

میزان تخصیص اصطلاح های اخص در اصطلاحنامه ERIC

اطلاع رسانی شماره ۲ دوره ۱۲ زمستان ۱۳۷۵

نوشته ویرژیل دیودیتو*

ترجمه لیلا مرتضایی

مرکز اطلاعات و مدارک علمی ایران

کلیدواژه ها:

اصطلاحنامه اریک (ERIC) □ کاربست توصیفگر □ نمایه سازی □ اصطلاح اخص

چکیده:

آیا پیوندی بین کاربست و شباهت ظاهری آن با اصطلاح اخص وجود دارد؟ این بررسی ۱۲۹۶ توصیفگر اصطلاحنامه اریک را که حداقل یک اصطلاح اخص دارد شامل می شود. تعداد مدارکی که این اصطلاح های مدخل در آنها بکار رفته نسبت به اصطلاحات اخصشان سنجیده شده است. تقریباً هیچ همبستگی بین این مسئله که اصطلاحی هر چند وقت بکار رفته و این مطلب و این مطلب که چند اصطلاح اخص با نشانه ظاهری داراست وجود نداشت. معهداً اصطلاح های اخص زیادی بودند که میزان کاربست آنها در مدارک بسیار بالاتر از اصطلاح های مدخلشان بود. بسیاری از این اصطلاح های اخص نسبت به اصطلاح مدخل خود در اصطلاحنامه، جدیدتر بودند.

مقدمه

زمانی که نمایه ساز اصطلاح هائی را به مدرک اختصاص می دهد این پرسش مطرح می شود که چرا این اصطلاحات انتخاب شده اند و چرا اصطلاح های دیگر نه؟ پاسخ مناسب این است که اصطلاح های انتخاب شده به بهترین صورت محتوای مدرک را نشان می دهند و متحمل ترین اصطلاح هایی است که محقق برای جستجوی محتوای اطلاعاتی مدرک به کار می برد.

هدف این تحقیق پاسخگویی به این پرسش است: آیا در نمایه سازی، اصطلاحی به این دلیل ساده انتخاب می شود که در اصطلاحنامه دارای واژه های اخص با تشابه ظاهری است؟ در مقاله حاضر و قبلی [۳] اصطلاح های به کار رفته در نظام بازیابی اطلاعات اریک را سنجیدیم. در تحقیق حاضر داده ها را از ویرایش دوازدهم اصطلاحنامه اریک [۵]، که در این مقاله برای سهولت "اصطلاحنامه" قید می شود گرد آورده ایم.

تعاریف:

اصطلاح مدخل^(۱): واژه یا عبارتی که در ترتیب الفبایی "اصطلاحنامه" مدخلی به آن اختصاص یافته است.

* - Virgil Diodato. "The rates of assignment of narrower Terms in the Thesaurus of ERIC Descriptors". Journal of Information Science. 19(1993) 137-141.

اصطلاح اخص^(۳): اصطلاحی که با نشانه NT در اصطلاحنامه مشخص شده است. برای مثال "تغییر نگرش"، "تغییر شغل"، "رویدادهای زندگی" جزء اصطلاح های اخص مدخل "تغییر" هستند.

اصطلاح اخص شامل^(۳) (NIT): اصطلاح اخصی که حداقل نیمی از واژه یا عبارت (خواه یا واژه) اصطلاح مدخل را شامل می شود. برای نمونه "تغییر نگرش"، "تغییر شغل" اصطلاح های اخص شامل (NIT) برای اصطلاح "تغییر" محسوب می شود. ولی "رویدادهای زندگی" اگر چه اصطلاح اخص است، اصطلاح اخص شامل نیست.

نسبت شمول اصطلاح مدخل^(۴): عبارتست از تعداد اصطلاح های اخص شامل (NIT) در هر مدخل تقسیم بر تعداد خانه های کل اصطلاحات اخص (NT) آن مدخل. بعنوان مثال در مدخل "تغییر" در میان ۱۴ اصطلاح اخص ۱۰ اصطلاح اخص شامل (NIT) وجود دارد. لذا نسبت شمول اصطلاح مدخل در این مورد ۱۰/۱۴ یا ۰/۷۱ است.

عمر اسنادی اصطلاح^(۵): تعداد مدارکی که در طول عمر اصطلاح در پایگاه اطلاعاتی اریک نمایه سازی شده است (خواه اصطلاح مورد نظر به مدرک اختصاص یافته باشد یا نباشد). پایگاه اطلاعات اریک مشتمل است بر فهرست جاری مجلات علوم تربیتی (CIJE)^(۶) و منابع علوم تربیتی^(۷) (RIE). برای ویرایش دوازدهم اصطلاحنامه اریک، عمر اسنادی عبارتست از فاصله زمانی بین تاریخ افزوده شدن اصطلاح به سیستم تا فاصله زمانی بین تاریخ افزوده شدن اصطلاح به سیستم تا آوریل ۱۹۹۰. به عنوان مثال اصطلاح "تغییر" در مارس ۱۹۸۰ وارد "اصطلاحنامه" شد، از آن زمان تا آوریل ۱۹۹۰ نظام اریک ۳۲۵۷۸۷ مدرک را نمایه سازی کرده است.

موقعیت اصطلاح^(۸): عبارت است از تعداد دفعاتی که یک اصطلاح در طول عمر خود به مدرکی که وارد پایگاه اطلاعاتی اریک شده اند، اختصاص یافته است.

اصلاح مدخل	الف	ب	ج	د	ه	و	ز	ح
عمر اسنادی اصطلاح	موقعیت اصطلاح	نسبت کاربرد اصطلاح مدخل به مدرک (PPD)	تعداد اصطلاح های اخص (NT)	تعداد اصطلاح های اخص شامل (NT)	نسبت میانگین شمول اصطلاح مدخل	PPD	نسبت قدرت مجموعه اصطلاح اخص	
۳۲۵۷۸۷	۶۳۹	۰/۰۰۱۹۶	۱۴	۱۰	۰/۷۱	۰/۰۰۵۵۲	۲/۸۲	
۷۰۷۴۴۳	۱۰۰۲۸	۰/۰۱۴۱۷	۲۹	۱۹	۰/۶۶	۰/۰۰۲۳۷	۰/۱۷	

جدول ۱- اطلاعات در مورد دو اصطلاح مدخل "تغییر" و "برنامه درسی" که در اینجا به عنوان نمونه انتخاب شده اند.

نسبت کاربرد اصطلاح به مدرک^(۹): موقعیت اصطلاح تقسیم بر عمر اسنادی اصطلاح است. نسبت کاربرد اصطلاح به مدرک (PPD) معیار سنجش ما از میزان تخصیص اصطلاح توسط نمایه ساز یا کاربرد آن اصطلاح است.

نسبت قدرت مجموعه اصطلاح های اخص^(۱۰): عبارت است از میانگین تعداد دفعاتی که کلیه اصطلاح های اخص یک اصطلاح مدخل به مدارک اختصاص یافته اند تقسیم بر تعداد دفعاتی که همان اصطلاح مدخل در مدارک بکار رفته است.

براساس تعاریف فوق جدول (۱) نشان دهنده اطلاعات به دست آمده در مورد ۲ نمونه از اصطلاحات اریک است.

سایر مطالعات انجام شده در این زمینه

واینبرگ^(۱۱) و کانینگهام^(۱۲) [۶] موقعیت اصطلاح و تاریخ افزوده^(۱۳) و سطوح سلسله مراتبی اصطلاح های "محیط زیست" و "بیماریهای غدد درون ریز" را در سلسله مراتب (به ترتیب فرعی و اصلی) سیستم مدلاین، و اصطلاح های "تکنولوژی" را به عنوان سلسله فرعی در سیستم اریک مورد بررسی قرار دادند. آنها در مورد "محیط زیست" همبستگی منفی چشم گیری بین موقعیت اصطلاح و سطح سلسله مراتبی آن و همبستگی مثبت بارزی بین سطح سلسله مراتبی و تاریخ افزوده مشاهده نمودند. در مورد "تکنولوژی" و "محیط زیست" همبستگی منفی قابل ملاحظه‌ای بین موقعیت اصطلاح و تاریخ افزوده، و در مورد "بیماریهای غدد درون ریز" همبستگی منفی عمده ای بین سطح سلسله مراتبی و موقعیت اصطلاح دیده شد.

بکاس^(۱۴) و دیگران [۱] هیچ تفاوت فاحشی از جهت موقعیت اصطلاح در مورد پانزده اصطلاح موجود و پانزده اصطلاح جدید در سیستم مدلاین مشاهده نکردند. ایست من^(۱۵) [۲] در سیستم مدلاین به دنبال تخصیص همزمان نمایه ها به مدرک گشته تا دریابد آیا اصطلاح ها نسبت به یکدیگر رابطه اعم و اخص دارند و یا همه اصطلاح های اخص یک اصطلاح اعم اند. معلوم شد در سلسله مراتب "محیط زیست" و "بیماریهای غدد درون ریز" رابطه به صورت اول (رابطه پدر فرزندی) بیشتر اتفاق می افتد تا به شکل دوم (رابطه خواهر برادری).

در مطالعه پیشین در سیستم اریک [۳] کلیه سلسله مراتبی را که بیش از یک اصطلاح داشتند آزمایش کردیم. مشخص شد میزان کاربست اصطلاح ها به مدارک به ندرت مربوط به سطحی است که اصطلاح در نظام سلسله مراتب اشغال کرده است. یافته هایی از این نوع ما را به این نتیجه رساند که صرف جای یک اصطلاح در نظام سلسله مراتبی در کاربست آن تاثیر ندارد. نمایه سازان در تخصیص اصطلاح از کلیه سطوحی سلسله مراتبی، بدون توجه به میزان اخص یا اعم بودن آن، بخوبی بهره می گیرند.

روش شناسی

پرسش:

آیا این واقعیت دارد که وجود یک یا چند اصطلاح اخص در یک مدخل که از جهت ظاهر شبیه واژه مدخل باشند، نمایه ساز را به سوی انتخاب واژه اخص، به جای واژه اعم، سوق می دهد؟ برای یافتن پاسخ، فرضیه هایی مطرح کردیم که جواب آنها مثبت بود. برای اثبات این مدعا که وجود اصطلاح اخص شامل "با کاهش تخصیص اصطلاح مدخل و افزایش کاربرد اصطلاح های اخص همراه است راههای مختلفی وجود دارد. این مطلب به طرح چهار فرضیه انجامید. هر فرضیه فقط اصطلاح هایی را شامل شد که در ترتیب الفبایی اصطلاحنامه، بعنوان مدخل، یک یا چند اصطلاح اخص داشتند.

فرضیه ها

فرضیه ۱- همبستگی منفی بین نسبت شمول اصطلاح مدخل و نسبت کاربست اصطلاح به مدرک (PPD) وجود دارد. (در جدول ۱- ستون "و" در مقابل ستون "ج")

فرضیه ۲- همبستگی منفی بین تعداد اصطلاح های اخص شامل (NIT) از یک اصطلاح مدخل و نسبت کاربست اصطلاح به مدرک (PPD) دیده می شود. (در جدول ۱- ستون "ه" در برابر ستون "ج")
فرضیه ۳- همبستگی مثبت بین نسبت شمول اصطلاح مدخل و نسبت قدرت مجموعه اصطلاح های اخص به چشم می خورد (در جدول ۱- ستون "و" در مقابل ستون "ح")
فرضیه ۴- همبستگی مثبت بین تعداد اصطلاح های اخص شامل (NIT) از یک اصطلاح مدخل و "نسبت قدرت اصطلاح های اخص" برقرار است. (در جدول ۱- ستون "ه" در برابر ستون "ح")

نتایج:

اطلاعات توصیفی

۱۲۹۶ اصطلاح که در اصطلاحنامه حداقل دارای یک اصطلاح اخص بودند تجزیه و تحلیل کردیم. میانگین "موقعیت اصطلاح" در مورد این ۱۲۹۶ اصطلاح ۰/۰۰۳۹۴ یا به بیان دیگر ۳۹۴ مورد کاربست اصطلاح در هر ۱۰۰۰۰۰ مدرک نمایه شده بود. همچنین میانگین تعداد اصطلاح های اخص این مدخلها ۴/۱۱ بود که ۲/۴۳ آنها اصطلاح های اخص شامل (NIT) بودند. این میزان معادل ۰/۵۹ نسبت شمول است. در پایان میانگین نسبت قدرت مجموعه اصطلاح های اخص ۱/۶۵ بود.
به بیانی نه چندان دقیق در مورد مدخلهایی با اصطلاح های اخص تقریباً ۶۰٪ اصطلاح های اخص شبیه اصطلاح های مدخل خود بودند. میزان استفاده نمایه سازان از اصطلاح های اخص یک مدخل، ۱/۶۵ بار بیشتر از میزان استفاده خود اصطلاح مدخل بود.

آزمون فرضیه

نتیجه اساسی بدست آمده آن بود که تقریباً هیچ تاییدی برای همبستگی های بین وجود اصطلاح های اخص شامل و تخصیص و یا اصطلاح مدخل آنها وجود نداشت.
چهار همبستگی مرتبط با چهار فرضیه، بسیار پایین و به ترتیب عبارت بودند از ۰/۱۰، ۰/۰۱ و ۰/۰۲. سطوح معنی دار برای همبستگی های اول و سوم و چهارم همگی بیش از ۰/۱۰ بود. به روشنی هیچ تاییدی برای چهار فرضیه، به لحاظ نشانه ها و یا همبستگی های ناچیز، وجود ندارد.
معهداً برای آزمون فرضیه دوم همبستگی مثبت ناچیز ۰/۱۰ متمایز از صفر برای سطح معنی دار ۰/۰۱ وجود داشت. لذا تمایل ناچیزی برای مدخل هایی که دارای اصطلاح های اخص بیشتری بودند وجود داشت که از اصطلاح مدخلشان استفاده شود. (در قیاس با استفاده از آن دسته از اصطلاح های مدخل که اصطلاح اخص کمتری دارند).

بحث:

مانند مطالعه پیشین نتوانستیم این مطلب را که استفاده از اصطلاح توسط نمایه ساز در ارتباط با ویژگی فیزیکی عادی آن اصطلاح است اثبات کنیم. در مطالعه حاضر تقریباً هیچ شاهدهی دال بر اینکه تشابه ظاهری بین اصطلاح اخص و اصطلاح مدخل ارتباطی با استفاده از آن اصطلاحها داشته باشد، موجود نبود.
جالب است که در میان یک مدخل با اصطلاح های اخص، تمایل به استفاده از اصطلاح های اخص بیش از اصطلاح مدخل بود. این مطلب دلالت بر میانگین نسبت قدرت ۱/۶۵ که در بالا ذکر شد، شده یافتند. مذکور بر این نکته گواهی می دهد که نمایه سازان دقیق ترین اصطلاح را که مناسب مفهوم است جستجو می کنند.

تجزیه و تحلیل ما ممکن است با این واقعیت که هر اصطلاح اخص در یک مدخل، خود همچنین، در جای دیگر اصطلاحنامه، یک اصطلاح مدخل است مغشوش شود. ماهیت بهم پیوسته اصطلاحات ممکن است ما را از جستجوی الگویی که فرض کرده ایم باز دارد. نقش شاخص ها در اصطلاحنامه اریک نیز ممکن است یک عامل باشد. معهدا این مطلب نتایج بدست آمده را روشن نمی کند. آسانتر آن است که معتقد شویم نتایج به سادگی مبین این حقیقت است که نمایه سازان اصطلاح ها را به دلایل خوبی غیر از تظاهر فیزیکی اصطلاح انتخاب می کنند.

نسبت قدرت نهایی

حتی وقتی فرضیه ها رد می شود ممکن است نتایج جالبی بدست آید به عنوان مثال: اصطلاح های مدخل با پائین ترین و بالاترین نسبت های توان ما را متعجب کرد. جدول ۲، ده اصطلاح را که دارای بالاترین و پایین ترین نسبت های قدرت بودند نشان می دهد. بیشترین عبارت بود از تنها اصطلاح اخص ذیل اصطلاح "آشنایی معلم" که ۸۴ بار بیشتر از خود اصطلاح "آشنایی معلم" بکار رفته بود. از سوی دیگر کمترین ها مربوط به اصطلاح های اخص ذیل "دست برتری" و سه اصطلاح دیگر بود که نسبت ۵ هزارم را نسبت به اصطلاح مدخلشان دارا بودند.

آنچه موجب حیرت ما در مورد این دامنه وسیع بین نسبت های قدرت شد، الگویی بود که ما به دنبال آن بودیم: رابطه بین تاریخ افزوده یک اصطلاح مدخل و تاریخ های افزوده اصطلاح های اخص. جدول ۲ نشان می دهد که در بین ۱۰ اصطلاحی که بیشترین نسبت های قدرت را دارا بودند، ۹ تای آنها دارای یک یا بیشتر اصطلاح اخص بودند که سالها (غالباً چند دهه) بعد از اصطلاح مدخلشان به اصطلاحنامه افزوده گردیده بودند. از سوی دیگر، در میان اصطلاحاتی که کمترین نسبت های قدرت را داشتند ۶ تا از ۱۰ اصطلاح، اصطلاح هایی بودند که اصطلاح اخصشان در طول یک سال بعد از ورودشان به اصطلاحنامه پدیدار شده بود.

به ویژه در مورد ۸ تا ۱۰ اصطلاحی که بالاترین نسبت های قدرت را دارا بودند، یک اصطلاح اخص وجود داشت که نخستین بار در آوریل ۱۹۹۰ وارد سیستم شده بود، حال آنکه اصطلاح مدخل نخستین بار جولای ۱۹۶۶ ویا ژوئن ۱۹۶۹ ظاهر گردید. چنین اصطلاح های اخصی قوی ترند زیرا تازه و زنده اند. به عنوان مثال اصطلاح "تمرین نوشتن" از جولای ۱۹۶۶ در اصطلاحنامه بوده است در ویرایش جدید اصطلاحنامه این مدخل دارای اصطلاح اخص "تکلیف نوشتن" شده است. این نوع خاص از "تمرین نوشتن" اخیراً ۳۲ بار بیش از دفعاتی که اصطلاح مدخل کهنه در طول عمرش، که یک ربع قرن از آن می گذرد، استفاده شده است. این مطلب متناسب است. زیرا اداره کنندگان واژگان اریک نیاز به یک اصطلاح خاص تر را احساس کردند. اطلاعات ما نشان می دهد که نمایه سازان نیز در این مورد توافق دارند.

تاریخ افزوده		نسبت توان	اصطلاح اخص شامل	اصطلاح اخص	اصطلاح مدخل
اصطلاح اخص	اصطلاح مدخل				
۹۰/۴	۶۶/۷	۸۴/۲۹	۱	۱	بالاترین ها آشنایی معلم
۹۰/۴ ۹۰/۴	۶۶/۷	۷۰/۹۶	۰	۲	تاریخ آسیا
۹۰/۴ ۶۶/۷	۶۹/۶	۶۷/۵۲	۱	۲	قصد
۹۰/۴ ۹۰/۴	۶۶/۷	۵۵/۴۳	۲	۲	تاریخ جهان
۹۰/۴ ۸۳/۱۰ ۷۲/۱۰ ۶۹/۷	۶۶/۷	۴۶/۶۵	۷	۱۲	واژگان
۸۰/۳ ۶۶/۷	۸۰/۳	۴۲/۵۹	۱	۲	پایان نامه
۹۰/۴	۶۶/۷	۳۷/۶۵	۰	۱	برتری جانبی
۹۰/۴	۶۶/۷	۳۴/۱۰	۰	۱	توانایی شناختی
۸۴/۳	۷۱/۸	۳۳/۶۴	۰	۱	زیستار شناختی
۹۰/۴	۶۶/۷	۳۲/۷۰	۱	۱	تمرین نوشتن پایین ترین ها
۶۶/۷	۹۰/۴	۰/۰۰۵	۱	۱	دست برتری
۶۶/۷	۹۰/۴	۰/۰۰۵	۰	۱	ژاپنی
۶۶/۷	۶۶/۷	۰/۰۰۵	۰	۱	فعالیت یادگیری
۷۱/۱	۷۰/۹	۰/۰۰۵	۰	۱	متغیرهای پیش‌بینی کننده
۸۰/۷	۶۶/۷	۰/۰۰۶	۱	۱	جدول (داده ها)
۸۰/۳	۶۶/۷	۰/۰۰۷	۱	۱	ارزشیابی دانش‌آموز
۶۶/۷	۶۶/۷	۰/۰۰۸	۱	۱	مهارت پایه
۶۹/۹	۶۶/۷	۰/۰۱۱	۰	۱	تاثیر زیستی
۶۶/۷	۸۰/۳	۰/۰۱۱	۰	۱	مدرس پاره وقت
۶۶/۷	۶۶/۷	۰/۰۱۳	۰	۱	آلمانی

جدول ۲- اصطلاحاتی که بالاترین و پایین ترین نسبت توان را دارا هستند

منابع:

- 1- J. E. B. Backus, S. Davidson and R. Rada, Searching for patterns in the MeSH vocabulary, *Bulletin of the Medical Library Association* 75 (1987) 221-227.
- 2- C. M. Eastman, Overlaps in postings to thesaurus terms: a preliminary study, *Proceedings of the Annual Meeting of the American Society for Information Science* 25 (1988) 181-184.
- 3- G. Henry and V. Diodato, The rates of assignment of thesaurus terms in the ERIC information retrieval system: an analysis of hierarchies and levels, *Journal of Documentation* 47 (1991) 276-283.
- 4- I. S. Simpon, *Basic Statistics for Librarians* (Library Association, London, 1988).

- 5- Thesaurus of ERIC Descriptors (Oryx, Phoenix, AZ, 1990).
- 6- B. H. Weinberg and J. A. Cunningham. The relationship between term specificity in MeSH and online postings in MEDLINE, *Bulletin of the Medical Library Association* 73 (1985) 365-372.

پی نویس:

- 1- Entry Term
- 2- Narrower Term
- 3- Narrower Including Term
- 4- Inclusion ratio of an entry Term
- 5- Lifetime documents of a Term
- 6- Current Index To Journals in Education
- 7- Resources in Education
- 8- Postings of a Term
- 9- Postings Per document ratio of a Term (PPD)
- 10- Strength ratio of a set of narrower Terms
- 11- Weinberg
- 12- Cunningham
- 13- Added data تاریخ افزوده نخستین تاریخی است که اصطلاح برای نخستین بار در سیستم اریک وارد شده است
- 14- Backus
- 15- Eastman