

بسمه تعالی

جزوه ی درس:

نمایه سازی

رشته:

فناوری اطلاعات

مقدمه:

نمایه از مصدر نمودن به معنای نشان دادن و راهنمایی کردن مشتق شده است. معادل آن در زبان انگلیسی واژه Index است که خود از واژه لاتین dicare به معنای نمایاندن و دلالت کردن گرفته شده که با پیشوند In به معنای "اشاره‌گر از بیرون" به کار می‌رود.

در زبان فارسی و در غالب رشته‌های تخصصی، واژه "شاخص" معادل Index به کار می‌رود، ولی در رشته کتابداری و اطلاع‌رسانی واژه "نمایه" برای آن انتخاب شده است. نمایه فهرستی از نام‌ها، موضوع‌ها، و دیگر مطالب موجود در یک سند یا مجموعه اسناد و مدارک است که به شیوه‌ای (معمولاً الفبایی) مرتب شده و از طریق جای‌نماها کاربر را به محلی که این نام یا موضوع در آنجا آمده راهنمایی می‌کند. به فرایند "تحلیل محتوای یک سند و یا مجموعه اسناد و برگردان نتیجه این تحلیل به‌واژه‌هایی که در نمایه می‌آید"، نمایه‌سازی می‌گویند.

از آنجا که نمایه‌سازی شیوه‌های دستیابی به اطلاعات مجموعه‌ای از مدارک یا یک مدرک و یا بخش‌هایی از یک مدرک را، صرف‌نظر از محل نگهداری آن، فراهم می‌کند، می‌توان گفت نمایه‌سازی روش بنیادین سازمان‌دهی دانش است و فهرستنویسی و رده‌بندی مصادیق آن هستند.

تعریف:

طبق استاندارد نمایه‌سازی بریتانیا، نمایه سیاهه نظام یافته مدخل‌هایی است که به منظور کمک به استفاده کنندگان در جایابی اطلاعات یک مدرک ساخته می‌شود.

طبق استاندارد ایزو ۵۹۶۳، عمل توصیف یا شناسایی محتوای موضوعی یک مدرک را نمایه‌سازی گویند.

ثبت و ضبط محتوای اطلاعاتی مدارک با استفاده از روشهای گوناگون به منظور سازماندهی اطلاعات به قصد سهولت بازیابی را نمایه‌سازی گویند (نوروزی، ۱۳۸۰).

به عبارت دیگر، نمایه‌سازی یعنی **تخصیص واژه‌ها یا اصطلاحات به مدارک به منظور توصیف محتوای موضوعی آنها برای بازیابی در مراحل بعد**. هدف نمایه‌سازی اصولاً آماده کردن مدرک برای بازیابی است.

هدف، کارکرد، و ضرورت:

نمایه در زندگی روزمره کاربردهای بسیاری دارد. بدون وجود نمایه‌ها، تلفن کردن، زمان‌بندی قرارهای ملاقات، جایابی مکان‌ها، یافتن خیابان‌ها، یافتن مواد کتابخانه‌ای، و انجام بسیاری از امور دیگر که ما اکنون انجام می‌دهیم به سادگی میسر نیست. نمایه‌ها راهنماهای مفیدی هستند که نه تنها ما را به موضوعات مورد علاقه هدایت می‌کنند، بلکه تصویری یکپارچه و دیدگاهی جامع از یک حوزه موضوعی می‌سازند، کاری که سایر ابزارهای کتابشناختی قادر به تدارک آن نیستند.

هدف نمایه‌ها تسهیل بازیابی اطلاعات است، بنابراین، ارزش هر نمایه به میزان توفیق آن در کمک به جست‌وجوگر برای بازیابی اطلاعات موردنیاز بستگی دارد. ضرورت تهیه نمایه با اندازه، ناهمگونی، و پیچیدگی مدارک موضوع نمایه‌سازی ارتباط مستقیم دارد. یک کتاب ساده و کم حجم، یک شماره از یک نشریه با حجم متوسط، و یا جزوه‌ای شامل ده یا بیست قطعه شعر می‌تواند بدون داشتن نمایه به سادگی جست‌وجو شود؛ اما یک متن درسی و سایر منابع اطلاعاتی غیرداستانی، شماره‌های متعدد یک یا چند نشریه، و گزیده‌ای مفصل از اشعار یا مقالات و مجموعه‌های بزرگ‌تر همه نیازمند نمایه‌اند.

امروزه با رشد روزافزون انتشارات در حوزه‌های مختلف علمی، وجود نمایه‌ها ضروری‌تر شده و پژوهش بدون آنها اساساً دشوار و گاه غیرممکن است. نمایه‌ها با سازماندهی اطلاعات جست‌وجو را تسهیل می‌کنند. اما کار نمایه، بیش از کم کردن تلاش جست‌وجوگر است: نمایه‌ها سبب اطمینان جوینده از جامعیت و مانعیت جست‌وجو می‌شوند؛ بدین معنا که اگر کار نمایه‌سازی درست انجام شده باشد و توصیف‌گرها مناسب انتخاب شده و جای‌نماها دقیق باشند، و ابزارهای کنترل مانند یادداشته‌ها و ارجاعات روشن و کامل باشند، آنگاه پژوهشگر می‌تواند مطمئن باشد که حداکثر بازیابی را با حداکثر دقت، در حداقل زمان انجام داده است. از این‌رو، می‌توان گفت نمایه‌ها نقش مهمی در افزایش بهره‌وری علمی دارند.

ساختار نمایه

نمایه از تعدادی مدخل تشکیل می‌شود، هر مدخل در فایل رکوردهای اطلاعات کتابشناختی یا سایر اطلاعات نظیر نام‌ها، عناوین، موضوعات، و مانند آن یک نقطه بازیابی محسوب می‌شود. اجزای مدخل عبارتند از:

۱. **شناسه** یا توصیف‌گر: کلید واژه مرجحی است که نمایه‌ساز آن را به‌عنوان موضوع یا یکی از موضوعات مدرک انتخاب می‌کند و جزء اصلی مدخل را تشکیل می‌دهد. مدخل‌ها بر اساس شناسه‌ها مرتب می‌شوند. شناسه‌ها انواع متعددی دارند؛ شناسه نام، شناسه موضوع، و شناسه عنوان.

۲. **بیانگر**: کلمه یا عبارتی است که جنبه خاصی از شناسه را نشان می‌دهد. بیانگرها می‌توانند انواعی چون بیانگرهای زمانی، مکانی، موضوعی، و جز آن داشته باشند. بعضی از نمایه‌ها فاقد بیانگرند. نحوه آمدن بیانگر ذیل شناسه متفاوت است، گاهی بیانگرها به‌صورت الفبایی ذیل شناسه می‌آیند (برون‌بافتی)، و گاه نیز بر اساس توالی حضور در متن به‌دنبال شناسه اصلی می‌آیند (درون‌بافتی).

۳. **جای‌نما Locator**. عدد یا هر علامت دیگری است که در نمایه در مقابل شناسه‌ها یا بیانگرها می‌آید و جست‌وجوگر را به محل آمدن اطلاعات هدایت می‌کند. البته گاهی در بعضی نمایه‌های الکترونیکی، نمایه فاقد جای‌نمای قابل رؤیت است، اما در فایل اصلی این جای‌نماها وجود دارند. در این‌گونه موارد کاربر با انتخاب شناسه و یا بیانگر مستقیماً به اطلاعات کتابشناختی، چکیده، و جز آن هدایت می‌شود. اینکه جای نما جست‌وجوگر را به کدام واحد اطلاعاتی هدایت کند، به سیاست‌گذاری اولیه در زمینه نوع واحد سندی بستگی دارد.

"واحد سندی" می‌تواند یک یا چند صفحه از مدرک؛ یک یا چند پاراگراف از یک متن؛ اطلاعات کتابشناختی یا چکیده؛ و یا متن کامل باشد.

۴. **ارجاعات:** ارجاع بخشی از مدخل است که استفاده‌کننده را از یک شناسه یا بیانگر به شناسه دیگر راهنمایی می‌کند. ارجاعات در نمایه‌سازی نیز همانند فهرستنویسی انواع مختلفی دارد، از جمله ارجاع "نگاه کنید"، و "نیز نگاه کنید". در نمایه‌سازی از ارجاع "نگاه کنید ذیل" نیز برای هدایت کاربر از یک شناسه به اطلاعات ذیل شناسه دیگر استفاده می‌شود. البته باید از ارجاع کاربر به‌جایی که در آن اطلاعاتی وجود ندارد خودداری شود، این نوع ارجاعات، ارجاع کور نامیده می‌شود. بعضی نمایه‌ها فقط از شناسه و جای‌نما تشکیل شده‌اند و سایر اجزا را ندارند.

انواع نمایه از نظر نوع اطلاعات (محتوا و پوشش)

نمایه از نظر محتوا و نوع اطلاعات به دو گروه عمده تقسیم می‌شوند:

الف. نمایه نامها (names Index): به این نوع نمایه "نمایه اعلام" یا "نامنامه" نیز گفته می‌شود، سیاهه الفبایی نامهای خاص است که در متن منبع از آنها نام برده شده است. نامهای خاص شامل نامهای اشخاص، اماکن، پدیده‌های جغرافیایی، اشیاء، سازمانها و نظایر آن است که هر منبع به تناسب موضوع مورد بحث می‌تواند دربردارنده‌ی تمام یا برخی از آنها باشد. نمایه نامها دارای انواع مختلف است:

۱- **نمایه پدید آور:** (Authors Index): سیاهه‌ای از نام پدید آورندگان آثاری که در متن به آنها استناد شده است.

۲- **نمایه عناوین:** (Title Index): سیاهه‌ای از عنوان آثاری که در متن به آنها استناد شده است.

۳- **نمایه نام های جغرافیایی:** (Geographical Index) سیاهه الفبایی از نام اماکن، مناطق، کشورها و پدیده‌های جغرافیایی

۴- **نمایه نام سازمان ها و مؤسسات:** نام سازمان ها و موسسه‌ها در آثار گوناگون و یا در ادوار مختلف یا حتی در یک منبع واحد ممکن است به شکل‌های مختلفی آورده شود. در این مورد لازم است اسامی سازمانها و مؤسسات با استفاده از منبع معیار به شکلی هماهنگ و یکنواخت در نمایه‌سازی استفاده شود. چون برخی سازمان ها در طول زمان دچار تغییرات شده اند. (مانند سازمان برنامه سپس سازمان برنامه و بودجه سپس سازمان مدیریت و برنامه ریزی).

۵- **نمایه اشخاص:** عبارتند از سیاهه الفبایی از نام های اشخاصی است که در منبع ذکر شده است. یکی از اهداف نمایه سازی نام ها آن است که کلیه اطلاعات راجع به یک نام و یا محل و یا حضور آن نام در متن با یک شناسه واحد نشان داده شود تا بازیابی آن سریع و آسان شود.

ب. نمایه موضوعی (Subject index): سیاهه‌ای الفبایی از مفاهیم و موضوع های مهمی است که در منبع درباره آن بحث شده است. در این نوع نمایه سازی مفاهیم مطرح شده در متن به وسیله ی یک اصطلاح بیان می‌شوند. موضوع بحث نمایه‌سازی موضوعی پرداختن به چگونگی توصیف محتوای مدرک است.

انواع نمایه از نظر بیان محتوای اطلاعاتی:

برای توصیف محتوای مدرک دو رهیافت کلی وجود دارد: نمایه‌سازی واژه‌ای؛ و نمایه‌سازی مفهومی

۱- **نمایه‌سازی واژه‌ای.** در این نوع نمایه بر اساس سیاست‌گذاری‌های اولیه واژه‌هایی از تمام یا بخشی از متن انتخاب و موضوع مدرک تلقی می‌شود. واژه‌ها می‌توانند از جای‌جای متن (عنوان یا چکیده و یا سراسر متن) استخراج شوند. این نوع نمایه‌سازی را که در آن واژگان نمایه از متن مدارک استخراج می‌شود، "نمایه‌سازی استخراجی (اشتقاقی)" می‌گویند. در مقابل، شیوه‌ای را که در آن نمایه‌ساز پس از مطالعه مدرک و درک مفاهیم آن، خود واژه‌هایی را به‌عنوان موضوع به مدرک منتسب می‌کند، "نمایه‌سازی انتسابی" گویند.

روال اصلی کار در تهیه نمایه‌های واژه‌ای عنوان، گردش دادن واژه‌های عنوان به گونه‌ای است که هر بار یکی از واژه‌ها در محل الفبایی مناسب، مدخل قرار می‌گیرد. از این نظر به این نمایه‌ها، "نمایه‌های گردان" هم گفته می‌شود. پیش‌فرض تهیه این نوع نمایه‌ها این است که عنوان بیانگر موضوع متن است. برای تهیه رایانه‌ای نمایه‌های واژه‌ای عنوان، ابتدا فهرستی از کلمات غیرمجاز (سیاهه بازدارنده) شامل کلمات فاقد بار اطلاعاتی، حروف اضافه، کلمات ربط، افعال، صفات، و مانند آن، تهیه و به برنامه رایانه‌ای معرفی می‌شود. رایانه این کلمات را از مجموعه متون (مثلا در این مورد عنوان) حذف و بقیه واژه‌ها را موضوع تلقی می‌کند و در قالب‌های تعریف شده، مانند نمایه‌های کوئیک، کوواک، و نمایه جایگشتی ارائه می‌کند.

نمایه کوئیک Keyword in Context

از قدیمی‌ترین و موفق‌ترین نمایه‌های واژه‌ای عنوان نمایه کوئیک است. در این نمایه عناوین گزارش‌ها، مقالات، و مانند آن به‌صورت ماشین‌خوان درمی‌آید و به‌وسیله رایانه تقطیع می‌شود، برنامه رایانه‌ای واژه‌های معنادار را از میان واژه‌ها انتخاب می‌کند و سپس هر عنوان را به گونه‌ای چاپ می‌کند که کلیدواژه در موقعیتی برجسته (معمولا در ستونی در وسط عناوین) قرار گیرد و تمام عناوین بر اساس نظم الفبایی کلیدواژه‌ها و حروف بعد از آن مرتب شوند. این شناسه‌ها با روش‌هایی مانند نوشته شدن به‌صورت ایرنیک (ایتالیک) یا سیاه، و یا فاصله‌گذاری از کلمات قبل و بعد در متن عنوان برجسته می‌شوند تا چشم هنگام بازبایی خسته نشود. از نشانه‌گذاری هم برای مشخص کردن ابتدا و انتهای عنوان استفاده می‌شود. برای مثال مدخل‌های نمایه کوئیک برای یک عنوان مثل "واژه‌نامه‌ها و دایره‌المعارف‌های کتابداری و اطلاع‌رسانی" به‌شکل ذیل در می‌آید.

های علوم کتابداری و	اطلاع‌رسانی#	واژه‌نامه‌ها و دایره‌المعارف‌ها
واژه‌نامه‌ها و	دایره‌المعارف‌های	علوم کتابداری و اطلاع‌رسانی#
واژه‌نامه‌ها و دایره‌المعارف‌های	علوم	کتابداری و اطلاع‌رسانی
دایره‌المعارف‌های علوم	کتابداری	و اطلاع‌رسانی# واژه‌نامه‌ها
کتابداری و اطلاع‌رسانی#	واژه‌نامه‌ها	ودایره‌المعارف‌های علوم

شکل ۱- نمونه‌ای از مدخل‌های یک نمایه کوئیک

نمایه کواک Keyword Out Of Context

نمایه کواک همان نمایه برون بافتی است. این نمایه برای رهایی از مسائل و مشکلات ناشی از ضرورت کوتاه کردن عنوان و نیز به منظور ساده ساختن خواندن مدخلها طراحی شده است. در این نمایه هر کلیدواژه به ترتیب از عنوان خود خارج شده و مقدم بر سایر اجزای عنوان قرار می‌گیرد. سپس عنوان مدرک به ترتیب طبیعی خود و به طور کامل در زیر این واژه یا به دنبال آن می‌آید. به این ترتیب برای هر واژه مهم یک مدخل ساخته می‌شود. در این نمایه نیز کلمات مجاز در متن عنوان وجود دارد و کلید واژه قرار نمی‌گیرد.

اطلاع‌رسانی

واژه‌نامه‌ها و دایره‌المعارف‌های علوم کتابداری و اطلاع‌رسانی

دایره‌المعارف‌ها

واژه‌نامه‌ها و دایره‌المعارف‌های علوم کتابداری و اطلاع‌رسانی

علوم

واژه‌نامه‌ها و دایره‌المعارف‌های علوم کتابداری و اطلاع‌رسانی

کتابداری

واژه‌نامه‌ها و دایره‌المعارف‌های علوم کتابداری و اطلاع‌رسانی

واژه‌نامه‌ها

واژه‌نامه‌ها و دایره‌المعارف‌های علوم کتابداری و اطلاع‌رسانی

شکل ۲ - نمونه‌ای از مدخل‌های یک نمایه کواک

از مزایای نمایه‌های واژه‌ای می‌توان به امکان تهیه سریع به کمک رایانه، عدم نیاز به تصمیم‌گیری موردی، یکدستی کار و سلیقه‌ای نبودن، ارزانی، طبیعی بودن زبان نمایه‌سازی، راحتی تهیه در هم‌کرد و وجود نوعی هم‌ارایی به دلیل حفظ بافت طبیعی عنوان اشاره کرد. عیب این نوع نمایه‌سازی آن است که همیشه عنوان، گویای موضوع مدرک نیست؛ عیب دیگر، نبود کنترل واژگانی و پراکندگی اطلاعات یک موضوع ذیل مدخل‌های مترادف و شبه مترادف، بدون وجود ارجاعات لازم است؛ و عیب سوم تلقی موضوعات هم املا به عنوان موضوع واحد است. به هر حال میزان مفید بودن این نمایه‌ها به واقع نما بودن و روشنی و معنا دار بودن واژه‌های عنوان بستگی دارد. نمایه‌سازی واژه‌ای معمولاً برای موادی نظیر کتاب و مقاله کاربرد دارد و به کار نمایه‌سازی مواد دیداری، اشیای هنری، و امثال آن نمی‌آید. نمایه‌سازی انواع اخیر معمولاً از نوع مفهومی است.

۲- **نمایه‌سازی مفهومی.** نمایه‌سازی مفهومی فرایندی پیچیده است و تشریح آن به سادگی میسر نیست. شیوه‌ای که توسط انسان به کار گرفته می‌شود تا موضوعات، ویژگی‌ها، معانی، و یا کاربردهای ممکن متون زبانی، تصاویر، و سایر

انواع پیام‌ها را درک کند و بشناسد، ریشه در فرایند شناختی ذهن و فکر انسان دارد. این فرایند به خوبی شناخته نشده و در عرصه‌هایی چون روان‌شناسی شناخت، علوم‌شناختی، و هوش مصنوعی موضوعی مهم به حساب می‌آید. همین اندازه می‌دانیم که این فرایند حداقل از دو مرحله تشکیل شده، شناخت موضوع و ویژگی‌های آن، و توصیف آنها در قالب واژگانی مناسب برای بازیابی. فرایند مذکور را می‌توان به مراحل جزئی‌تر ذیل تفکیک کرد:

- **تدوین سیاست‌ها و راهبردهای کلی.** قبل از شروع به کار، نمایه‌ساز باید با نگاه به تمامی مجموعه‌ای که باید نمایه شود، در مورد این مسائل تصمیم‌گیری کند: نوع نمایه (به‌طور مثال دستی یا ماشینی بودن یا نحوه انتشار و ارائه آن)، ساختار مدخل‌ها (مثلاً استفاده یا عدم استفاده از بیانگرها و شکل آنها)، میزان اطلاعاتی که قابلیت مدخل شدن را ایجاد می‌کند (عمق نمایه)، واحد سندی‌ای که جای‌نماها به آن رهنمون می‌شوند (پاراگراف، صفحه، کل مدرک، و جز آن)، نحوه کنترل واژگانی، نحوه همارایی، مسئله ارجاعات، و سایر سیاست‌ها که در تدوین نمایه ضروری است.

سه عامل مهم در همه این تصمیم‌گیری‌ها مؤثر است: اول ماهیت مجموعه یا نوع موادی که باید نمایه شود، به‌طور مثال تجانس یا عدم تجانس آنها، فنی یا غیرفنی بودن منابع، یک زبانه یا چندزبانه بودن آنها؛ دوم ویژگی‌های استفاده‌کنندگان است، نمایه‌ساز بداند که آیا استفاده‌کنندگان از مجموعه و نمایه، گروهی متجانس هستند یا نه، جست‌وجوگران موردی‌اند یا پژوهشگران حرفه‌ای، کم سوادند یا تحصیل‌کرده، متخصص هستند یا غیرمتخصص؛ و سوم شرایط و محیط فیزیکی است، مثلاً تعیین اینکه آیا نمایه به‌صورت مستقیم و به همراه منبع اصلی (مثل نمایه انتهای کتاب) و یا جداگانه منتشر می‌شود، و یا اینکه آیا نمایه در دست تهیه مستقل و خودکفاست یا بخشی از یک نمایه دنباله‌دار است که بعداً با سایر قسمت‌ها در هم‌مکرد می‌شود، و آیا به‌صورت دستی ذخیره و بازیابی می‌شود یا به‌صورت ماشینی. همچنین قبل از شروع باید درباره آنچه که نباید نمایه شود (مانند تبلیغات، یادداشت‌ها، و منابع و مآخذ)؛ یا آنچه که می‌تواند به‌صورت گذرا و سطحی نمایه شود (مانند نامه‌ها، یادداشت‌های کوتاه، پانویس‌ها، و مانند آن)؛ و آنچه که در مورد آن نیاز به استفاده از واژگان فنی تخصصی و یا واژگان سایر زبان‌ها وجود دارد، تصمیم‌گیری شود. پس از پاسخگویی به چنین سؤالاتی، طراحی کلی و معیارهای مهم ساخت نمایه تدوین شده و می‌توان کار تهیه نمایه را آغاز کرد. بهتر است اگر نمایه به‌صورت نمایش دادنی منتشر می‌شود، این سیاست‌ها در مقدمه نمایه بیاید.

- **آشنایی و شناخت.** نمایه‌ساز باید تصویر روشنی از محتوای اطلاعاتی مدرک به‌دست آورد. به این منظور مطالعه عنوان، رئوس مطالب، چکیده، نتیجه‌گیری، پیشنهادات، و یا تمام متن (بسته به عمق نمایه) ضروری است. از سوی دیگر نمایه‌ساز باید با نیازهای جامعه استفاده‌کننده (بالقوه و بالفعل) آشنا باشد تا بتواند تمامی راه‌های دستیابی موضوعی به مدرک را در نمایه پیش‌بینی کند. داشتن دید هدفمند اما بی‌طرف، به همراه توجه به نیازهای احتمالی جامعه، برای نمایه‌سازی ضروری است.

از این منظر نمایه‌ساز باید توجه کند که استفاده‌کننده از نمایه ممکن است با مخاطبانی که اصل اثر برای آنها نگاشته شده کاملاً متفاوت باشد، مثلاً اگر قطعاتی از یک متن فلسفی به‌عنوان گزیده نثر در یک کتاب آیین‌نگارش گنجانیده شده، در این حالت نمایه‌ساز متن را از بُعد آموزش متن نیز مورد توجه قرار می‌دهد و سعی در نمایش این ویژگی‌ها در نمایه دارد.

- **تحلیل محتوا و انتخاب اولیه مفاهیم.** ارائه تعریفی جامع از "مفهوم" در نمایه‌سازی مفهومی، مشکل است. در تحلیل محتوا شاید بتوان چکیده یک فکر یا خمیر مایه موجود در مدرک و یا روابط موجود میان اشیا یا رخدادها و سایر پدیده‌ها را "مفهوم" نامید و یا ممکن است به دسته‌ها و یا رده‌هایی که این پدیده‌ها به آنها تعلق دارند مفهوم اطلاق کرد. در این مرحله نمایه‌ساز باید در مورد موضوعاتی تصمیم‌گیری کند که از نظر نمایه‌سازی دارای اهمیت است. می‌توان اطلاعات هر مدرک را گروه‌بندی و هر گروه را در قالب یک عبارت کوتاه یادداشت کرد، تقریباً همه صاحب‌نظرانی که درباره تحلیل محتوای مدرک سخن گفته‌اند، بر اهمیت نقش تجربه در این امر تأکید دارند. طبیعی است که موضوعات اصلی باید مدخل قرار گیرند، اما در مورد موضوعات فرعی و اینکه تا چه حد در نمایه حضور یابند باید تصمیم‌گیری کرد. بعضی از مراکز نمایه‌سازی، سیاست خاصی در مورد نحوه انتخاب موضوع، تدوین و به نمایه‌سازان ارائه می‌کنند، مثلاً به نمایه‌سازان "چکیده‌نامه شیمی" گفته شده هر نوع روش اندازه‌گیری، ابزار کار، پیشنهاد و نظریه جدید، و تمام ترکیبات جدید شیمیایی که در متن آمده باید نمایه شود.

استخراج مفاهیم و واژه‌ها از یک متن و انتساب آن به مدرک بر مبنای اصول معینی صورت می‌گیرد: **الف) نیاز استفاده‌کننده،** یعنی آشنایی نمایه‌ساز با نیازهای اطلاعاتی استفاده‌کننده بالفعل و بالقوه و همچنین واژگان رایج در حوزه مورد نمایه‌سازی حائز اهمیت است؛ **ب) حفظ جامعیت نمایه،** بدین معنا که در نمایه‌سازی سعی بر آن است که تمام مفاهیم اصلی موجود در مدرک که دارای ارزش اطلاعاتی است به‌عنوان شناسه در نمایه گنجانده شود. انتخاب مفاهیم فرعی که در مدرک به‌طور جانبی به آن پرداخته شده، سبب حجیم شدن نمایه و انحراف کاربر از مفاهیم اصلی است. در این زمینه توجه به عمق نمایه که در سیاست‌گذاری‌های اولیه انجام می‌شود می‌تواند راهگشا باشد؛ و **ج) اخص بودن کلیدواژه،** بدین معنا که واژه‌هایی که به‌عنوان توصیفگر انتخاب می‌شوند باید اخص‌ترین واژه‌های ممکن برای توصیف موضوع باشد، مثلاً اگر مدرکی راجع به "مدارس ابتدایی" است، اختصاص توصیفگر کلی مدارس به آن خلاف این اصل است. البته با اخص شدن موضوع، میزان بازیابی کاهش می‌یابد، با این حال استانداردهای نمایه‌سازی به‌کارگیری واژه‌های اخص را به‌عنوان توصیفگر توصیه می‌کند، چرا که اگر کنترل واژگانی به خوبی صورت گرفته باشد، در صورت نیاز امکان گسترش و عمومی کردن جست‌وجو وجود دارد. در نظام‌های نمایه‌سازی واژه‌ای، مانند بیشتر شیوه‌های نمایه‌سازی خودکار که از کلیدواژه‌های متن استفاده می‌کنند، امکان انتخابی در این زمینه وجود ندارد، اما نظام‌های نمایه‌سازی که ضمن تحلیل موضوع، کلیدواژه به آن منتسب می‌کنند، باید این اصل را مدنظر داشته باشند.

- **ارزیابی و انتخاب نهایی مفاهیم.** در این مرحله نمایه‌ساز سعی می‌کند همه عناصری را که ممکن است استفاده‌کنندگان مایل باشند در جست‌وجو بازیابی کنند، از میان مفاهیم انتخاب شده در مرحله قبل گزینش کند و موارد بی‌ربط یا کم ربط از نظر اطلاعات را حذف کند. در این مرحله نمایه‌ساز تصمیم می‌گیرد که یک پاره اطلاع تا چه حد می‌تواند برای کاربر مفید باشد و حذف آن تا چه میزان از جامعیت نمایه می‌کاهد.

تصمیم‌گیری در این مورد غالباً کاری شاق است و ممکن است در سطوح مختلف کار بارها و بارها تکرار شود، نمایه‌ساز در این مرحله در سطحی خاص‌تر تصمیم‌گیری می‌کند، او ضمن حرکت از بخشی از مجموعه به بخشی دیگر، از موردی به مورد دیگر، و یا از پاراگرافی به پاراگراف دیگر باید تصمیم بگیرد که اصولاً مطلبی در نمایه بیاید یا نه و اگر بیاید در قالب چه توصیفگری.

مسئله عمق نمایه‌سازی، میزان ریز شدن و اخص بودن پیام‌هایی که باید در نمایه بیاید، نوع توصیفگر با توجه به سیاست‌گذاری‌های اولیه، و تعداد واژه‌هایی که هر مطلب باید با آن بیان شود، در اینجا مورد توجه قرار می‌گیرد. نمایه‌ساز برای تصمیم‌گیری صحیح باید همواره توجه کند که اولاً کاربر برای یافتن مطلبی خاص احتمالاً از چه واژه‌هایی استفاده خواهد کرد؛ ثانیاً آیا هنگامی که کاربر ذیل یک مدخل منتخب از نمایه به دنبال اطلاعات موردنظر می‌گردد، آیا اطلاعات موردنیاز برایش در نمایه گنجانده شده است. توجه به جدید و مفید بودن اطلاع برای گروه‌های بزرگی از استفاده‌کنندگان، و تخصصی یا عمومی بودن آن در حوزه مورد نمایه‌سازی، از ملاک‌های دیگر ارزیابی و تصمیم‌گیری است.

- **تبدیل مفاهیم انتخاب شده به اصطلاحات و توصیفگرهای نمایه و گنجاندن جای نماها.**
پس از انتخاب نهایی مفاهیم، برای تبدیل این مفاهیم به توصیفگرهای نمایه مسائل متعددی مطرح است که عمده‌ترین آنها عبارتند از: الف) نحوه کنترل واژگانی به‌منظور رعایت اصل وحدت [۲۶] در نمایه‌سازی؛ و ب) نحوه هم‌ارایی شناسه‌ها یا نظام‌های پیش‌همارا و پس‌همارا، که هر یک در ادامه مورد بحث قرار خواهد گرفت.

- **سایر مراحل.** در اینجا سه مرحله باید طی شود: الف) مرتب کردن مدخل‌های نمایه بر اساس نظام پیش‌بینی شده مثلاً به‌صورت الفبایی، تاریخی، رده‌ای، و یا هر نظم شناخته‌شده دیگر؛ ب) مقایسه و ویرایش مدخل‌ها و مشخص کردن روابط میان آنها؛ و ج) مرتب کردن اطلاعات توصیفی و برقراری ارتباط لازم بین بخش توصیفی و تحلیلی (در نمایه‌هایی که مستقل از متن اصلی تهیه می‌شوند).