نام مقاله: نمايه‌سازان در پايگاه اطلاعاتي «چكيدة پايان‌نامه‌هاي ايران»

نام نشريه: فصلنامه كتابداري و اطلاع رساني (اين نشريه در www.isc.gov.ir نمايه مي شود)

شماره نشريه: 27 \_ شماره سوم، جلد 7

پديدآور: محمد ايرانشاهي

چكيده

پژوهش حاضر با هدف ارزيابي ميزان همخواني كليدواژه‌هاي عنوان در پايان‌نامه‌هاي مندرج در «پايگاه چكيده‌نامه‌هاي ايران» با توصيفگرهاي اختصاص‌يافته به اين پايان‌نامه‌ها صورت پذيرفت. جامعة مورد بررسي، كلية ركوردهاي موجود در اين پايگاه (65536 ركورد) از ابتدا تا پايان سال 1381 است. با استفاده از فرمول تعيين حجم نمونه از هر حوزة علوم انساني، علوم پايه، علوم پزشكي، فني و مهندسي، علوم كشاورزي و هنر، 100 ركورد به شيوة تصادفي انتخاب شد. نتايج پژوهش نشان داد كه بين تعداد متوسط كليدواژه‌هاي عنوان در حوزه‌هاي مورد بررسي تفاوت معناداري وجود دارد و 67 درصد توصيفگرهاي نمايه‌سازان را كليدواژه‌هاي عنوان تشكيل مي‌دهند. همبستگي بالايي بين تعداد كليدواژه‌هاي عنوان و تعداد توصيفگرها و ميزان همخواني آن‌ها وجود دارد. از نظر شيوة نگارش 48 درصد توصيفگرها به صورت تك‌واژه، 52 درصد تركيب‌واژه، 87 درصد توصيفگرها مفرد و 13 درصد آن‌ها به صورت جمع تهيه شده‌اند. بازيافت بيشتر از طريق توصيفگرها در مقايسه با كليدواژه‌هاي عنوان، اثربخشي طرح‌هايي را نشان داد كه به منظور افزايش ارتباط بين توصيفگرهاي هم‌موضوع در اين پايگاه صورت پذيرفته است.

كليدواژه‌ها: كليدواژه‌ها، توصيفگرها، چكيدة پايان‌نامه‌هاي ايران، پايان‌نامه‌هاي ايران

مقدمه

«عنوان» نخستين بخش از هر اثر است كه نظر خواننده را به خود جلب مي‌كند. از ديدگاه خواننده يا پژوهشگر، «عنوان» داراي ارزش زيادي است. «عنوان» هر اثر جلوه‌اي از هويت آن اثر مي‌باشد و به تعبيري نخستين جلوه از متن است كه خواننده با آن روبرو مي‌شود. هر كتاب، مقاله، داستان، شعر، فيلم يا وب‌سايت، جدا از نام پديدآور، با نام عنوانش نيز شناخته مي‌شود. عنوان ظرفي است كه مظروف آن، انديشة اصلي متن است. گاه عنوان نقش اصلي خود را ايفا نمي‌كند و رسانندة مباحث متن نيست؛ حتي ممكن است به صورت استعاري بيان شده باشد. اين امر در ارتباط با آثار هنري يا برخي مباحث علوم انساني بيشتر مشاهده مي‌شود. عناويني چون «هنر عشق ورزيدن»، «رنج و سرمستي»، يا «خسي در ميقات» اطلاعات خاصي دربارة متن به دست نمي‌دهند؛ يا از عنوان وب‌سايت What the Thunder said معلوم نمي‌شود كه حاوي مطلبي دربارة زندگي و آثار «تي.اس. اليوت» است (تيلور، 1381، 167). «نيومارك» (Newmark, 1988: 57) عناوين را به دو گروه «عناوين توضيحي»*[[3]](http://www.aqlibrary.org/editor/main.htm" \l "_ftn3" \o ")* و «تلميحي»*[[4]](http://www.aqlibrary.org/editor/main.htm" \l "_ftn4" \o ")* تقسيم كرده است. به نظر وي در متون غيرادبي، «عنوان توضيحي» به اختصار به موضوع اثر اشاره مي‌كند و هدف متن را به نحوي درخور بيان مي‌دارد و عنوان تلميحي نيز نوعي ارتباط كنايي يا استعاري با موضوع دارد. براي آثار علمي، پژوهشي و اطلاعاتي، استفاده از «عنوان تلميحي» مطلوب نيست. بسياري از ناشران متون علمي مي‌كوشند اين دو نقش عنوان- يعني زيبايي و اطلاع‌رساني- را تلفيق كنند، و اين كار را با استفاده از عنوان فرعي انجام مي‌دهند. در اين‌گونه تلفيق‌ها، عنوان اصلي عمدتاً نقش زيبايي‌شناسانه و عنوان فرعي نقش اطلاعاتي دارد. مجلات معتبر دنيا نيز يكي از شرايط قبول مقاله را برابري عنوان با محتواي متن تعيين مي‌كنند، و اين كار را به دو دليل انجام مي‌دهند:

1. به دليل آن كه صورت مسئلة تحقيق بايد به گوياترين وجه، در عنوان متجلي باشد؛

2. به اين دليل كه عنوان، پايگاه موضوعي بسياري از سازمان‌هايي است كه به كار نمايه‌سازي ماشيني مقاله‌ها و مدارك ديگر مي‌پردازند. فرضي كه وراي چنين گزينشي نهفته، اين است كه مجموعه كليدواژه‌هاي عنوان، جلوه‌اي از موضوع كلي مدرك است. روش استفاده از كليدواژه‌هاي عنوان براي تعيين و منعكس‌كردن موضوع مدارك، هنوز هم در انواع نمايه‌سازي‌ها معمول است. حتي اگر عنوان رسانندة موضوع متن نباشد، براي اين منظور از عنوان ساختگي (كه به چكيدة عنواني معروف است) استفاده مي‌كنند؛ يعني عنواني را كنار عنوان ساختة مؤلف به كار مي‌برند كه تجلي مباحث متن باشد (حري، 1372، 179). حال، مسائلي كه در اين خصوص مطرح مي‌شود اين است كه در نمايه‌سازي پايان‌نامه‌ها، «عنوان» تا چه حد مي‌تواند در ارائة محتواي اين مدارك نقش داشته باشد؟ نمايه‌‌سازان «مركز اطلاعات و مدارك علمي ايران» تا چه حد از عنوان پايان‌نامه‌ها در نمايه‌سازي بهره مي‌گيرند؟ اين بهره‌گيري در حوزه‌هاي مختلف، چه تفاوتي دارد؟ آيا با افزايش كليدواژه‌هاي عنوان ميزان همخواني بين كليدواژه‌ها و توصيفگرها افزايش مي‌يابد؟

هدف‌هاي پژوهش

هدف اين پژوهش بررسي ساختار و اجزاي تشكيل‌دهندة عنوان پايان‌نامه‌ها در حوزه‌هاي مختلف، نشان دادن تشابهات و تمايزها در عنوان پايان‌نامه‌هاي هر حوزه، و نيز بررسي ارتباط اين اجزا (بويژه كليدواژه‌هاي آن) با توصيفگرهايي است كه به منظور نمايه‌سازي در پايگاه «چكيدة پايان‌نامه‌هاي ايران» از آن‌ها استفاده شده. نتايج حاصل از اين تحقيق نشان خواهد داد كه عنوان پايان‌نامه‌ها تا چه حد بيانگر محتواي موضوعي است، و به منظور نمايه‌سازي ماشيني پايان‌نامه‌ها- كه عمدتاً با استفاده از كليدواژه‌هاي عنوان صورت مي‌گيرد و در مقايسه با نمايه‌سازي كنترل‌شده، از هزينة كمتر و سرعت بيشتري برخوردار است- تا چه حد مي‌توان به عنوان متكي بود. علاوه بر اين، يافته‌هاي اين پژوهش متخصصان نمايه‌سازي را در اتخاذ تصميم‌هاي مناسب براي افزايش نقاط دستيابي به منابع اطلاعاتي و نيز بهبود شيوه‌هاي ذخيره و بازيابي اطلاعات پايان‌نامه‌هاي كشور ياري مي‌دهد.

سؤالهاي پژوهش

1. چه تفاوتي بين تعداد متوسط كليدواژه‌هاي عنوان پايان‌نامه‌ها در حوزه‌هاي علوم انساني، علوم پايه، علوم پزشكي، فني و مهندسي، كشاورزي و هنر وجود دارد؟

2. در حوزه‌هاي علوم انساني، علوم پايه، علوم پزشكي، فني و مهندسي، كشاورزي و هنر چه تفاوتي بين ميزان همخواني كليدواژه‌هاي عنوان با توصيفگرهاي نمايه‌سازان وجود دارد؟

3. آيا با افزايش تعداد كليدواژه‌هاي عنوان، ميزان همخواني آن‌ها با توصيفگرها نيز افزايش مي‌يابد؟

4. بين تعداد كليدواژه‌هاي عنوان پايان‌نامه‌ها با تعداد توصيفگرهاي اختصاص‌يافته، چه رابطه‌اي وجود دارد؟

5. ساختار توصيفگرها از نظر ساده و مركب، و مفرد و جمع‌بودن چگونه است؟

6. در پايگاه «چكيده پايان‌نامه‌هاي ايران» بين ميزان بازيافت با استفاده از كليدواژه‌هاي عنوان و توصيفگرهاي نمايه‌سازان، چه تفاوتي وجود دارد؟

پيشينة پژوهش

«باتل» و «پرايبيش» (Bottle and Preibish, 1970) در پژوهشي به بررسي سودمندي نمايه‌سازي كوئيك در مورد مقالات روانشناسي پرداختند و دريافتند در «چكيده‌نامه روانشناسي»*[[5]](http://www.aqlibrary.org/editor/main.htm" \l "_ftn5" \o ")* 60 درصد مقالات با استفاده از كليدواژه‌هاي عنوان قابل بازيابي است.

«هنزلر» (Henzler, 1978) در پژوهشي با بررسي پايگاه اطلاعاتي «كنسرنت»*[[6]](http://www.aqlibrary.org/editor/main.htm" \l "_ftn6" \o ")* به اين نتيجه رسيد كه 35 درصد از كليدواژه‌هاي عناوين، فاقد توصيفگرهاي مناسب در اصطلاحنامه‌ هستند. از نظر او استفادة همزمان از كليدواژه‌ها و توصيفگرها منجر به بازيابي مطلوب خواهد شد.

پژوهش «انانسون» (Ananson, 1987) به ارزيابي جستجوي كليدواژه‌اي پرداخت و نشان داد نه‌تنها جستجوي كليدواژه‌اي عنوان باعث بازيابي مدارك مرتبط مي‌شود، بلكه با چنين جستجويي، ركوردهايي بازيابي مي‌شوند كه با جستجوي موضوعي قابل بازيابي نيستند.

«شفرد» (Shepherd, 1987) در تحقيقي به بررسي ميزان همخواني كليدواژه‌هاي عنوان و توصيفگرهاي موضوعي در پايگاه اطلاعاتي «اُ اُ اِن»*[[7]](http://www.aqlibrary.org/editor/main.htm" \l "_ftn7" \o ")* پرداخت. در اين پايگاه كه وابسته به مؤسسه اطلاعات علمي و فني كانادا*[[8]](http://www.aqlibrary.org/editor/main.htm" \l "_ftn8" \o ")* مي‌باشد، گزارش كنفرانس‌ها، تك‌نگاشتها، و گزارش‌هاي فني در حوزة علوم پزشكي و فناوري نگهداري مي‌شوند. نتايج اين بررسي نشان داد كه استفاده‌كنندگان اين پايگاه براي دسترسي به بيشترين اطلاعات، بايد از هر دو روش بازيابي استفاده نمايند.

«فراست» (Frost, 1989) در پژوهشي به بررسي همپوشاني بين سرعنوان‌هاي برگزيده از «سرعنوان‌هاي موضوعي كتابخانة كنگره»*[[9]](http://www.aqlibrary.org/editor/main.htm" \l "_ftn9" \o ")* و واژه‌هاي استخراج‌شده از عنوان، در فهرست كتابخانة دانشگاه ميشيگان پرداخت. يافته‌هاي اين پژوهش حاكي از آن بود كه 73 درصد واژه‌هاي عنوان، با تمام يا حداقل با بخشي از سرعنوان يا تقسيمات فرعي، همپوشاني دارند. در 11 درصد موارد نيز واژه‌هاي عنوان، دقيقاً با سرعنوان برابر است و در بيش از 53 درصد موارد، واژه‌هاي عنوان حداقل با يك واژه از سرعنوان‌ همپوشاني دارند.

«كلر» (Keller, 1992) در پژوهشي بر روي پايان‌نامه‌هاي مقطع كارشناسي ارشد كتابخانه‌ دانشگاه ايالتي اينديانا به مقايسة نخستين سرعنوان موضوعي كنگره- كه به پايان‌نامه اختصاص داده شده بود- با واژه‌هاي عنوان پرداخت تا مشخص كند احتمال وجود كليدواژه‌هاي عنوان در سرعنوان‌هاي موضوعي چقدر است. نتايج حاصل از اين بررسي نشان داد كه از 371 پايان‌نامه، در 238 عنوان (64 درصد) بين واژه‌هاي عنوان و سرعنوان موضوعي، همپوشاني وجود دارد؛ يعني در پايان‌نامه‌هايي كه داراي 6-1 واژه در عنوان هستند 13 درصد، پايان‌نامه‌هاي داراي 12-7 واژه در عنوان 47 درصد، 18-13 واژه در عنوان 30 درصد، و 24-19 واژه در عنوان 10 درصد همپوشاني ميان واژه‌هاي عنوان و نخستين سرعنوان‌ موضوعي وجود دارد. بدين‌ترتيب مشخص شد كه با افزايش تعداد واژه‌هاي تشكيل‌دهندة عنوان، ميزان همپوشاني سرعنوان‌هاي موضوعي با واژه‌هاي عنوان كاهش مي‌يابد.»

«اسپينك» و «ساراسويك» (Spink and Saracevic, 1993) در پژوهشي قابليت كليدواژه‌ها در بازيابي اطلاعات مرتبط در پايگاه «ديالوگ»*[[10]](http://www.aqlibrary.org/editor/main.htm" \l "_ftn10" \o ")* را مورد بررسي قرار دادند. نتايج پژوهش نشان داد 40 درصد واژه‌هاي مورد استفاده براي جستجو توسط مراجعه‌كننده ارائه شده بودند و در بازيابي موضوعات مرتبط، نقش مثبتي داشتند. همچنين 67 درصد از واژه‌هايي كه از طريق مصاحبه با مراجعه‌كننده استخراج شده بودند باعث بازيافت مدارك مرتبط شدند. درواقع اين واژه‌ها كه 10 درصد از كل واژه‌هاي استخراج‌شده را تشكيل داده بودند گرچه خيلي مفيد نبودند، ولي هنگامي كه استفاده مي‌شدند، در به دست آوردن نتايج بهتر، مؤثر بودند.

«لنكستر» و همكارانش (Lancaster, et al. 1991) نيز از طريق «فهرست‌هاي همگاني پيوسته» اطلاعات مربوط به 51 موضوع را استخراج كردند. آن‌ها از بين عناوين بازيابي شده، 607 مدرك را كه در كتابخانة دانشگاه ايلينويز موجود بود به عنوان «سياهة وارسي» در نظر گرفتند؛ سپس از جستجوگران باتجربه درخواست كردند كه از طريق «فهرست همگاني پيوسته» و تنها با استفاده از سرعنوان‌هاي موضوعي كتابخانة كنگره، ركوردهاي موجود دررابطه با 51 موضوع يادشده را استخراج كنند. بدين طريق تنها 53 درصد از 607 مدرك موجود در دانشگاه ايلينويز بازيابي شد.

«حميدي» (Hmeidi, 1997) در پژوهشي دربارة نمايه‌سازي ماشيني مدارك به زبان عربي، نتايج بازيابي از طريق كليدواژه‌ها و ريشة واژه‌ها در زبان عربي را مورد بررسي و ارزيابي قرار داد. وي دريافت كه در زبان عربي استفاده از ريشة لغات از نظر ميزان بازيافت ودقت، نتايج بهتري را در مقايسه با بازيابي از طريق كليدواژه‌ها به دست مي‌دهد.

«ووربيج» (Voorbij, 1998) نيز بر روي تك‌نگاشت‌هاي موجود در كتابخانة ملي هلند، در حوزه علوم انساني و اجتماعي كه از طريق فهرست همگاني پيوسته قابل بازيابي بود، تحقيقي انجام داد. در اين تحقيق مشخص شد 86 درصد منابع با استفاده از توصيفگرها و 48 درصد با استفاده از كليدواژه‌هاي عنوان بازيابي شدند. «ووربيج» علت بازيافت كم از طريق كليدواژه‌ در مقايسه با توصيفگرها را محدود بودن دامنة توصيفگرها، و تنوع عناوين گسترة موضوعي دانست كه مطالب علوم انساني و اجتماعي در قالب آن عرضه مي‌شوند.

«معرف‌زاده» (1366) در پژوهشي با عنوان «بررسي وضعيت ابزار بازيابي اطلاعات پزشكي فارسي و پيشنهادي براي تهية نماية پزشكي ايران» به بررسي اين پرسش پرداخت كه كدام شيوة نمايه‌سازي گردان براي ذخيره و بازيابي اطلاعات پزشكي مفيدتر است. يافته‌ها نشان داد كليدواژه‌هاي استخراج‌شده از عنوان مقالات، كتب و پايان‌نامه‌هاي مورد بررسي، 58 درصد از كل واژه‌هاي عناوين اين مدارك است كه اين نسبت در پايان‌نامه‌هاي پزشكي به 69 درصد افزايش يافته است. در اين پژوهش روش «نوع افزوده به عنوان» براي تهية نماية پزشكي ايران پيشنهاد گرديد.

«بزرگي» (1377) در پژوهشي ميزان سازگاري عنوان‌ها و موضوع‌هاي كتاب‌هاي فارسي منتشر شده در ايران در فاصلة سال‌هاي 1375-1365 را مورد بررسي قرار داد. يافته‌هاي پژوهش نشان داد كه ميان موضوعات و عناوين كتاب‌هاي فارسي، بطور ميانگين 73 درصد سازگاري وجود دارد. بالاترين ميزان سازگاري (يعني 87 درصد) به ردة «كليات» تعلق دارد و پس از آن ردة «علوم كاربردي» (فناوري) با برخورداري از 83 درصد سازگاري در مقام دوم قرار دارد. كمترين ميزان سازگاري در ردة «ادبيات» (41 درصد) مشاهده شد.

«داورپناه» (1375) نيز در پژوهشي ميزان همخواني عناوين مقالات فارسي با محتواي آن‌ها را در گروه‌هاي مختلف علمي بررسي كرد و به اين نتيجه رسيد كه عناوين مقالات در حوزة علوم انساني، در مقايسه با عناوين مقالات علوم پايه و مهندسي و مقالات علوم پزشكي، از سازگاري كمتري با محتوا برخوردار هستند.

«انصاري» (1379) در پايان‌نامة خود به «بررسي انطباق ميان توصيفگرهاي نمايه‌سازي و كليدواژه‌هاي عنوان در پايان‌نامه‌هاي دكتراي تخصصي كودكان، زنان، قلب و عروق، و روانپزشكي» پرداخت و نشان داد كه در حوزة مورد بررسي به ميزان 3/70 درصد ميان توصيفگرها و كليدواژه‌هاي عنوان انطباق وجود دارد.

روش پژوهش، جامعة آماري و روش نمونه‌گيري

اين تحقيق با روش تحليل محتوا انجام شده؛ زيرا در آن به گونه‌اي كم‍ّي، استفاده از كليدواژه‌ها در عنوان پايان‌نامه‌ها و ميزان همخواني آن‌ها با توصيفگرها تحليل و ارزيابي شده است.

جامعة آماري پژوهش، 65536 عنوان پايان‌نامه‌ است كه در پايگاه «چكيده پايان‌نامه‌هاي ايران» تا پايان سال 1381 معرفي شده‌اند. در اولين مرحله از تحقيق، اطلاعات مربوط به عنوان پايان‌نامه‌ها، توصيفگرهاي اختصاص داده شده به هر عنوان، تاريخ نگارش، حوزة اصلي، و رشتة فرعي پايان‌نامه‌هاي موجود در «پايگاه چكيده پايان‌نامه‌هاي ايران» از ابتدا تا پايان سال 1381، از محيط نرم‌افزاري «سي دي اس آيسيس»*[[11]](http://www.aqlibrary.org/editor/main.htm" \l "_ftn11" \o ")* به محيط نرم‌افزاري «اكسل»*[[12]](http://www.aqlibrary.org/editor/main.htm" \l "_ftn12" \o ")* انتقال يافت و هريك از فيلدهاي ذكرشده، در ستون جداگانه‌اي قرار گرفت. در ادامه به منظور يكدستي جامعة مورد تحقيق، قبل از تفكيك‌ پايان‌نامه‌ها به رشته‌هاي اصلي، عناويني كه داراي نقايص آشكاري در فيلدهاي مورد اشاره بودند از جامعه حذف شدند.

پس از طبقه‌بندي پايان‌نامه‌ها به رشته‌هاي اصلي، عناوين تكراري يا فاقد رشتة فرعي نيز حذف شد. بدين‌ترتيب با توجه به موارد يادشده، 9785 عنوان پايان‌نامه از جامعة اولية تحقيق به دليل داشتن مشكلات فاحش حذف شد. در ادامه، جامعة اصلي تحقيق در هر طبقه مشخص شد و با استفاده از فرمول n=NZpq/(d2(N-1)+Z2pq) (سرايي، 1372، 128) حجم نمونه مربوط به هر طبقه محاسبه شد. چون از تحقيقات صورت گرفتة قبلي تخمين مناسبي براي p وجود نداشت، بنابراين مقدار p مساوي 05/0 درنظر گرفته شد كه بدين‌ترتيب، بيشترين حجم نمونة ممكن براي انجام پژوهش به دست آمده است. همچنين مقدار خطاي حاصل از برآورد آماري 05/0، Z معادل 96/1 و مقدار d نيز برابر 01/0 درنظر گرفته شد. بر اين اساس حجم نمونه در هر طبقه به طور جداگانه محاسبه شده و نتايج زير به دست آمد:

علوم انساني 96، علوم پايه 93، علوم پزشكي 95، فني و مهندسي 100، كشاورزي 94، هنر 88 .

به نظر «ببي» (Babbie, 1979, 274) اگر نمونة احتمالي كمتر از 100 مورد باشد، به احتمال زياد معرف و نمايندة واقعي جامعه نخواهد بود. از طرف ديگر به منظور اين كه در مراحل بعدي تحقيق بتوان يافته‌هاي مربوط به نمونه‌هاي هر حوزه را از ديدگاه‌هاي مختلف با يكديگر مقايسه نمود، حجم نمونة مربوط به هر طبقه با اندكي تغيير، 100 ركورد درنظر گرفته شد. در ادامه با استفاده از دستور «نمونه‌گيري» در نرم‌افزار «اكسل» از بين جامعه آماري هر حوزه، 100 ركورد به شكل كاملاً تصادفي انتخاب گرديد.

داده‌هاي مورد نياز پژوهش با استفاده از فهرست‌ كنترلي (چك‌ليست) محقق‌ساخته‌اي گردآوري شد.

يافته‌ها

با استفاده از داده‌هاي گردآوري‌شده به سؤالهاي پژوهش پاسخ داده شده كه توضيح آن به شرح زير است:

1. چه تفاوتي بين تعداد متوسط كليدواژه‌هاي عنوان پايان‌نامه در حوزه‌هاي علوم انساني، علوم پايه، علوم پزشكي، فني و مهندسي، كشاورزي و هنر وجود دارد؟

براي پاسخگويي به اين پرسش، ابتدا تعداد كليدواژه‌هاي هر يك از پايان‌نامه‌ها به تفكيك حوزة موضوعي شمارش شده، سپس برابر جدول شماره 1، تعداد متوسط كليدواژه‌ها در حوزه‌هاي مختلف شمارش شده است.

جدول 1. تعداد متوسط كليدواژه‌ها در عناوين پايان‌نامه‌هاي هر يك از حوزه‌هاي شش‌گانه

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| حوزه | علوم انساني | علوم پايه | علوم پزشكي | فني و مهندسي | كشاورزي | هنر | جمع |
| تعداد پايان‌نامه‌ | 100 | 100 | 100 | 100 | 100 | 100 | 600 |
| تعداد كليدواژه‌هاي عنوان | 744 | 730 | 515 | 804 | 765 | 534 | 4082 |
| تعداد متوسط كليدواژه‌هاي عنوان | 7 | 7 | 5 | 8 | 8 | 5 | 7 |
| نما | 6 | 4 | 3 | 6 | 6 | 4 | 6 |

در اين جدول نشان داده شده كه به طور كلي (با درنظرگرفتن كلية حوزه‌ها)، هر عنوان پايان‌نامه داراي 7 كليدواژه است. اين تعداد در عناوين پايان‌نامه‌هاي حوزة علوم پزشكي و هنر كمتر از حد متوسط و در ديگر حوزه‌ها تقريباً مساوي با تعداد متوسط كليدواژه‌هاي عنوان است. شايد يكي از دلايل كوتاه‌بودن عنوان در آثار و تأليفات هنري اين باشد كه معمولاً هنرمندان عنوان را نيز چون رنگ‌ها، سايه‌روشن‌ها، اشكال و صور مختلف، جزئي از جريان هنري مي‌دانند و با نگرشي فلسفي، براي يك عنوان مفاهيم بسياري را تصور مي‌كنند و از مخاطبان نيز مي‌خواهند كه با كنكاش ذهني، از يك اثر يا عنوان كوتاه به يافته‌هاي بديعي دست يابند. در خصوص عنوان پايان‌نامه‌هاي حوزة علوم پزشكي نيز مشاهده شد كه عمدتاً مطالعه يا بررسي بيماري خاصي مطرح شده و محقق با استفاده از چند كليدواژه در عنوان پايان‌نامه، مقصود خود را بيان داشته است. در ادامه با استفاده از آزمون «تحليل واريانس» تفاوت تعداد متوسط كليدواژه‌هاي عنوان در حوزه‌هاي مختلف بررسي شد كه رقم محاسبه شده (P=0/000) بيانگر معناداري تفاوت بين تعداد متوسط كليدواژه‌هاي عنوان در حوزه‌هاي مختلف است. به منظور دريافت تصوير روشن‌تري از ساختار واژگاني عناوين پايان‌نامه‌ها در حوزه‌هاي مختلف تعداد واژه‌هاي عنوان پايان‌نامه‌ها (اعم از كلمات با بار معنايي*[[13]](http://www.aqlibrary.org/editor/main.htm" \l "_ftn13" \o ")* و بدون بار معنايي*[[14]](http://www.aqlibrary.org/editor/main.htm" \l "_ftn14" \o ")*) بررسي شد. درواقع سؤالي كه مطرح مي‌شود اين است كه در حوزه‌هاي مختلف، بين ميزان عناصر تشكيل‌دهندة عنوان (اعم از كليدواژه‌ يا واژگان غيرموضوعي) چه تفاوتي وجود دارد؟ جدول شماره 2 تصويري كلي از تفاوت‌هاي تعداد كليدواژه‌ها و واژه‌هاي غيرموضوعي به دست مي‌دهد.

جدول 2. مقايسة تعداد كليدواژه‌ها و واژه‌هاي غيرموضوعي در عناوين پايان‌نامه‌هاي حوزه‌هاي شش‌گانه

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| حوزه | علوم انساني | علوم پايه | علوم پزشكي | فني و مهندسي | كشاورزي | هنر | جمع |
| تعداد واژه‌ | 1003 | 1121 | 835 | 1041 | 1001 | 668 | 5669 |
| تعداد كليدواژه‌ | 744 (74%) | 730 (65%) | 515 (62%) | 804 (77%) | 765  (80%) | 534 (80%) | 4092 (72%) |
| تعداد واژه‌هاي غيرموضوعي | 259 (26%) | 391 (35%) | 320 (38%) | 237 (23%) | 236  (24%) | 134 (20%) | 1577 (28%) |

بر اساس آزمون مجذور كاي، مشخص شد كه بين تعداد كليدواژه‌ها و واژه‌هاي عنوان در حوزه‌هاي مختلف تفاوت معناداري وجود دارد.

2. بين تعداد كليدواژه‌هاي عنوان پايان‌نامه‌ها با تعداد توصيفگرهاي اختصاص يافته چه رابطه‌اي وجود دارد؟

پاسخگويي به اين سؤال مستلزم اين بود كه توزيع فراواني تعداد كليدواژه‌ها و ميزان تبعيت توصيفگرها از اين توزيع مشخص شود.

بدين منظور پايان‌نامه‌ها بر اساس تعداد كليدواژه‌هاي عنوان به 6 طبقه دسته‌بندي شدند و پس از آن، مجموع كليدواژه‌ها و مجموع توصيفگرهاي هر طبقه شمارش شد تا توزيع فراواني هر يك مشخص شود. نتايج به‌دست آمده نشان داد كه به طور كلي (با احتساب كلية حوزه‌ها) تعداد توصيفگرهاي اختصاص يافته، از تعداد كليدواژه‌هاي عنوان پايان‌نامه‌ها تأثير مي‌پذيرد و در موقعي كه عنوان پايان‌نامه‌ها داراي تعداد كليدواژه‌هايي در حد متوسط (9 كليدواژه در عنوان) است بيشترين ميزان همخواني با تعداد توصيفگرها وجود دارد.

علاوه بر بررسي فوق، با استفاده از آزمون ضريب همبستگي پيرسون ميزان همبستگي تعداد كليدواژه‌هاي عنوان و تعداد توصيفگرها به تفكيك حوزه‌هاي شش‌گانه محاسبه شد كه نتايج به دست آمده در زير مشاهده مي‌شود:

جدول 3. ضريب همبستگي تعداد كليدواژه‌هاي عنوان و تعداد توصيفگرها به تفكيك حوزه

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| حوزه | علوم انساني | | | علوم پايه | | | علوم پزشكي | | | فني و مهندسي | | | كشاورزي | | | هنر | | |
| محاسبه پيرسون | تعداد | سطح معناداري | ضريب همبستگي | تعداد | سطح معناداري | ضريب همبستگي | تعداد | سطح معناداري | ضريب همبستگي | تعداد | سطح معناداري | ضريب همبستگي | تعداد | سطح معناداري | ضريب همبستگي | تعداد | سطح معناداري | ضريب همبستگي |
| تعداد كليدواژه‌ها و تعداد توصيفگرها | 6 | 006/0 | (\*\*)934/0 | 6 | 005/0 | (\*\*)939/0 | 6 | 005/0 | (\*\*)941/0 | 6 | 000/0 | (\*\*)982/0 | 6 | 002/0 | (\*\*)963/0 | 6 | 014/0 | (\*)902/0 |

همان‌گونه كه داده‌هاي جدول فوق نيز نشان مي‌دهد در كليه حوزه‌ها همبستگي مثبتي بين تعداد كليدواژه‌هاي عنوان و تعداد توصيفگرها وجود دارد.

3. در حوزه‌هاي علوم انساني، علوم پايه، علوم پزشكي، فني و مهندسي، كشاورزي وهنر چه تفاوتي بين ميزان همخواني كليدواژه‌هاي عنوان با توصيفگرهاي نمايه‌سازان وجود دارد؟

به منظور پاسخگويي به سؤال فوق، وضعيت همخواني كليدواژه‌هاي عنوان با توصيفگرها، در قالب 3 حالت زير بررسي شد:

- همخواني كامل، مانند: توصيفگر «جرائم مطبوعاتي» براي پايان‌نامه‌اي با عنوان «تحقيقي پيرامون جرائم مطبوعاتي». در اين حالت توصيفگر دقيقاً با يكي از كليدواژه‌هاي عنوان همخوان است.

- همخواني نسبي، مانند: توصيفگر «مطبوعاتي» براي پايان‌نامه‌اي با عنوان «تحقيقي پيرامون جرائم مطبوعاتي». در اين حالت توصيفگر با يكي از كليدواژه‌هاي عنوان، هم‌ريشه است و تنها حالت وصفي آن تغيير كرده است.

- ناهمخواني، مانند: توصيفگر «نظا‌م‌هاي حقوقي» براي پايان‌نامه‌اي با عنوان «تحقيقي پيرامون جرائم مطبوعاتي». در اين حالت عبارت «نظام‌هاي حقوقي» به عنوان توصيفگر به شكلي با عنوان فوق مربوط است، ولي به هيچ شكلي با كليدواژه‌هاي اين عنوان همخواني ندارد.

پس از بررسي حالت‌هاي ذكرشده در نمونة تحقيق، نتايج به دست آمده در جدول شماره 4 گرد آمده است.

جدول 4. ميزان همخواني كليدواژه‌هاي عنوان با توصيفگرها، به تفكيك حوزه

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| حوزه | علوم انساني | علوم پايه | علوم پزشكي | فني و مهندسي | كشاورزي | هنر | جمع |
| وضعيت همخواني | درصد | درصد | درصد | درصد | درصد | درصد | درصد |
| همخواني كامل | 49% | 41% | 50% | 35% | 50% | 52% | 45% |
| همخواني نسبي | 22% | 24% | 19% | 23% | 21% | 22% | 22% |
| ناهمخواني | 29% | 35% | 31% | 42% | 29% | 26% | 33% |
| جمع | 100% | 100% | 100% | 100% | 100% | 100% | 100% |

داده‌هاي اين جدول نشان مي‌دهد كه به طور كلي، 45 درصد از توصيفگرهاي نمايه‌سازان را كليدواژه‌هاي عنوان تشكيل مي‌دهند. به عبارت ديگر در 45 درصد موارد، كليدواژه‌هاي عنوان و توصيفگرهاي انتخاب شده توسط نمايه‌سازان همخواني كامل وجود دارد. بيشترين ميزان همخواني كامل به ترتيب در حوزه‌هاي هنر (52%)، كشاورزي (50%)، علوم پزشكي (50%) و علوم انساني (49%)، و كمترين ميزان همخواني كامل واژه‌هاي عنوان با توصيفگرهاي نمايه‌سازان در حوزه علوم پايه (41%) و فني و مهندسي (35%) است.

با توجه به اين كه توصيفگرهاي داراي همخواني نسبي با كليدواژه‌هاي عنوان نيز (همانند موارد همخواني كامل) به شكلي از عنوان پايان‌نامه‌ها استخراج شده‌اند، بنابراين احتمالاً مي‌توان موارد همخواني (اعم از همخواني كامل و نسبي) را با موارد ناهمخواني مقايسه كرد. همان‌گونه كه در جدول 4 مشاهده مي‌شود 67 درصد توصيفگرهاي نمايه‌سازان- به طور كامل يا نسبي- با كليدواژه‌هاي عنوان همخواني دارند. حوزه‌هاي هنر (74%)، كشاورزي (71%)، علوم انساني (71%) و علوم پزشكي (69%) داراي همخواني بيشتري در مقايسه با حوزه‌هاي علوم پايه (65%) و فني و مهندسي (58%) هستند. در عين‌حال برخلاف تصور محققين، ميزان همخواني كامل در حوزه‌هاي فني و مهندسي و علوم پايه كمتر از ديگر حوزه‌ها بود. از اين‌رو براي اين كه بتوان تأثير خطاهاي انساني را به حداقل كاهش داد و به شكلي صحت يافته‌هاي تحقيق حاضر را ارزيابي نمود، و نيز زمينة بهره‌گيري از امكانات نرم‌افزاري رايج را براي انجام تحقيقات مشابه فراهم ساخت، سعي شد ميزان همخواني كامل توصيفگرها و كليدواژه‌هاي عنوان به شيوة ماشيني محاسبه شود. براي انجام اين كار از امكانات برنامة نرم‌افزاري «اكسل» استفاده شد. در جدول شماره 5 نتايج حاصل از محاسبة ماشيني، با محاسبه دستي در نمونة تحقيق مقايسه شده است.

جدول 5. مقايسة بررسي همخواني كامل كليدواژه‌هاي عنوان با توصيفگرها، به شيوة دستي و ماشيني

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| حوزه | علوم انساني | علوم پايه | علوم پزشكي | فني و مهندسي | كشاورزي | هنر | جمع |
| محاسبة دستي تعداد توصيفگر | 427 | 505 | 320 | 565 | 427 | 365 | 2609 |
| محاسبة ماشيني تعداد توصيفگر | 419 | 490 | 301 | 557 | 410 | 358 | 2535 |
| محاسبة دستي تعداد همخواني كامل | (49%)  210 | (41%)  209 | (50%)   160 | (35%)  200 | (50%)   215 | (52%)  190 | (45%)  1184 |
| محاسبة ماشيني تعداد همخواني كامل | (44%)  183 | (39%)  193 | (42%)   125 | (33%)  183 | (49%)   199 | (47%)  169 | (41%)  1052 |

جدول شماره 5 به طور كلي 4 درصد اختلاف را در محاسبة ماشيني و محاسبة دستي مربوط به ميزان همخواني كامل توصيفگرها و كليدواژه‌هاي عنوان نشان مي‌دهد. با استفاده از آزمون مجذور كاي، اختلاف بين محاسبة دستي و محاسبة ماشيني تعداد همخواني كامل ارزيابي شد و نتايج به‌دست‌آمده نشان داد كه تفاوت معناداري بين نتايج حاصل از شيوة دستي و ماشيني وجود ندارد.

4. ميزان همخواني كليدواژه‌هاي عنوان با توصيفگرهاي نمايه‌سازان، در سال‌هاي مختلف چه تفاوتي دارند؟

به منظور پاسخگويي به اين سؤال، ستون مربوط به تاريخ تهية پايان‌نامه‌ها در محيط نرم‌افزاري «اكسل» از كوچك به بزرگ مرتب شد. با توجه به اين كه مجموع پايان‌نامه‌هاي حضوريافته در نمونة تحقيق، مربوط به سال‌هاي قبل از 1335 كمتر از 50 ركورد بود پايان‌نامه‌هاي مربوط به اين سال‌ها از نمونة تحقيق حذف شدند. در ادامه، وضعيت همخواني كليدواژه‌هاي عنوان با توصيفگرها در سال ‌هاي مختلف بررسي شد كه نتايج حاصل از آن در جدول شماره 6 درج شده است.

جدول 6. توزيع فراواني همخواني كليدواژه‌هاي عنوان با توصيفگرها به تفكيك سال نگارش

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| سال نگارش  همخواني | -1335  1341 | -1342  1348 | -1349  1355 | -1356  1362 | -1363  1369 | -1370  1376 | -1377  1381 |
| همخواني كامل | 38  (69%) | 108   (58%) | 126   (59%) | 112   (56%) | 140   (54%) | 498   (38%) | 134   (41%) |
| همخواني نسبي | 10   (18%) | 34  (18%) | 55   (26%) | 56   (28%) | 58  (22%) | 257   (19%) | 92   (28%) |
| ناهمخواني | 7   (13%) | 43   (23%) | 33   (15%) | 32   (16%) | 61   (24%) | 567   (43%) | 101   (31%) |
| جمع | 55 | 185 | 224 | 200 | 259 | 1322 | 327 |

بر اساس داده‌هاي جدول فوق ميزان همخواني (كامل، نسبي، ناهمخواني) كليدواژه‌هاي عنوان با توصيفگرها در دوره‌هاي مختلف با استفاده از آزمون كاي مورد مقايسه قرار گرفت. نتايج حاصله نشان داد كه ميزان همخواني كامل، همخواني نسبي و ناهمخواني در دوره‌هاي مختلف داراي اختلاف معناداري است.

5. آيا با افزايش تعداد كليدواژه‌هاي عنوان، ميزان همخواني آن‌ها با توصيفگرها نيز افزايش مي‌يابد؟

براي پاسخگويي به اين سؤال در 6 طبقة توزيع فراواني، كليدواژه‌هاي عنوان و موارد همخواني (اعم از همخواني كامل و نسبي) مشخص شد تا رابطة اين دو متغير با يكديگر سنجيده شود. در عين حال بر اساس آزمون ضريب همبستگي پيرسون نيز ارتباط تعداد كليدواژه‌هاي عنوان و موارد همخواني كامل مورد ارزيابي قرار گرفت كه نتايج آزمون نشان داد در كلية حوزه‌ها (در سطح معناداري 1%) همبستگي بالايي بين تعداد كليدواژه‌هاي عنوان و موارد همخواني وجود دارد (جدول شماره 7).

جدول 7. ضريب همبستگي تعداد كليدواژه‌هاي عنوان و تعداد همخواني كامل در حوزه‌هاي شش‌گانه

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| حوزه | علوم انساني | | | علوم پايه | | | علوم پزشكي | | | فني و مهندسي | | | كشاورزي | | | هنر | | |
| محاسبه پيرسون | تعداد | سطح معناداري | ضريب | تعداد | سطح معناداري | ضريب | تعداد | سطح معناداري | ضريب | تعداد | سطح معناداري | ضريب | تعداد | سطح معناداري | ضريب | تعداد | سطح معناداري | ضريب |
| تعداد كليدواژه و تعداد همخواني كامل | 6 | 009/0 | (\*\*)922/0 | 6 | 002/0 | (\*\*)964/0 | 6 | 004/0 | (\*\*)948/0 | 6 | 000/0 | (\*\*)984/0 | 6 | 001/0 | (\*\*)969/0 | 6 | 004/0 | (\*)947/0 |

6. از نظر شيوة نگارش، توصيفگرها چگونه هستند؟

«مورتيمرتاب» در سال 1952 «نمايه‌سازي همارا» را به منظور بررسي گزارش‌هاي نظامي و مدارك ديگر ابداع كرد. وي در سيستم خود از تك‌واژه‌ها يا واژه‌هايي كه بيانگر يك مفهوم بودند استفاده كرد. فرض وي بر اين بود كه محتواي مدارك را مي‌توان در قالب چند مفهوم اصلي ريخت و مفاهيم به‌دست‌آمده را به كمك تك‌واژه‌ها بيان كرد. انتخاب تركيبات نامحدودي از واژه‌ها در هنگام بازيابي، عدم ضرورت رعايت ترتيب خاص واژه‌ها، سهولت استفاده از رايانه در ذخيره و بازيابي، از مزايا، و ريزش كاذب و نامشخص‌بودن روابط بين واژه‌ها جزو معايب اين شيوة نمايه‌سازي به شمار مي‌روند. در پايگاه اطلاعاتي توجه به واژه‌هاي تركيبي در ذخيره و بازيابي اطلاعات، امري ضروري به شمار مي‌رود. در ادامة اين تحقيق سعي شد تا تصويري از وضعيت توصيفگرهاي تك‌واژه و تركيب- واژه به دست داده شود. نتايج حاصل از اين بررسي در جدول شماره 8 آمده است.

جدول 8. نسبت تعداد توصيفگرهاي تك‌واژه به تعدادواژه‌هاي تركيبي، به تفكيك حوزه

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| حوزه | علوم انساني | علوم پايه | علوم پزشكي | فني و مهندسي | كشاورزي | هنر | جمع |
| مجموع توصيفگرها | 427 | 505 | 320 | 565 | 427 | 365 | 2609 |
| تعداد توصيفگر تك‌واژه | (46%) | (42%) | (47%) | (42%) | (59%) | (59%) | (48%) |
| تعداد توصيفگر با واژه‌هاي تركيبي | (54%) | (58%) | (53%) | (58%) | (41%) | (41%) | (52%) |

در جدول شماره8 نشان داده شده كه به طور كلي نسبت تقريباً يكساني در اختصاص تك‌‌واژه‌ها و تركيب- واژه‌ها وجود دارد. در عين حال حوزه‌هاي فني و مهندسي و علوم پايه (هر دو با 58 درصد) داراي تعداد تركيب- واژه‌هاي بيشتري در مقايسه با ديگر حوزه‌ها از جمله حوزة كشاورزي و هنر (هردو با 41 درصد) هستند. يكي از دلايل اين اختلاف احتمالاً به ساختمان زبانشناختي موضوعات مربوط مي‌شود كه حاصل اختلاف ساختمان اطلاعاتي آن‌ها است. مثلاً زبان يك شيميدان در حوزة شيمي آلي كاملاً با زبان مورد استفادة يك جانورشناس متفاوت است، زيرا اطلاعاتي كه از نظر يك متخصص شيمي اهميت دارد نام تركيبات آلي، آرايش ساختماني، و خواص فيزيكي و شيميايي است، ولي از ديدگاه يك جانورشناس توصيف و تعيين جنس و نوع حيوانات مطرح است (بهاتاچاريا، 1356ص. 14). از طرف ديگر، با توجه به اين كه مفاهيم بازنمودهاي اشيا، منفرد هستند (هلموت فلبر، 1381ص. 155) و محققين حوزه‌‌هاي فني و مهندسي و علوم پايه معمولاً با تركيب اشيا، مواد و موضوعات مختلف سروكار دارند، ناگزير عبارات نيز به عنوان نمادهاي زباني اين اشيا، به شكل تركيبي ظاهر مي‌شوند.

يكي ديگر از مواردي كه در خصوص شيوة نگارش توصيفگرها مورد بررسي قرار گرفت شكل «مفرد» يا «جمع» آن‌ها است. معمولاً جمع و مفرد در سرعنوانهاي موضوعي همة زبان‌ها، مشكلاتي را ايجاد مي‌كند. هميشه اين سؤال پيش مي‌آيد كه چه چيزهايي را جمع به كار بريم و چه چيزهايي را مفرد (پوري سلطاني، 1373، ص. 29). در زبان انگليسي معمولاً استفاده از شكل جمع كلمات (تيلور، 1381ص. 187؛ Rydings, 1999) و در زبان فارسي استفاده از شكل مفرد در نمايه‌سازي پيشنهاد مي‌شود (سمايي، 1379ص. 31). در تحقيق حاضر كلماتي كه با «ان»، «ها»، «ات»، «جات» و كلماتي كه طبق قاعدة زبان عربي به شكل مكسر جمع بسته شده بودند بررسي شدند، كه نتايج حاصل از آن، برابر داده‌هاي جدول شماره 9 است.

جدول 9. نسبت توصيفگرهاي مفرد به توصيفگرهاي جمع، به تفكيك حوزه

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| حوزه | علوم انساني | علوم پايه | علوم پزشكي | فني و مهندسي | كشاورزي | هنر | جمع |
| كل توصيفگرها | 427 | 505 | 320 | 565 | 427 | 365 | 2609 |
| توصيفگر مفرد | (83%) | (87%) | (83%) | (92%) | (90%) | (85%) | (87%) |
| توصيفگر جمع | (17%) | (13%) | (18%) | (8%) | (10%) | (15%) | (13%) |

به طور كلي 87 درصد توصيفگرها به شكل مفردو 13 درصد به شكل جمع به كار رفته‌اند.

7. در پايگاه «چكيده پايان‌نامه‌هاي ايران» بين ميزان بازيافت با استفاده از كليدواژه‌هاي عنوان و با استفاده از توصيفگرهاي نمايه‌سازان چه تفاوتي وجود دارد؟

براي دريافت اطلاعات از يك نظام، استفاده‌كننده بايد با مكانيسم‌هاي مربوط به نظام نمايه ارتباط برقرار كند (ويسمان، 1373ص. 106). هرچه اين ارتباط سهل‌تر و عاري از پيچيدگي باشد و بين زبان متقاضي اطلاعات و زبان نمايه‌ساز، هماهنگي بيشتري پديد آيد، ضريب بازيافت و دقت بالاتر خواهد بود. پيوستگي بين نظام نمايه‌سازي و بازيابي اطلاعات به حدي است كه اساساً يكي از معيارهاي ارزشيابي فعاليت‌هاي نمايه‌سازي مدارك را، بازيابي جامع اطلاعات موردنظر در جستجو محسوب نموده‌اند (پائو، 1378ص. 115).

استفاده از كليدواژه‌ها و توصيفگرها در نمايه‌سازي، هر يك اثر خاصي بر نتايج بازيابي خواهد داشت. در هر پايگاه اطلاعاتي مداركي وجود دارد كه تعدادي از آنها مرتبط با موضوع مورد جستجو و بسياري نيز نامرتبط هستند. از ميان مدارك مرتبط نيز برخي با نياز جستجوگر مغايرت دارند. توانايي نظام در بازيابي مدارك مرتبط را «جامعيت» و توانايي نظام در كنار گذاشتن مدارك نامرتبط را «مانعيت» گويند. براي اين‌كه پايگاه «چكيدة پايان‌نامه‌هاي ايران» از منظر جامعيت ارزيابي شود، طي سه مرحله موارد بازيافت از طريق كليدواژه‌ها و توصيفگرها، مقايسه شدند. در مرحله نخست جستجو از طريق كليدواژه‌ها و توصيفگرهايي كه داراي همخواني نسبي بودند صورت گرفت (مواردي چون كليدواژة «شيعه» با توصيفگر «تشيع» در يك ركورد، يا كليدواژه‌ «هيئت علمي» با توصيفگر «هيئت علمي» در يك ركورد). پس از شناسايي كليدواژه‌ها و توصيفگرهاي يادشده در 292 ركورد، جستجو در پايگاه «چكيدة پايان‌نامه‌هاي ايران» صورت گرفت كه نتايج حاصل از آن در جدول شماره 10 آمده است.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| حوزه | علوم انساني | علوم پايه | علوم پزشكي | فني و مهندسي | كشاورزي | هنر | جمع |
| تعداد بازيابي با استفاده از كليدواژه | (47%) | (39%) | (44%) | (37%) | (38%) | (46%) | (44%) |
| تعداد بازيابي با استفاده از توصيفگر | (53%) | (61%) | (56%) | (63%) | (62%) | (54%) | (56%) |

بدين ترتيب در پاسخ به سؤال اخير بايد اذعان داشت كه بازيابي با استفاده از توصيفگرها (56%) بيش از موارد بازيابي شده با استفاده از كليدواژه‌هاي عنوان (44%) است. با استفاده از آزمون مجذور كاي، تفاوت معناداري بازيافت با استفاده از كليدواژه‌‍‌هاي عنوان و توصيفگرها محاسبه شد كه يافتة حاصل از اين محاسبه، تفاوت معناداري را در بازيافت از طريق كليدواژه‌هاي عنوان و توصيفگرهاي نمايه‌سازان نشان داد. دليل بازيابي و جامعيت بيشتر از طريق توصيفگرها، انجام طرح پژوهشي، «سازماندهي توصيفگرهاي پايگاه‌هاي اطلاعاتي مركز اطلاعات و مدارك علمي ايران (جستجوي هوشمند)»*[[15]](http://www.aqlibrary.org/editor/main.htm" \l "_ftn15" \o ")* است كه توسط واحد خدمات ماشيني مركز در سال 1380 صورت پذيرفت. قبل از انجام اين طرح به دلايلي چون فقدان دستورالعمل مشخص و مدون براي نمايه‌سازي، تغيير كارشناسان و نيز وجود سليقه‌هاي مختلف در امر نمايه‌سازي، استفاده از اصطلاحنامه‌ها، واژگان و واژه‌نامه‌هاي مختلف و ... باعث ناهمگوني و نابساماني بسيار در توصيفگرها شده بود و پس از هر جستجو، ركوردهاي مرتبط بسياري بازيابي نمي‌شد. در مرحلة بعد، بازيافت از طريق كليدواژه و توصيفگرهايي مقايسه شد كه داراي همخواني كامل بودند. تعداد 521 ركورد كه حداقل داراي يك كليدواژه و توصيفگر با همخواني كامل بودند، شناسايي شدند و جستجو در آن‌ها صورت گرفت. داده‌هاي به‌دست آمده نشان داد كه ميزان بازيافت از طريق كليدواژه‌ها و توصيفگرهايي كه داراي همخواني كامل هستند تفاوتي ندارد، هرچند انتظار مي‌رفت به دلايلي چون فاصلة خالي اضافي در بين كليدواژه‌ها يا توصيفگرهاي مركب، ميزان بازيافت بين كليدواژه‌ها و توصيفگرهايي كه همخواني كامل دارند نيز تفاوت داشته باشد؛ مثلاً تعداد بازيافت از طريق كليدواژه «قارچكشها» با تعداد بازيافت از طريق توصيفگر «قارچ كش‌ها» تفاوت كند. ولي در يكي از بخش‌هاي طرح «ساماندهي توصيفگرهاي پايگاه‌هاي اطلاعاتي مركز اطلاعات و مدارك علمي ايران (جستجوي هوشمند)» امكاناتي در نظر گرفته شده كه فاصلة خالي اضافي بين واژه‌هاي مركب حذف شود. در مرحلة پاياني، سنجش ميزان بازيافت از كليدواژه‌ها و توصيفگرهايي صورت گرفت كه در ظاهر هيچ‌گونه همخواني با يكديگر نداشتند، ولي مترادف بودند؛ مانند كليدواژه‌ «سوبسيد» با توصيفگر «يارانه» يا كليدواژه «آبستني» با توصيفگر «حاملگي». در اين مرحله 102 ركورد با ويژگي مورد نظر شناسايي و عمليات جستجو بر روي آن‌ها صورت گرفت (جدول شماره 11).

جدول 11. ميزان بازيافت از طريق كليدواژه‌ها و توصيفگرهاي مترادف به تفكيك حوزه

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| حوزه | علوم انساني | علوم پايه | علوم پزشكي | فني و مهندسي | كشاورزي | هنر | جمع |
| تعداد بازيابي با استفاده از كليدواژه | (44%) | (41%) | (47%) | (38%) | (42%) | (25%) | (39%) |
| تعداد بازيابي با استفاده از توصيفگر | (56%) | (59%) | (53%) | (62%) | (58%) | (75%) | (61%) |

داده‌هاي اين جدول نشان مي‌دهد كه بازيافت از طريق توصيفگرهاي مترادف بيش از موارد بازيافت از طريق كليدواژه‌ها است. صورت پذيرفتن طرح «ساماندهي توصيفگرهاي پايگاه‌هاي اطلاعاتي مركز اطلاعات و مدارك علمي ايران (جستجوي هوشمند)»، اجتناب نمايه‌سازان از اختصاص واژه‌هاي مهجور و نامتداول (مانند «كتابخانه همگاني» به جاي «كتابخانه عمومي» يا «بندر پهلوي» به جاي «بندر انزلي»)، اجتناب از اختصاص معادل‌هاي لاتين (مانند «آناليز» به جاي «تجزيه و تحليل») و مانند آن را مي‌توان از جمله دلايل ايجاد يكپارچگي بيشتر بين توصيفگرها و افزايش ميزان بازيافت از طريق توصيفگرهاي مترادف با كليدواژه‌ها دانست.

نتيجه‌گيري

عنوان پايان‌نامه‌ها و كليدواژه‌هاي تشكيل‌دهندة آن‌ها از منظر خوانندگان و پژوهشگران داراي اهميت است و ايشان را در اولين برخورد، در تشخيص مدرك مربوط يا نامربوط ياري مي‌دهد. از طرف ديگر نمايه‌سازان نيز از نقش و اهميت «عنوان» در ساماندهي اطلاعات غافل نبوده‌اند و از ديرباز تاكنون به عنوان يكي از مهم‌‍‌ترين ملاك‌هاي توصيف‌كنندة مدارك، در ايجاد پايگاه‌هاي اطلاعاتي از آن بهره جسته‌اند. با توجه به اين نكته در تحقيق حاضر نخست ساختار و اجزاي تشكيل‌دهندة عنوان پايان‌نامه‌ها در حوزه‌هاي مختلف بررسي و اين نتيجه حاصل شد كه به طور كلي عنوان پايان‌نامه‌ها داراي 7 كليدواژه است كه در ذخيره و بازيابي اطلاعات مربوط به پايان‌نامه‌ها مي‌توان از آن بهره جست؛ هرچند در اين خصوص بين تعداد متوسط كليدواژه‌هاي عنوان پايان‌نامه‌ها در حوزه‌هاي مختلف تفاوت معناداري وجود دارد (جدول شماره 2). همچنين نشان داده شد كه در حوزه‌هاي مختلف، توزيع فراواني تعداد كليدواژه‌هاي عنوان در پايان‌نامه‌ها يكسان نيست. به عنوان مثال در حوزه‌هاي هنر و علوم پزشكي برخلاف حوزه‌هاي فني و مهندسي و علوم پايه بيشتر پايان‌نامه‌ها داراي 1 تا 5 كليدواژه هستند. از طرف ديگر به طور متوسط در هر عنوان پايان‌نامه 2 تا 3 واژه بدون بار معنايي وجود دارد كه به منظور جلوگيري از حجيم شدن پايگاه اطلاعاتي و فراهم آوردن زمينة نمايه‌سازي ماشيني، بايد انواع آن‌ها را شناسايي كرد و به عنوان مجموعه‌اي از واژگان بدون بار معنايي، در امر ذخيره و بازيابي اطلاعات مورد توجه قرار داد (جدول شماره 3). 45 درصد همخواني كامل بين كليدواژه‌هاي عنوان و توصيفگرهاي نمايه‌سازان يا به عبارت ديگر اتخاذ نيمي از كليدواژه‌هاي عنوان به منزلة توصيفگر، و به طور كلي وابستگي 67 درصد توصيفگرهاي نمايه‌سازان (همخواني كامل و نسبي) به كليدواژه‌هاي عنوان (جدول شماره 6) تأييد ديگري بر نقش و اهميت عنوان مدارك در ساماندهي اطلاعات مربوط به آن‌ها است. محققان ديگري چون «فراست» (Frost, 1989) و «بزرگي» (1377) نيز هر دو در بررسي خود نشان داده‌اند كه 73 درصد واژه‌هاي عنوان دقيقاً يا حداقل با بخشي از واژه‌هاي سرعنوان‌هاي موضوعي همپوشاني دارند. از طرف ديگر نشان داده شد كه بين تعداد كليدواژه‌هاي عنوان و تعداد توصيفگرها و همچنين ميزان همخواني آن‌ها همبستگي بالايي وجود دارد (جداول شماره 5 و 9). در عين حال مشاهده شد كه با گذشت زمان و احتمالاً بنا به دلايلي چون تدوين اصطلاحنامه‌ها و آموزش نمايه‌سازان، ميزان وابستگي به كليدواژه‌هاي عنوان كاهش يافته و از اين مراجع در اختصاص توصيفگرها استفاده شده است (جدول شماره 8). البته استفاده از اين منابع در نمايه‌سازي از نقش و اهميت عناوين نمي‌كاهد، بلكه استفاده توأمان از كليدواژه‌هاي عنوان و اصطلاحنامه‌ها منجر به افزايش جامعيت و مانعيت در ذخيره و بازيابي اطلاعات خواهد شد.

بررسي كليدواژه‌هاي عنوان و توصيفگرها از نظر شيوة نگارش نيز يكي ديگر از اجزاي اين تحقيق بود و مشخص شد كه به طور كلي با درنظر گرفتن كلية حوزه‌ها، نسبت تقريباً يكساني بين توصيفگرهاي تكواژه (48 درصد) و تركيب- واژه (52 درصد) وجود دارد. ام‍ّا درصد تركيب‌واژه‌ها در حوزه‌هاي فني و مهندسي و علوم پايه بيش از ديگر حوزه‌ها است. علاوه بر اين مشخص شد كه توصيفگرها از منظر مفرد (87 درصد) و جمع (13درصد) بودن تا حد‌ّ زيادي با مفرد (84 درصد) و جمع (16 درصد) كليدواژه‌ها همخوان هستند. و شايد به همين دليل در زبان فارسي برخلاف زبان انگليسي بيشتر از واژه‌هاي مفرد در امر نمايه‌سازي استفاده مي‌شود (جدول شماره 11). در بخش پاياني اين تحقيق ميزان بازيافت از طريق كليدواژه‌هاي عنوان و توصيفگرهاي نمايه‌سازان در سه حالت همخواني نسبي، همخواني كامل و ناهمخواني، به عنوان يكي از ملاك‌هاي ارزيابي پايگاه‌هاي اطلاعاتي بررسي شد. در حالت همخواني نسبي، فزوني بازيافت از طريق توصيفگرها (56 درصد) در مقايسه با كليدواژه‌ها و نيز بازيافت مساوي در موارد همخواني كامل، نقش و اهميت انجام فرايندهايي به منظور ايجاد ارتباط منطقي بين توصيفگرها يا كليدواژه‌ها را آشكار ساخت. در خصوص كليدواژه‌ها و توصيفگرهاي مترادف نيز به دليل يكپارچگي بيشتر و برقرارشدن ارتباط منطقي بين توصيفگرهاي هم مفهوم (در پي اجراي ساماندهي توصيفگرها)، موارد بازيافت از طريق توصيفگرها بيش از كليدواژه‌ها بود.

منابع

انصاري، مريم (1379). بررسي انطباق ميان توصيفگرهاي نمايه‌سازي و كليدواژه‌هاي عنوان پايان‌نامه‌هاي دكتراي تخصصي كودكان، زنان، قلب و عروق و روانپزشكي. پايان‌نامه كارشناسي ارشد كتابداري و اطلاع‌رساني پزشكي، دانشكده مديريت و اطلاع‌رساني پزشكي، دانشگاه علوم پزشكي و خدمات بهداشتي- درماني ايران.

بزرگي، اشرف السادات (1377). بررسي ميزان سازگاري عنوان‌ها و موضوع‌هاي كتاب‌هاي فارسي منتشرشده در ايران در فاصله سال‌هاي 1375-1365. پايان‌نامه دكترا، واحد علوم و تحقيقات دانشگاه آزاد اسلامي.

بهاتاچاريا (1974). «تأثير زبان طبيعي در نمايه‌سازي و بازيابي علوم». ترجمه مصطفي سيد صادق. 1356. فصلنامه اطلاع‌رساني. 4 (1و2): 20-16.

پائو، ميراندا لي (1989). مفاهيم بازيابي اطلاعات. ترجمه اسدالله آزاد و رحمت‌الله فتاحي. 1379. مشهد: دانشگاه فردوسي.

تيلور، آرلين جي (1996). سازماندهي اطلاعات. ترجمه محمد حسين دياني. 1381. مشهد: كتابخانه رايانه‌اي.

حر‌ّي، عباس (1372). مروري بر اطلاعات و اطلاع‌رساني. تهران: دبيرخانه هيئت امناي كتابخانه‌هاي كشور، نشر كتابخانه.

حريري، مهرانگيز، مسعودة توفيق (1362). «نمايه گردان». اطلاع‌رساني، نشريه فني مركز اسناد و مدارك علمي. 8 (2): 5-4.

داورپناه، محمدرضا (1375). «بررسي ميزان سازگاري عناوين مقالات فارسي با محتواي آن‌ها». فصلنامه اطلاع‌رساني، 13 (2): 12-1.

سرايي، حسن (1372). مقدمه‌اي بر نمونه‌گيري در تحقيق. تهران: سازمان مطالعه و تدوين كتب علوم انساني.

سلطاني، پوري (1379). دانشنامه كتابداري و اطلاع‌رساني. تهران: فرهنگ معاصر.

سمايي، سيد مهدي (1379). «مفرد و جمع در نمايه‌سازي». فصلنامه اطلاع‌رساني. 16 (1و2): 31-27.

فلبر، هلموت (1984). مباني اصطلاح‌شناسي. ترجمه محسن عزيزي. 1381، تهران: مركز اطلاعات و مدارك علمي ايران.

معرف‌زاده، عبدالحميد (1366). بررسي وضعيت ابزار بازيابي اطلاعات پزشكي فارسي و پيشنهادي براي تهيه نمايه پزشكي ايران. پايان‌نامه كارشناسي ارشد كتابداري و اطلاع‌رساني پزشكي، دانشكده مديريت و اطلاع‌رساني پزشكي، دانشگاه علوم پزشكي و خدمات بهداشتي درماني ايران.

ويسمان، هرمان. (بي تا). سيستم‌ها و خدمات اطلاع‌رساني كامپيوتر. ترجمه جعفر مهراد 1373. شيراز: انتشارات نويد.

Aanonson, John (1987). "Precision and Recall in Title keyword searchers". Information technology and libraries, 14(3): 162-170.

Babbie, Earl R (1979). The practice of Social Research. 2nd ed. Belmont, CA: Wadsworth, P.247.

Bottle, Robert, Preibish (1970). "Proposed KWIC index for Psychology". Journal of the American Society for information science, (11,12): 427-8.

Frost, Carolyn (1989). Title word as entry vocabulary LCSH: Correlation between Assigned LCSH terms and derived terms from Title". Cataloging & Classification Quartrly, 10 (1,2); 165-179.

Henzler, R. G (1978). "Free or controlled vocabularies". International Classification, 5.21-26.

Hmeidi, Ismail. (1997). Design and Implementation of Automatic Indexing for Information Retrieval with Arabic Documentation". Journal of the American Society for Information Science, 48 (10): 867-88.

Keller, Barbara. (1992). "Subject content through Title: A Master Matching study in Indian State University". Cataloging & classification Quarterly, 15 (3): 69-80.

Lancaster, F. w. et al. (1991). "Identifying barriers to effective subject access in Library catalogs". Library Resources and Thechnical Services, 35. 377-391.

Newmark, Peter. (1988). Approaches to translation, Prentice Hall International.

Rydings, Tony. (1999)."Idexing of periodicals-format of entries". Society of indexers Genealogical Group newsletter.NO. 20: p. 5-8. [Online]

[*http://www.sigg.org.uk/arts/arts2001.htm.[22sep2003]*](http://www.sigg.org.uk/arts/arts2001.htm.%5b22sep2003)

Shepherd, Michael; A., Philips; WJ. Lui. (1983). "Overlap of title keyword and Assigned Descriptors in the OON Database". Canadian Journal of Information Service, 12 (1): 58-66.

Spink, Amanda; Saracevic, Tefko. (1993). "Interaction in Information Retrieval: Selection & Effectiveness of Search Terms". Journal of the American Society for Information Science, 48 (8): 741-761.

Voorbij, Henk. (1998). Title kewords and subject descriptors A Comparison of subject search entries of books in the humanities and Social sciencse". Journal of Documentation. 54 (4): 466-476.

نكته:

با پوزش از محضر جناب آقاي دكتر جعفر مهراد؛ پانوشت مقالة «ميزان رضايت مراجعان از خدمات كتابخانه و مراكز اطلاع‌رساني صنعت نفت» در صفحة 95 فصلنامة كتابداري و اطلاع‌رساني شمارة 26 (تابستان 1383) بدين‌نحو تصحيح مي‌گردد:

دكتر جعفر مهراد، استاد گروه كتابداري و اطلاع‌رساني دانشگاه شيراز.

محمدعلي جهانيان، كارشناس ارشد علوم كتابداري و اطلاع‌رساني.

1. كارشناس ارشد كتابداري و اطلاع‌رساني

2. استاديار گروه كتابداري و اطلاع‌رساني دانشگاه فردوسي مشهد

*[[3]](http://www.aqlibrary.org/editor/main.htm" \l "_ftnref3" \o ")*. Descriptive

*[[4]](http://www.aqlibrary.org/editor/main.htm" \l "_ftnref4" \o ")*. Allusive

*[[5]](http://www.aqlibrary.org/editor/main.htm" \l "_ftnref5" \o ")*. Psychological Abstracts

*[[6]](http://www.aqlibrary.org/editor/main.htm" \l "_ftnref6" \o ")*. Cancer net

*[[7]](http://www.aqlibrary.org/editor/main.htm" \l "_ftnref7" \o ")*. OON= Object Oriented Networking

*[[8]](http://www.aqlibrary.org/editor/main.htm" \l "_ftnref8" \o ")*. CISTI= Canada institute for Scienctific and Technical information

*[[9]](http://www.aqlibrary.org/editor/main.htm" \l "_ftnref9" \o ")*. LCSH= Library of Congress Subject Headings

*[[10]](http://www.aqlibrary.org/editor/main.htm" \l "_ftnref10" \o ")*. DIALOG

*[[11]](http://www.aqlibrary.org/editor/main.htm" \l "_ftnref11" \o ")*. CDS-ISIS

*[[12]](http://www.aqlibrary.org/editor/main.htm" \l "_ftnref12" \o ")*. EXCEL

*[[13]](http://www.aqlibrary.org/editor/main.htm" \l "_ftnref13" \o ")*. Golist

*[[14]](http://www.aqlibrary.org/editor/main.htm" \l "_ftnref14" \o ")*. Stoplist

1. پارسي اصفهاني، امير (1380) ساماندهي توصيفگرهاي پايگاه‌هاي اطلاعاتي مركز اطلاعات و مدارك علمي ايران (جستجوي هوشمند). طرح پژوهشي. تهران: مركز اطلاعات و مدارك علمي ايران.