

اصطلاحشناسی و کاربرد آن در فعالیتهای اطلاع رسانی

نوشته: فیروزه مولاپرست*

اطلاعات بیش از همه از طریق زبان، بیان و مبادله می گردد و اکثر امکانات بالقوه زبان نیز در جریان انتقال اطلاعات، از طریق تعیین واژه ها و اصطلاحات، بکار گرفته می شود. بنابراین علوم اطلاع رسانی و اصطلاحشناسی^۱ تأثیر فراوان بر یکدیگر دارند. عبارت دیگر، اطلاع رسانی و اصطلاحشناسی شاخه هایی از علوم را تشکیل می دهند که نقاط اشتراک و تجانس بسیاری با یکدیگر دارند و به دلیل داشتن روابط و مناسبات فراوان و برخوردهای مداوم، همواره در هم نفوذ می کنند، از هم می آموزند و پیوندی مستمر بین آنها برقرار است. حداقل می توان گفت که هر دوی آنها در محور اصلی فعالیتهای خود به بررسی متون زبانهای ویژه^۲ می پردازند، و لذا قویاً هدف مشترکی را پی گیری می کنند. هر چند تا قبل از فراگیر شدن استفاده از کامپیوتر ارتباط بین این دو علم، علی رغم وجود حوزه فعالیت مشترک، نسبتاً کم بوده و افراد معدودی در هر دو فعال بوده اند؛ اما ورود به عصر کامپیوتر این ارتباط را بطریقی پویا بارور و متجلی ساخته است بطوری که بازتاب آن را می توان در تعداد افرادی مشاهده نمود که این روزها فعالانه به هر دو رشته اشتغال دارند. البته علی رغم وجود مرزها و موارد اشتراک فراوان در موضوع تحت بررسی و مطالعه این دو علم، هنوز بسیاری از متخصصان اطلاعاتی ای که با نظریه ها، اصول، و راهنماهای اصطلاحشناختی آشنایی لازم را نیافته اند.

ارتباط و پیوستگی بین این دو رشته را می توان بر این اساس تشریح و تبیین نمود که دانش اطلاع رسانی در حوزه فعالیتهای خود عمدتاً با نوعی مدارک ثبت و ضبط شده یا اسناد و مدارکی سر و کار دارد که متون آنها، از نظر پرداختن به موضوعی خاص، اساساً دارای ماهیتی ویژه است. اینها همان متون موضوعی هستند که مشمول بررسی و مطالعه علم اصطلاحشناختی نیز قرار می گیرند و در واقع موضوع پژوهش این رشته نیز می باشند.

در اینجا علم اطلاع رسانی در وسیعترین مفهوم خود مطرح می گردد و دلالت بر کلیه فعالیتهایی دارد که بنحوی در فرایند هر گونه ذخیره، بازیابی، و انتقال اطلاعات انجام می پذیرند. از اینرو این

* عضو هیئت علمی دانشکده کتابداری پزشکی مرکز علوم پزشکی ایران.

^۱ - Terminology.

^۲ - Special Languages.

علم با یکی از فعالیت‌های اساسی بشر سر و کار دارد. به گفته شرا^۱ «... ما در طول زندگی‌مان همواره بدون اینکه خود بدانیم به انتقال اطلاعات پرداختیم»^۱

اما از آنجایی که زبان وسیله اصلی این انتقال است علم اطلاع رسانی می‌تواند بعنوان آنچه عمدتاً در قالب رمز و علائم زبانشناختی با اطلاعات سر و کار دارد مورد توجه و مطالعه قرار گیرد.^۲ در نظام‌های اطلاعاتی، بمنظور برقرار ساختن ارتباطی ثمربخش میان تولیدکنندگان و مصرف‌کنندگان اطلاعات و انتقال دقیق یافته‌ها و مطالب، محتوای مجموعه، می‌بایستی بنحوی تحلیل شود و به زبانی روشن، مشخص و رسا مجهز گردد. عبارت دیگر، به دلیل اهمیت ویژه امر انتقال صحیح و دقیق اطلاعات در این نظامها، لازم است؛ مدارک از لحاظ محتوا بدقت توصیف شده باشند، ... باید نیاز استفاده‌کننده اطلاعات هم به شکلی به نظام عرضه شود که عمل انطباق بین مدارک ذخیره شده و پرسش مطرح شده^۳، و از آنجا امر بازیابی اطلاعات مورد نظر تسهیل و تسریع گردد. این امر، با استفاده از مجموعه‌های اصطلاحات و قوانینی _ به نام زبانهای نمایه _ که بمنظور بیان دقیق موضوع و محتوای مدارک، نحوه بکار بردن اصطلاحات را تعیین می‌کنند انجام می‌پذیرد.

محور اصلی و تمرکز خاص این گزارش بر روی شناخت ویژگیهای مسائل و مشکلاتی است که در زبانهای نمایه و در فرایند توصیف و مشخص ساختن محتوای مدارک از یک طرف، و تشریح نیازهای جستجوکننده و سئوالهای اطلاعاتی از سوی دیگر، و نیز در جریان تطبیق مدارک مجموعه با اطلاع مورد نیاز پیش می‌آید. در این رابطه کوشش می‌شود که طرح مسائل به گونه‌ای باشد که به حل مشکلات مربوط به تحلیل و توصیف اختصاصات محتوایی مدارک، بویژه در نظام‌های اطلاعاتی خودکار، کمک شود. در این زمینه است که دانش اطلاع رسانی به بیشترین رقابت در مقابل علم اصطلاحشناسی بر می‌خیزد؛ زیرا زبان در اطلاع رسانی نیز بعنوان وسیله‌ای عمده در جریان تنظیم و انتقال خودکار اطلاعات و در نتیجه برای طبقه‌بندی، و چکیده‌نویسی ماشینی عمل می‌نماید. علت تأکید بر جریان خودکار این فعالیتها و نظام‌های اطلاعاتی ماشینی آن است که بدون شک ضرورت‌های بسیاری، حرکت جوامع انسانی پیشرفته را در دهه ۱۹۷۰ و ۱۹۸۰ به سوی انفورماتیکی کردن جامعه متحول نموده است. سروان شرایبر با نگاهی به چگونگی رشد تکنولوژی، گذر از جامعه صنعتی به جامعه انفورماتیکی را حتی الوقوع می‌داند^۴ و در این رابطه به شوک انفورماتیکی^۵ اشاره می‌کند و با رشد صنعت کامپیوتر و کوچکتر شدن و ارزانتر

^۱ - J. H. Shera, "The Sociological Relationships of Information Science". *Journal of ASIS*, 22 (1971), p.77.

^۲ - K. Sparck Jones, and M. Kay, *Linguistics and Information Science*, ed. H. Borko (New York: Academic Press, 1973), p.2.

^۳ - مهرانگیز حریری، «اصطلاحنامه در نظام ذخیره و بازیابی اطلاعات»، *اطلاع رسانی*، نشریه فنی مرکز اسناد و مدارک علمی، دوره ششم، شماره ۱ و ۲ (۱۳۶۱)، ص. ۶.

^۴ - ژان ژاک سروان شرایبر، *تکاپوی جهانی*، ترجمه عبدالحسین نیک گهر، چاپ دوم (تهران: نشر نو، ۱۳۶۲)، ص. ۳۵۲.

^۵ - Le choc Informatique; Computer Shock, *idid*.

گشتن آنها و رسیدن به میکروپروسرها،^۱ جهانی شدن جامعه انفورماتیکی را بعنوان یک هدف، مورد پیش بینی قرار می دهد.^۲ بهر حال، در موقعیت کنونی - حتی برای کشورهای در مرحله فعلی رشد اقتصادی، اجتماعی، و صنعتی - به دلیل حجم وسیع و سیر صعودی انتشارات جهان و ناتوانی نظامهای دستی در مهار کردن و در دسترس قرار دادن بهنگام آنها بمنظور کاربرد نتایج تحقیقات و در جریان سازنده و پویای پژوهشهای علمی، حداقل حرکت به سوی خودکار نمودن فعالیتهای اطلاع رسانی امری اجتناب ناپذیر می نماید و لذا آشنایی با برخی مشکلات و مسائل فنی که در هر یک از مراحل و فعالیتهای اطلاع رسانی خودکار پیش می آید و پیش بینی های لازم در انطباق تلاشهای ذهنی و فکری با قابلیتهای ماشینی و فنی بسیار ضروری است.

هر چند وجود اهداف مشترک و علایق یکسان برای توجیه توجهی که اصطلاحشناسان و دکومانتالیستها به حوزه فعالیتهای یکدیگر دارند کافی بنظر می رسد؛ اما در واقع ابداع و ظهور کامپیوتر، پیشرفتها و نوآوریهای سریع، و استفاده وسیع فراگیر از آن بوده است که بهره برداری از این علاقه مشترک و لزوم هماهنگی مطالعات این دو رشته را از اهمیتی خاص برخوردار گردانده است. در هر حال، ذکر این نکته از لحاظ بحث کنونی مهم است که بیشتر آنچه که می بایستی در رابطه با تطبیق دادن دانش اطلاع رسانی با روشهای کامپیوتری و خودکار گفته شود، در محدوده مطالعات، پژوهشها و فعالیتهایی قرار می گیرد که اصطلاحشناسان بمنظور کشف آگاهیهای جدید در زمینه برپایی بانکهای اطلاعاتی اصطلاحشناختی بعهده گرفته اند.^۳

۱- زبانشناسی اطلاعات و نظریه اصطلاحشناسی

متخصصان اطلاعاتی در بادی امر، و تا همین اواخر، امید داشتند که علم زبانشناسی^۴ راه حل مشکلات دکومانتالیسیون را به آنها بنمایاند و به مدد آن، در پیچ و خمها و فراز و نشیبهای تجربه و کار عملی، پرده از روی بسیاری از ابهامات برگرفته شود، گره ها و دشواریهای این دانش جوان و نوحاسته شکافته گردد و این بهره گیری به مرور زمان و در جریان متکامل شدن و وسعت و عمق یافتن علم اطلاع رسانی به وجهی بسیار منظم و سنجیده به تحولاتی زایا در این علم بینجامد. ولی این تمایل، آرزویی بیهوده بود زیرا زبانشناسی در مرحله ای از رشد و تکامل نبود که بتواند پایه ای محکم برای فعالیتهای کاربردی مورد بحث و استفاده متخصصان اطلاعات فراهم آورد. نظریه های زبانشناسی همواره در حال تحول و تغییرات پی در پی و دائم است. این نظریه ها ماهیتاً بیشتر انتزاعی اند و کمتر جنبه عملی و کاربردی دارند. بهر جهت، عاملی که همچون سدی مانع از بکارگیری نظریه های زبانشناختی در صحنه نظری و عملی دانش اطلاع رسانی است ریشه در این

^۱ - Microprocessors.

^۲ - سروان شرایبر، همان منبع، ص ۳۳۴ و ۳۷۵.

^۳ - INFOTERM, "Terminological Data Banks", *INFOTERM International Conference, Ist., Vienna, April 2-3 1979, Proceedings* (Munchen: Verlag Dokumentation, 1980), (INFOTERM Series, no.5).

^۴ - Linguistics.

واقعیت دارد که علم زبان‌شناسی به مطالعه زبان عادی^۱ می‌پردازد. در حالی که در فعالیتهای دانش اطلاع‌رسانی بررسی زبانهای ویژه – که قوانین کاملاً معین و مشخص بر آنها حکمفرماست – هدف و موضوع عمده مطالعه می‌گیرد.

هسته اصلی و مهم هر زبان ویژه، اصطلاحات علمی و فنی یا واژگان آن است. این واژگان معانی یا مفاهیمی را انتقال می‌دهد که نتیجه و اساس آن زبان را تشکیل می‌دهند. اصطلاحات زبانهای ویژه، که در واقع تسهیل‌کننده جریان ارتباطات علمی و انتقال صحیح اطلاعات می‌باشند، پایه‌هایی برای اقدامات و فعالیتهای زیر بشمار می‌روند^۲:

۱- تنظیم مفاهیم علمی از طریق ابداع نظامهای خاص مفهومی؛

۲- انتقال دانش از طریق آموزش و پرورش؛

۳- انتقال دانش از یک زبان به زبانی دیگر با استفاده از واسطه‌های زبانی؛

۴- صورتبندی، فشردگی و چگالش اطلاعات علمی و حرفه‌ای؛

۵- ذخیره و بازیابی اطلاعات (زبانهای نمایه‌سازی و جستجو بنحوی که در گنج‌واژه‌ها یا اصطلاحنامه‌های دکومانتاسیون تنظیم و تدوین شده است).

در زبان عادی معنی خاص هر یک از اجزاء زبان و سازه‌های تکلم (واژه و جمله) همواره وابسته و در ارتباطی تنگاتنگ با زمینه کاربرد آن قرار می‌گیرد.

در زبانهای ویژه، یک اصطلاح به یک معنی خاص، یعنی به مفهومی تعریف شده، اختصاص می‌یابد. در این زبانها، بمنظور باز نمایاندن مفاهیم علم مورد نظر، نخست هر مفهومی دقیقاً تعریف می‌گردد و سپس اصطلاح (واژه یا واژه‌هایی) معینی برای آن مفهوم در نظر گرفته می‌شود و بدان استناد می‌گردد. اصطلاح تعیین شده در سراسر حیطه عمل و در کلیه ابزارهای ارتباطی آن علم در آن معنی خاص بکار برده می‌شود. این مفهوم که مناسبات و حد و مرزهای معینی با دیگر مفاهیم دارد – به دلیل آنکه ایجاد ابهام می‌کند و در غیر معنی تعیین شده خود بکار نمی‌رود – در نظامی از مجموعه مفاهیم آن علم، به صورتی پیوسته و بهم بافته با آنها، جای می‌گیرد. این مفهوم پیوند و وابستگی پویایی با آن نظام مفهومی دارد و چنین ارتباطی بنا‌گزیر و بطور غیر مستقیم در مورد وابستگی نظام مفاهیم به هر یک از این واژه‌ها یا اصطلاحات نیز برقرار است.

در زبانهای نمایه‌سازی، یک واژه اصطلاحنامه‌ای – یعنی واژه‌ای مهار شده و پذیرفته شده در یک نظام اطلاعاتی – به مفهومی اختصاص می‌یابد که توصیف‌کننده و نشان‌دهنده محتوای مدارک ذخیره شده در آن نظام است. معنی و یا عبارتی مفهوم یک واژه پذیرفته شده و موجود در یک اصطلاحنامه بستگی تام به ساختار و محتوای نظام اطلاعاتی مربوط دارد^۳. ورود مفاهیم جدید

^۱ - Common Language.

^۲ - H. Felber. "The General Theory of Terminology-A Theoretical Basis for Information" paper presented at the 41st FID Congress, 1982 (Wien: Infotern, 1982), p.8

^۳ - H. Felber, "Some Basic Issues of Terminology", *The Incorporated Linguist* 21, no. 1 (1982), p. 14.

در نظام "موجب تغییرات لازم در ساخت مفهومی زبان نظام"^۱ می شود و بازتاب تطور زبانی نظام اطلاعاتی، در جریان نمایه سازی بمنظور ذخیره نظام یافته اطلاعات و درجریان جستجو بمنظور بازیابی دقیق و کارا، به تسهیل مطابقت اطلاعات موجود در مجموعه با نیاز اطلاعاتی استفاده کنندگان، و مآلاً به حفظ پویایی نظام می انجامد.

اصطلاحشناسی به عنوان یک علم، معنی و مفاهیم زبانه‌های ویژه را می کاود و به اکتشاف و تبیین زوایا و جلوه های گوناگون آنها می پردازد. برای نیل به این مقاصد، علم اصطلاحشناسی روشها و اصول خاصی را برای خود ترتیب و گسترش داده است. بسیاری از این اصول و شیوه ها از علم منطق و بخصوص از علوم مفهوم شناسی^۲ و معرفت شناسی^۳ اخذ و اقتباس شده اند.

بطور کلی می توان گفت که مشکل اساسی علم اطلاع رسانی آگاهی از چگونگی استخراج معانی (یا مفاهیم) از یک متن یا اختصاص مفاهیم به آن است. برای حل این مسأله، بنظر می رسد که می بایستی شناخت دقیقی از معنی^۴ و صفات و مرزهای آن داشت. بعبارت دیگر، عمیقاً باید دانست که مراد از معنی چیست، متعلقات و اختصاصات آن کدام است، و بالاخره شیوه های نامگذاری خصوصیات و واکنشهای فکری یا وضع الفاظ و انتقال معانی چگونه است.

در زبانه‌های ویژه، معانی بوسیله مفاهیم صورتبندی می گردند و از طریق واژه ها یا اصطلاحاتی خاص به دیگران رسانده می شوند و انتقال می یابند. مفاهیم، خود نماینده اشیا و عناصر جهان بیرونی و یا پدیده های درونی و باطنی هستند. یک پدیده می تواند همانند یک میز، عینی و ملموس و یا همچون دردی که احساس می شود، مجرد و انتزاعی باشد. مفاهیم را می توان نه تنها به اشیا و رویدادها، بلکه به صفات، وابستگیها و کارها نیز منسوب ساخت.

تحریک و انعکاس محیط اطراف در مغز انسان ابتدا به احساس^۵ و سپس به صورتی مشخص به ادراک^۶ منجر می گردد. اثر ادراک، که در نتیجه آن انسان به یک نمود جزئی پی می برد، موجود نگار یا تصویر ذهنی می شود.

در مرحله شناخت منطقی، ادراکها یا نگارهای ذهنی که نماینده صریح نموده‌های جزئی جهان بیرونی هستند، به سبب برخورد با ادراکات یا نگارهای ذهنی پیشین، مقایسه، سنجیده و رده بندی می شوند در نتیجه، ادراک یا نگار جزئی و سطحی که متعلق به یک نمود معین و حاکی از ظواهر آن نمود است، به یاری نگارها با ادراکهای پیشین، تعمیم می پذیرد، تحت نامی عام در می آید و ذات یا ماهیت آن نمود و نظایر آن را نمایش می دهد. ادراک یا نگار ذهنی پس از طی این جریان، مفهوم (Concept) نامیده می شود^۷

^۱ - حریری، همان منبع، ص ۶۶.

^۲ - Conceptology.

^۳ - Epistemology.

^۴ - Meaning.

^۵ - Sensation.

^۶ - Perception.

^۷ - آگ برن و نیم کف، زمینه جامعه شناسی، اقتباس ا. ح. آریان پور (تهران: کتابهای جیبی، ۱۳۵۷)، ص ۲-۳.

بهر حال، مفهوم تنها یک ساخت ذهنی است که از انعکاس واقعیت یا شئی خارجی در ذهن پدید می‌آید. بمنظور تبادل و انتقال این ساخت ذهنی یک نماد یا علامت به آن مفهوم اختصاص می‌یابد که در واقع نمایانگر آن است. این نماد در ارتباطات علمی و فنی، معمولاً یک اصطلاح یا واژه خاص است. برای اینکه انتقال معنی و مفهوم صورت گیرد، علاوه بر وسایل انتقال پیام، باید کلمات یا اصوات که سازندگان نمادها هستند برای گوینده (یا فرستنده پیام) و شنونده (یا گیرنده پیام) روشن و دارای ارزش مشابه باشند "آنچه گوینده اظهار می‌کند تقریباً همان تصور ذهنی را برای شنونده نیز پدید می‌آورد در غیر اینصورت انتقال معنی بوجود نمی‌آید".^۱

عبارتی دیگر، آنچه در مجموعه‌ای از اشیاء مشترکاً یافت می‌شود در ذهن از مرحله تلخیص یا تجرید می‌گذرد، تعمیم می‌پذیرد و از طریق یک مفهوم بیان می‌گردد. به این ترتیب، مفهوم، رکن و عنصر تفکر است و اختصاصاتی را در بر می‌گیرد که در پاره‌ای از اشیاء مشترک است. اجتماع اختصاصاتی که یک مفهوم تبیین می‌کند مفاد یا معنای آن نام دارد. هر مفهوم عضوی از یک مجموعه یا یک طبقه از مفاهیم است که خود نیز می‌تواند تشکیل دهنده طبقه‌ای مشتعل بر مفاهیم دیگر باشد. مجموع کلیه مفاهیم وابسته و فرعی (گونه یا نوع)^۲ در همان سطح تجرید و یا مجموع کلیه اشیاء مجزا و منفردی که در حیطه تعلق به مفهوم مورد نظر قرار دارند، شاهد یا مصدق(های)^۳ آن مفهوم نامیده می‌شود. اجتماع تمام اشیاء و موضوعات منفرد مربوط به یک مفهوم و یا اجتماع عضوهای همانندی که یک مفهوم، اختصاصات آنها را بیان می‌کند طبقه^۴ نیز خوانده می‌شود. "بعضی مفهومیها مانند انسان" ... مصداقهای روشنی دارند. اما بعضی دیگر در واقع بیشتر مفهومیها مانند ... "باهوش" ... مصداقهای روشن و دقیقی ندارند، یعنی دایره مصداقهای آنها مرز قاطعی ندارد. این امر بویژه در نظامهای بازیابی اطلاعات که در آنها مدرکها را با رشته‌ای از مفهومیها می‌شناسند و بازیابی می‌کنند پی آمدهای ناخواسته‌ای دارد.^۵ در رابطه با رفع این گونه دشواریها در دانش اطلاع‌رسانی از نظریه‌های مختلفی از جمله نظریه مجموعه‌ها استفاده می‌شود. شناخت اختصاصات مفاهیم به ما کمک می‌کند که حد و مرزهای مفاهیم و دایره مصداق آنها را حتی المقدور روشن سازیم؛ "آنها را طبقه بندی کنیم و نظامهای مفهومی گوناگون را بنا نهیم"^۶ و از نتایج این فرایند نظری در انتقال دانش و سازماندهی نظامهای انتقال و اشاعه اطلاعات بهره‌گیریم.

^۱ - منصور اختیار، معنی‌شناسی (تهران: دانشگاه تهران، ۱۳۴۸)، ص ۹۲.

^۲ - Intension.

^۳ - Species.

^۴ - Extension.

^۵ - Class.

^۶ - ضیاء موحد، "طرح ابتدایی منطق بول و کاربرد آن در نظامهای بازیابی پس‌همارا"، اطلاع‌رسانی، نشریه فنی مرکز اسناد و مدارک علمی، دوره هفتم، شماره ۱ (۱۳۶۲)، ص ۳۱.

^۷ - Felber, "Some Basic Issues of Terminology", *Op. Cit.*, p. 20.

ذکر این نکته ضروری است که مفاهیم نه تنها با روش استقرائی (یعنی از طریق مشاهده نظمها و قواعد محیط بیرونی، تحریک محیط و ایجاد نگارهای ذهنی از عوامل و نمودهای محیطی، فراهم آمدن ادراک و بدست دادن مفهوم، و بالاخره جمع شدن مفاهیم متعدد، تعمیم آنها و حصول مفهومی وسیعتر یا عبارتی "رسیدن از نمودهای جزئی به مفهوم کلی"^۱) و نه تنها با روش قیاسی (یعنی از طریق بکارگیری و تطبیق ساختهای شناختی^۲ از پیش شکل گرفته با رویدادها و اشیاء اطراف یعنی شمول مفهوم وسیعتر و کلیتر بر مصادیق آن)، بلکه از تعامل و کنش متقابل و دو جانبه رویدادهای محیطی و ساختهای شناختی حاصل می گردند. ما مفاهیم را بوسیله آزمون و تطابق آنها با واقعیات محیط اطراف عمق و وسعت می بخشیم و به کمک ابداع مفاهیم و نظامهای مفهومی و تبیین و تدفیق آنها قادر می شویم به شناخت علمی و تعمیق و گسترش دانش خود در اطراف هستی و جهان بیرونی و درونی نائل آئیم.^۳

اگر متخصصان اطلاعاتی و دست اندرکاران علم اطلاع رسانی از این اصول و مبانی علم مفهوم شناسی بخوبی آگاه باشند، مطمئناً در رابطه با تصمیم گیری صحیح در زمینه سازماندهی نظامهای اطلاعاتی به ابزار مؤثر و مفیدی مجهز خواهند بود. مفهوم شناسی، علمی چند شاخه ای یا چند رشته ای است که کلیه زمینه های موضوعی و فعالیتهای زندگی انسان را که در آنها ارتباط از طریق زبان با هدفهایی خاص و بعنوان یک وسیله انتقال معنی، نقش مهمی را ایفا می کند، در حیطه توجه و مطالعه خود دارد. عبارت دیگر مفهوم شناسی از تلفیق و ترکیب علوم منطق، شناخت شناسی، فرمان شناسی^۴ (سیبرنتیک)، و معنی شناسی^۵ بوجود آمده است و با زبانشناسی، روانشناسی، جامعه شناسی، مردمشناسی و فلسفه ارتباط و پیوستگی دارد و از روشها، نظریه ها و مبانی آنها بهره می گیرد. مطالعات در زمینه مفهوم شناسی نیازمند تلاشی پیگیر در رابطه با انجام تحقیقات میان رشته ای با استفاده از دستاوردهای علوم فوق و بکار بستن مبانی نظری در صحنه تجارب عملی است. اگر متخصصان اطلاعاتی مایل به انجام چنین فعالیتهایی باشند می توانند بینشها و دیدگاههای مهمی در زمینه کار انتقال اطلاعات کسب نمایند و از نتایج آن در فعالیتهای جاری خویش چون طبقه بندی، نمایه سازی و چکیده نویسی سود جویند.

۲- فرایند دکوماناسیون از دید اصطلاحشناختی

دیدگاهی سنتی که در علم اطلاع رسانی نسبتاً پذیرفته شده است، و در آن کما بیش اتفاق رای وجود دارد، فرایند دکوماناسیون را شامل فعالیتهای زیر می داند:

^۱ - آریانپور، همان منبع، ص ۳.

^۲ - Cognitive Structures.

^۳ - N. Bolton, et al. *Concept Formation* (Oxford: Pergamon Press, 1977), p. 10-14.

^۴ - Cybernetics.

^۵ - Semiotics of Semantics.

تحلیل محتوای هر یک از مدارک، صورتبندی این محتوا در قالب مجموعه ای از توصیفگرها، و تنظیم و سازماندهی این توصیفگرها بنحوی که جویندگان اطلاعات بتوانند (بسهولت و بسرعت) سئوالهای مورد جستجوی خود را با مدارک موجود در مجموعه منطبق سازند و پاسخ خویش را بدون از دست دادن هیچیک از مدارک موجود و مرتبط با سئوال مورد نظر، بیابند.^۱

براساس این تعریف، سه گونه فعالیت را می توانیم تشخیص دهیم:

۱- تشریح و توصیف مدرک به صورت آزاد و بدون تبعیت از قانون و قاعده خاص؛

۲- ارائه این توصیف به صورتی مهار شده؛

۳- بکار بردن این ارائه در جریان جستجو.

نکات مورد توجه از نظر زبانشناسی و اصطلاحشناسی در مرحله کار با متن مدرک، در صورتبندی متن توصیف (توصیفگرها)، و در نحوه استفاده از متن توصیف، در جریان جستجو بروز می کند.^۲

از دیدگاه دیگر، فرایند دکومانتاسیون به صورت دو جریان همانند، موازی و همسو است. این دو جریان عبارتند از: یکی فرایند تشریح و باز نمودن محتوای مدرک، و توصیف و قابل بازیابی کردن آن با استفاده از عین متن اصلی مدرک، یا چکیده و یا عنوان متن بمنزله جانشین آن، بدون اینکه هیچگونه آماده سازی و تغییری در آن صورت گیرد. و دیگر، فرایند توصیف مدرک به کمک نوعی آماده سازی، یعنی از طریق استفاده از متن (یا چکیده یا عنوان آن) بمنزله سند اولیه و وضع کردن توصیفی برای آن، یا عبارتی دیگر تحلیل سند اولیه، استخراج کلیدواژه‌ها مبین موضوع(های) مورد پوشش آن؛ بررسی نمایه واژه های تعیین شده و کنترل زبانی آنها یعنی انطباق کلیدواژه ها با یک واژه نامه (اصطلاحنامه تخصصی نظام مربوط) و انتخاب توصیفگرها یا واژه های مجاز نظام که در جریان جستجو نیز بکار برده می شوند. در مورد اول، که با شکلی از متن اصلی مدرک بعنوان توصیف محتوای آن سر و کار داریم، متن و زبان نمایه به صورت یک زبان طبیعی است.^۳ در مورد دوم متن به صورت یک زبان کمابیش مصنوعی و ساختگی^۴ است که به دلیل کنترل زبانی و یکدست شدن کلمات آن، هر چند دقیقاً مبین مفاهیم اصلی است، ممکن است ابداعی بوده و عیناً در متن بکار نرفته باشد. بنابراین ممکن است فعالیت‌های زبانشناختی که در این دو فرایند انجام می پذیرد عیناً مشابه هم نباشد و از این نظر هر یک از آنها موضوع عملیات یا مراحل متفاوتی قرار گیرد.

^۱ - W. J Hutchins, "Automatic Document Selection Without Indexing", *Journal of Documentation*, 23(1967), p. 273-290, quoted in K. Sparck Jones, and M. Kay, *Linguistics and Information Science*, ed. H. Borko, (New York: Academic Press, 1973), p. 45.

^۲ - Ibid., p. 46.

^۳ - Natural Language:

زبانی که بدان تکلم می کنیم و با آن می نویسیم.

^۴ - Artificial Language:

زبانی که برای مقاصد اطلاعاتی، اصول و قواعد آن پیش از کاربرد وضع شده باشد. (آذرنگ، عبدالحسین. فرهنگ اصطلاحات دکومانتاسیون. تهران، مرکز مدارک علمی، ۱۳۵۹.)

صرفنظر از اینکه کدامیک از این دیدگاهها بکار گرفته شوند، وظیفه متخصصین اطلاعات است که تقابل و ارتباط پویا و صحیحی بین واژگان مدرک، واژگان نظام اطلاعاتی، و واژگان جستجو کننده اطلاعات برقرار سازند. اگر این سه زبان همسنگ و سازگار با یکدیگر نباشند انتقال معنی و پیام بدرستی صورت نمی گیرد و کارایی فرآیند اطلاعات و جریان پویای نقل و انتقال آن دچار اشکال می شود.

خصیصه اصلی دیدگان سنتی در زمینه فرایند دکومانتاسیون و بازیابی اطلاعات، تعویض و جایگزینی یک واحد زبانشناسی پیچیده و طولانی _ یعنی تمام متن مدرک _ با یک توصیف فشرده و خلاصه شده از آن _ یعنی یک مفهوم _ است. این مفهوم معمولاً بخشی از یک واژگان مهار شده است و بدین ترتیب، توصیفگر حامل اطلاعاتی می گردد که نوعاً بمراتب از آنچه مستقیماً توسط خود مدرک به تنهایی منتقل می شود یکدستتر و یکنواختتر است. این مطلب در مورد آن دسته از نظامهای بازیابی اطلاعات بازیابی اطلاعات که مبتنی بر بازیابی متن کامل مدارک هستند، و این متون خود بعنوان ابزار توصیف عمل می کنند، صادق نیست و چون ثابت شده است جستجوی متن کامل مدارک فقط در برخی موارد، مانند متون حقوقی کارایی دارد، این نوع نظامها از محدوده پوشش این مطالعه حذف شده اند.

بهر جهت، موارد زیر می بایستی از لحاظ اجزاء اصطلاحشناسی مورد بررسی قرار گیرد:

- ۱- روشهای تشخیص و بازشناسی واحدها و اجزاء مهم و معنی دار متن اصلی مدرک، یا چکیده یا عنوان آن، از نظر بار مفهومی که در رابطه با موضوع کلی تحت بررسی و نظام اطلاعاتی مربوط دارد؛
- ۲- استخراج واژه های توصیف کننده مدرک از این واحدها با توجه به اختصاصات نحوی آنها؛
- ۳- ساختن و بکار بردن طبقه بندیها و سایر اشکال سازماندهی زبانهای نمایه بمنظور استفاده از توصیفها در جریان جستجو.

۲_۱. تحلیل اصطلاحشناختی متون

در بررسی زبانشناختی و اصطلاحشناختی متون مدارک عملاً به این سؤال بر می خوریم که در جریان تحلیل، کدامیک از واحدها و بخشهای متن ورودی به نظام اطلاعاتی می بایستی مورد توجه و تعمق قرار گیرد و کدام یک _ از اختصاصات معنایی و نحوی آنها می باید به حساب آورده شود. یک پاسخ احتمالی آن است که گفته شود قبل از هر چیز لازم است به اصطلاحات موضوعی خاص یا واژگانی که در یک متن به کار رفته است توجه گردد؛ زیرا لغات و اصطلاحات بکار رفته در یک متن، شاخص قوی و معتبر برای موضوعهای مورد بحث در آن است. سپس، به واژگانهای موجود (مانند فرهنگهای نظام یافته و سلسله مرتبه ای) مراجعه شود تا از امکان شمول مفاهیم بکار رفته تحت واژه جنس یا طبقه عامتر که بتواند به عنوان یک توصیفگر مناسب بکار رود، اطلاع حاصل گردد.

این تحلیل، که هدفش تعیین معانی توصیفگرها است، ما را به صحنه رقابتی که روشهای معنی شناسی و مفهوم شناسی با یکدیگر دارند می کشاند. همانگونه که قبلاً ذکر شد، برای مطالعات و بررسیهای زبانهای ویژه، روشهای علم مفهوم شناسی مناسبتر از روشهای دیگر است.

ارتباط بین منبع و واحدها و بخشهای توصیف کننده آن، ارتباطی است که می بایستی قبل از هر چیز مورد توجه قرار گیرد. سپس، ملاحظه ارتباطهای ثانوی که میان توصیفهای یک مدرک با دیگری و بین یک مدرک و یک سؤال برقرار است نیز امری ضروری است. اگر چه این ارتباطها نوعاً و ماهیتاً با یکدیگر متفاوتند؛ اما متأسفانه نمی توان ارتباطهای بین مدارک و توصیفهای آنها، و ارتباط بین یک توصیف و توصیف دیگر را مستقل از هم خواند بویژه آنکه نحوه ارتباط توصیفهای مدارک با یکدیگر و نحوه ارتباط مدارک و سئوالات به گونه ای است که آنها را به صورت دو روی یک سکه، در کلیتی و یگانه و بهم بافته با یکدیگر قرار می دهد. زیرا بازیابی به طور کلی و ضمنی دلالت و اشاره در انتخاب مدارک دارد و از آنجاییکه گزینش یک مدرک و پذیرفتن و رد کردن مدرک دیگر با استفاده از توصیفهای آنها عمدتاً در صورتی به وقوع می پیوندد که توصیفها، اختلافات ماهوی و درونی میان مدارک را به خوبی منعکس سازند، امر بازیابی، به دلیل مصنوعی بودن ساخت زبان نمایه تحت تأثیر تصمیم های هوشیارانه در زمینه نحوه بکار بردن زبان در جریان مشخص ساختن اختصاصات مجموعه نیز قرار می گیرد. این مسأله با یک نگاه به روشی که غالباً طی آن واژگانها گسترش می یابند - نظیر ساخت واژه نامه بر مبنای بررسی نمونه ای از مدارک مجموعه و ایجاد اسکلت اولیه توصیفگرهای نظام - روشنتر می گردد. هر چند محتوای هر یک از مدارک عمدتاً تعیین کننده واژه هایی است که باید بعنوان نمایه واژه انتخاب شوند و بعنوان توصیفگر اختصاص یابند، مع هذا، در این انتخاب ارتباطهای بین مدارک مجموعه نیز باید در نظر گرفته شود^۱.

جریان تعیین و اختصاص توصیفگرها را می توان بعنوان یک فرایند ترجمه از زبان اصلی مدرک به زبان قراردادی اصطلاحنامه نیز تلقی نمود. بنابراین آنچه بولر در مورد ترجمه بیان می کند، در علم اطلاع رسانی نیز صادق است.

... بنابراین در فرایند ترجمه، شایسته است به مرحله ای که شامل درک کردن و فهمیدن متن منبع اصلی است توجه خاص مبذول گردد. زیرا در این مرحله است که احتمالاً مشکلات خاص ترجمه بروز می کند. یک مترجم در نوع خود یک گیرنده است و می توان تصور کرد که کار باز نمودن و کشف معانی محتوای متن یک منبع که بر ذمه مترجمین یا مفسرین است، از برخی جهات با آنچه یک خواننده یا شنونده معمولی انجام می دهد متفاوت است^۲.

^۱ - Sparck Jones and Kay, *Op. Cit.*, P. 53.

^۲ - H. Buhler, "Suprasentential Semantics and Translation," *Meta*, 24, no. 4 (1976), p. 455.

علم اصطلاحشناسی به گسترش ابزارهای بسیاری نائل آمده است که هم به مترجمین و هم به متخصصین اطلاعات در فهمیدن متون منابع کمک می کنند. این ابزارها، که شامل عناصر داده ای مفیدی برای این منظور هستند، اصطلاحاً بعنوان واژگانه‌های توضیحی^۱، شناخته شده اند. در رابطه با این ابزارها قبل از هر چیز می بایستی به ذکر تعاریف لازم پرداخت. این تعاریف، که اکثراً توصیفهای تحت الفظی یک مفهوم با استفاده از مفاهیم شناخته شده هستند، موقع و مکان هر مفهوم معین را در نظام مفاهیمی که به آن تعلق دارد تعیین می نمایند.

۳- طبقه بندی

اصطلاحشناسی و دکومانتاسیون نه تنها اهداف مشابهی را دنبال می کنند؛ بلکه نتیجتاً روشهای همانندی را نیز برای نیل به آن اهداف بکار می گیرند. یکی از این روشها تحت نام طبقه بندی شناخته شده است.

در دکومانتاسیون، گاه لازم می شود مدارکی در زمینه موضوعهای وابسته به موضوع اصلی نیز مورد تحقیق و بررسی قرار گیرند و برای این منظور استفاده از نوعی طبقه بندی، خواه از نوع طبقه بندیهای شمارشی^۲ نسبتاً قدیمی و خواه از نوع رده بندی چهریزه ای^۳ که جدیدتر است ضروری می باشد. طبقه بندی، که تشکیل دهنده نوعی ترتیب و تنظیم موضوعات است، بعنوان ابزاری برای نشان دادن ارتباط بین این موضوعات عمل می نماید. در اینجا باید به این نکته اشاره کرد که طبقه بندی بعنوان یکی از ابزارهای نمایه سازی نیز بکار می رود؛ زیرا در یک طبقه بندی سلسله مرتبه ای^۴ به هنگام حرکت و رسیدن به یک سطح تجرید بالاتر، موضوعهایی که قبلاً از هم جدا بوده اند همانند موضوعهایی یکسان قلمداد می گردند. معذالک، طبقه بندی در اکثر موارد نشانگر ارتباطهای موجود بین موضوعاتی است که بیشتر به صورت مجزا شناخته می شوند.

در اصطلاحشناسی، طبقه بندی برای پی ریزی ساختار واژگان بکار می رود. در اینجا صحبت از طبقه بندی مفاهیم است که بعنوان مثال ارتباطهای بین جنس و گونه هایش (انواعش) یا میان یک کلیت و اعضاء و اجزایش یا میان مفاهیم همبسته از طریق ارتباط معنایی و ذهنی را تعیین می کند. حریری^۵ در مقاله خود به تفاوت‌های قاطع میان طبقه بندی موضوعی (رده بندیهای

^۱ - Defining Vocabularies.

^۲ - Enumerative Classification.

^۳ - Faceted Classification.

^۴ - Hierarchical Classification.

^۵ - مهرانگیز حریری "اصطلاحنامه در نظام ذخیره و بازیابی اطلاعات (۲)", اطلاع رسانی، نشریه فنی مرکز اسناد و مدارک علمی، دوره هفتم، شماره ۱ (۱۳۶۲)، ص ۵۲-۵۳.

برای بررسی ساختار اصطلاحنامه و انواع روابط معادل، مرتبه ای و همبستگی بین اصطلاحات بویژه در زبان فارسی به همین منبع رجوع کنید.

سنتی) و طبقه بندی مفاهیم (اصطلاحنامه) اجمالاً اشاره کرده است.

علی رغم وجود تفاوت‌های بسیار و تعیین کننده میان طبقه بندی مفهومی و طبقه بندی موضوعی، می توان گفت که واژگان نظام یافته و مهار شده، پایه سودمندی برای ایجاد طرح طبقه بندی مناسب برای دکومانتاسیون است. بعضی از ارتباطات و مفاهیم ممکن است مستقیماً بکار گرفته شوند و برخی دیگر باید با توجه به ماهیت مدارک موجود، یا منابعی که احتمال ورود به نظام اطلاعاتی را دارند، جهت استفاده ساده کردند.

۴- نمایه سازی

برای نمایه سازی، و بمنظور نشان دادن موضوعهای یک مدرک در قالبی فشرده و خلاصه شده، مجموعه ای از اصطلاحات - که از متن یک زبان ویژه انتخاب و استخراج گشته اند - بکار برده می شود. در نظامهای اطلاعاتی پس همواره، این اصطلاحات، معمولاً بدون اشاره به ارتباطهای نحوی آنها، بعنوان "کلیدواژه" مورد استفاده قرار می گیرند.^۱

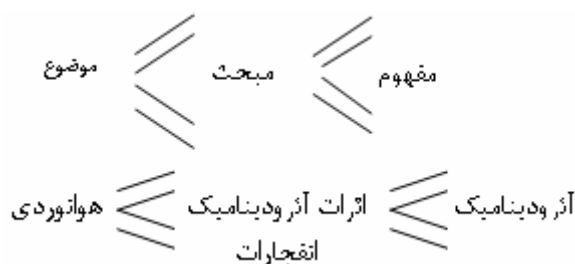
در نظام اطلاعاتی، اصطلاحات استخراج شده را می توان به همان صورت کلیدواژه قبول و حفظ نمود و آنها را عیناً بکار برد؛ ولی نظر به وظیفه ای که نمایه سازی در رابطه با مهار کردن واژه ها دارد، اعمال یکسان سازی و کنترل زبانشناختی در فرایند نمایه سازی - بویژه در زمینه تعیین معادلها و رسم ارجاعات لازم، تدقیق حد و مرز مفهومی اصطلاحات و نوع ارتباطات معنایی و ترکیبی آنها با یکدیگر و پیش بینی روابط مرتبه ای و همبستگی میان آنها، و جلوگیری از ابهام و ناسازگاری زبان ورودی نظام با زبان جستجو - معمولتر است. در این رابطه، رایجترین شیوه ترکیب و تلفیق شکل‌های مفرد و جمع یک اصطلاح و سپس تقلیل شکل‌های متنوع و گوناگون یک واژه به یک شکل پایه واحد از طریق کوتاه سازی^۲ واژه ها و حذف بخشی از حروف از ابتدا یا انتهای آن یا تعویض و جایگزینی وندهاست.

^۱ - D. Austin and J. Waters, *Guidelines for the Establishment and Development of Multilingual Thesauri*. (Paris: Unesco, 1980). P. 4-5.

^۲ - Truncation:

از آنجایی که بسیاری از اصطلاحات ریشه های مشترک دارند، در جریان جستجو در نظامهای خودکار، گاه انطباق فقط ریشه یا بخشی از آن اصطلاحات کافی و سودمند بنظر می رسد. بعنوان مثال چنانچه جستجوگری طالب اطلاعات در زمینه "کتابدوستی"، "کتابداری"، "کتابداران"، و امثال آن است می تواند به جای وارد کردن هر یک از این واژه ها تنها به حروف کتابد در جستجو اکتفا نماید. این کار، تحت عنوان برش یا کوتاه سازی واژه ها، در بسیاری از نظامها با استفاده از یک علامت مثل ستاره، نقطه، ممیز، علامت سؤال یا علایمی دیگر، که با توجه به نوع تفکیک در ابتدا یا انتهای واژه شکسته شده گذاشته می شود، انجام می گیرد و نشان می دهد که در کدام نقطه الگوی انطباق می باید قطع شود. بنابراین جستجوکننده می تواند اصطلاح کتابد* یا کتابد/ یا کتابد. یا کتابد؟ را بمنظور جستجو وارد نظام کند و با هر عنصر داده ای دیگر که مورد نیاز است (چون ایران برای کتابداری در ایران یا کتابداران ایران) آن را ترکیب و هماهنگ نماید. البته با این کار دامنه جستجو وسیعتر می شود و در نتیجه منابع بیشتری با دقت نسبی کمتری بازیابی خواهد شد. باید توجه کرد که در اینجا انتخاب نقطه قطع بسیار مهم است و حروف کتاب بعنوان مرز قطع، انتخاب درستی نیست؛ زیرا در اینصورت هر چند مدارک برخی دیگر از ترکیبات نسبتاً وابسته و مربوط چون کتابخانه و کتابخوان بازیابی می گردد، جستجو به بازیابی مدارکی در زمینه کتابت،

توصیف کامل مدرک می تواند شامل چند بخش یا واحد مستقل یا وابسته به هم باشد که این بخشها یا واحدها بیانگر یک یا چند مفهوم، مبحث^۱، یا موضوع^۲ می باشند. بنظر می رسد تشخیص وجوه تمایز میان این سه واحدی که برای توصیف مدرک بکار می روند مهم و ضروری باشد. بعنوان مثال، در یک مدرک خاص، واژه^۳ "آئرودینامیک"^۳ اصطلاحی است که به یک مفهوم مشخص اختصاص یافته است، "اثرات آئرودینامیک انفجارات"^۳ مفهومی است که نمایانگر یک مبحث است، و هوانوردی مفهومی کلی است که بعنوان موضوع مدرک بکار رفته است. زبانهای نمایه سازی عمدتاً با موضوعات سر و کار دارند که خود به مباحث و در مواردی خاص به تک تک مفاهیم تقسیم و تجزیه می گردند.



از نقطه نظر بازیابی، می توان وجوه تمایزی میان توصیفگرها و غیر توصیفگرها قائل شد. از مجموعه هر دوی آنها با هم واژگان یک زمینه موضوع خاص، شکل می گیرد. رعایت نسبت در بکار بردن این دو دسته اصطلاح از اهمیتی قابل ملاحظه برخوردار است. هر چه تعداد دفعات بکار رفتن یک توصیفگر در نظام اطلاعاتی خاص بیشتر باشد، حالت مشخص بودن آن نسبت به سایر توصیفگرها کمتر می شود و احتمال تطابقهایی که به بازیابی مدارک نامربوط منجر می شوند افزایش می یابد. عین این مشکل در مورد کلیدواژه های استخراج شده از متون مجموعه های نسبتاً همگون پیش می آید. بمنظور حفظ حدود و وجوه تمایز میان توصیفگرها و بار معنایی آنها، در مورد انتخاب و اختصاص توصیفگرهایی که کاملاً بجا و مناسب بنظر می رسند، ولی بیش از حد کلی یا در مجموعه خیلی تکراری هستند، می بایستی دقت فراوان مبذول گردد. اصطلاحات منتخب کاملاً مربوط، ممکن است مبهم باشند و از جانب دیگر، پیش بینی و تدارک لازم برای تفکیک کلیه مفاهیم گوناگون واژه های متشابه یا همانم بسیار مشکل است. از جهت دیگر، در مواردی ممکن است برخی واژه ها نسبت به هم مترادف باشند و ورود این نوع کلمات با هم به

کتابچه، کتاب ساختن، کتابفروشی، کتابه و امثال آن نیز منجر می شود که نسبت به موضوع مورد درخواست بیربط یا بیش از حد وسیع هستند. انتخاب نقطه قطع برای ابتدای اصطلاح (مانند* سیلین برای هر گونه جستجویی در زمینه آنتی بیوتیکهای آمپی سیلین، آموکسی سیلین، کاربنی سیلین، کلوکساسیلین، پنی سیلین و امثال آن) و یا همزمان در هر دو طرف (مانند: **album** برای، **hemialbumin**, **myoalbumin**, **hemialbumose**, **myoalbumose**, **hemialbumosuria** و امثال آن) نیز امکان پذیر است. در حال حاضر ابزارهایی گسترش یافته اند (مانند: Key-Letter-In-Context:KLIC) که به جستجوگر در انتخاب حد اصطلاحات کوتاه شده کمک می نمایند.

^۱ - Topic.

^۲ - Subject.

^۳ - علم حرکت گازها و هوا و حرکت اجسام در آنها.

نظام اطلاعاتی نادرست است. اصطلاحات استخراج شده گاه به اندازه کافی اخص نیستند و گاه بیش از حد اخص می باشند. عبارت دیگر استخراج اصطلاحات نه تنها چندان مشکل نیست؛ بلکه نسبتاً ساده هم هست، ولی در مرحله بعدی که به تشریح و تفسیر آنها می نشینیم مواجه با مشکلاتی می گردیم. در بیشتر مواقع، حل این مشکلات را می توان در وجود فرهنگی از توصیفگرهای مجاز و کنترل و تطابق هر یک از واژه های استخراج شده با این توصیفگرها دانست. در نتیجه، واری واری های ورودی از طریق مراجعه به فرهنگهای توصیفگرها (اصطلاحنامه ها) و ایجاد تطابق و یکدستیهای لازم در واژه ها به این معنی است که اختصاص توصیفگرها کاری است که اساساً ماهیتی واژه شناختی دارد.

صرفنظر از اینکه یک زبان نمایه تا چه درجه ممکن است مهار شده و ساختگی باشد، بهر حال می بایستی کاملاً با اختصاصات زبان ویژه ای که توسط متخصصین رشته مورد نظر بکار می رود، هماهنگ بوده و قادر باشد استفاده جاری از آن زبان را منعکس سازد. عبارت دیگر، کنترل زبانی نظام اطلاعاتی می باید به گونه ای انجام گیرد که به بیگانگی و گسستگی متون ویژه از نظام منتهی نگردد.

اصطلاحی که در یک نظام، بعنوان اصطلاح مرجح و توصیفگر شناخته شده است، در نظامی دیگر ممکن است با توجه به شرایط و نیاز غیر توصیفگر تلقی گردد. این تشخیص یک ساخت و فرایند فکری و ذهنی است که کامپیوتر و ابزارهای خودکار را نمی توان جایگزین آن ساخت. بنابراین در نظامهای خودکاری که کلیدواژه ها را تنها بر مبنای ترتیب و درجه بندی میزان بسامد آنچه در مدرک یا میزان بسامد مدارک حاوی آنها بکار می برند، مشکل در انتخاب نقطه و حدی بوجود می آید که طی آن در فهرست درجه بندی شده حاصله واژه ها با اصطلاحات نمایه سازی یا مرجح و واژه های غیرنمایه یا غیرمرجح تقسیم می گردند. با توجه به این واقعیات، بسیاری از پیش کسوتان توصیه می کنند که بمنظور جلوگیری از اختلال و اشکال، هم برای نمایه سازی خودکار و هم برای چکیده نویسی خودکار، از واژگانه های مهار شده دستی که البته کارهای فکری اولیه آن لزوماً توسط انسان انجام گرفته است نه کارهای محاسبه ای، شکلی و چاپی آن استفاده گردد.

۲_۴. استفاده از کامپیوتر در نمایه سازی

در زمینه مبحث فرایند نمایه سازی، می توان گفت که چون تحلیل مدارک شامل مشخص ساختن صریح محتوا یا پیام یک مدرک است - یعنی به کشف اینکه مدرک واقعاً درباره چیست می پردازد - یعنی به کشف اینکه واقعاً درباره چیست می پردازد - عمل تحلیل چنان مشکل و پیچیده می نماید که فقط انسان قادر است با موفقیتی نسبی به انجام آن نائل آید و ماشین باندازه کافی صلاحیت و توانایی انجام کامل و همه جانبه آن را ندارد. برای تشریح مسأله بطریقی ساده تر، ممکن است بپرسیم که چگونه می توان از کامپیوتر انتظار داشت که با در نظر گرفتن تمام محاسبات ذهنی بشر، با توجه به معنی و مفهوم و بنحوی که موضوعات اصلی مدرک کاملاً رسانده شود، پنج واژه را جایگزین پنج هزار کلمه سازد. در نظامهای خودکار، کلیدواژه ها معمولاً بر مبنای

بسامدشان در مدرک یا مدارک مجموعه انتخاب می شوند، یعنی میزان دفعات بکار رفتن کلیدواژه‌ها محاسبه، و کلیدواژه‌هایی که بیشتر تکرار شده باشند استخراج می گردند. ولی دقت این روش در حد قابل قبولی نیست. به جای آن، معیار گزینش کلیدواژه‌ها می بایستی میزان تأکید آنها در مدرک و اهمیت و ارتباط آنها نسبت به موضوع مورد پوشش مدرک و مجموعه باشد. روش دیگر، که در تهیه نمایه‌های گردان کوئیک نیز بکار برده می شود، کلمات عنوان را نمایندگان و جانشینان قابل قبولی برای متن اصلی مدرک در تعیین و تبیین موضوع مورد پوشش آن می داند و توصیفگرها را از میان آنها انتخاب می نماید. مشکلاتی که در این روش بروز می کند از یک طرف مربوط به نحوه حذف کلمات بی فایده و بدون بار مفهومی لازم از نظر نظام است و از سوی دیگر به چگونگی تنظیم و آمایش کلمات انتخاب شده مربوط می گردد. در این رابطه، استفاده از فهرست کلمات غیرمجاز^۱ و فهرست کلمات نیمه مجاز^۲ بمنظور جلوگیری و حذف عملیات نحوی بسیار معمول است؛ ولی حذف و وجین کردن اصطلاحاتی مانند گسترش و توسعه، تجزیه و تحلیل و امثال آنها که گاه محتوی بار اطلاعاتی و مفهومی ثمربخشی در توصیف موضوع مورد پوشش هستند و گاه از این نظر نقش و ارزش معنایی چندانی ندارند مشکلتر بنظر می رسد. بهر حال، هر گونه تلاش بمنظور کمک اساسی به انسان در جریان تشریح و توصیف صریح محتوای مدارک یا در جهت جایگزین کردن ماشین به جای نیروی انسانی، مسلماً و ضرورتاً مستلزم یک روش تحلیل اصطلاحشناختی نسبتاً پیشرفته است. در این رابطه، برای بانکهای اطلاعاتی اصطلاحشناختی، گسترش روشهای مناسب و شمول عناصر داده ای لازم که فعالیتهای اعمال مربوط به چکیده نویسی و نمایه سازی خودکار را تسهیل نماید وظیفه ای بسیار مهم محسوب می گردد.

در پرتو این تلاشها، تولید اصطلاحات نمایه به صورت کلیدواژه و بر مبنای نظامهای مفهومی که در متن مدرک بکار برده شده است، می بایستی امکان پذیر گردد. بر همین اساس، در چکیده نویسی نیز کامپیوتر می تواند جملات کلیدی و مهم را استخراج نماید.

بهر حال، در جریان هر گونه ارزشیابی روشهای آماری جدید و پیشرفته و استفاده از مدلهای ریاضی بمنظور کاربرد در استخراج و طبقه بندی خودکار اصطلاحات و تعیین و کنترل ارتباطهای معنایی و ساختاری میان آنها به کمک کامپیوتر، باید نیازها و هدفهای اصلی تهیه توصیفهای نمایه در ارائه دقیق محتوای مدرک همواره مد نظر قرار داشته باشد.

^۱ - Stop List:

فهرستی از واژه‌ها یا اصطلاحات یا ریشه کلمات که در رابطه با اهداف بازیابی بی معنی یا بی اهمیت هستند و در نمایه سازی حذف می گردند، مثل حروف اضافه، ربط، تعریف، ضماير و امثال آنها

^۲ - Semi-Stop List:

فهرستی از واژه‌ها که بکار بردنشان بعنوان واژه‌های موضوعی اصلی مجاز نیست؛ ولی استفاده از آنها بعنوان موضوعات فرعی و ترکیبی با موضوعات اصلی بلامانع است.

۵- چکیده نویسی

واژگانها و علم اصطلاحشناسی بویژه در رابطه با چکیده نویسی که بطریق خودکار و با استفاده از کامپیوتر صورت می گیرد اهمیتی خاص می یابد.

ویلیز در تعریف و تشریح آنچه از چکیده نویسی خودکار مراد می گردد می گوید: "هدف نهایی بررسی و تحقیق در زمینه چکیده نویسی خودکار آن است که برنامه کامپیوتر را قادر سازد تا مدرک را بخواند و چکیده آن را در قالب نثر معمول و قراردادی بنویسد. البته نیل به چنین هدفی هنوز مملو از بسیاری موانع لاینحل است."^۱

بعبارت دیگر، نظام چکیده نویسی مبتنی بر کامپیوتر می باید:

۱- مدرک مورد نظر را بخواند؛

۲- مدرک را تحلیل کند؛

۳- برای تهیه چکیده از تعدادی قوانین مربوط به انتخاب واژه ها و جملات و یا تغییر شکل آنها استفاده نماید؛

۴- چکیده تهیه شده را قالب بندی کند؛

۵- چکیده را چاپ کند.^۲

خواندن مدرک اصلی توسط کامپیوتر به اندازه سایر مراحل تهیه چکیده پیچیده است. گزارشهای فنی معمولاً علایم و حروف متفاوت بسیاری را به صورتهای مختلف بکار می برند که گاه ابزارهای ورودی کامپیوتر معمولی قادر به ضبط و ارائه تمام آنها نیستند و ابزارهای نوری نیز همیشه در دسترس نمی باشند. بنابراین یا با ویرایش قبلی متن و حذف حروف و بخشهای غیرقابل انتقال به ماشین، یا با بکار بردن الگویی از علایم و حروفی که کامپیوتر قادر به خواندن مستقیم آنها نیست و یا با اعمال هر دو روش با هم است که این کار صورت می گیرد.

بورکو و برنیر، از چهره های شناخته شده در مطالعات مربوط به نمایه سازی و چکیده نویسی که کتاب جامعی نیز در زمینه مفاهیم و روشهای چکیده نویسی دارند، ضمن بررسی پیشرفتهایی که تاکنون در نیل به هدف فوق صورت گرفته و موانعی که هنوز بر جای مانده است، از چکیده نویسی خودکار اساساً بعنوان یک نوع روش استخراج یاد می کنند. آنها معتقدند که "محصول و بازده آمایش کامپیوتر یک استخراج^۳ است. بعبارت دیگر، در این جریان تعدادی از جملات مهم و نماینده (از نظر بار مفهومی که در رابطه با موضوع کلی دارند) مشخص گشته، از مقاله اصلی

^۱ - R. E. WYLLYS, "Extracting and Abstracting by Computer", in H. Borko, *Automated Language Processing* (New York: Wiley, 1967), p.128

^۲ - C. H. Davis and J. E. Rush, *Guide to Information Science* (Westport, Conn.: Greenwood Press, 1979), p. 41.

^۳ - Extract.

استخراج می گردند و بعنوان چکیده استخراجی از مقاله که توسط کامپیوتر تولید شده است چاپ می شوند.^۱

جریان تهیه چکیده و استخراج بخشهای کلیدی متن اصلی توسط کامپیوتر طی مراحل زیر صورت می گیرد:

۱- تبدیل مقاله مورد چکیده به اطلاعات ماشین خوان، از طریق منگنه کردن برگه های کامپیوتر یا سایر روشهای ضبط بر روی نوار، یا دیسک و یا استوانه مغناطیسی؛

۲- تعیین معیارهای مشخص کننده اهمیت جملات یا کلمات و یا نماینده و معرف بودن هر یک از آنها؛

۳- تحلیل متن وارد شده به حافظه کامپیوتر از طریق شمارش نمره ترتیبی اهمیت و نمایندگی هر یک از کلمات و جملات و گزینش جملاتی که تشکیل دهنده متن چکیده هستند. روشهای تحلیل متفاوتی در نظامهای مختلف بکار می رود. معیارهای گزینش بخشهایی از متن اصلی جهت ورود به چکیده نیز بنا به نظام و نیازهای آن متفاوت است؛

۴- قالببندی و چاپ چکیده استخراجی.

مشکل اساسی در این جریان، یافتن ضوابط و معیارهایی است که بر مبنای آنها کامپیوتر بتواند جملات کلیدی و با ارزشی را که بخوبی رساننده اطلاعات و مفاهیم مهم موجود در مقاله باشند مشخص و انتخاب نماید. تا کنون، بررسیها و مطالعات بسیاری در زمینه چکیده نویسی خودکار و شیوه های گزینش واژه ها و جملات بعمل آمده است که بوروکو و برنیر^۲ و دیویس و راش^۳ به خلاصه ای از آنها اشاره می کنند. اکثر این شیوه ها مبتنی بر کاربرد علم آمار است،^۴ به این ترتیب که طبق آنها جملات مورد بررسی قرار می گیرند و در هر جمله نسبت میان کلمات "معرف" و "نماینده" موضوع و کلمات "غیرنماینده" تعیین می گردد. جملاتی که حاوی بیشترین تعداد کلمات نماینده و دارای بالاترین درصد از این کلمات باشند، بعنوان جملات کلیدی تلقی می شوند و پیکره چکیده را تشکیل می دهند. لورن،^۵ اولین کسی که نشان داد چکیده ها را می شود با استفاده از کامپیوتر تهیه نمود، کلمات نماینده را بعنوان "واژه های محتوایی" می خواند و فرایند تهیه چکیده را چنین تشریح می کند:

بمنظور پردازش با کامپیوتر، متن مقاله ابتدا به شکل ماشین خوان منگنه و ضبط می شود. سپس برتریبی که کامپیوتر هر یک از کلمات متن ورودی را می خواند، آنها را درد و گروه به نام "واژه

^۱ - H. Borko, and C. Bernier. *Abstracting Concepts and Methods* (New York: Academic Press, 1975), p. 164.

^۲ - Ibid. p. 165-179.

^۳ - Davis and Rush, *Op. Cit.*, p. 41-54.

^۴ - Borko and Bernier, *Op. Cit.*, p.172.

^۵ - Luhn.

عادی^۱ یا "واژه محتوایی"^۲ همزمان طبقه بندی و از یکدیگر متمایز می سازد. واژه های عادی که گاه "واژه های وظیفه ای" یا "واژه های غیراساسی" نیز خوانده می شوند شامل حروف ربط، ضمائر، حروف اضافه، حروف تعریف، افعال معین و برخی صفات و قیود می باشند. فهرستی از این کلمات در فرهنگگی از پیش ساخته شده عنوان "جدول مراجعه ای کامپیوتر"^۳ (این فهرست همان فهرست کلمات غیرمجاز است که قبلاً اشاره رفت) تهیه و در حافظه آن، جهت انطباق مرتب واژه های هر متن با آنها و حذف این واژه ها از جریان آمایش و تحلیلها ضبط می گردد.^۴

در اینجا می باید به تعریف کلماتی که اصطلاحاً "واژه های محتوایی" نامیده می شوند و نقش اساسی و مهم در چکیده نویسی دارند اشاره نمود. در یک متن موضوع گرای معین، واژه های محتوایی به آن اصطلاحات فنی اطلاق می گردد که در رابطه با رشته موضوعی خاص آن متن قرار دارند. از این لحاظ است که در رابطه با فرایند چکیده نویسی خودکار، علم اصطلاحشناسی بیشترین اهمیت و ارتباط را پیدا می کند. چون اصطلاحات مربوط به موضوع مورد پوشش مدرک، بعنوان معیار گزینش عمل می نماید، پس در چکیده نویسی خودکار استفاده از واژگانه های موضوعی ماشینی خوان الزامی و برای اعمال کنترل اجتناب ناپذیر است. واژه های محتوایی پس از استخراج به صورت الفبایی ردیف می گردند، کلمات حرف به حرف مقایسه می شوند و بسامد کلماتی که هویت مفهومی مشخص با توجه به واژگان موضوعی و اصطلاحات نماینده نظام دارند از نظر تکرار محاسبه می گردد. سپس جملات نماینده بر حسب نسبت تعداد کلمات و اصطلاحات مهم در هر جمله به کل تعداد کلمات آن تعیین می شوند.^۵ جملاتی که دارای بیشترین تعداد مفاهیم و اصطلاحات مشخصه یک موضوع معین می باشند، یعنی تعداد واژه های نماینده آنها بیشتر است و نمره بالاتری دارند، بعنوان جملاتی تلقی می گردند که بیشترین میزان اطلاعات مورد نظر را در بر دارند و لذا از نظر بار مفهومی با اهمیت هستند و برای شمول در متن چکیده یا خلاصه انتخاب می شوند. چکیده سپس به شکل مجموعه ای از جملات، که به صورت پاراگراف قالببندی شده است، توسط کامپیوتر چاپ می گردد.

رابطه با اجرای مراحل فوق، پیش بینی و تهیه فهرستهایی از اصطلاحات نماینده جهت ضبط قبلی در حافظه کامپیوتر ضروری است. این فهرستها، که صورتهای کوتاه شده اصطلاحات و عناصر و عبارات اصطلاحی را نیز در بر می گیرند، دقیقاً طبقه بندی می شوند بطوری که بتوان انطباق لازم بین آنها و مقالاتی را که می بایستی چکیده گردند بوجود آورد. علاوه بر آن، برنامه هایی نیز لازم است تهیه شوند که سهولت میزان امتیاز و نمره هر جمله را بر مبنای تعداد گروه واژه های نماینده موجود در آن جمله محاسبه نمایند. برای نیل به اهداف چکیده نویسی، می بایستی عناصر

^۱ - Common Words.

^۲ - Content Words.

^۳ - Look up tables.

^۴ - Borko and Bernier, *Op. Cit.*, p. 165.

^۵ - Davis and Rush, *Op, cit.*, p. 43.

داده ای جدید برای بانکهای اطلاعاتی اصطلاحشناختی ایجاد گردد. چکیده نویسی خودکار تنها یک روش است که در آن بمنظور تراکم و اشاعه کم خرجتر و کاراتر اطلاعات متون مقالات به صورت ماشین خوان همراه با واژگانه‌های موضوعی کامپیوتری بکار برده می شود. در خاتمه، می باید اذعان داشت که "مشکلات پژوهشگران دست اندرکاران چکیده نویسی کامپیوتری، و همچنین شاغل در سایر زمینه های آمایش متون، را می توان عمدتاً بعنوان مشکلات زبانی تلقی نمود. بعبارت دیگر، مشکلات اینان یا ماهیت معنایی و تفسیری دارند یا ماهیت نحوی و فیزیکی. برای مقابله کارا و مؤثر با این مشکلات، این گونه محققین می باید خود را به دانش و اطلاعات زبانی عمیقتری مجهز نمایند".^۱ نتایج مطالعات علم اصطلاحشناسی، هر چند که نسبتاً جوان است، مطمئناً در تفهیم بهتر و حل مشکلات زبانی که در جریان دکومانتاسیون بروز می کند بمیزان قابل ملاحظه ای مؤثر خواهد افتاد.

^۱ - Ibid. p. 58.

“منابع”

- آگ برن و نیم کف. زمینه جامعه شناسی. اقتباس ا. ح. آریانپور. تهران: کتابهای جیبی، ۱۳۵۷.
- اختیار، منصور. معنی شناسی. تهران: دانشگاه تهران، ۱۳۴۸.
- حریری، مهرانگیز. “اصطلاحنامه در نظام ذخیره و بازیابی اطلاعات”. اطلاع رسانی، نشریه فنی مرکز اسناد و مدارک علمی. دوره ششم، شماره ۱ و ۲ (۱۳۶۱) ص. ۶۷_۶۰.
- _____ “اصطلاحنامه در نظام ذخیره و بازیابی اطلاعات (۲)”. اطلاع رسانی، نشریه فنی مرکز اسناد و مدارک علمی، دوره هفتم، شماره ۱ (۱۳۶۲)، ص. ۷۶_۴۸.
- دستور ساختن واژه نامه، استانداردهای بین المللی. ترجمه پرویز مهاجر. تهران: مؤسسه تحقیقات و برنامه ریزی علمی و آموزشی، مرکز مدارک علمی، ۱۳۵۴.
- زبان، تفکر و شناخت در روند تکامل اجتماعی. ترجمه فیروز شیروانلو. تهران: توس، ۱۳۵۰.
- سروان شرایبر، ژان ژاک. تکاپوی جهانی. ترجمه عبدالحسین نیک گهر. چاپ دوم. تهران: نشر نو ۱۳۶۲.
- قواعد و قراردادهای ساختن اصطلاحنامه. ترجمه گیتی هندسی افشار. تهران: مؤسسه تحقیقات و برنامه ریزی علمی و آموزشی، مرکز اسناد و مدارک علمی ایران، ۱۳۵۰.
- موحد، ضیاء. “طرح ابتدایی منطق بول و کاربرد آن در نظامهای بازیابی همارا”. اطلاع رسانی، نشریه فنی مرکز اسناد و مدارک علمی. دوره هفتم، شماره ۱ (۱۳۶۲)، ص ۴۲_۲۸.
- واترمن، جان تی. سیری در زبانشناسی. ترجمه فریدون بدره ای. تهران: کتابهای جیبی، فرانکلین، ۱۳۴۷.

Aitchison, J. and Gilchrist, A. *Thesaurus Construction: a Practical Manual*. London: Aslib, 1972.

Austin, D. and Waters, J. *Guidelines for the Establishment and Development of Multilingual Thesauri*. Paris: Unesco, 1980.

Bolton, N. et al. *Concept Formation*. Oxford: Pergamon Press, 1977.

Borko, H. and Bernier, C. *Abstracting Concepts and Methods*. New York: Academic Press, 1975.

Buhler, H. “Suprasentential Semantics and Translation “. *Meta*, 24, no. (1979): 451_8.

Davis, C. H. and Rush, J. E. *Guide to Information Science*. Westport, Conn.: Greenwood Press, 1979.

Felber, H. “The General Theory of Terminology_ A Theoretical Basis for Information”. Paper Presented at the 41 st FID Congress 1982. Wien: INFOTERM, 1982.

_____. “Some Basic Issues of Terminology”. *The Incorporated Linguist*, 21, no. 1 (1982).

Fosket, A. C. *The subject Approach to Information*. 4th ed. Hamden, Conn.: Linnet Books, 1982.

Gilchrist, A. *The Thesaurus in Retrieval*. London: Aslib, 1971.

Hutchins, w. J. "Automatic Document Selection Without Indexing." *Journal of Documentation*, 23 (1967): 273_290.

INFOTERM. "Terminological Data Banks". INFOTERM International Conference, 1st, Vienna "April 2_3, 1979. *Proceedings*. Munchen: Verlag Dokumentation, 1980. INFOTERM Series, no. 5.

Lancaster, F. W. "Vocabulary Control in Information Retrieval System", *Advances in Librarinship*, vol. 7 (1977): 2_40.

Shera, J. H. "The Sociological Relationship of Information Science", *Journal of ASIS*, 22 (1971): 76_80.

Soergel, D. *Indexing Languages and Thesauri: Construction and Maintenance*. Los Angeles, California: Melville, 1974.

Sparck Jones, K. and Kay, M. *Linguistics and Information Science*. ed. H. Broko. New York: Academic Press, 1973.