، نرم‌افزارهای اجتماعی و نشانه‌گذاری اجتماعی^۱ است و پشتوانه کاربری^۲ دارد، با کمک سایر فناوری‌های نسل دوم وب راه‌حلی نو برای خودسازمان‌دهی^۳ منابع اطلاعاتی وبی در هر جای جهان و با هر زبانی است. این نظام سازمان‌دهی اجتماعی دانش^۴ فرصت مناسبی را برای کاربران فراهم نموده تا منابع دلخواه خود را به‌صورت برخط از طریق اینترنت دسترس‌پذیر سازند و همچنین امکان شناسایی برچسب‌های^۵ پرسامد را برای کاربران و کتابداران (نمایه‌سازان) فراهم نموده و اجازه می‌دهد تا برچسب‌های داده‌شده به یک اثر را سایر کاربران شناسایی کنند (نوروزی ۱۳۹۵).

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

قابل ذکر است که یکی از مهم‌ترین و قدیمی‌ترین پایگاه‌های اطلاعاتی جهان، پایگاه اطلاعاتی علوم تربیتی (اریک)^۶ است. این پایگاه به لحاظ گردآوری، سازمان‌دهی و نمایه‌سازی اطلاعات، به‌ویژه مجله‌ها و منابع الکترونیکی، نسبت به سایر پایگاه‌ها ساختارمندتر و نظام‌مندتر بوده و دارای اصطلاحنامه تخصصی است. همچنین، پایگاه اطلاعاتی «مندلی»^۷ یکی از پایگاه‌ها و خدمات اطلاعاتی مبتنی بر رده‌بندی و نمایه‌سازی اجتماعی است که برای سازمان‌دهی اجتماعی، به‌اشتراک‌گذاری مقاله‌های علمی، یافتن داده‌های پژوهشی و همکاری برخط از آن استفاده می‌شود. عدم تطابق ساختار زبانی نویسنده و نمایه‌ساز (Bruce 2008) ما را بر آن داشت تا ساختار زبان کاربر (برچسب‌گذار)^۸ را با دو ساختار پیش‌گفته مورد مطالعه قرار دهیم. بنابراین، مسئله اصلی پژوهش حاضر این

1. social bookmarking
2. user warrant
3. self-organization of information
4. social organization of knowledge
5. tag
6. ERIC
7. Mendeley
8. tagger

است که آیا میان سه نوع ساختار زبانی مهارشده اصطلاحنامه‌محور، مهارنشده کاربرمدار و کلیدواژه‌های تخصیص داده‌شده به مدارک توسط نویسندگان تطابق و همخوانی وجود دارد؟ برای پاسخ به این پرسش، پژوهش حاضر به بررسی میزان تطابق زبان نمایه‌سازان، نویسندگان و برجسب‌گذاران در پایگاه اطلاعاتی «اریک» و وبگاه «مندلی» می‌پردازد.

۲. پرسش‌های اساسی پژوهش

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

۱. میزان تطابق کلیدواژه‌های تخصیص داده‌شده توسط نویسندگان مدارک با برجسب‌های تخصیص داده‌شده توسط برجسب‌گذاران به همان مدارک در وبگاه «مندلی» چگونه است؟
۲. میزان تطابق توصیفگرهای تخصیص داده‌شده توسط نمایه‌سازان به مدارک در پایگاه «اریک» با برجسب‌های تخصیص داده‌شده توسط برجسب‌گذاران به همان مدارک در وبگاه «مندلی» تا چه اندازه است؟
۳. میزان تطابق توصیفگرهای تخصیص داده‌شده توسط نمایه‌سازان به مدارک در پایگاه «اریک» با کلیدواژه‌های تخصیص داده‌شده توسط نویسندگان مدارک به همان مدارک چگونه است؟
۴. میزان تطابق کلیدواژه‌های تخصیص داده‌شده توسط نویسندگان مدارک و توصیفگرهای تخصیص داده‌شده توسط نمایه‌سازان به مدارک در پایگاه «اریک» و برجسب‌های تخصیص داده‌شده توسط برجسب‌گذاران به همان مدارک در وبگاه «مندلی» چگونه است؟
۵. میزان حضور کلیدواژه‌های تخصیص داده‌شده توسط نویسندگان مدارک و توصیفگرهای تخصیص داده‌شده توسط نمایه‌سازان و برجسب‌های تخصیص داده‌شده توسط برجسب‌گذاران در اصطلاحنامه «اریک» چگونه است؟
۶. میزان فراوانی کلیدواژه‌های به‌کاررفته توسط سه گروه نویسنده، نمایه‌ساز، و برجسب‌گذار 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» 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 برچسب‌های اختصاص یافته به مقاله‌های کتابداری و توصیفگرهای اختصاص یافته به همان مقاله‌ها همپوشانی وجود دارد (۱۳۹۰).

«بهادری و بنی‌اقبال» 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، در توصیفگرهای انگلیسی ۹۹/۷ درصد، در کلیدواژه‌های فارسی ۱۸ درصد و در کلیدواژه‌های انگلیسی ۲۶/۸ درصد همخوانی وجود داشته است (۱۳۹۲).

«کیپ» در پژوهش خود به «بررسی کاربر، نویسنده و نمایه‌سازان حرفه‌ای در محتوا: پژوهشی در عملکرد برچسب‌گذاری در سایت یولایک» 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 (Kipp 2011).

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، چشمه‌سهرابی، بنی‌اقبال» به «مطالعه تطبیقی کلیدواژه‌های فارسی و انگلیسی پایان‌نامه‌های دانشگاه علوم پزشکی اصفهان با اصطلاحنامه‌ها و سرعنوان‌های موضوعی پزشکی فارسی و انگلیسی» پرداختند. یافته‌ها نشان داد که بین میزان تطابق کلیدواژه‌های فارسی و انگلیسی با پزشکی فارسی و MeSH رابطه‌ای معنادار وجود دارد

1. LISTA

2. CiteUlike

و در این میان ساختار کلیدواژه‌ها بیشتر به سمت ساختار ترکیبی و تک‌واژه‌ای و بیشتر کلیدواژه‌ها نیز از چکیده انتخاب شده‌اند. همچنین، بیشتر کلیدواژه‌های انتخابی انگلیسی بودند (۱۳۹۱).

«رجیان» به «بررسی میزان همخوانی عبارت‌های جست‌وجوی 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»^۲، بیشترین میزان همخوانی در اصطلاحات «آی‌تریپل‌ای»، همخوانی نسبی (۵۱/۵ درصد)، در اصطلاحات کنترل‌نشده پایگاه اطلاعاتی «اینسپک»^۳، همخوانی کامل (۵۵/۹ درصد) و در کلیدواژه‌های نویسندگان، همخوانی نسبی (۴۸/۲ درصد) است. بین میزان همخوانی کلیدواژه‌های جست‌وجوی کاربران با کلیدواژه‌های نویسندگان، در پایگاه «ابسکو» و «آی‌تریپل‌ای»، تفاوت معناداری مشاهده شد (۱۳۹۱).

«لی» به «مقایسه تحلیلی اصطلاح‌های نمایه‌ای و برجسب‌های اجتماعی: سرعنوان موضوعی پزشکی در مقایسه با بیبسونومی»^۴ و دلشز^۵ پرداخت. یافته‌ها نشان داد که میزان تطابق «مش» با برجسب‌های داده‌شده کاربران در کل، ۳۴ درصد و میزان تطابق با «بیبسونومی» ۲۱/۷۷ درصد و میزان تطابق با «دلشز» ۸/۲۲ درصد است (Lee 2015).

«مینت» و همکاران «کیفیت سرعنوان‌های موضوعی پزشکی (مش) در حوزه داروسازی و تطابق آن با مجله‌های داروسازی که در مدلاین^۶ نمایه‌سازی شده‌اند» را مورد بررسی قرار دادند. نتایج جامعه آماری این پژوهش شامل ۶۹۸۹ مقاله از ۱۰ مجله داروسازی بود که از این میان ۳۲۸ مقاله (۴/۷ درصد) به‌طور کامل نمایه نشده بودند. در نتیجه، دارای اصطلاحی از «مش» نبودند. 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 به‌طور خاص در مورد داروسازی است، برای نمایه‌سازی آن از «مش» استفاده شده که

1. EBSCO
2. IEEE
3. INSPEC
4. bibSonomy
5. Delicious
6. Medline

حدود ۰/۹ درصد تطابق دارد (Minguet et al. 2015).

جدول ۱. جمع‌بندی پیشینه‌ها

مشخصات منبع	عنوان / هدف	جامعه	روش	یافته‌ها
محمدپور (۱۳۹۰)	بررسی میزان همپوشانی میان توصیفگرهای مقاله‌های کتابداری در پایگاه لیستا و پایگاه فولکسونومی سایت یولایک	مقاله‌های کتابداری منتشر شده در پایگاه لیستا و پایگاه فولکسونومی سایت یولایک	توصیفی-پیمایشی و تطبیقی مبتنی بر مشاهده	کمتر از ۲ درصد میان برچسب‌های اختصاص یافته به مقاله‌های کتابداری و توصیفگرهای اختصاص یافته به همان مقاله‌ها همپوشانی وجود دارد.
بهادری و بنی‌اقبال (۱۳۹۲)	میزان انطباق توصیفگرهای نمایه‌سازی و کلیدواژه‌های پایان‌نامه‌ای رشته پزشکی از سال ۱۳۸۵-۱۳۸۹	کلیه پایان‌نامه‌های رشته پزشکی از سال ۱۳۸۵-۱۳۸۹	پیمایشی به شیوه کتابسنجی مقایسه‌ای و کتابخانه‌ای است و ابزار گردآوری داده‌های چک‌لیستی.	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.
Kipp (2011)	کاربر، نویسنده و نمایه‌سازان حرفه‌ای در محتوا؛ پژوهشی در عملکرد برچسب‌گذاری در سایت یولایک	کتابداری و علم اطلاع‌رسانی	کتابداری و علم اطلاع‌رسانی روش‌های اطلاع‌سنجی	میزان انطباق بین برچسب‌گذاران ۳۳ درصد، میزان انطباق بین توصیفگرها و برچسب‌های برچسب‌گذاران ۱۶ درصد و میزان انطباق بین کلیدواژه نویسنده و توصیفگرها ۱۹ درصد بود.

۴. روش پژوهش

روش این پژوهش از لحاظ هدف، 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 دربردارد 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 «آ‌اس‌آ‌وب‌آ‌وساینس» 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 این که دارای کلیدواژه‌های نویسنده باشند. هنگام انتخاب مجله‌ها از بین ۳۰ مجله که بر اساس ضریب تأثیر اولویت‌بندی شده بودند، در سال ۲۰۱۴، ده مجله در پایگاه اطلاعاتی «اریک» نمایه‌سازی نمی‌شدند و فقط شماره‌های مربوط به سال ۲۰۱۳ و قبل از آن موجود بودند که در نهایت، ۲۰ مجله در سال ۲۰۱۴ انتخاب شد که هم در پایگاه «اریک» بودند 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 برجسی 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 بودند. در واقع، در مبحث مجموعه‌ها در علم ریاضی، پژوهش حاضر شامل سه مجموعه است. این سه مجموعه عبارت‌اند از:

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

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

$C =$ کلیدواژه‌های مجموعه مقاله‌های دارای برجسب‌های برجسب‌گذار که Z تعداد کلیدواژه دارد.

در این پژوهش با توجه به محدودیت‌های شرح داده‌شده 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 داده‌شده 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 داده‌شده به مقاله‌ها توسط نویسندگان، از خود مقاله‌ها استخراج گردید. توصیفگرهای منتسب 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 شی‌اگر «سی‌شارپ» مورد تجزیه و تحلیل قرار گرفت. برای یکپارچه‌سازی داده‌ها و برای این که داده‌ها به شکل مرتب و منظم در یک قالب و فرمت یکپارچه شوند، بایستی از زبان شی‌اگر 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 چگونگی بررسی میزان تطابق سه زبان

برچسب‌های برچسب‌گذار	توصیفگرهای نمایه‌ساز	کلیدواژه نویسنده	عنوان مقاله
- Articles	- Health Personnel	- Assessment	What Counts as Validity
- Assessment	- Professional Education	- Educational technology	Evidence? Examples
- MS	- Validity	- Evaluation	and Prevalence in a
- Simulation-Based	- Evidence	- Medical education	Systematic Review
Education	- Simulation	- Methods quantitative	of Simulation-Based
- Simulation-Based	- Evaluation	- Reliability	Assessment
Medical Educat	- Predictor Variables	- Reporting quality	
- Systematic Review		- Simulation	
- Validation_ Studies		- Validity	
- Validity			
- Validity Evidence			

برجسب‌های برجسب‌گذار	توصیفگرهای نمایه‌ساز	کلیدواژه نویسنده	عنوان مقاله
- CSCL	- Cooperative Learning	- Collaborative learning	Toward collaboration
- Dual Eye-Tracking	- Network Analysis	- Dual eye-tracking	sensing
- eye-tracking	- Eye Movements	- Network analysis	
- Gaze Visualization	- Learning Activities		
- Network Analysis	- Graphs		
	- Students		
	- Group Dynamics		
	- Prediction		
	- Correlation		
- Case study	- Teacher Effectiveness	- Beginning teachers	Toward More Joyful
- Chapter 1	- Middle School Teachers	- Middle grades teaching	Learning: Integrating Play
- Education	- Middle School Students	- Play	into Frameworks of Middle
- Guided play	- Teaching Methods	- Social studies teaching	Grades Teaching
- Joy	- Play		
- Learning	- Social Studies		
- Literature review	- Student Motivation		
- Middle school	- Student Development		
- Mise en place	- Context Effect		
- Play	- Adolescent Development		
	- Imagination		
	- Creativity		
	- Role Playing		
	- Class Activities		
	- Independent Study		
	- Humor		
	- Classroom Environment		
	- Interviews		
	- Student Surveys		
	- Observation		

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

عنوان مقاله	اشتراک کلیدواژه نویسنده و توصیفگرهای نمایه‌ساز	اشتراک کلیدواژه نویسنده و برچسب‌های	اشتراک توصیفگرهای نمایه‌ساز و برچسب‌گذار	اشتراک کلیدواژه نویسنده و توصیفگرهای نمایه‌ساز
What Counts as Validity Evidence? Examples and Prevalence in a Systematic Review of Simulation-Based Assessment	Validity	Assessment Validity	Validity	Evaluation Simulation Validity
Toward collaboration sensing	Network Analysis	Dual Eye-Tracking Network Analysis	Network Analysis	Network Analysis
Toward More Joyful Learning: Integrating Play into Frameworks of Middle Grades Teaching	Play	Play	Play	Play

۵. تجزیه و تحلیل یافته‌ها

الف. پرسش اول. میزان تطابق کلیدواژه‌های تخصیص داده‌شده توسط نویسندگان مدارک با برچسب‌های تخصیص داده‌شده توسط برچسب‌گذاران به همان مدارک در وبگاه «مندی» 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 داده‌شده توسط نمایه‌سازان به مدارک در پایگاه «اریک» با برچسب‌های تخصیص داده‌شده 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 داده‌شده توسط نمایه‌سازان به مدارک در پایگاه «اریک» با کلیدواژه‌های تخصیص داده‌شده 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 واژه‌ها تقسیم شده و نسبت اشتراک آن‌ها به صورت (۴ درصد) به دست آمده است.

جدول ۵. تطابق بین کلید واژه نویسنده و توصیفگرهای نمایه‌سازان

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 داده‌شده توسط نویسندگان مدارک و توصیفگرهای تخصیص داده‌شده توسط نمایه‌سازان به مدارک در پایگاه «اریک» و برچسب‌های تخصیص داده‌شده 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

Archive of SID

کلیدواژه‌های نویسنده و توصیفگرهای نمایه‌سازان و برچسب‌های گذاران 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 گذاران

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

۵. پرسش پنجم. میزان حضور کلیدواژه‌های تخصیص داده‌شده توسط نویسندگان مدارک و توصیفگرهای تخصیص داده‌شده توسط نمایه‌سازان و برچسب‌های تخصیص داده‌شده توسط برچسب‌گذاران در اصطلاحنامه «اریک» چگونه است؟
برای پاسخ به این پرسش با توجه به جدول ۷، میزان حضور توصیفگرهای تخصیص داده‌شده توسط نمایه‌سازان ۳۴ درصد است. که تعداد کل توصیفگرهای مقاله‌های بررسی‌شده این پژوهش ۱۷۰۶ واژه یکتا بوده است که تعداد ۵۸۶ واژه آن در اصطلاحنامه «اریک» حضور داشته‌اند.

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

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

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

برای پاسخ به این پرسش بسیاری از کلیدواژه‌های به‌کاررفته توسط نویسندگان و توصیفگرهای مورد استفاده نمایه‌سازان و برچسب‌های داده‌شده به مدارک توسط برچسب‌گذاران که در این پژوهش گردآوری شده، دارای فراوانی است. با توجه به جدول ۹، کلیدواژه "Educational Policy" دارای بیشترین فراوانی (۲۰ مرتبه) بوده که نویسندگان مدارک به‌کار 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 به‌کاررفته نویسندگان مدارک

کلیدواژه‌های نویسنده	فراوانی
Educational Policy	۲۰
Assessment	۱۵
Higher Education	۱۵
College Students	۱۳

کلیدواژه‌های نویسنده	فراوانی
Professional Development	۱۲
Equity	۱۲
Research Methodology	۱۱
Science Education	۱۰

جدول ۱۰. بیشترین توصیفگرهای 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 کاررفته مدارک در پایگاه «اریک» و وبگاه «مندی»

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» در مطالعه‌اش نشان داد که انطباق بین کلیدواژه‌های نویسنده و برچسب‌های برچسب‌گذاران به میزان ۳۳ درصد وجود دارد (Kipp 2011). 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» انجام داده بود، نشان داد که کمتر از ۲ درصد میان برچسب‌های اختصاص یافته به مقاله‌های علم اطلاعات و دانش‌شناسی (کتابداری) و توصیفگرهای تخصصی یافته به همان مقاله‌ها انطباق و اشتراک وجود دارد (۱۳۹۰). نتایج این پژوهش با

پژوهش حاضر همخوانی دارد.

«کیپ» در پژوهش خود نشان داد که انطباق بین توصیفگرها و برجسب‌های برجسب‌گذاران به میزان ۱۶ درصد وجود دارد (Kipp 2011). 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» نشان داده شد که انطباق بین کلیدواژه‌های نویسنده و توصیفگرهای نمایه‌سازی به میزان ۱۹ درصد وجود دارد (Kipp 2011). 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 داده‌شده به آن‌ها توسط نمایه‌سازان انجام شده، نتایجی بر خلاف نتایج پژوهش‌ها در خصوص میزان تطابق و همخوانی کلیدواژه‌های مقاله‌های مجله‌ها و توصیفگرهای اختصاص داده‌شده به آن‌ها توسط نمایه‌سازان وجود دارد. برای نمونه، در مطالعه‌ای که «بنی‌اقبال، خسروی و پیرهادی» با عنوان «مقایسه‌ی واژه‌های عنوان و چکیده‌ی پایان‌نامه‌ها با توصیفگرهای تعیین‌شده 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 (۱۳۸۹). 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.

مقایسه‌ی سه زبان نویسنده، نمایه‌ساز، و برجسب‌گذار نشان داد که میزان انطباق زبان نویسنده و نمایه‌ساز و برجسب‌گذار به میزان ۱/۱ درصد وجود دارد. همچنین، به‌منظور مشخص نمودن میزان حضور واژه‌های تخصیص داده‌شده مقاله‌ها توسط سه گروه نویسنده، نمایه‌ساز، و برجسب‌گذار پس از بررسی مشخص شد که میزان حضور توصیفگرهای تخصیص داده‌شده توسط نمایه‌سازان در اصطلاحنامه «اریک» ۳۴ درصد و از دو گروه نویسنده و برجسب‌گذار بیشتر است.

فراوانی برجسب‌های برجسب‌گذاران امکان شناسایی برجسب‌های پربسامد را برای کاربران و کتابداران (نمایه‌سازان) فراهم نموده و اجازه می‌دهد که برجسب‌های داده‌شده به یک اثر توسط سایر کاربران نیز شناسایی شود که این می‌تواند نقش مهم و تعیین‌کننده‌ای 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 داده‌شده 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 برچسب‌گذار» از زبان مورد استفاده همدیگر و به تبع آن ناهمخوانی در بازنمایی اطلاعات است. از دیگر اصطلاح‌هایی که تطابق نداشتند می‌توان در بسط اصطلاحنامه به‌عنوان ابزار نمایه‌سازی استفاده کرد تا این سه زبان را بیشتر به هم نزدیک نمود. همچنین، از این اصطلاح‌ها برای تسهیل جست‌وجو و پژوهش در مقاله‌های پایگاه‌ها می‌توان استفاده نمود؛ زیرا در برچسب‌گذاری اجتماعی به برچسب‌گذاران اجازه داده می‌شود که عموماً و آزادانه با اختصاص کلیدواژه یا هر کلمه‌ای به منابع برچسب داده و همچنین نظرات شخصی خود درباره همین منابع برچسب‌گذاری 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 جست‌وجو 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، جست‌وجوی 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 کنترل‌شده و طبیعی می‌تواند فقدان اخص بودن در زبان کنترل‌شده را جبران کند و مانعیت را بهبود بخشد، در حالی که زبان کنترل‌شده می‌تواند جامعیت را از طریق کنترل مترادف‌ها و نمایش رابطه‌ها بهبود بخشد.

فهرست منابع

- بنی‌اقبال، ناهید، فریبرز خسروی، و صدیق پیرهادی. ۱۳۸۹. مقایسهٔ واژه‌های عنوان و چکیدهٔ پایان‌نامه‌ها با توصیفگرهای تعیین‌شده در نمایهٔ سازمان اسناد و کتابخانه ملی ایران. *فصلنامه مطالعات ملی کتابداری و سازماندهی اطلاعات* ۸۶: ۱۳۴-۱۴۶.
- بهادری، لیلا و ناهید بنی‌اقبال. ۱۳۹۲. میزان انطباق توصیفگرهای نمایه‌سازی و کلیدواژه‌های پایان‌نامه‌ای رشتهٔ پزشکی دانشگاه علوم پزشکی تهران با اصطلاحنامه پزشکی فارسی و MESH. *فصلنامهٔ دانش‌شناسی (علوم کتابداری و اطلاع‌رسانی و فناوری اطلاعات)* ۷ (۲۴): ۵۳-۶۸.
- حرّی، عباس. ۱۳۹۱. تقریرات درس روش تحقیق. نوشتهٔ علیرضا نوروزی. دانشکدهٔ علوم اطلاعات و دانش‌شناسی، دانشگاه تهران.
- رجیبیان، حمیده. ۱۳۹۱. بررسی میزان همخوانی عبارت‌های جستجوی کاربران با اصطلاحات پیشنهادی مقاله‌ها پایگاه‌های اطلاعاتی لاتین و تحلیل ساختار واژگانی این اصطلاحات بر اساس روابط اصطلاحنامه‌ای. پایان‌نامه کارشناسی ارشد کتابداری و اطلاع‌رسانی. دانشگاه فردوسی مشهد. مشهد.
- کلیولند، دونالد بی.، و آنا دی. کلیولند. ۱۳۸۰. *درآمدی بر نمایه‌سازی و چکیده‌نویسی*. ترجمه مهدی حسینی. تهران: چاپار.
- محمدپور، معصومه. ۱۳۹۰. بررسی میزان همپوشانی میان برچسب‌های مقاله‌های علم اطلاعات و دانش‌شناسی (کتابداری) سایت فولکسونومی CiteULike و توصیفگرهای اختصاص یافته به همان مقاله‌ها در پایگاه ISTA. پایان‌نامهٔ کارشناسی ارشد کتابداری و اطلاع‌رسانی. دانشگاه علامه طباطبائی.
- نقته‌اصفهان، معصومه و مظفر چشمه‌سهرابی، و ناهید بنی‌اقبال. ۱۳۹۱. مطالعه تطبیقی کلیدواژه‌های فارسی و انگلیسی پایان‌نامه‌های دانشگاه علوم پزشکی اصفهان با اصطلاحنامه‌ها و سرعنوان‌های موضوعی پزشکی فارسی و انگلیسی. *مدیریت اطلاعات سلامت*. ۹ (۶): ۸۰۲-۸۱۳.
- نوروزی، علیرضا. ۱۳۹۵. برچسب‌گذاری اجتماعی. دایره‌المعارف کتابداری و اطلاع‌رسانی (داکا). ویرایش دوم. تهران: کتابخانه ملی جمهوری اسلامی ایران.
- _____، علی منصور، و مهدی حسینی. ۱۳۸۶. رده‌بندی مردمی (فوکسونومی): سازماندهی دانش بر اساس خرد جمعی. *اطلاعات‌شناسی* ۵ (۲۱): ۱۷ و ۱۸: ۱۵۱-۱۶۶.
- نوروزی، علیرضا، و خالد ولایتی. ۱۳۸۹. *نمایه‌سازی موضوعی: نمایه‌سازی مفهومی*. تهران: چاپار.
- Bruce, R. 2008. Descriptor and folksonomy concurrence in education related scholarly research. *Webology* 5 (3), Article 59. Retrieved from <http://www.webology.ir/2008/v5n3/a59.html> (accessed Aug. 25, 2016)
- ERIC. 2016. ERIC. Retrieved from <https://www.eric.ed.gov> (accessed Aug. 25, 2016)
- 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 from <https://www.mendeley.com>. (accessed Aug. 25, 2016)

Minguet, F., T. M. Salgado, L. van den Boogerd, & F. Fernandez-Llimos. 2015. Quality of pharmacy-specific Medical Subject Headings (MeSH) assignment in pharmacy journals indexed in MEDLINE. *Research in Social and Administrative Pharmacy* 686-695 : (5) 1 .

Wikipedia Encyclopedia. 2015. Social network. Wikipedia, *Free Encyclopedia*. Retrieved from http://en.m.wikipedia.org/wiki/Social_network. (accessed Aug. 25, 2016)

مریم قنواتی

دارای مدرک کارشناس ارشد مدیریت اطلاعات از دانشگاه تهران است. سازماندهی دانش، نمایه‌سازی، علم‌سنجی از جمله علایق پژوهشی ایشان است.



علیرضا نوروزی

متولد سال ۱۳۵۵، دارای مدرک تحصیلی دکتری در رشته علوم اطلاعات و ارتباطات از دانشگاه اکس مارسی (پاول سزان) فرانسه است. ایشان هم‌اکنون استادیار گروه علوم اطلاعات و دانش‌شناسی دانشگاه تهران است. بازنمایی اطلاعات، نمایه‌سازی، ابر داده، علم‌سنجی، و وب‌شناسی از جمله علایق پژوهشی وی است.



مریم ناخدا

دارای مدرک دکتری در رشته علم اطلاعات و دانش‌شناسی از دانشگاه تهران است. وی هم‌اکنون به‌عنوان عضو هیئت علمی و استادیار در گروه علم اطلاعات و دانش‌شناسی دانشگاه تهران فعالیت دارد. سازماندهی اطلاعات علمی تخصصی، مدیریت مراکز و خدمات اطلاع‌رسانی 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 است. پردازش متن، داده‌کاوی و علم‌سنجی از جمله علایق پژوهشی ایشان است.

