



رویکردی معناشناسانه به ترجمه اصطلاحات علمی و تأثیر آن بر بازیابی اطلاعات:
مطالعه موردی حوزه علم اطلاعات و دانش‌شناسی

شعله ارسطوپور^۱

گروه ارزیابی و توسعه منابع، مرکز منطبق‌های اطلاع‌رسانی علوم و فناوری

فاطمه احمدی‌نسب

گروه زبان و ادبیات انگلیسی، دانشگاه شهید چمران اهواز

چکیده. زبان فارسی در رابطه با اکثر حوزه‌های دانش نقش مصرف‌کننده را داشته و لذا بسیاری از واژگان تخصصی رشته‌های مختلف از طریق ترجمه، غالباً به صورت انفرادی و بدون توجه به روابط مفهومی توسط متخصصان و مترجمان وارد این زبان شده‌اند، و لذا تعدد و آشفتگی معنایی نیز در سطح اصطلاحات علمی قابل مشاهده است. چالش‌های معنایی اصطلاح‌گزینی را به صورت خلاصه می‌توان در قالب هم‌معنایی فراوان، چندمعنایی، تداخل معنایی و... مشاهده نمود. بدیهی است بروز چنین چالش‌هایی نه تنها به خودی خود منجر به بروز مشکلاتی در انتقال مفاهیم ارائه شده در متون تخصصی توسط مخاطبان این دست از نوشتارها می‌شود، بلکه در عمل تأثیر جدی بر فرایندهای مرتبط با بازیابی اطلاعات نیز دارد. طراحان نظام‌های کنونی بازیابی اطلاعات تلاش می‌کنند نظام‌هایی را توسعه دهند که ضمن رفع حداکثری ابهام‌های واژگانی و معنایی، بهترین و مرتبط‌ترین نتایج را به کاربران ارائه نماید. در حال حاضر چالش‌های حاصل از این دست ابهام‌های معنایی، به صورت مستقیم در دو حوزه اصلی بازیابی موضوعی اطلاعات و نظام‌های بازیابی بین‌زبانی تأثیر گذارند. در این مطالعه تلاش بر آن است تا تأثیر هریک از این چالش‌های معنایی بر این دو حوزه تبیین شده و از دو منظر فناوری اطلاعات و مباحث معنایی راهکارهایی را ارائه نماید. همچنین در بخش دوم این نوشتار، به عنوان نمونه، گزارشی از یک مطالعه موردی در حوزه علم اطلاعات و دانش‌شناسی ارائه می‌شود که ضمن بررسی ۱۵۱۰ واژه، مشکلات معنایی موجود در میان اصطلاحات ترجمه شده این حوزه، از زبان انگلیسی به فارسی را گزارش می‌کند. بدیهی است این مطالعه در پیوند با آنچه در بخش اول نوشتار حاضر ارائه شد، شمایی از حقایق موجود در خصوص مشکلات معادل‌گزینی در حوزه‌های تخصصی و تأثیرات نامطلوب آن بر بازیابی اطلاعات را ارائه می‌دهد.

کلیدواژگان: بازیابی اطلاعات، بازیابی اطلاعات بین‌زبانی، روابط معنایی، واژه‌گزینی، معناشناسی

۱- نویسنده مسئول. رایانامه: arastoopoor@ricest.ac.ir

لایحه

مقصد
ردیف ۱





۱. مقدمه

واژه مهم‌ترین واحد زبانی برای انتقال مفاهیم است و بدون آن امکان صحبت‌کردن در مورد اشیاء، پدیده‌ها و مفاهیم جهان ذهنی و عینی وجود ندارد. مک‌کونن (نقل از نعمت‌زاده، ۱۳۸۲) واژه را ابزار اساسی تفکر می‌داند. انسان برای اندیشیدن به جهان و مسائل پیرامون آن ابتدا باید جهان را طبقه‌بندی نماید. این طبقه‌بندی براساس تفاوت‌ها و شباهت‌ها صورت می‌گیرد. بنابراین، در حقیقت ظرافت‌ها و تفاوت‌های معنایی حاصل از واژگان نه تنها در سازمان‌دهی اطلاعات بلکه در سامان‌دهی مفاهیم و پدیده‌های ذهنی نیز کاربرد دارد. زبان علم زیرمجموعه‌ای از زبان انسانی به‌شمار می‌آید و به همین دلیل در حوزه‌های مختلف علم، واژگان از اهمیت خاصی برخوردار است، چرا که بخشی از انتقال موفقیت‌آمیز دانش به دقت واژگان آن حوزه از دانش بستگی داشته و متکی بر یک‌دستی و تا حد امکان وجود تناظر یک به یک بین مفهوم و واژه است. از این رو، انتخاب معادل نامناسب و وجود تعدد در معادل‌های گوناگون برای اصطلاحات یکسان مسئولیت‌خطیر گزینش اصطلاحات معیار و معادل‌های یکسان به زبان فارسی توسط متخصصین و دانشمندان هر رشته‌ی علمی را آشکار می‌سازد (بهشتی، ۱۳۷۵)؛ زیرا سهولت‌انگاری در این زمینه موجب مشکلاتی همچون مسائل ذخیره و بازیابی اطلاعات، مشکلات فرهنگی، مشکل سردرگمی خواننده و اختلال تعامل در جریان جهانی اطلاعات می‌شود (حسن‌زاده و نوروزی چاکلی، ۱۳۸۲). علاوه بر این باید توجه داشت که زبانی زبان علم است که بتوان به آن زبان درباره حوزه‌های مختلف علم با واژگانی شفاف و روان به راحتی مطلب نوشت. بنابراین برای آن که زبان فارسی به زبان علم تبدیل شود، باید در واژه‌گزینی علمی نیز نهایت دقت و توجه را روا داشت؛ چرا که در غیر این‌صورت جویندگان علم موفق به دستیابی به کلیه منابع مرتبط با نیاز اطلاعاتی خود به این زبان نشده و نظام‌های جستجو و بازیابی اطلاعات نیز قادر به رسیدن به حداکثر جامعیت و مانعیت در جستجو نخواهند بود. از این رو، در این نوشتار ابتدا به مباحثی چند در خصوص تأثیر ابهام‌های معنایی حاصل از انتخاب نادرست و غیردقیق واژگان در دو حوزه بازیابی موضوعی و سپس بازیابی بین‌زبانی پرداخته و در نهایت نتیجه مطالعاتی موردی گزارش خواهد شد.

۲. روابط معنایی و چالش بازیابی اطلاعات

بازیابی اطلاعات در اکثر نظام‌های اطلاعاتی مبتنی بر واژه است. صفوی (۱۳۸۷) روابط مفهومی موجود در سطح واژگان را به ۱۰ گروه مختلف تقسیم‌بندی می‌کند که عبارتند از شمول معنایی، جزءواژگی، عضوواژگی، واحدواژگی، هم‌معنایی، هم‌آوا-هم‌نویسی، چندمعنایی، تقابل معنایی، تبیین معنایی و نشان‌داری و بی‌نشانی. نمایش روابط میان واژگان یکی از مهم‌ترین و در عین حال دشوارترین فعالیت‌هایی بوده که در طی زمان توسط متخصصان حوزه علم اطلاعات و دانش‌شناسی و هم‌چنین حوزه علوم رایانه مورد توجه قرار گرفته‌است؛ چرا که در نظرگرفتن هر یک از این روابط در جایگاه خود می‌تواند امکان بسط و یا محدودسازی جستجو و نتایج آن را فراهم نماید. به عنوان نمونه، رابطه‌هایی همچون شمول معنایی، جزءواژگی، عضوواژگی، واحدواژگی سال‌هاست که در قالب تعریف روابط

سلسله‌مراتبی میان مفاهیم در نظام‌های مختلف رده‌بندی و یا سازمان‌دهی هم‌چون سرعنوان‌های موضوعی و یا اصطلاح‌نامه‌ها مورد توجه قرار گرفته‌است. نظام‌های کنترل واژگانی تاکنون در خصوص مهار واژگان و کنترل این روابط موفق عمل کرده‌اند، در حالی که در حوزه بازیابی اطلاعات خودکار، چنین کاری دشوارتر بوده‌است؛ چرا که فعالیتی که تا پیش از این توسط متخصصان مهار واژگانی صورت می‌گرفته را بایستی به ماشین واگذار کرد. دشواری این مسئله در نهایت به چالش‌هایی برای بازیابی اطلاعات منجر شده است. به صورت مشخص، اصطلاح‌گزینی نادرست در حوزه‌های مختلف علمی تأثیر منفی خود را بر دو زمینه بازیابی موضوعی منابع اطلاعاتی و نیز بازیابی بین زبانی می‌گذارد که در ادامه در این خصوص به صورت مفصل‌تر بحث خواهد شد.

۱.۲ ابهام‌های معنایی در جستجوی موضوعی

زبان طبیعی از بعد معنایی آکنده از ابهام است. نگرش‌های متفاوتی به دسته‌بندی‌های مختلف ابهام‌های معنایی می‌توان داشت. ابهام مقوله‌ای در پیوند با معانی مختلف یک واژه در نقش‌های دستوری متفاوت، یکی از این گونه‌ها به‌شمار می‌آید. در زبان فارسی برخی از این ابهام‌ها را با استفاده از علائم واکه‌های کوتاه برطرف می‌کنند که در متون مخصوص بزرگسالان و افراد باسواد معمول نیست و به همین دلیل در نظام‌های نمایه‌سازی، چه به صورت دستی و چه به صورت خودکار نیز کاربرد ندارد. به عنوان نمونه واژه «بردار» در پایگاه مقالات فارسی مرکز منطقه‌ای اطلاع‌رسانی علوم و فناوری جستجو شد که در صفحه اول نتایج، سه مقاله با عنوان‌های زیر قابل مشاهده بود:

بردار: وزنه بردار دیابتی! قوی‌ترین مرد جهان (ورزش)

بردار: پیشرفتهایی در کنترل بردار رانش (ریاضیات)

بردار: داستان بر دار کردن حسنگ وزیر (ادبیات)

نظام‌های بازیابی اطلاعات کنونی ماهیتاً با توجه به نوع داده‌های درون‌دهی شده با محدودیت‌هایی رو به رو هستند. این در حالی است که کاربران، معمولاً بدون آگاهی از این محدودیت‌ها دست به جستجوی اطلاعات می‌زنند. به همین دلیل است که یک واژه مانند «بردار» نتیجه‌ای غیرقابل پیش‌بینی در سه حوزه کاملاً متفاوت به همراه دارد. چرا که الگوریتم‌های بازیابی اطلاعات در نظام‌های اطلاعاتی در شرایط معمول به میزان حضور واژه در متن و محل آن برای رتبه‌بندی نتایج توجه دارند. چنین روشی در پیوند با حالت‌های معمول زبانی مشکلی ایجاد نمی‌کند اما در مواجهه با ابهام‌های زبانی هم‌چون نمونه پیش‌گفته می‌تواند منجر به ریزش کاذب شود. گروه دیگر ابهام‌ها شامل هم‌نگاشتی، هم‌معنایی و چند معنایی می‌شوند. در نظام‌های بازیابی مبتنی بر جستجوی متنی، ابهام‌های حاصل از هم‌نگاشتی مشکلاتی را برای جستجوگران اطلاعات به همراه دارد، که در پاره‌ای از اوقات با ابهام مقوله‌ای نیز هم‌پوشانی پیدا می‌کنند. به عنوان نمونه می‌توان به واژه‌های «گرم» و «گرم» و یا «شیر» و «شیر» و یا «آتش» اشاره کرد. تصویر ۱

نتیجه جستجوی واژه «گرم» را در پایگاه تخصصی نور نشان می‌دهد. همان‌گونه که در نمایش نتایج خلاصه می‌توان مشاهده کرد، این واژه در بسترهای مختلف با معانی متفاوتی به کار رفته است.

وضعیت محتوا	3
دارای نسخه تمام متن	فاجاق تسلیحات گرم به عنوان یک جنایت سازمان یافته فزاملی
دارای چکیده و کلیدواژه	نویسنده: سلیمی، صادق نشریه: علوم انسانی «مدرس علوم انسانی» زمستان 1382 - شماره 3 (علمی پژوهش) (22 صفحه - از 101 تا 122)
مرتب سازی	جدیدترین
میزان ارتباط	جدیدترین
قدیمی ترین	4
نشریه	بررسی آثار روش گرم کردن بر نتایج یک برنامه تمرین قدرتی
موضوع	نویسنده: محمد زاده، حسن نشریه: تربیت بدنی «حرکت» پاییز 1380 - شماره 9 (12 صفحه - از 115 تا 126)
	5
	ابزار محتوای برنامه گرم کردن فعال بر تغییرات چابکی و انعطاف پذیری
	نویسندگان: لطفی حسین آباد، غلامرضا - گابینی، عباس علی نشریه: تربیت بدنی «المنیك» تابستان 1383 - شماره 26 (علمی پژوهش) (12 صفحه - از 85 تا 96)

شکل 1. نمایی از نتایج جستجو در پایگاه مجلات تخصصی نور

بدیهی است بروز چنین مشکلاتی در یک زبان باعث افزایش ابهامات زبانی و غیردقیق شدن معنای واژگان خواهد شد. با وجود این، نیابستی فراموش کرد که در زبان‌های مختلف، طیف وسیعی از ابهام‌های واژگانی به واسطه قرار گرفتن واژگان در بافت جملات برطرف می‌شوند. به عبارت بهتر، گویشوران زبان دشواری چندانی در رمزگشایی این ابهامات ندارند؛ در حالی که در سایر زمینه‌هایی که با واژگان سر و کار دارند، نمی‌توان با چنین قطعیتی سخن گفت. همان‌گونه که وایلدنموث¹ (۲۰۰۴) نیز اشاره می‌کند، کاربران معمولاً بدون توجه به حواشی یک جستجو از عبارت‌های اسمی برای جستجوی موضوعی استفاده می‌کنند و عبارات را نیز با توجه به پرکاربرد بودن آن‌ها انتخاب می‌کنند. به

1. Wildermuth

همین دلیل همان گونه که چری^۱ (۱۹۹۲) نیز بیان داشته است، کاربران ممکن است با استفاده از کلیدواژه‌های انتخابی خود به جستجو در عنوان، متن چکیده، و یا تمام متن منابع بپردازند. در چنین شرایطی، کاربر با واژگان موجود در ذهن خود به دنبال واژگان نمایه شده در نظامی می‌گردد که یا برخاسته از ذهن نویسنده و یا نمایه‌ساز است. بنابراین حداقل سه برداشت متفاوت در امر بازیابی دخیل شده و به زعم بگتول^۲ (۲۰۰۴، به نقل از ساپرل^۳، ۱۹۸۶) کار کاربر را دشوار می‌سازد. بنابراین، نظام بازیابی اطلاعاتی که صرفاً بر پایه واژه عمل کرده و ربط را می‌سنجد در ابهام‌زدایی از واژگان مشکل‌ساز زبان فارسی ناتوان شده و با ریزش کاذب همراه خواهد بود.

۲.۲ ابهام‌های معنایی و بازیابی بین زبانی

در همین راستا در حوزه معنایی، چالش دیگری تحت عنوان چندمعنایی و هم‌معنایی مطرح می‌شود. هم‌معنایی به وضعیتی گفته می‌شود که یک واژه به چند مفهوم اشاره داشته باشد (صفوی ۱۳۸۷، ص ۱۶۰). به صورت معمول در هر زبانی مسئله هم‌معنایی و چندمعنایی به چشم می‌خورد که در این میان زبان فارسی نیز استثنا نبوده و نمونه‌های متفاوتی از این موارد را می‌توان در این زبان مشاهده کرد؛ مانند معانی مختلف واژه «روشن» و یا «آتش» در جملات مختلف. اما در زبان علم، عوامل تشدیدکننده دیگری نیز قابل مشاهده است. بسیاری از علوم در ایران بومی نبوده و پژوهشگران و دانشجویان غالباً از طریق ترجمه آثار غربی با آنها آشنا می‌شوند. بنابراین، اصطلاح‌گزینی برای واژگان علمی توسط یک مرجع واحد صورت نگرفته و گاه مترجمان و پژوهشگران با سلیقه شخصی و با توجه به توانمندی‌های متفاوت واژگانی، برابر نهاده‌های متفاوتی را برای یک مفهوم و یا حتی یک واژه برمی‌گزینند. برای مثال، احمدی‌نسب (۱۳۹۰، ص ۲۷۶) طی مطالعه‌ای به این نتیجه رسید که در حوزه زبان‌شناسی، به طور متوسط برای هر مفهوم سه برابر نهاده فارسی وجود دارد. حال نظر به آن که بازیابی اطلاعات بر پایه واژگان موجود در متن صورت می‌پذیرد و از آنجا که در متون مختلف مفاهیم به صورت‌های متفاوتی ثبت می‌شوند، لذا با جستجوی یک صورت از مفهوم، ممکن است دسترسی به صورت‌های دیگر همان مفهوم که متطابقاً مرتبط با نیاز اطلاعاتی است، از بین برود. برای مثال برای مفهوم ontology در فارسی سه معادل هستی‌شناسی، هستان‌شناسی و آنتولوژی بکار می‌رود که جستجوی هر کدام از این کلیدواژه‌ها تنها مدارکی را بازیابی می‌کند که صرفاً همان صورت واژگانی در آنها به کار رفته است در حالی که هر سه واژه به یک مفهوم اشاره دارد (جدول ۱).

1. Cherry
2. Beghtol
3. Sauperl



جدول ۱. تفاوت باز یافت‌ها در پیوند با سه صورت واژگانی متفاوت از یک مفهوم واحد (ارسطوپور و احمدی‌نسب، ۱۳۹۱)

پایگاه مجلات تخصصی نور	بانک اطلاعات نشریات کشور (مگا ایران)	پایگاه مقالات فارسی مرکز منطقه‌ای (Ricest)	
۲۸۷۷	۴۲۰	۲۷۶	هستی‌شناسی
۶۹	۷	۸	هستان‌شناسی
۱۰۰	۵	۸	آنتولوژی

در نظام‌های باز یابی اطلاعات بین‌زبانی وجود این دست از هم‌معنایی‌ها به بروز مشکلاتی می‌انجامد. عمده تلاش نظام‌های باز یابی بین زبانی ورود اطلاعات به یک زبان و ایجاد امکان باز یابی منابع به زبان دیگر است. در چنین شرایطی، نظام ناگزیر از ترجمه ماشینی است و برای این مهم لازم است که از واژگان دو یا چندزبانه استفاده کند. بنابراین دقت جستجو به دقت واژگان گره می‌خورد. هر چه تعداد برابر نهاده‌ها در چنین شرایطی بیشتر باشد، قواعد تدوین شده برای واژگان مورد استفاده بیشتر می‌شود. به عنوان نمونه، ۴ واژه *Transpiration* (تعرق)، *Sweating* (تعرق)، *Evapotranspiration* (تعرق و تبخیر)، و *Guttation* (تعریق) هر کدام در زبان انگلیسی با معنایی متفاوت به کار برده می‌شوند، مثلاً در زبان مبدأ *Transpiration* برای گیاهان در حوزه کشاورزی، و *Sweating* در پیوند با انسان به کار می‌رود؛ این در حالی است که واژه‌های پیش‌گفته در زبان فارسی معادل نسبتاً یکسانی دارند. هم‌اکنون گوگل تا حدی قابلیت جستجوی بین زبانی در سطح وب را فراهم آورده است. اما زمانی که معادل واژه *Transpiration* در صفحات فارسی جستجو شود، نتایجی مشابه با شکل ۲ و زمانی که واژه تعرق به فارسی برای باز یابی صفحات انگلیسی جستجو شود، نتایجی مشابه با شکل ۳ به دست می‌آید که مورد دوم به مراتب از پراکندگی بیشتری برخوردار است.

تعرق - ویکیپدیا

fa.wikipedia.org/wiki/تعرق - Translate this page

این پدیده فیزیکی-بیولوژیکی را ترانسپن (تعرق) گویند که از نظر ماهیت شبیه تبخیر است و در واقع یکی از فرآیندهای
جانسی عمل قوسنتر به فعل می‌رود. گیاهان به دلیل ...

تبخیر-تعرق - ویکیپدیا

fa.wikipedia.org/wiki/تبخیر-تعرق - Translate this page

به دلیل این که در حوضه‌های آبریز نو پدیده تبخیر از سطح مرطوب خاک و تعرق از سطح گیاهان را می‌توان از
همدیگر مجزا ساخت غالباً این دو فرآیند توأم با یکدیگر و به نام ...

SiteRooz.com :: علل و نحوه درمان تعرق زیاد

www.siterooz.com/elmi/past/036245.php - Translate this page

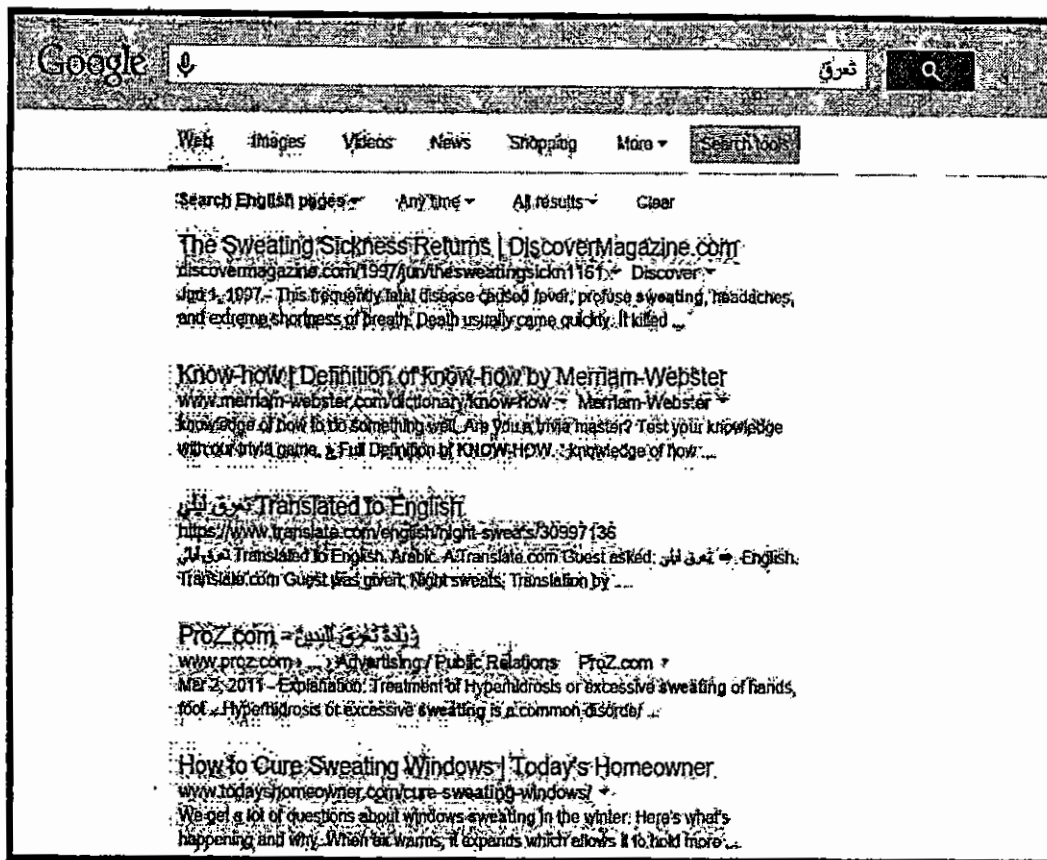
تعرق بیش از اندازه خیلی بیشتر از آن چیزی که شما تصور می‌کنید رایج است به همین دلیل هیچ دلیلی برای نگرانی و
ناامیدی وجود ندارد و تبدیل تصور کنید که در دنیا تنها ...

تعرق:

danesh.roshd.ir/mavara/mavara-index.php?... - Translate this page

این عمل بوسیله دو مکانیزم مختلف تعرق (Transpiration) و تعریق (Guttation) انجام می‌شود که اولی اساساً
بوده و در نتیجه انجام آن در اتمسفر اشباع شده آب به صورت ...

شکل ۲. نمایی از صفحه اول نتایج بازیابی شده بر اساس اولین برابر نهاده پیشنهادی گوگل برای واژه Transpiration



شکل ۳. نمایی از صفحه دوم نتایج بازیابی شده برای واژه تعرق در صفحات انگلیسی زبان

حال با توجه به آن چه پیش‌تر نیز بیان شد و نیز نظر به آن که در بسیاری از حوزه‌ها با واردات علم رو به رو هستیم، کاربران فارسی‌زبان مشکلات جدی در بازیابی اطلاعات تخصصی مورد نظر خود خواهند داشت. پیش‌تر به صورت گذرا به نتیجه پژوهش احمدی نسب (۱۳۹۰) در خصوص واژگان زبان‌شناسی و بروز مشکل چندمعنایی و یا هم‌معنایی در آن اشاره شد. در ادامه گزارشی از یک پژوهش دیگر در حوزه علم اطلاعات و دانش‌شناسی ارائه می‌شود. در این پژوهش، گزارشی از نحوه واژه‌سازی مرسوم در این حوزه تخصصی و تأثیر آن بر مشکلات معنایی مرتبط با بازیابی اطلاعات مورد توجه قرار گرفته‌است.

۳. گزارشی از یک مطالعه موردی در حوزه علم اطلاعات و دانش‌شناسی

با در نظر داشتن این واقعیت که جامعه علمی ایران در غالب حوزه‌ها مصرف‌کننده علم است و زبان فارسی با ورود انبوه واژگان علمی انگلیسی رو به روست و با توجه به اینکه معادل‌گزینی برای این واژه‌ها اغلب به صورت انفرادی و به

وسیله‌ی متخصصان موضوعی و یا حتی مترجمان غیرمتخصص صورت می‌گیرد، در اکثر حوزه‌های دانش ابهام معنایی و عدم شفافیت در متون علمی بروز می‌کند و همان‌گونه که پیش‌تر نیز بیان شد این مسئله در جای خود می‌تواند تأثیر مستقیمی بر کیفیت بازیابی اطلاعات داشته باشد. به همین منظور رشته علم اطلاعات و دانش‌شناسی به منزله یک مورد، برای بررسی انتخاب شد.

۱.۳ یافته‌های مطالعه

در این مطالعه با انتخاب ۱۵۱۰ واژه و اصطلاح تخصصی حوزه علم اطلاعات و دانش‌شناسی به روش تصادفی نظام‌مند از فرهنگ واژگان کتابداری و اطلاع‌رسانی هاشمی (۱۳۸۷)، ساختمان برابرنده‌های فارسی واژگان این حوزه از لحاظ معنایی، صرفی و کاربردشناختی مورد بررسی قرار گرفت. لازم به ذکر است که بررسی کمی برابرنده‌ها نشانگر آن است که ۶۶٪ اصطلاحات انگلیسی دارای یک برابرنده فارسی، ۱۶٪ اصطلاحات دارای دو برابرنده، ۷٪ اصطلاحات دارای سه برابرنده، ۳٪ اصطلاحات دارای چهار برابرنده، ۲٪ اصطلاحات دارای پنج برابرنده و بالاخره ۶٪ اصطلاحات دارای ۶ تا ۳۸ برابرنده فارسی هستند. این آمار و ارقام تلویحاً بدین معناست که پژوهشگران و مؤلفان حوزه علم اطلاعات و دانش‌شناسی در تولید علم به زبان فارسی از برابرنده‌های مختلفی استفاده می‌کنند که این امر نه تنها در انتقال مفاهیم علمی، بلکه به علت هم‌معنایی و چندمعنایی فراوان، در بازیابی اطلاعات از نظام‌های بازیابی نیز اختلال ایجاد خواهد نمود. این درحالی است که معادل‌یابی و ترجمه اصطلاحات علمی غالباً توسط متخصص ناآشنا به ساختمان و صرف زبان فارسی صورت می‌گیرد، برخی از این برابرنده‌ها با قواعد معنایی، صرفی و واژگانی زبان فارسی نیز مطابقت ندارند. این امر از یک طرف باعث آشفتگی در معادل‌گزینی، عدم مقبولیت برابرنده‌ها و ضربه به ساختمان زبان فارسی شده و از طرف دیگر بر روی بازیابی منابع اطلاعاتی در این حوزه تأثیر منفی می‌گذارد. در ادامه به برخی از این مشکلات اشاره می‌شود:

۱.۱.۳ عدم توجه به هسته و وابسته واژگانی

در اصطلاحات و عبارات پیچیده، برخی از واژگان نقش هسته و واژگان دیگر نقش وابسته را ایفاء می‌کنند و معادل‌یابی نیز باید با توجه به این مهم انجام شود. به عنوان مثال در برابر اصطلاح *academic information management system* معادل «نظام مدیریت آکادمیک اطلاعات» انتخاب شده‌است. در این مثال، معادل‌گزینی بدون توجه به هسته و وابسته‌های واژگانی صورت گرفته‌است. با اغماض از ظرایف زبان‌شناختی مبحث، باید گفت که در این اصطلاح *management system* هسته واژگانی و *academic information* وابسته واژگانی است و بنابراین، معادل این اصطلاح «نظام مدیریت اطلاعات آکادمیک» خواهد بود. لازم به توضیح است نه تنها توجه به ترتیب واژگان یکی از



مقوله‌های بسیار مهم در نظام‌های بازیابی اطلاعات است؛ بلکه در جستجوهای موضوعی به صورت مشخص عدم توجه به هسته واژگانی می‌تواند نتایجی کاملاً بی‌ارتباط به نیاز اطلاعاتی کاربر را به همراه داشته باشد. در مثال پیش گفته، حتی با تعیین قوی‌ترین قواعد نیز، در بازیابی بین‌زبانی دچار مشکل خواهیم بود.

۳.۱.۲ عدم توجه به خوشه واژگانی

عدم توجه به نقش اجزای اصطلاحات هنگام معادل‌گزینی به تعدد در معادل‌یابی منجر شده و ارتباط واژگانی و معنایی معادل‌های حاصل از معادل‌گزینی قطع می‌شود. این امر اتفاقی است که در معادل‌یابی برای واژه access در سه اصطلاح مرتبط access control، access copy و access counter رخ داده است. برای این اصطلاحات به ترتیب سه معادل «کنترل بازیابی»، «نسخه دسترسی» و «دستیاب‌سنج» انتخاب شده‌است که هیچ ارتباط واژگانی بین برابرنهاده‌های فارسی آن‌ها وجود ندارد؛ در حالی که در زبان مبدأ، این واژگان از یک خوشه واژگانی بوده و با انتخاب سه برابرنهاده متفاوت برای واژه access این ارتباط واژگانی نادیده گرفته شده‌است. نبایستی از نظر دورداشت که یکی از توانمندی‌های غالب در نظام‌های بازیابی اطلاعات قابلیت ریشه‌سازی است. با استفاده از این قابلیت امکان بسط جستجو با استفاده از ریشه واژگان برای جستجوی عبارات مختلف فراهم می‌شود. این درحالی است که در چنین شرایطی به علت دوری صوری واژگان، کاربر از استفاده مطلوب از این قابلیت محروم می‌شود.

۳.۱.۳ عدم توجه به بار عاطفی واژگان

در برخی از معادل‌گزینی‌ها به بار عاطفی واژه‌ها توجهی نشده‌است. برای مثال، برای دو اصطلاح monitoring و broadcast services access point and search به ترتیب دو اصطلاح «مانیتور خدمات سخن‌پراکنی» و «قابلیت بازیابی و تجسس» برگزیده شده‌است، این در حالی است که دو واژه «سخن‌پراکنی» و «تجسس» دارای بار عاطفی منفی است و باید به جای آن واژه‌هایی با بار عاطفی مناسب انتخاب شود که مناسب واژه‌گزینی علمی باشد. استفاده از واژگان دارای بار عاطفی منفی برای واژگان علمی خنثی باعث می‌شود که به مفهوم اصلی و مورد نظر نویسنده صدمه خورده و در عین حال ارتباط غیرواقعی متون با یکدیگر شود که بر روند جستجو نیز تأثیر بگذارد. برای مثال، به هنگام بازیابی بین‌زبانی با جستجوی عبارت «مانیتور خدمات سخن‌پراکنی»، به احتمال قوی متون علمی مرتبط به monitoring broadcast services بازیابی نخواهد شد، بلکه متونی بازیابی خواهد شد که با رسانه، سیاست‌گذاری رسانه‌ای، اخبار و ... در ارتباط باشد.

۳.۱.۴ وجود برابر نهاده‌های فراوان در برابر یک مفهوم

از چهار برابر نهاده «رمزبرداری»، «رمزخوانی»، «رمزگشایی» و «کد برگردانی» برای اصطلاح decoding استفاده شده است. برابر نهاده‌های فوق از لحاظ ساخت و معنا در زبان فارسی مشکلی نداشته و معنای مورد نظر را به خوبی بیان می‌کنند. تنها ایراد، آن است که در برابر تنها یک اصطلاح، از چهار برابر نهاده فارسی استفاده شده است که این امر باعث ایجاد بار اضافی بر واژگان حوزه تخصصی شده است. این در حالی است که وجود یک برابر نهاده، ارتباط علمی را بین پژوهشگران فارسی‌زبان تسهیل می‌کند و قابلیت زبان‌فارسی را برای تبدیل به زبان علم افزایش می‌دهد. ناگفته پیداست بدون اعمال مهار واژگانی، کاربران ناچارند کلیه برابر نهاده‌های یک واژه را در نظام جستجو نمایند تا مبادا اطلاعات خاصی را از دست دهند.

۳.۱.۵ عدم توجه به هماهنگی معنایی بین اصطلاح و برابر نهاده

با مراجعه به متون مرتبط به علم اطلاعات و دانش‌شناسی به راحتی می‌توان دید که اکثر پژوهشگران این حوزه از اصطلاح «پیوسته» در برابر online استفاده می‌کنند. باید توجه داشت که اصطلاح «پیوسته» یک اصطلاح رایج در بین عموم مردم البته با معنایی متفاوت است. بنابراین به منظور جلوگیری از کاربرد اصطلاحی عام و با معنایی کاملاً غیرمرتبط و هم‌چنین پیروی از مصوبات فرهنگستان، استفاده از اصطلاح «بر خط» پیشنهاد می‌شود که اصطلاحی علمی با معنایی تخصصی است و در حوزه‌های دیگر دانش نیز کاربرد دارد. این در حالی است که اتخاذ این دست از راهبردها به مرور زمان و با تداول کاربرد موجب تمایز متون تخصصی از متون عمومی‌تر شده که خود به بهبود جستجو در نظام‌های بازیابی منجر می‌شود.

۳.۲ استنتاج و پیشنهادهای پژوهش

در این راستا، به نظر می‌رسد علی‌رغم آن که فرهنگستان زبان و ادب فارسی متولی و سیاست‌گذار در حوزه زبان و واژه‌گزینی است، پژوهشگران حوزه‌های موضوعی مختلف، در معادل‌گزینی به مصوبات فرهنگستان توجه کمتری دارند. برای مثال، در حوزه علم اطلاعات و دانش‌شناسی از سه اصطلاح «رقمی»، «رقومی» و «دیجیتال» در برابر اصطلاح digital استفاده شده است که در راستای یکسان‌سازی توصیه می‌شود که واژه مصوب فرهنگستان یعنی اصطلاح «رقمی» به کار گرفته شود. پرواضح است که یکسان‌سازی واژگان به یکسان‌سازی متون علمی منجر شده و در نتیجه به طور غیرمستقیم بر عملکرد نظام‌های بازیابی تأثیر مثبت خواهد گذارد. این در حالی است که توجه متخصصان حوزه

علوم رایانه به بحث یکسان‌سازی با بهره‌گیری از مصوبات فرهنگستان، و یا تعریف قواعدی برای جستجو به صورت متمرکز با استفاده از این مصوبات می‌تواند در این زمینه راهگشا باشد.

موارد فوق برخی از ایرادات معنایی، صرفی و کاربردشناختی ترجمه اصطلاحات علمی است که اگر چه از لحاظ زبان‌شناختی حائز اهمیت بوده، اما اهمیت و تأثیر آن از لحاظ بازیابی اطلاعات کمتر از بعد زبان‌شناختی آن نیست. بی‌شک مطالعه ساختمان واژه‌ها و اصطلاحات ترجمه‌شده و رفع کاستی‌ها و تصحیح زبانی آن‌ها می‌تواند به معادل‌گزینی بهتر، مقبولیت بیشتر در بین متخصصان حوزه و تبدیل زبان فارسی به عنوان زبان تولید علم کمک کند. این درحالی است که یافته‌های این دست از مطالعات به متخصصان حوزه علوم رایانه در راستای بهبود و به‌سازی الگوریتم‌های جستجو و تدوین قواعد لازم برای کمک به یادگیری ماشینی مفید خواهد بود.

۴. نتیجه

- بازیابی‌های مبتنی بر واژه همواره کاربران را با مشکلاتی هم‌چون عدم آشنایی با واژگان مناسب برای جستجو و یا ابهام‌های ماهوی واژگان تخصصی موجود در متون رو به رو ساخته‌است. به صورت کلاسیک جستجوهای کاربران را در دو گروه جستجوهای مرتبط با یک منبع اطلاعاتی مشخص و جستجوهای موضوعی دسته‌بندی می‌کنند. در گروه دوم، کاربران معمولاً با استفاده از عبارت‌های اسمی سعی در یافتن مدارکی دارند که به موضوع مورد نظر آنها نزدیک بوده و نیاز اطلاعاتی آنها را برآورده می‌سازد. از آنجا که در طی این دست از جستجوها مدارک مشخصی مدنظر کاربر نبوده، لذا تخمین ربط در نظام‌های بازیابی اطلاعات در آن‌ها نقش حیاتی ایفا می‌کند. طی سالها، نظام‌های بازیابی اطلاعات سعی کرده‌اند با استفاده از الگوریتم‌های پیشرفته‌تر و پیچیده‌تر، از فضای برداری گرفته تا شبکه‌های ژنتیکی، مدارک مرتبط‌تری را در اختیار کاربران قرار دهند. با وجود این، مسائل ویژه زبانی متخصصان را بر آن داشته‌است تا مطالعاتی نیز در خصوص احتمال تأثیرگذاری زبان بر فرایند بازیابی اطلاعات داشته باشند. بدیهی است هر یک از زبان‌ها بر اساس رده زبانی و نوع خط دارای مسائل مخصوص به خود بوده که بر روی جستجو در نظام‌های بازیابی تأثیرگذار است. برای مثال، عدم نمایش واژه کوتاه در خط فارسی باعث ابهام معنایی شده و ریزش کاذب را به دنبال دارد. در ضمن به علت آن که زبان فارسی در اکثر علوم مصرف‌کننده است، بیشترین واژه‌سازی در این زبان بر اساس وام‌گیری و یا گرده‌برداری انجام شده و از آنجا که این کار غالباً توسط فرد ناآشنا به زبان صورت می‌گیرد، برابر نهاده‌ها بعضاً با ساختمان معنایی و صرفی زبان فارسی سازگار نمی‌باشد. از دیگر سو، عدم سیاست‌گذاری درست در حوزه واژه‌گزینی، هم‌معنایی و چندمعنایی فراوان در حوزه‌های مختلف علوم را در پی دارد که نمونه‌ای از آن در بخش سوم همین نوشتار در قالب مورد کاوی علم اطلاعات و دانش‌شناسی گزارش گردید. برخی از راه‌های برون رفت از این معضل در سطح زبان، پیروی از مصوبات نهاد سیاست‌گذار حوزه زبان یعنی فرهنگستان زبان و ادب فارسی، ایجاد کارگروه‌های واژه‌گزین متشکل از زبان‌شناس و متخصص حوزه و در نظام‌های بازیابی، مجهز نمودن نظام به انواع

واژه‌نامه، اصطلاح‌نامه و دیگر ابزارهای زبانی بهبود جستجو و هم‌چنین ارتقاء نظام بر اساس توانایی‌های درک زبان بر مبنای شمّ زبانی کاربر با استفاده از فناوری‌ها و دستاوردهای هوش مصنوعی است.

منابع

- احمدی نسب، ف. (۱۳۹۰). تهیه و تدوین اصطلاح‌نامه تک‌زبان فارسی زبان‌شناسی. رساله دکتر، دانشگاه علامه طباطبایی، بهشتی، م. (۱۳۷۵). ترجمه فارسی اصطلاح‌نامه اسپینز. *اطلاع‌رسانی*. ۱۲(۱).
- ارسطویور، ش. و احمدی‌نسب، ف. (۱۳۹۱). آسیب‌شناسی زبان و خط فارسی در بازیابی اطلاعات: نگاهی به موتورهای کاوش و پایگاه‌های برخط. در غ. ع. منتظر و ف. شادان‌پور (ویراستار)، *مجموعه مقالات مدیریت منابع اطلاعاتی وب: مبانی و تجربه‌های جهانی* (۱۴-۳۰). تهران: سازمان اسناد و کتابخانه ملی جمهوری اسلامی ایران.
- حسن‌زاده، م. و نوروزی‌چاکلی، ع. (۱۳۸۲). غلامک‌شناسی تن‌شناسی، مخاطبه: نگاهی به جریان جهانی اطلاعات و واژه‌سازی مترجمان ایرانی. *همشهری*. ۲ مرداد.
- صفوی، ک. (۱۳۸۷). *درآمدی بر معنی‌شناسی*. تهران: انتشارات سوره مهر (حوزه هنری سازمان تبلیغات اسلامی).
- فرهنگستان زبان و ادب فارسی (۱۳۸۸). *اصول و ضوابط واژه‌گزینی (ویرایش سوم)*. تهران: فرهنگستان ادب و زبان فارسی (نشر آثار).
- نعمت‌زاده، ش. (۱۳۸۲). واژه‌گزینی، تلاش برای آشنایی با جهان. *فصلنامه فرهنگستان*. ۶ (۱۴)، ۲-۴.
- هاشمی، ا. (۱۳۸۷). *واژگان کتابداری و اطلاع‌رسانی (ویرایش سوم)*. تهران، ایران: انتشارات دبیزش.
- Cherry, J. M. (1992). Improving subject access in OPACs: An explanatory study of conversion of users' queries. *The Journal of Academic Librarianship*, 18(2), 95 – 99.
- Sauperl, A. (2004). Catalogers' common ground and shared knowledge. *Journal of the American Society for Information Science and Technology*, 55(1), 55 -63.
- Wildermuth, B. M. (2004). The effects of domain knowledge on search tactic formulation. *Journal of the American Society for Information Science and Technology*, 55(3), 246–258.