

## بررسی میزان همخوانی کلیدواژه‌های عنوان و توصیفگرهای نمایه‌سازان در پایگاه اطلاعاتی «چکیده پایانه‌های ایران»

تألیف: محمد ایرانشاهی<sup>1</sup>

دکتر محمدرضا داورپناه<sup>2</sup>

### چکیده

پژوهش حاضر با هدف ارزیابی میزان همخوانی کلیدواژه‌های عنوان در پایانه‌های مندرج در «پایگاه چکیده‌نامه‌های ایران» با توصیفگرهای اختصاص یافته به این پایانه‌ها صورت پذیرفت. جامعه مورد بررسی، کلیه رکوردهای موجود در این پایگاه (65536 رکورد) از ابتدا تا پایان سال 1381 است. با استفاده از فرمول تعیین حجم نمونه از هر حوزه علوم انسانی، علوم پایه، علوم پزشکی، فنی و مهندسی، علوم کشاورزی و هنر، 100 رکورد به شیوه تصادفی انتخاب شد. نتایج پژوهش نشان داد که بین تعداد متوسط کلیدواژه‌های عنوان در حوزه‌های مورد بررسی تفاوت معناداری وجود دارد و 67 درصد توصیفگرهای نمایه‌سازان را کلیدواژه‌های عنوان تشکیل می‌دهند. همبستگی بالایی بین تعداد کلیدواژه‌های عنوان و تعداد توصیفگرها و میزان همخوانی آن‌ها وجود دارد. از نظر شیوه نگارش 48 درصد توصیفگرها به صورت تک‌واژه، 52 درصد ترکیب‌واژه، 87 درصد توصیفگرها مفرد و 13 درصد آن‌ها به صورت جمع تهیه شده‌اند. باز یافت بیشتر از طریق توصیفگرها در مقایسه با کلیدواژه‌های عنوان، اثربخشی طرح‌هایی را نشان داد که به منظور افزایش ارتباط بین توصیفگرهای هم‌موضوع در این پایگاه صورت پذیرفته است.

کلیدواژه‌ها: کلیدواژه‌ها، توصیفگرها، چکیده پایانه‌های ایران، پایانه‌های ایران

<sup>1</sup>. کارشناس ارشد کتابداری و اطلاع‌رسانی

<sup>2</sup>. استادیار گروه کتابداری و اطلاع‌رسانی دانشگاه فردوسی مشهد

## مقدمه

«عنوان» نخستین بخش از هر اثر است که نظر خواننده را به خود جلب می‌کند. از دیدگاه خواننده یا پژوهشگر، «عنوان» دارای ارزش زیادی است. «عنوان» هر اثر جلوه‌ای از هویت آن اثر می‌باشد و به تعبیری نخستین جلوه از متن است که خواننده با آن روبرو می‌شود. هر کتاب، مقاله، داستان، شعر، فیلم یا وب‌سایت، جدا از نام پدیدآور، با نام عنوانش نیز شناخته می‌شود. عنوان ظرفی است که مظروف آن، اندیشه اصلی متن است. گاه عنوان نقش اصلی خود را ایفا نمی‌کند و رساننده مباحث متن نیست؛ حتی ممکن است به صورت استعاری بیان شده باشد. این امر در ارتباط با آثار هنری یا برخی مباحث علوم انسانی بیشتر مشاهده می‌شود. عناوینی چون «هنر عشق ورزیدن»، «رنج و سرمستی»، یا «خسی در میقات» اطلاعات خاصی درباره متن به دست نمی‌دهند؛ یا از عنوان وب‌سایت *What the Thunder said* معلوم نمی‌شود که حاوی مطلبی درباره زندگی و آثار «تی. اس. الیوت» است (تیلور، 1381، 167). «نیومارک» (Newmark, 1988: 57) عناوین را به دو گروه «عناوین توضیحی»<sup>1</sup> و «تلمیحی»<sup>2</sup> تقسیم کرده است. به نظر وی در متون غیرادبی، «عنوان توضیحی» به اختصار به موضوع اثر اشاره می‌کند و هدف متن را به نحوی درخور بیان می‌دارد و عنوان تلمیحی نیز نوعی ارتباط کنایی یا استعاری با موضوع دارد. برای آثار علمی، پژوهشی و اطلاعاتی، استفاده از «عنوان تلمیحی» مطلوب نیست. بسیاری از ناشران متون علمی می‌کوشند این دو نقش عنوان - یعنی زیبایی و اطلاع‌رسانی - را تلفیق کنند، و این کار را با استفاده از عنوان فرعی انجام می‌دهند. در این گونه تلفیق‌ها، عنوان اصلی عمدتاً نقش زیبایی‌شناسانه و عنوان فرعی نقش اطلاعاتی دارد. مجلات معتبر دنیا نیز یکی از شرایط قبول مقاله را برابری عنوان با محتوای متن تعیین می‌کنند، و این کار را به دو دلیل انجام می‌دهند:

1. به دلیل آن که صورت مسئله تحقیق باید به گویاترین وجه، در عنوان متجلی باشد؛
2. به این دلیل که عنوان، پایگاه موضوعی بسیاری از سازمان‌هایی است که به کار نمایه‌سازی ماشینی مقاله‌ها و مدارک دیگر می‌پردازند. فرضی که ورای چنین گزینشی نهفته، این است که مجموعه کلیدواژه‌های عنوان، جلوه‌ای از موضوع کلی مدرک است.

---

<sup>1</sup>. Descriptive

<sup>2</sup>. Allusive

روش استفاده از کلیدواژه‌های عنوان برای تعیین و منعکس کردن موضوع مدارک، هنوز هم در انواع نمایه‌سازی‌ها معمول است. حتی اگر عنوان رساننده موضوع متن نباشد، برای این منظور از عنوان ساختگی (که به چکیده عنوانی معروف است) استفاده می‌کنند؛ یعنی عنوانی را کنار عنوان ساخته مؤلف به کار می‌برند که تجلی مباحث متن باشد (حری، 1372، 179). حال، مسائلی که در این خصوص مطرح می‌شود این است که در نمایه‌سازی پایان‌نامه‌ها، «عنوان» تا چه حد می‌تواند در ارائه محتوای این مدارک نقش داشته باشد؟ نمایه‌سازان «مرکز اطلاعات و مدارک علمی ایران» تا چه حد از عنوان پایان‌نامه‌ها در نمایه‌سازی بهره می‌گیرند؟ این بهره‌گیری در حوزه‌های مختلف، چه تفاوتی دارد؟ آیا با افزایش کلیدواژه‌های عنوان میزان همخوانی بین کلیدواژه‌ها و توصیفگرها افزایش می‌یابد؟

### هدف‌های پژوهش

هدف این پژوهش بررسی ساختار و اجزای تشکیل‌دهنده عنوان پایان‌نامه‌ها در حوزه‌های مختلف، نشان دادن تشابهات و تمایزها در عنوان پایان‌نامه‌های هر حوزه، و نیز بررسی ارتباط این اجزا (بوژه کلیدواژه‌های آن) با توصیفگرهایی است که به منظور نمایه‌سازی در پایگاه «چکیده پایان‌نامه‌های ایران» از آن‌ها استفاده شده. نتایج حاصل از این تحقیق نشان خواهد داد که عنوان پایان‌نامه‌ها تا چه حد بیانگر محتوای موضوعی است، و به منظور نمایه‌سازی ماشینی پایان‌نامه‌ها - که عمدتاً با استفاده از کلیدواژه‌های عنوان صورت می‌گیرد و در مقایسه با نمایه‌سازی کنترل‌شده، از هزینه کمتر و سرعت بیشتری برخوردار است - تا چه حد می‌توان به عنوان متکی بود. علاوه بر این، یافته‌های این پژوهش متخصصان نمایه‌سازی را در اتخاذ تصمیم‌های مناسب برای افزایش نقاط دستیابی به منابع اطلاعاتی و نیز بهبود شیوه‌های ذخیره و بازیابی اطلاعات پایان‌نامه‌های کشور یاری می‌دهد.

### سؤالات پژوهش

1. چه تفاوتی بین تعداد متوسط کلیدواژه‌های عنوان پایان‌نامه‌ها در حوزه‌های علوم انسانی، علوم پایه، علوم پزشکی، فنی و مهندسی، کشاورزی و هنر وجود دارد؟
2. در حوزه‌های علوم انسانی، علوم پایه، علوم پزشکی، فنی و مهندسی، کشاورزی و هنر چه تفاوتی بین میزان همخوانی کلیدواژه‌های عنوان با توصیفگرهای نمایه‌سازان

وجود دارد؟

3. آیا با افزایش تعداد کلیدواژه‌های عنوان، میزان همخوانی آن‌ها با توصیفگرها نیز افزایش می‌یابد؟
4. بین تعداد کلیدواژه‌های عنوان پایان‌نامه‌ها با تعداد توصیفگرهای اختصاص یافته، چه رابطه‌ای وجود دارد؟
5. ساختار توصیفگرها از نظر ساده و مرکب، و مفرد و جمع بودن چگونه است؟
6. در پایگاه «چکیده پایان‌نامه‌های ایران» بین میزان بازیافت با استفاده از کلیدواژه‌های عنوان و توصیفگرهای نمایه‌سازان، چه تفاوتی وجود دارد؟

### پیشینه پژوهش

«باتل» و «پرایبیش» (Bottle and Preibish, 1970) در پژوهشی به بررسی سودمندی نمایه‌سازی کوئیک در مورد مقالات روانشناسی پرداختند و دریافتند در «چکیده‌نامه روانشناسی»<sup>1</sup> 60 درصد مقالات با استفاده از کلیدواژه‌های عنوان قابل بازیابی است.

«هنزler» (Henzler, 1978) در پژوهشی با بررسی پایگاه اطلاعاتی «کنسرت»<sup>2</sup> به این نتیجه رسید که 35 درصد از کلیدواژه‌های عناوین، فاقد توصیفگرهای مناسب در اصطلاحنامه هستند. از نظر او استفاده همزمان از کلیدواژه‌ها و توصیفگرها منجر به بازیابی مطلوب خواهد شد.

پژوهش «انانسون» (Ananson, 1987) به ارزیابی جستجوی کلیدواژه‌ای پرداخت و نشان داد نه تنها جستجوی کلیدواژه‌ای عنوان باعث بازیابی مدارک مرتبط می‌شود، بلکه با چنین جستجویی، رکوردهایی بازیابی می‌شوند که با جستجوی موضوعی قابل بازیابی نیستند.

«شپرد» (Shepherd, 1987) در تحقیقی به بررسی میزان همخوانی کلیدواژه‌های عنوان و توصیفگرهای موضوعی در پایگاه اطلاعاتی «اُان»<sup>3</sup> پرداخت. در این پایگاه که

---

<sup>1</sup> . Psychological Abstracts

<sup>2</sup> . Cancer net

<sup>3</sup> . OON= Object Oriented Networking

وابسته به مؤسسه اطلاعات علمی و فنی کانادا<sup>1</sup> می‌باشد، گزارش کنفرانس‌ها، تک‌نگاشتها، و گزارش‌های فنی در حوزه علوم پزشکی و فناوری نگهداری می‌شوند. نتایج این بررسی نشان داد که استفاده کنندگان این پایگاه برای دسترسی به بیشترین اطلاعات، باید از هر دو روش بازیابی استفاده نمایند.

«فراست» (Frost, 1989) در پژوهشی به بررسی همپوشانی بین سرعنوان‌های برگزیده از «سرعنوان‌های موضوعی کتابخانه کنگره»<sup>2</sup> و واژه‌های استخراج‌شده از عنوان، در فهرست کتابخانه دانشگاه میشیگان پرداخت. یافته‌های این پژوهش حاکی از آن بود که 73 درصد واژه‌های عنوان، با تمام یا حداقل با بخشی از سرعنوان یا تقسیمات فرعی، همپوشانی دارند. در 11 درصد موارد نیز واژه‌های عنوان، دقیقاً با سرعنوان برابر است و در بیش از 53 درصد موارد، واژه‌های عنوان حداقل با یک واژه از سرعنوان همپوشانی دارند. «کلر» (Keller, 1992) در پژوهشی بر روی پایان‌نامه‌های مقطع کارشناسی ارشد کتابخانه دانشگاه ایالتی ایندیانا به مقایسه نخستین سرعنوان موضوعی کنگره - که به پایان‌نامه اختصاص داده شده بود - با واژه‌های عنوان پرداخت تا مشخص کند احتمال وجود کلیدواژه‌های عنوان در سرعنوان‌های موضوعی چقدر است. نتایج حاصل از این بررسی نشان داد که از 371 پایان‌نامه، در 238 عنوان (64 درصد) بین واژه‌های عنوان و سرعنوان موضوعی، همپوشانی وجود دارد؛ یعنی در پایان‌نامه‌هایی که دارای 6-1 واژه در عنوان هستند 13 درصد، پایان‌نامه‌های دارای 7-12 واژه در عنوان 47 درصد، 13-18 واژه در عنوان 30 درصد، و 19-24 واژه در عنوان 10 درصد همپوشانی میان واژه‌های عنوان و نخستین سرعنوان موضوعی وجود دارد. بدین ترتیب مشخص شد که با افزایش تعداد واژه‌های تشکیل‌دهنده عنوان، میزان همپوشانی سرعنوان‌های موضوعی با واژه‌های عنوان کاهش می‌یابد.»

«اسپینک» و «ساراسویک» (Spink and Saracevic, 1993) در پژوهشی قابلیت کلیدواژه‌ها در بازیابی اطلاعات مرتبط در پایگاه «دیالوگ»<sup>3</sup> را مورد بررسی قرار دادند. نتایج پژوهش نشان داد 40 درصد واژه‌های مورد استفاده برای جستجو توسط مراجعه‌کننده ارائه شده بودند و در بازیابی موضوعات مرتبط، نقش مثبتی داشتند. همچنین 67 درصد از

<sup>1</sup>. CISTI= Canada institute for Scientific and Technical information

<sup>2</sup>. LCSH= Library of Congress Subject Headings

<sup>3</sup>. DIALOG

واژه‌هایی که از طریق مصاحبه با مراجعه‌کننده استخراج شده بودند باعث بازیافت مدارک مرتبط شدند. در واقع این واژه‌ها که 10 درصد از کل واژه‌های استخراج‌شده را تشکیل داده بودند گرچه خیلی مفید نبودند، ولی هنگامی که استفاده می‌شدند، در به دست آوردن نتایج بهتر، مؤثر بودند.

«لنکستر» و همکارانش (Lancaster, et al. 1991) نیز از طریق «فهرست‌های همگانی پیوسته» اطلاعات مربوط به 51 موضوع را استخراج کردند. آن‌ها از بین عناوین بازیابی شده، 607 مدرک را که در کتابخانه دانشگاه ایلینویز موجود بود به عنوان «سیاهه واری» در نظر گرفتند؛ سپس از جستجوگران باتجربه درخواست کردند که از طریق «فهرست همگانی پیوسته» و تنها با استفاده از سرعنوان‌های موضوعی کتابخانه کنگره، رکوردهای موجود در رابطه با 51 موضوع یادشده را استخراج کنند. بدین طریق تنها 53 درصد از 607 مدرک موجود در دانشگاه ایلینویز بازیابی شد.

«حمیدی» (Hmeidi, 1997) در پژوهشی درباره‌ی نمایه‌سازی ماشینی مدارک به زبان عربی، نتایج بازیابی از طریق کلیدواژه‌ها و ریشه‌واژه‌ها در زبان عربی را مورد بررسی و ارزیابی قرار داد. وی دریافت که در زبان عربی استفاده از ریشه‌ لغات از نظر میزان بازیافت ودقت، نتایج بهتری را در مقایسه با بازیابی از طریق کلیدواژه‌ها به دست می‌دهد.

«ووریج» (Voorbij, 1998) نیز بر روی تک‌نگاشت‌های موجود در کتابخانه ملی هلند، در حوزه علوم انسانی و اجتماعی که از طریق فهرست همگانی پیوسته قابل بازیابی بود، تحقیقی انجام داد. در این تحقیق مشخص شد 86 درصد منابع با استفاده از توصیفگرها و 48 درصد با استفاده از کلیدواژه‌های عنوان بازیابی شدند. «ووریج» علت بازیافت کم از طریق کلیدواژه در مقایسه با توصیفگرها را محدود بودن دامنه‌ی توصیفگرها، و تنوع عناوین گسترده‌ی موضوعی دانست که مطالب علوم انسانی و اجتماعی در قالب آن عرضه می‌شوند.

«معرف‌زاده» (1366) در پژوهشی با عنوان «بررسی وضعیت ابزار بازیابی اطلاعات پزشکی فارسی و پیشنهادی برای تهیه‌ی نمایه‌ی پزشکی ایران» به بررسی این پرسش پرداخت که کدام شیوه‌ی نمایه‌سازی گردان برای ذخیره و بازیابی اطلاعات پزشکی مفیدتر است. یافته‌ها نشان داد کلیدواژه‌های استخراج‌شده از عنوان مقالات، کتب و پایان‌نامه‌های مورد بررسی، 58 درصد از کل واژه‌های عناوین این مدارک است که این نسبت در پایان‌نامه‌های پزشکی به 69 درصد افزایش یافته است. در این پژوهش روش «نوع افزوده به

عنوان» برای تهیه نمایه پزشکی ایران پیشنهاد گردید.

«بزرگی» (1377) در پژوهشی میزان سازگاری عنوان‌ها و موضوع‌های کتاب‌های فارسی منتشر شده در ایران در فاصله سال‌های 1365-1375 را مورد بررسی قرار داد. یافته‌های پژوهش نشان داد که میان موضوعات و عناوین کتاب‌های فارسی، بطور میانگین 73 درصد سازگاری وجود دارد. بالاترین میزان سازگاری (یعنی 87 درصد) به رده «کلیات» تعلق دارد و پس از آن رده «علوم کاربردی» (فناوری) با برخورداری از 83 درصد سازگاری در مقام دوم قرار دارد. کمترین میزان سازگاری در رده «ادبیات» (41 درصد) مشاهده شد.

«داورپناه» (1375) نیز در پژوهشی میزان همخوانی عناوین مقالات فارسی با محتوای آن‌ها را در گروه‌های مختلف علمی بررسی کرد و به این نتیجه رسید که عناوین مقالات در حوزه علوم انسانی، در مقایسه با عناوین مقالات علوم پایه و مهندسی و مقالات علوم پزشکی، از سازگاری کمتری با محتوا برخوردار هستند.

«انصاری» (1379) در پایان‌نامه خود به «بررسی انطباق میان توصیفگرهای نمایه‌سازی و کلیدواژه‌های عنوان در پایان‌نامه‌های دکترای تخصصی کودکان، زنان، قلب و عروق، و روانپزشکی» پرداخت و نشان داد که در حوزه مورد بررسی به میزان 70/3 درصد میان توصیفگرها و کلیدواژه‌های عنوان انطباق وجود دارد.

### روش پژوهش، جامعه آماری و روش نمونه‌گیری

این تحقیق با روش تحلیل محتوا انجام شده؛ زیرا در آن به گونه‌ای کمی، استفاده از کلیدواژه‌ها در عنوان پایان‌نامه‌ها و میزان همخوانی آن‌ها با توصیفگرها تحلیل و ارزیابی شده است.

جامعه آماری پژوهش، 65536 عنوان پایان‌نامه است که در پایگاه «چکیده پایان‌نامه‌های ایران» تا پایان سال 1381 معرفی شده‌اند. در اولین مرحله از تحقیق، اطلاعات مربوط به عنوان پایان‌نامه‌ها، توصیفگرهای اختصاص داده شده به هر عنوان، تاریخ نگارش، حوزه اصلی، و رشته فرعی پایان‌نامه‌های موجود در «پایگاه چکیده پایان‌نامه‌های ایران» از ابتدا تا پایان سال 1381، از محیط نرم‌افزاری «سی دی اس آیسس»<sup>1</sup> به محیط نرم‌افزاری

---

<sup>1</sup>. CDS-ISIS

«اکسل»<sup>1</sup> انتقال یافت و هریک از فیلدهای ذکر شده، در ستون جداگانه‌ای قرار گرفت. در ادامه به منظور یکدستی جامعه مورد تحقیق، قبل از تفکیک پایان‌نامه‌ها به رشته‌های اصلی، عناوینی که دارای نقایص آشکاری در فیلدهای مورد اشاره بودند از جامعه حذف شدند. پس از طبقه‌بندی پایان‌نامه‌ها به رشته‌های اصلی، عناوین تکراری یا فاقد رشته فرعی نیز حذف شد. بدین ترتیب با توجه به موارد یادشده، 9785 عنوان پایان‌نامه از جامعه اولیه تحقیق به دلیل داشتن مشکلات فاحش حذف شد. در ادامه، جامعه اصلی تحقیق در هر طبقه مشخص شد و با استفاده از فرمول  $n = \frac{NZpq}{(d2(N-1) + Z2pq)}$  (سرای، 1372، 128) حجم نمونه مربوط به هر طبقه محاسبه شد. چون از تحقیقات صورت گرفته قبلی تخمین مناسبی برای  $p$  وجود نداشت، بنابراین مقدار  $p$  مساوی 0/05 در نظر گرفته شد که بدین ترتیب، بیشترین حجم نمونه ممکن برای انجام پژوهش به دست آمده است. همچنین مقدار خطای حاصل از برآورد آماری 0/05،  $Z$  معادل 1/96 و مقدار  $d$  نیز برابر 0/01 در نظر گرفته شد. بر این اساس حجم نمونه در هر طبقه به طور جداگانه محاسبه شده و نتایج زیر به دست آمد:

علوم انسانی 96، علوم پایه 93، علوم پزشکی 95، فنی و مهندسی 100، کشاورزی 94، هنر 88.

به نظر «ببی» (Babbie, 1979, 274) اگر نمونه احتمالی کمتر از 100 مورد باشد، به احتمال زیاد معرف و نماینده واقعی جامعه نخواهد بود. از طرف دیگر به منظور این که در مراحل بعدی تحقیق بتوان یافته‌های مربوط به نمونه‌های هر حوزه را از دیدگاه‌های مختلف با یکدیگر مقایسه نمود، حجم نمونه مربوط به هر طبقه با اندکی تغییر، 100 رکورد در نظر گرفته شد. در ادامه با استفاده از دستور «نمونه‌گیری» در نرم‌افزار «اکسل» از بین جامعه آماری هر حوزه، 100 رکورد به شکل کاملاً تصادفی انتخاب گردید. داده‌های مورد نیاز پژوهش با استفاده از فهرست کنترلی (چک‌لیست) محقق ساخته‌ای گردآوری شد.

## یافته‌ها

با استفاده از داده‌های گردآوری شده به سؤالهای پژوهش پاسخ داده شده که

---

<sup>1</sup>. EXCEL



توضیح آن به شرح زیر است:

1. چه تفاوتی بین تعداد متوسط کلیدواژه‌های عنوان پایان‌نامه در حوزه‌های علوم انسانی، علوم پایه، علوم پزشکی، فنی و مهندسی، کشاورزی و هنر وجود دارد؟  
 برای پاسخگویی به این پرسش، ابتدا تعداد کلیدواژه‌های هر یک از پایان‌نامه‌ها به تفکیک حوزه موضوعی شمارش شده، سپس برابر جدول شماره 1، تعداد متوسط کلیدواژه‌ها در حوزه‌های مختلف شمارش شده است.

جدول 1. تعداد متوسط کلیدواژه‌ها در عناوین پایان‌نامه‌های هر یک از حوزه‌های شش‌گانه

حوزه	علوم انسانی	علوم پایه	علوم پزشکی	فنی و مهندسی	کشاورزی	هنر	جمع
تعداد پایان‌نامه	100	100	100	100	100	100	600
تعداد کلیدواژه‌های عنوان	744	730	515	804	765	534	4082
تعداد متوسط کلیدواژه‌های عنوان	7	7	5	8	8	5	7
نما	6	4	3	6	6	4	6

در این جدول نشان داده شده که به طور کلی (با در نظر گرفتن کلیه حوزه‌ها)، هر عنوان پایان‌نامه دارای 7 کلیدواژه است. این تعداد در عناوین پایان‌نامه‌های حوزه علوم پزشکی و هنر کمتر از حد متوسط و در دیگر حوزه‌ها تقریباً مساوی با تعداد متوسط کلیدواژه‌های عنوان است. شاید یکی از دلایل کوتاه‌بودن عنوان در آثار و تألیفات هنری این باشد که معمولاً هنرمندان عنوان را نیز چون رنگ‌ها، سایه‌روشن‌ها، اشکال و صور مختلف، جزئی از جریان هنری می‌دانند و با نگرشی فلسفی، برای یک عنوان مفاهیم بسیاری را تصور می‌کنند و از مخاطبان نیز می‌خواهند که با کنکاش ذهنی، از یک اثر یا عنوان کوتاه به یافته‌های بدیعی دست یابند. در خصوص عنوان پایان‌نامه‌های حوزه علوم پزشکی نیز مشاهده شد که عمدتاً مطالعه یا بررسی بیماری خاصی مطرح شده و محقق با استفاده از چند کلیدواژه در عنوان پایان‌نامه، مقصود خود را بیان داشته است. در ادامه با

استفاده از آزمون «تحلیل واریانس» تفاوت تعداد متوسط کلیدواژه‌های عنوان در حوزه‌های مختلف بررسی شد که رقم محاسبه شده ( $P=0/000$ ) بیانگر معناداری تفاوت بین تعداد متوسط کلیدواژه‌های عنوان در حوزه‌های مختلف است. به منظور دریافت تصویر روشن‌تری از ساختار واژگانی عناوین پایان‌نامه‌ها در حوزه‌های مختلف تعداد واژه‌های عنوان پایان‌نامه‌ها (اعم از کلمات با بار معنایی<sup>1</sup> و بدون بار معنایی<sup>2</sup>) بررسی شد. در واقع سؤالی که مطرح می‌شود این است که در حوزه‌های مختلف، بین میزان عناصر تشکیل‌دهنده عنوان (اعم از کلیدواژه یا واژگان غیرموضوعی) چه تفاوتی وجود دارد؟ جدول شماره 2 تصویری کلی از تفاوت‌های تعداد کلیدواژه‌ها و واژه‌های غیرموضوعی به دست می‌دهد.

جدول 2. مقایسه تعداد کلیدواژه‌ها و واژه‌های غیرموضوعی در عناوین پایان‌نامه‌های حوزه‌های شش‌گانه

حوزه	علوم انسانی	علوم پایه	علوم پزشکی	فنی و مهندسی	کشاورزی	هنر	جمع
تعداد واژه	1003	1121	835	1041	1001	668	5669
تعداد کلیدواژه	744 (74%)	730 (65%)	515 (62%)	804 (77%)	765 (80%)	534 (80%)	4092 (72%)
تعداد واژه‌های غیرموضوعی	259 (26%)	391 (35%)	320 (38%)	237 (23%)	236 (24%)	134 (20%)	1577 (28%)

بر اساس آزمون مجذور کای، مشخص شد که بین تعداد کلیدواژه‌ها و واژه‌های عنوان در حوزه‌های مختلف تفاوت معناداری وجود دارد.

2. بین تعداد کلیدواژه‌های عنوان پایان‌نامه‌ها با تعداد توصیفگرهای اختصاص یافته چه رابطه‌ای وجود دارد؟

پاسخگویی به این سؤال مستلزم این بود که توزیع فراوانی تعداد کلیدواژه‌ها و میزان تبعیت توصیفگرها از این توزیع مشخص شود.

بدین منظور پایان‌نامه‌ها بر اساس تعداد کلیدواژه‌های عنوان به 6 طبقه دسته‌بندی شدند

<sup>1</sup>. Golist

<sup>2</sup>. Stoplist

و پس از آن، مجموع کلیدواژه‌ها و مجموع توصیفگرهای هر طبقه شمارش شد تا توزیع فراوانی هر یک مشخص شود. نتایج به دست آمده نشان داد که به طور کلی (با احتساب کلیه حوزه‌ها) تعداد توصیفگرهای اختصاص یافته، از تعداد کلیدواژه‌های عنوان پایان‌نامه‌ها تأثیر می‌پذیرد و در موقعی که عنوان پایان‌نامه‌ها دارای تعداد کلیدواژه‌هایی در حد متوسط (9 کلیدواژه در عنوان) است بیشترین میزان همخوانی با تعداد توصیفگرها وجود دارد. علاوه بر بررسی فوق، با استفاده از آزمون ضریب همبستگی پیرسون میزان همبستگی تعداد کلیدواژه‌های عنوان و تعداد توصیفگرها به تفکیک حوزه‌های شش‌گانه محاسبه شد که نتایج به دست آمده در زیر مشاهده می‌شود:

جدول 3. ضریب همبستگی تعداد کلیدواژه‌های عنوان و تعداد توصیفگرها به تفکیک حوزه

رده	علوم انسانی			علوم پایه			علوم پزشکی			فنی و مهندسی			کشاورزی			هنر		
	ضریب همبستگی	سطح معناداری	تعداد	ضریب همبستگی	سطح معناداری	تعداد	ضریب همبستگی	سطح معناداری	تعداد	ضریب همبستگی	سطح معناداری	تعداد	ضریب همبستگی	سطح معناداری	تعداد	ضریب همبستگی	سطح معناداری	تعداد
محاسبه پیرسون																		
تعداد کلیدواژه‌ها و تعداد توصیفگرها	0/934(**)	0/005	6	0/939(**)	0/005	6	0/941(**)	0/005	6	0/982(**)	0/000	6	0/963(**)	0/002	6	0/902(*)	0/014	6

همان‌گونه که داده‌های جدول فوق نیز نشان می‌دهد در کلیه حوزه‌ها همبستگی مثبتی بین تعداد کلیدواژه‌های عنوان و تعداد توصیفگرها وجود دارد. 3. در حوزه‌های علوم انسانی، علوم پایه، علوم پزشکی، فنی و مهندسی، کشاورزی و هنر چه تفاوتی بین میزان همخوانی کلیدواژه‌های عنوان با توصیفگرهای نمایه‌سازان وجود دارد؟ به منظور پاسخگویی به سؤال فوق، وضعیت همخوانی کلیدواژه‌های عنوان با توصیفگرها، در قالب 3 حالت زیر بررسی شد:

- همخوانی کامل، مانند: توصیفگر «جرائم مطبوعاتی» برای پایان‌نامه‌ای با عنوان «تحقیقی پیرامون جرائم مطبوعاتی». در این حالت توصیفگر دقیقاً با یکی از کلیدواژه‌های عنوان همخوان است.

- همخوانی نسبی، مانند: توصیفگر «مطبوعاتی» برای پایان‌نامه‌ای با عنوان «تحقیقی پیرامون جرائم مطبوعاتی». در این حالت توصیفگر با یکی از کلیدواژه‌های عنوان، هم‌ریشه است و تنها حالت وصفی آن تغییر کرده است.

- ناهمخوانی، مانند: توصیفگر «نظام‌های حقوقی» برای پایان‌نامه‌ای با عنوان «تحقیقی پیرامون جرائم مطبوعاتی». در این حالت عبارت «نظام‌های حقوقی» به عنوان توصیفگر به شکلی با عنوان فوق مربوط است، ولی به هیچ شکلی با کلیدواژه‌های این عنوان همخوانی ندارد.

پس از بررسی حالت‌های ذکر شده در نمونه تحقیق، نتایج به دست آمده در جدول شماره 4 گرد آمده است.

جدول 4. میزان همخوانی کلیدواژه‌های عنوان با توصیفگرها، به تفکیک حوزه

حوزه	علوم انسانی	علوم پایه	علوم پزشکی	فنی و مهندسی	کشاورزی	هنر	جمع
وضعیت همخوانی	درصد	درصد	درصد	درصد	درصد	درصد	درصد
همخوانی کامل	49%	41%	50%	35%	50%	52%	45%
همخوانی نسبی	22%	24%	19%	23%	21%	22%	22%
ناهمخوانی	29%	35%	31%	42%	29%	26%	33%
جمع	100%	100%	100%	100%	100%	100%	100%

داده‌های این جدول نشان می‌دهد که به طور کلی، 45 درصد از توصیفگرهای نمایه‌سازان را کلیدواژه‌های عنوان تشکیل می‌دهند. به عبارت دیگر در 45 درصد موارد، کلیدواژه‌های عنوان و توصیفگرهای انتخاب شده توسط نمایه‌سازان همخوانی کامل وجود دارد. بیشترین میزان همخوانی کامل به ترتیب در حوزه‌های هنر (52%)، کشاورزی (50%)، علوم پزشکی (50%) و علوم انسانی (49%)، و کمترین میزان همخوانی کامل واژه‌های عنوان با توصیفگرهای نمایه‌سازان در حوزه علوم پایه (41%) و فنی و مهندسی (35%) است. با توجه به این که توصیفگرهای دارای همخوانی نسبی با کلیدواژه‌های عنوان نیز همانند موارد همخوانی کامل) به شکلی از عنوان پایان‌نامه‌ها استخراج شده‌اند، بنابراین

احتمالاً می‌توان موارد همخوانی (اعم از همخوانی کامل و نسبی) را با موارد ناهمخوانی مقایسه کرد. همان‌گونه که در جدول 4 مشاهده می‌شود 67 درصد توصیفگرهای نمایه‌سازان - به طور کامل یا نسبی - با کلیدواژه‌های عنوان همخوانی دارند. حوزه‌های هنر (74%)، کشاورزی (71%)، علوم انسانی (71%) و علوم پزشکی (69%) دارای همخوانی بیشتری در مقایسه با حوزه‌های علوم پایه (65%) و فنی و مهندسی (58%) هستند. در عین حال برخلاف تصور محققین، میزان همخوانی کامل در حوزه‌های فنی و مهندسی و علوم پایه کمتر از دیگر حوزه‌ها بود. از این رو برای این که بتوان تأثیر خطاهای انسانی را به حداقل کاهش داد و به شکلی صحت یافته‌های تحقیق حاضر را ارزیابی نمود، و نیز زمینه بهره‌گیری از امکانات نرم‌افزاری رایج را برای انجام تحقیقات مشابه فراهم ساخت، سعی شد میزان همخوانی کامل توصیفگرها و کلیدواژه‌های عنوان به شیوه ماشینی محاسبه شود. برای انجام این کار از امکانات برنامه نرم‌افزاری «اکسل» استفاده شد. در جدول شماره 5 نتایج حاصل از محاسبه ماشینی، با محاسبه دستی در نمونه تحقیق مقایسه شده است.

جدول 5. مقایسه بررسی همخوانی کامل کلیدواژه‌های عنوان با توصیفگرها، به شیوه دستی و ماشینی

حوزه	علوم انسانی	علوم پایه	علوم پزشکی	فنی و مهندسی	کشاورزی	هنر	جمع
محاسبه دستی تعداد توصیفگر	427	505	320	565	427	365	2609
محاسبه ماشینی تعداد توصیفگر	419	490	301	557	410	358	2535
محاسبه دستی تعداد همخوانی کامل	(49%) 210	(41%) 209	(50%) 160	(35%) 200	(50%) 215	(52%) 190	(45%) 1184
محاسبه ماشینی تعداد همخوانی کامل	(44%) 183	(39%) 193	(42%) 125	(33%) 183	(49%) 199	(47%) 169	(41%) 1052

جدول شماره 5 به طور کلی 4 درصد اختلاف را در محاسبه ماشینی و محاسبه دستی مربوط به میزان همخوانی کامل توصیفگرها و کلیدواژه‌های عنوان نشان می‌دهد. با استفاده از آزمون مجذور کای، اختلاف بین محاسبه دستی و محاسبه ماشینی تعداد همخوانی کامل ارزیابی شد و نتایج به دست آمده نشان داد که تفاوت معناداری بین نتایج حاصل از شیوه دستی و ماشینی وجود ندارد.

4. میزان همخوانی کلیدواژه‌های عنوان با توصیفگرهای نمایه‌سازان، در سال‌های

مختلف چه تفاوتی دارند؟

به منظور پاسخگویی به این سؤال، ستون مربوط به تاریخ تهیه پایان‌نامه‌ها در محیط نرم‌افزاری «اکسل» از کوچک به بزرگ مرتب شد. با توجه به این که مجموع پایان‌نامه‌های حضور یافته در نمونه تحقیق، مربوط به سال‌های قبل از 1335 کمتر از 50 رکورد بود پایان‌نامه‌های مربوط به این سال‌ها از نمونه تحقیق حذف شدند. در ادامه، وضعیت همخوانی کلیدواژه‌های عنوان با توصیفگرها در سال‌های مختلف بررسی شد که نتایج حاصل از آن در جدول شماره 6 درج شده است.

جدول 6. توزیع فراوانی همخوانی کلیدواژه‌های عنوان با توصیفگرها به تکنیک سال نگارش

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

بر اساس داده‌های جدول فوق میزان همخوانی (کامل، نسبی، ناهمخوانی) کلیدواژه‌های عنوان با توصیفگرها در دوره‌های مختلف با استفاده از آزمون کای مورد مقایسه قرار گرفت. نتایج حاصله نشان داد که میزان همخوانی کامل، همخوانی نسبی و ناهمخوانی در دوره‌های مختلف دارای اختلاف معناداری است.

5. آیا با افزایش تعداد کلیدواژه‌های عنوان، میزان همخوانی آن‌ها با توصیفگرها نیز

افزایش می‌یابد؟

برای پاسخگویی به این سؤال در 6 طبقه توزیع فراوانی، کلیدواژه‌های عنوان و موارد همخوانی (اعم از همخوانی کامل و نسبی) مشخص شد تا رابطه این دو متغیر با

یکدیگر سنجیده شود. در عین حال بر اساس آزمون ضریب همبستگی پیرسون نیز ارتباط تعداد کلیدواژه‌های عنوان و موارد همخوانی کامل مورد ارزیابی قرار گرفت که نتایج آزمون نشان داد در کلیه حوزه‌ها (در سطح معناداری 1%) همبستگی بالایی بین تعداد کلیدواژه‌های عنوان و موارد همخوانی وجود دارد (جدول شماره 7).

جدول 7. ضریب همبستگی تعداد کلیدواژه‌های عنوان و تعداد همخوانی کامل در حوزه‌های شش‌گانه

حوزه	علوم انسانی			علوم پایه			علوم پزشکی			فنی و مهندسی			کشاورزی			هنر		
	ضریب	سطح معناداری	تعداد	ضریب	سطح معناداری	تعداد	ضریب	سطح معناداری	تعداد	ضریب	سطح معناداری	تعداد	ضریب	سطح معناداری	تعداد	ضریب	سطح معناداری	تعداد
محاسبه پیرسون																		
تعداد کلیدواژه و تعداد همخوانی کامل	0/9227(**)	0/0009	6	0/9664(**)	0/0002	6	0/9448(**)	0/0004	6	0/984(**)	0/0000	6	0/969(**)	0/0001	6	0/947(*)	0/0004	6

#### 6. از نظر شیوه نگارش، توصیفگرها چگونه هستند؟

«مورتیمر تاب» در سال 1952 «نمایه‌سازی همارا» را به منظور بررسی گزارش‌های نظامی و مدارک دیگر ابداع کرد. وی در سیستم خود از تک‌واژه‌ها یا واژه‌هایی که بیانگر یک مفهوم بودند استفاده کرد. فرض وی بر این بود که محتوای مدارک را می‌توان در قالب چند مفهوم اصلی ریخت و مفاهیم به دست آمده را به کمک تک‌واژه‌ها بیان کرد. انتخاب ترکیبات نامحدودی از واژه‌ها در هنگام بازیابی، عدم ضرورت رعایت ترتیب خاص واژه‌ها، سهولت استفاده از رایانه در ذخیره و بازیابی، از مزایا، و ریزش کاذب و نامشخص بودن روابط بین واژه‌ها جزو معایب این شیوه نمایه‌سازی به شمار می‌روند. در پایگاه اطلاعاتی توجه به واژه‌های ترکیبی در ذخیره و بازیابی اطلاعات، امری ضروری به شمار می‌رود. در ادامه این تحقیق سعی شد تا تصویری از وضعیت توصیفگرهای تک‌واژه و ترکیب - واژه به دست داده شود. نتایج حاصل از این بررسی در جدول شماره 8 آمده است.

جدول 8. نسبت تعداد توصیفگرهای تک‌واژه به تعدادواژه‌های ترکیبی، به تفکیک حوزه

حوزه	علوم انسانی	علوم پایه	علوم پزشکی	فنی و مهندسی	کشاورزی	هنر	جمع
مجموع توصیفگرها	427	505	320	565	427	365	2609
تعداد توصیفگر تک‌واژه	(%46)	(%42)	(%47)	(%42)	(%59)	(%59)	(%48)
تعداد توصیفگر با واژه‌های ترکیبی	(%54)	(%58)	(%53)	(%58)	(%41)	(%41)	(%52)

در جدول شماره 8 نشان داده شده که به طور کلی نسبت تقریباً یکسانی در اختصاص تک‌واژه‌ها و ترکیب - واژه‌ها وجود دارد. در عین حال حوزه‌های فنی و مهندسی و علوم پایه (هر دو با 58 درصد) دارای تعداد ترکیب - واژه‌های بیشتری در مقایسه با دیگر حوزه‌ها از جمله حوزه کشاورزی و هنر (هر دو با 41 درصد) هستند. یکی از دلایل این اختلاف احتمالاً به ساختمان زبانشناختی موضوعات مربوط می‌شود که حاصل اختلاف ساختمان اطلاعاتی آن‌ها است. مثلاً زبان یک شیمیدان در حوزه شیمی آلی کاملاً با زبان مورد استفاده یک جانورشناس متفاوت است، زیرا اطلاعاتی که از نظر یک متخصص شیمی اهمیت دارد نام ترکیبات آلی، آرایش ساختمانی، و خواص فیزیکی و شیمیایی است، ولی از دیدگاه یک جانورشناس توصیف و تعیین جنس و نوع حیوانات مطرح است (بهاتاچاریا، 1356 ص. 14). از طرف دیگر، با توجه به این که مفاهیم بازنمودهای اشیا، منفرد هستند (هلموت فلبر، 1381 ص. 155) و محققین حوزه‌های فنی و مهندسی و علوم پایه معمولاً با ترکیب اشیا، مواد و موضوعات مختلف سروکار دارند، ناگزیر عبارات نیز به عنوان نمادهای زبانی این اشیا، به شکل ترکیبی ظاهر می‌شوند.

یکی دیگر از مواردی که در خصوص شیوه نگارش توصیفگرها مورد بررسی قرار گرفت شکل «مفرد» یا «جمع» آن‌ها است. معمولاً جمع و مفرد در سرعنوانهای موضوعی همه زبان‌ها، مشکلاتی را ایجاد می‌کند. همیشه این سؤال پیش می‌آید که چه چیزهایی را جمع به کار بریم و چه چیزهایی را مفرد (پوری سلطانی، 1373، ص. 29). در زبان انگلیسی معمولاً استفاده از شکل جمع کلمات (تیلور، 1381 ص. 187؛ Rydings, 1999) و در زبان فارسی استفاده از شکل مفرد در نمایه‌سازی پیشنهاد می‌شود (سمایی، 1379 ص. 31). در تحقیق حاضر کلماتی که با «ان»، «ها»، «ات»، «جات» و کلماتی که طبق قاعده زبان



عربی به شکل مکسر جمع بسته شده بودند بررسی شدند، که نتایج حاصل از آن، برابر داده‌های جدول شماره 9 است.

جدول 9. نسبت توصیفگرهای مفرد به توصیفگرهای جمع، به تفکیک حوزه

حوزه	علوم انسانی	علوم پایه	علوم پزشکی	فنی و مهندسی	کشاورزی	هنر	جمع
کل توصیفگرها	427	505	320	565	427	365	2609
توصیفگر مفرد	(%83)	(%87)	(%83)	(%92)	(%90)	(%85)	(%87)
توصیفگر جمع	(%17)	(%13)	(%18)	(%8)	(%10)	(%15)	(%13)

به طور کلی 87 درصد توصیفگرها به شکل مفرد و 13 درصد به شکل جمع به کار رفته‌اند. 7. در پایگاه «چکیده پایان‌نامه‌های ایران» بین میزان بازیافت با استفاده از کلیدواژه‌های عنوان و با استفاده از توصیفگرهای نمایه‌سازان چه تفاوتی وجود دارد؟ برای دریافت اطلاعات از یک نظام، استفاده کننده باید با مکانیسم‌های مربوط به نظام نمایه ارتباط برقرار کند (ویسمان، 1373 ص. 106). هرچه این ارتباط سهل تر و عاری از پیچیدگی باشد و بین زبان متقاضی اطلاعات و زبان نمایه‌ساز، هماهنگی بیشتری پدید آید، ضریب بازیافت و دقت بالاتر خواهد بود. پیوستگی بین نظام نمایه‌سازی و بازیابی اطلاعات به حدی است که اساساً یکی از معیارهای ارزشیابی فعالیت‌های نمایه‌سازی مدارک را، بازیابی جامع اطلاعات موردنظر در جستجو محسوب نموده‌اند (پائو، 1378 ص. 115). استفاده از کلیدواژه‌ها و توصیفگرها در نمایه‌سازی، هر یک اثر خاصی بر نتایج بازیابی خواهد داشت. در هر پایگاه اطلاعاتی مدارکی وجود دارد که تعدادی از آنها مرتبط با موضوع مورد جستجو و بسیاری نیز نامرتبط هستند. از میان مدارک مرتبط نیز برخی با نیاز جستجوگر مغایرت دارند. توانایی نظام در بازیابی مدارک مرتبط را «جامعیت» و توانایی نظام در کنار گذاشتن مدارک نامرتبط را «مانعیت» گویند. برای این که پایگاه «چکیده پایان‌نامه‌های ایران» از منظر جامعیت ارزیابی شود، طی سه مرحله موارد بازیافت از طریق کلیدواژه‌ها و توصیفگرها، مقایسه شدند. در مرحله نخست جستجو از طریق کلیدواژه‌ها و توصیفگرهایی که دارای همخوانی نسبی بودند صورت گرفت (مواردی

چون کلیدواژه «شیعه» با توصیفگر «تشیع» در یک رکورد، یا کلیدواژه «هیئت علمی» با توصیفگر «هیئت علمی» در یک رکورد). پس از شناسایی کلیدواژه‌ها و توصیفگرهای یادشده در 292 رکورد، جستجو در پایگاه «چکیده پایان‌نامه‌های ایران» صورت گرفت که نتایج حاصل از آن در جدول شماره 10 آمده است.

حوزه	علوم انسانی	علوم پایه	علوم پزشکی	فنی و مهندسی	کشاورزی	هنر	جمع
تعداد بازایی با استفاده از کلیدواژه	(47%)	(39%)	(44%)	(37%)	(38%)	(46%)	(44%)
تعداد بازایی با استفاده از توصیفگر	(53%)	(61%)	(56%)	(63%)	(62%)	(54%)	(56%)

بدین ترتیب در پاسخ به سؤال اخیر باید اذعان داشت که بازایی با استفاده از توصیفگرها (56%) بیش از موارد بازایی شده با استفاده از کلیدواژه‌های عنوان (44%) است. با استفاده از آزمون مجذور کای، تفاوت معناداری باز یافت با استفاده از کلیدواژه‌های عنوان و توصیفگرها محاسبه شد که یافته حاصل از این محاسبه، تفاوت معناداری را در باز یافت از طریق کلیدواژه‌های عنوان و توصیفگرهای نمایه‌سازان نشان داد. دلیل بازایی و جامعیت بیشتر از طریق توصیفگرها، انجام طرح پژوهشی، «سازماندهی توصیفگرهای پایگاه‌های اطلاعاتی مرکز اطلاعات و مدارک علمی ایران (جستجوی هوشمند)»<sup>1</sup> است که توسط واحد خدمات ماشینی مرکز در سال 1380 صورت پذیرفت. قبل از انجام این طرح به دلایلی چون فقدان دستورالعمل مشخص و مدون برای نمایه‌سازی، تغییر کارشناسان و نیز وجود سلیقه‌های مختلف در امر نمایه‌سازی، استفاده از اصطلاحنامه‌ها، واژگان و واژه‌نامه‌های مختلف و ... باعث ناهمگونی و نابسامانی بسیار در توصیفگرها شده بود و پس از هر جستجو، رکوردهای مرتبط بسیاری بازایی نمی‌شد. در مرحله بعد، باز یافت از طریق کلیدواژه و توصیفگرهایی مقایسه شد که دارای همخوانی کامل بودند. تعداد 521 رکورد که حداقل دارای یک کلیدواژه و توصیفگر با همخوانی کامل بودند، شناسایی شدند و جستجو در آنها صورت گرفت. داده‌های به دست آمده نشان داد که میزان باز یافت از طریق کلیدواژه‌ها و توصیفگرهایی که دارای همخوانی کامل هستند تفاوتی ندارد، هر چند انتظار می‌رفت به دلایلی چون فاصله خالی اضافی در بین

۱. پارسی اصفهانی، امیر (۱۳۸۰) سازماندهی توصیفگرهای پایگاه‌های اطلاعاتی مرکز اطلاعات و مدارک علمی ایران (جستجوی هوشمند). طرح پژوهشی. تهران: مرکز اطلاعات و مدارک علمی ایران.

کلیدواژه‌ها یا توصیفگرهای مرکب، میزان بازیافت بین کلیدواژه‌ها و توصیفگرهایی که همخوانی کامل دارند نیز تفاوت داشته باشد؛ مثلاً تعداد بازیافت از طریق کلیدواژه «قارچ‌کشها» با تعداد بازیافت از طریق توصیفگر «قارچ‌کش‌ها» تفاوت کند. ولی در یکی از بخش‌های طرح «ساماندهی توصیفگرهای پایگاه‌های اطلاعاتی مرکز اطلاعات و مدارک علمی ایران (جستجوی هوشمند)» امکاناتی در نظر گرفته شده که فاصله خالی اضافی بین واژه‌های مرکب حذف شود. در مرحله پایانی، سنجش میزان بازیافت از کلیدواژه‌ها و توصیفگرهایی صورت گرفت که در ظاهر هیچ‌گونه همخوانی با یکدیگر نداشتند، ولی مترادف بودند؛ مانند کلیدواژه «سوسید» با توصیفگر «یارانه» یا کلیدواژه «آبستنی» با توصیفگر «حاملگی». در این مرحله 102 رکورد با ویژگی مورد نظر شناسایی و عملیات جستجو بر روی آن‌ها صورت گرفت (جدول شماره 11).

جدول 11. میزان بازیافت از طریق کلیدواژه‌ها و توصیفگرهای مترادف به تکنیک حوزه

حوزه	علوم انسانی	علوم پایه	علوم پزشکی	فنی و مهندسی	کشاورزی	هنر	جمع
تعداد بازیابی با استفاده از کلیدواژه	(%44)	(%41)	(%47)	(%38)	(%42)	(%25)	(%39)
تعداد بازیابی با استفاده از توصیفگر	(%56)	(%59)	(%53)	(%62)	(%58)	(%75)	(%61)

داده‌های این جدول نشان می‌دهد که بازیافت از طریق توصیفگرهای مترادف بیش از موارد بازیافت از طریق کلیدواژه‌ها است. صورت پذیرفتن طرح «ساماندهی توصیفگرهای پایگاه‌های اطلاعاتی مرکز اطلاعات و مدارک علمی ایران (جستجوی هوشمند)»، اجتناب نمایه‌سازان از اختصاص واژه‌های مهجور و نامتداول (مانند «کتابخانه همگانی» به جای «کتابخانه عمومی» یا «بندر پهلوی» به جای «بندر انزلی»)، اجتناب از اختصاص معادل‌های لاتین (مانند «آنالیز» به جای «تجزیه و تحلیل») و مانند آن را می‌توان از جمله دلایل ایجاد یکپارچگی بیشتر بین توصیفگرها و افزایش میزان بازیافت از طریق توصیفگرهای مترادف با کلیدواژه‌ها دانست.

### نتیجه‌گیری

عنوان پایان‌نامه‌ها و کلیدواژه‌های تشکیل‌دهنده آن‌ها از منظر خوانندگان و

پژوهشگران دارای اهمیت است و ایشان را در اولین برخورد، در تشخیص مدرک مربوط یا نامربوط یاری می‌دهد. از طرف دیگر نمایه‌سازان نیز از نقش و اهمیت «عنوان» در ساماندهی اطلاعات غافل نبوده‌اند و از دیرباز تاکنون به عنوان یکی از مهم‌ترین ملاک‌های توصیف‌کننده مدارک، در ایجاد پایگاه‌های اطلاعاتی از آن بهره‌جسته‌اند. با توجه به این نکته در تحقیق حاضر نخست ساختار و اجزای تشکیل‌دهنده عنوان پایان‌نامه‌ها در حوزه‌های مختلف بررسی و این نتیجه حاصل شد که به طور کلی عنوان پایان‌نامه‌ها دارای 7 کلیدواژه است که در ذخیره و بازیابی اطلاعات مربوط به پایان‌نامه‌ها می‌توان از آن بهره‌جست؛ هرچند در این خصوص بین تعداد متوسط کلیدواژه‌های عنوان پایان‌نامه‌ها در حوزه‌های مختلف تفاوت معناداری وجود دارد (جدول شماره 2). همچنین نشان داده شد که در حوزه‌های مختلف، توزیع فراوانی تعداد کلیدواژه‌های عنوان در پایان‌نامه‌ها یکسان نیست. به عنوان مثال در حوزه‌های هنر و علوم پزشکی برخلاف حوزه‌های فنی و مهندسی و علوم پایه بیشتر پایان‌نامه‌ها دارای 1 تا 5 کلیدواژه هستند. از طرف دیگر به طور متوسط در هر عنوان پایان‌نامه 2 تا 3 واژه بدون بار معنایی وجود دارد که به منظور جلوگیری از حجیم شدن پایگاه اطلاعاتی و فراهم آوردن زمینه نمایه‌سازی ماشینی، باید انواع آن‌ها را شناسایی کرد و به عنوان مجموعه‌ای از واژگان بدون بار معنایی، در امر ذخیره و بازیابی اطلاعات مورد توجه قرار داد (جدول شماره 3). 45 درصد همخوانی کامل بین کلیدواژه‌های عنوان و توصیفگرهای نمایه‌سازان یا به عبارت دیگر اتخاذ نیمی از کلیدواژه‌های عنوان به منزله توصیفگر، و به طور کلی وابستگی 67 درصد توصیفگرهای نمایه‌سازان (همخوانی کامل و نسبی) به کلیدواژه‌های عنوان (جدول شماره 6) تأیید دیگری بر نقش و اهمیت عنوان مدارک در ساماندهی اطلاعات مربوط به آن‌ها است. محققان دیگری چون «فراست» (Frost, 1989) و «بزرگی» (1377) نیز هر دو در بررسی خود نشان داده‌اند که 73 درصد واژه‌های عنوان دقیقاً یا حداقل با بخشی از واژه‌های سرعنوان‌های موضوعی همپوشانی دارند. از طرف دیگر نشان داده شد که بین تعداد کلیدواژه‌های عنوان و تعداد توصیفگرها و همچنین میزان همخوانی آن‌ها همبستگی بالایی وجود دارد (جدول شماره 5 و 9). در عین حال مشاهده شد که با گذشت زمان و احتمالاً بنا به دلایلی چون تدوین اصطلاحنامه‌ها و آموزش نمایه‌سازان، میزان وابستگی به کلیدواژه‌های عنوان کاهش یافته و از این مراجع در اختصاص توصیفگرها استفاده شده

است (جدول شماره 8). البته استفاده از این منابع در نمایه‌سازی از نقش و اهمیت عناوین نمی‌کاهد، بلکه استفاده توأمان از کلیدواژه‌های عنوان و اصطلاحنامه‌ها منجر به افزایش جامعیت و مانعیت در ذخیره و بازیابی اطلاعات خواهد شد.

بررسی کلیدواژه‌های عنوان و توصیفگرها از نظر شیوه نگارش نیز یکی دیگر از اجزای این تحقیق بود و مشخص شد که به طور کلی با در نظر گرفتن کلیه حوزه‌ها، نسبت تقریباً یکسانی بین توصیفگرهای تکواژه (48 درصد) و ترکیب-واژه (52 درصد) وجود دارد. اما درصد ترکیب‌واژه‌ها در حوزه‌های فنی و مهندسی و علوم پایه بیش از دیگر حوزه‌ها است. علاوه بر این مشخص شد که توصیفگرها از منظر مفرد (87 درصد) و جمع (13 درصد) بودن تا حد زیادی با مفرد (84 درصد) و جمع (16 درصد) کلیدواژه‌ها همخوان هستند. و شاید به همین دلیل در زبان فارسی برخلاف زبان انگلیسی بیشتر از واژه‌های مفرد در امر نمایه‌سازی استفاده می‌شود (جدول شماره 11). در بخش پایانی این تحقیق میزان بازیافت از طریق کلیدواژه‌های عنوان و توصیفگرهای نمایه‌سازان در سه حالت همخوانی نسبی، همخوانی کامل و ناهمخوانی، به عنوان یکی از ملاک‌های ارزیابی پایگاه‌های اطلاعاتی بررسی شد. در حالت همخوانی نسبی، فزونی بازیافت از طریق توصیفگرها (56 درصد) در مقایسه با کلیدواژه‌ها و نیز بازیافت مساوی در موارد همخوانی کامل، نقش و اهمیت انجام فرایندهایی به منظور ایجاد ارتباط منطقی بین توصیفگرها یا کلیدواژه‌ها را آشکار ساخت. در خصوص کلیدواژه‌ها و توصیفگرهای مترادف نیز به دلیل یکپارچگی بیشتر و برقرار شدن ارتباط منطقی بین توصیفگرهای هم مفهوم (در پی اجرای ساماندهی توصیفگرها)، موارد بازیافت از طریق توصیفگرها بیش از کلیدواژه‌ها بود.

### منابع

- انصاری، مریم (1379). بررسی انطباق میان توصیفگرهای نمایه‌سازی و کلیدواژه‌های عنوان پایان‌نامه‌های دکترای تخصصی کودکان، زنان، قلب و عروق و روانپزشکی. پایان‌نامه کارشناسی ارشد کتابداری و اطلاع‌رسانی پزشکی، دانشکده مدیریت و اطلاع‌رسانی پزشکی، دانشگاه علوم پزشکی و خدمات بهداشتی - درمانی ایران.
- بزرگی، اشرف السادات (1377). بررسی میزان سازگاری عنوان‌ها و موضوع‌های کتاب‌های فارسی منتشر شده در ایران در فاصله سال‌های 1375-1365. پایان‌نامه دکترا، واحد علوم و تحقیقات دانشگاه آزاد اسلامی.

بها تاجاریا (1974). «تأثیر زبان طبیعی در نمایه‌سازی و بازیابی علوم». ترجمه مصطفی سید صادق. 1356. فصلنامه اطلاع‌رسانی. 4 (2و1): 16-20.

پائو، میراندالی (1989). مفاهیم بازیابی اطلاعات. ترجمه اسدالله آزاد و رحمت‌الله فتاحی. 1379. مشهد: دانشگاه فردوسی.

تیلور، آرلین جی (1996). سازماندهی اطلاعات. ترجمه محمد حسین دیانی. 1381. مشهد: کتابخانه رایانه‌ای.

حرّی، عباس (1372). مروری بر اطلاعات و اطلاع‌رسانی. تهران: دبیرخانه هیئت امنای کتابخانه‌های کشور، نشر کتابخانه.

حریری، مهرانگیز، مسعوده توفیق (1362). «نمایه گردان». اطلاع‌رسانی، نشریه فنی مرکز اسناد و مدارک علمی. 8 (2): 4-5.

داورپناه، محمدرضا (1375). «بررسی میزان سازگاری عناوین مقالات فارسی با محتوای آن‌ها». فصلنامه اطلاع‌رسانی، 13 (2): 1-12.

سرای، حسن (1372). مقدمه‌ای بر نمونه‌گیری در تحقیق. تهران: سازمان مطالعه و تدوین کتب علوم انسانی.

سلطانی، پوری (1379). دانشنامه کتابداری و اطلاع‌رسانی. تهران: فرهنگ معاصر.

سمایی، سید مهدی (1379). «مفرد و جمع در نمایه‌سازی». فصلنامه اطلاع‌رسانی. 16 (2و1): 27-31.

فلبر، هلموت (1984). مبانی اصطلاح‌شناسی. ترجمه محسن عزیزی. 1381، تهران: مرکز اطلاعات و مدارک علمی ایران.

معرف‌زاده، عبدالحمید (1366). بررسی وضعیت ابزار بازیابی اطلاعات پزشکی فارسی و پیشنهادی برای تهیه نمایه پزشکی ایران. پایان‌نامه کارشناسی ارشد کتابداری و اطلاع‌رسانی پزشکی، دانشکده مدیریت و اطلاع‌رسانی پزشکی، دانشگاه علوم پزشکی و خدمات بهداشتی درمانی ایران.

ویسمان، هرمان. (بی تا). سیستم‌ها و خدمات اطلاع‌رسانی کامپیوتر. ترجمه جعفر مهاد 1373. شیراز: انتشارات نوید.

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*. 2<sup>nd</sup> 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 Quarterly**, 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 Technical Services**, 35. 377-391.

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

Rydings, Tony. (1999). "Indexing 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]

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 keywords and subject descriptors A Comparison of subject search entries of books in the humanities and Social science". **Journal of Documentation**. 54 (4): 466-476.